Knowing Where You're Going: Information Systems for Agricultural Research Management

Richard Vernon

About ISNAR (www.isnar.cgiar.org)

Mission

The International Service for National Agricultural Research (ISNAR) assists developing countries in improving the performance of their national agricultural research systems and organizations. It does this by promoting appropriate agricultural research policies, sustainable research institutions, and improved research management. ISNAR's services to national research are ultimately intended to benefit producers and consumers in developing countries and to safeguard the natural environment for future generations.

Impact

To maximize the impact of its work in developing countries, ISNAR focuses on three objectives:

- enhancing the capacity of agricultural research organizations to respond to their clients' needs and to emerging challenges
- expanding global knowledge on agricultural research policy, organization, and management
- improving developing countries' access to knowledge on agricultural research policy, organization, and management

Background

ISNAR was established in 1979 by the Consultative Group on International Agricultural Research (CGIAR), on the basis of recommendations from an international task force. It began operating its headquarters in The Hague, the Netherlands, on September 1, 1980.

ISNAR is a nonprofit autonomous institution, international in character and apolitical in its management, staffing, and operations. It is financially supported by a number of the members of the CGIAR, an informal group of donors that includes countries, development banks, international organizations, and foundations. Of the 16 centers in the CGIAR system of international centers, ISNAR is the only one that focuses specifically on institutional development within national agricultural research systems.

About CTA (www.agricta.org)

The Technical Centre for Agricultural and Rural Cooperation (CTA) was established in 1983 under the Lomé Convention between the ACP (African, Caribbean, and Pacific) Group of States and the European Union Member States. Since 2000 it has operated within the framework of the ACP-EC Cotonou Agreement.

CTA's tasks are to develop and provide services that improve access to information for agricultural and rural development, and to strengthen the capacity of ACP countries to produce, acquire, exchange, and utilize information in this area. CTA's programs are organized around four principal themes: developing information management and partnership strategies needed for policy formulation and implementation; promoting contact and exchange of experience; providing ACP partners with information on demand; and strengthening their information and communication capacities.

Knowing Where You're Going: Information Systems for Agricultural Research Management

Richard Vernon

November 2001

ISNAR appreciates the support from the Technical Centre for Agricultural and Rural Cooperation (CTA) in the production of this publication.

About the Author

Richard Vernon worked for 13 years as an agricultural research scientist and research manager in the tropics. He then returned to college to take up formal studies in information systems. For the past 16 years he has worked on information system development and training within agricultural development programs. This includes eight years with ISNAR. He currently leads a five-person Information and Extension team in the Plant Protection Services program of the Secretariat of the Pacific Community.

Citation

Vernon, R. 2001. Knowing Where You're Going: Information Systems for Agricultural Research Management. The Hague: International Service for National Agricultural Research.

AGROVOC descriptors

agriculture; case studies; management information systems; research institutions

CABI descriptors

agricultural research; case studies; information systems management; research institutes

ISBN 92-9118-054-8

Contents

Part 2. For the Information Manager

Annexes

List of figures

List of tables

Contributors

Allmand, Mónica
Librarian
ISNAR
P.O. Box 93375, 2509 AJ The Hague, The Netherlands
E-mail: m.allmand@cgiar.org

Baguma, Dickson
National Agricultural Research Organization (NARO)
P.O. Box 295, Entebbe, Uganda
Tel: +256-42-20341, or 20512. Fax: +256-42-21070

Ballantyne, Peter
Manager, Knowledge Sharing
International Institute for Communication and Development (IICD)
P.O. Box 11586, 2502 AN The Hague, The Netherlands
E-mail: pballantyne@iicd.org

Besemer, Hugo
Wageningen Agricultural University (Library WAU/PUDOC-DLO)
P.O. Box 9100, 6700 HA Wageningen, The Netherlands
E-mail: hugo.besemer@id.bib.wau.nl

Brush, Ted
Consultant
P.O. Box 785, 22503 Lancaster, Virginia, USA
E-mail: brush@rivnet.net

Devaux, André
Coordinador Proyecto Papa Andina (CIP-COSUDE)
International Potato Center—Centro Internacional de la Papa (CIP)
Apartado 1558, Lima 12, Peru
Tel: +51-1-349 6017. Fax: +51-1-3175326
E-mail: a.devaux@cgiar.org

Fraser, Bruce
ISNAR
Financial Manager
P.O. Box 93375, 2509 AJ The Hague, The Netherlands
E-mail: b.fraser@cgiar.org

Gijsbers, Govert
Research Officer
ISNAR
P.O. Box 93375, 2509 AJ The Hague, The Netherlands
E-mail: g.gijsbers@cgiar.org

Gomes de Castro, Antônio Maria
Researcher
EMBRAPA-SEDE (Embresa Brasileira de Perquisa Agropecuaria)
P.O. Box 040315, 70770-901 Brasilia, D.F. Brazil
E-mail: amcastro@sede.embrapa.br

Hartkamp, A. Dewi
Formerly with Natural Resources Group, CIMMYT (Centro Internacional de Mejoramiento de Maiz y Trigo).
Current address: Oude Houtensepad 18, 3582 CW Utrecht, The Netherlands
E-mail: d.hartkamp@wisint.org

Hobbs, Huntingdon
Director, ISNAR Global Associates Program
c/o IICA, Apdo. 55-2200, Coronado, Costa Rica
E-mail: h.hobbs@cgiar.org

Hodson, David P.
Natural Resources Group
CIMMYT (Centro Internacional de Mejoramiento de Maiz y Trigo)
Lisboa 27, Apartado Postal 6-641, 06600 Mexico, D.F. Mexico
E-mail: d.hodson@cgiar.org

Kisauzi, Dan
Formerly with National Agricultural Research Organization (NARO), Uganda
Current tel. no: +256 77 408593. Fax: +256 41 342082
E-mail: buddies@swiftuganda.com

Lima, Suzana Maria Valle
Researcher
EMBRAPA-SEDE (Embresa Brasileira de Perquisa Agropecuaria)
P.O. Box 040315, 70770-901 Brasilia D.F., Brazil
E-mail: svlima@sede.embrapa.br

Mook, Byron
Senior Research Officer
ISNAR
P.O. Box 93375, 2509 AJ The Hague, The Netherlands
E-mail: b.mook@cgiar.org

Ngwira, Margaret
Senior Librarian
University of Namibia, Faculty of Agriculture and Natural Resources
P. Bag 13301,Windhoek, Namibia
E-mail: mengwira@unam.na

Niang, Thiendou
CTA (Technical Centre for Agricultural and Rural Cooperation)
P.O. Box 380, 6700 AJ Wageningen, The Netherlands
E-mail: niang@cta.nl

Parker, David
Biometrician
Pfizer Limited
Ramsgate Road, Sandwich, Kent CT13 9NJ, United Kingdom
E-mail: david_parker@sandwich.pfizer.com

Pedroso, Moacia
Head of Center of Information Technology for Agriculture
EMBRAPA-SEDE (Embresa Brasileira de Perquisa Agropecuaria)
P.O. Box 040315, 70770-901 Brasilia, D.F. Brazil
E-mail: pedroso@cnptia.embrapa.br

Philpot, Paul
Multimedia Consultant
International Institute of Tropical Agriculture (IITA)
PMB 5320, Oyo Road, Ibadan, Nigeria
E-mail: p.philpot@cgiar.org

Rodriguez-Aguila, Adriana
Natural Resources Group
CIMMYT (Centro Internacional de Mejoramiento de Maiz y Trigo)
Lisboa 27, Apartado Postal 6-641, 06600 Mexico, D.F. Mexico
E-mail: a.d.rodriguez@cgiar.org

Setsoafia, Kwasi M.
CSIR (Council for Scientific and Industrial Research)
P.O. Box M.32, Accra, Ghana
Fax: +233-21-779809

Webber, Hope
Research Analyst
ISNAR
P.O. Box 93375, 2509 AJ The Hague, The Netherlands
E-mail: h.webber@cgiar.org

White, Jeffrey W.
Natural Resources Group
CIMMYT (Centro Internacional de Mejoramiento de Maiz y Trigo)
Lisboa 27, Apdo. Postal 6-641, 06600 Mexico, D.F., MEXICO
E-mail: j.white@cgiar.org

Preface

The objective of this book is to provide agricultural research managers at all levels and information specialists within agricultural research organizations with a source of ideas, concepts, methodologies, explanations, and guidance in managing information.

Agricultural research successfully identified new technologies for farmers long before computers were introduced. Research managers in many countries were served well enough with information on the research program to manage it effectively. Scientists were able to check previous research from research journals: agricultural research libraries took a wide range of journal titles supporting all the major disciplines pursued within the research program. Regular annual reports provided information on recent and past local research. A large amount of information was therefore processed, mostly on paper, in the business of conducting and managing agricultural research.

A number of things have happened in recent decades that have increased the information management load. Agricultural research organizations, particularly in developing countries, have expanded their scientific workforce and organizational complexity. They are under increased pressure for new food-production technologies to meet the increased demand for food for expanding populations. At the same time they are being called on to broaden their scope from traditional preoccupation with yields to issues of sustainability and natural resource management. They are also facing demands for increased reporting requirements from donors. And yet the research organizations have to meet these demands for ever increasing management with diminishing budgets—average spending per researcher in the developing countries has been declining since the early 1970s. Pressure for greater efficiency in the management of limited resources is therefore severe.

Information is crucial to enhanced efficiency, in particular in the following two areas. The first is information on all aspects of the research program, such as resource allocation among commodities, disciplines, and agroecological regions, comparisons of the research program with agreed priorities, ensuring that the strength and specialisms of the research cadre are in line with the needs of the research program and with the problems of the farming community. The second area is scientific information: details of research that has been done before and research in progress elsewhere. Good information helps avoid unnecessary duplication and allows new research to proceed with maximum advantage from what has already been accomplished. Keeping up with current information is not a trivial task; it is estimated that the amount of published scientific information approximately doubles every 15 years.

A serious side effect of the deteriorating budgets of many research services in recent years has been a severe reduction in access to scientific information as library funds —and hence subscriptions to scientific journals—have been among the first cost centers targeted by budget cuts. Ironically, it seems as though libraries were some kind of luxury item when in fact they provide information vital to efficient research. On the other hand, the last 20 years have seen extraordinary gains in those technologies that store, process, and communicate information. Capacities and speeds of computers are increasing rapidly while prices are falling. It is a remarkable situation, but it offers hope to the otherwise rather bleak perspective of the preceding paragraphs.

The management of information and information technology has proved to be difficult and complicated, and there have been some spectacular failures. But lessons have been steadily learned. This book intends to provide guidance to those in agricultural research management setting out to harness the potentially enormous benefits of information and communications technology, to minimize the risks and maximize the benefits. It draws on the experience of many information system implementations and implementers, both within and beyond agricultural research.

The book has two parts. The first is primarily for the research manager. It explains what management information systems are, what they offer, and how to get one, and it discusses information management and strategy. The second part is primarily for information managers—those responsible for developing and managing management information systems. However, it also contains material that scientists and research managers with an interest in information and communications technologies may find useful. The annexes provide information on several related topics.

The book is illustrated with several examples of outputs and reports generated by a management information system entitled INFORM-R, which was developed by ISNAR for its partner agricultural research organizations. It is also used to illustrate the structure of a management information system in the annexes. Most of the case studies in chapter 3 are based on an earlier version of this system. We have used INFORM-R out of convenience and do not mean to imply that this system is superior to any other. We do believe, however, that there is a strong argument for taking on, and adapting as necessary, a system that already exists, complete with its inevitable faults, rather than attempt to build one from scratch. This is particularly the case where such a product has received the inputs from many research managers in many countries over many years. It would be difficult for an institution to gather that collection of experience for itself. In addition, a common standard facilitates the exchange of research information between research institutions. The development and mass marketing of a standard personal computer by IBM in the early 1980s and a standard operating system by Microsoft have, despite the criticism of the commercial exploitation of a near-monopoly situation by each company, greatly facilitated the spread of information technologies.

Acknowledgments

A book of this nature is necessarily the product of the ideas of many. First, I wish to acknowledge a special debt to the scientists, research managers, and development workers in many countries with whom I have been working and who have collaborated in designing, developing, implementing, and using management information systems. This has proved a rich source of experience and knowledge.

The contributing authors, Monica Allmand, Dickson Baguma, Peter Ballantyne, Hugo Besemer, Ted Brush, André Devaux, Govert Gijsbers, Bruce Fraser, A. Dewi Hartkamp, Hunt Hobbs, David P. Hodson, Antônio Maria Gomes de Castro, Dan Kisauzi, Suzana Maria Valle Lima, Byron Mook, Thiendou Niang, Margaret Ngwira, David Parker, Moacia Pedroso, Paul Philpot, Adriana Rodriguez-Aguilar, Kwasi M. Setsoafia, Hope Webber, and Jeffrey W. White, bring valuable expertise and experience from their specialist areas.

ISNAR's work on its original management information system, INFORM, provided a sound basis and much original thinking from which the current model, INFORM-R, was developed. The original system was developed by Barry Nestel and Govert Gijsbers, with assistance from Byron Mook and in collaboration with Asian agricultural research systems, particularly in Sri Lanka and Indonesia. The development of INFORM-R and, consequently, this book, would not have occurred without the inspiration of Paul Perrault, former director of ISNAR's management program. The quality of INFORM-R owes much to the professional skill and the dedication of Peter van Veen of Blue Line Development consultancy. Hope Webber has explored the prospects for integrating parts of the data sets from different countries into, for example, a regional information resource, and how the Internet might be harnessed to provide access to such a resource. He has also been instrumental in the development of the smaller version named INFORM-R Light and making both versions available on the Internet.

I am indebted to my colleagues Rudolf Contant and Larry Zuidema as well as Janice Reid, member of ISNAR's Board of Trustees, for their useful comments on some sections. Peter Ballantyne provided valuable comment, and he, Barry Nestel, and M. Sompo-Ceesay helped restructure much of the original draft into what is hopefully a much more readable and useful text. Barry Nestel has greatly enriched the text with numerous and valuable ideas and constructive criticisms from a lifetime's experience of agricultural research and management. Jörg Edsen has been meticulous in reviewing the draft text and adding ideas and support from his own field experience in implementing a management information system.

Hugo Besemer, Mark van Meulen, Mike Jensen, and Peter van Veen—specialists in the fast-moving field of information and communications technology—provided valuable scrutiny of and suggestions for chapter 7 "Aspects of Information Science and Technology." The layout and design of this book has been a large and complex task and is credited to the skill, patience, and capacity for hard work of Jan van

Dongen, whom I was fortunate to have as my editor. Richard Claase designed the cover.

Finally I would like to put on record the unstinting patience and support of my wife, Johanna Vernon-de Jong, throughout the gestation of this book, which made its preparation so much easier.

Richard Vernon

The Hague, November 2001

Acronyms

ACP	Africa, Caribbean and Pacific States
AGRIS	Agricultural Research Information System (FAO)
ASARECA	Association for Strengthening Agricultural Research in Eastern and Central Africa
CABI	Centres for Applied BioSciences – International
CARDI	Caribbean Agricultural Research and Development Institute
CARIS	Current Agricultural Research Information System (FAO)
CIRAD	Centre de coopération internationale en recherche agronomique pour le développement
CGIAR	Consultative Group on International Agricultural Research
CRIS	Current Research Information System (USDA)
CTA	Technical Centre for Agricultural and Rural Cooperation
ECAPAPA	Eastern and Central Africa Programme for Agricultural Policy Analysis
EPMR	external program and management review
FAO	Food and Agriculture Organization of the United Nations
ICRAF	International Centre for Research in Agroforestry
ICT	information and communication technology
ILRI	International Livestock Research Institute
IPGRI	International Plant Genetic Resources Institute
ISNAR	International Service for National Agricultural Research
MIS	management information system
NARI	national agricultural research institute
NARO	national agricultural research organization
NARS	national agricultural research system(s)
NGO	nongovernmental organization
PC	personal computer
PM&E	planning, monitoring, and evaluation
PRAP	Pacific Regional Agricultural Program
REPA	Réseau d’Etude et d’Analyse des Politiques Agricoles
SDI	Selective Dissemination of Information
SPAAR	Special Program for African Agricultural Research (World Bank, Washington)
SPC	South Pacific Commission
USDA	US Department of Agriculture
WARDA	West Africa Rice Development Association

Executive Summary

In the present economic climate, research managers are increasingly faced with the challenge of raising the efficiency of agricultural research to maximize the output of enhanced technology to farmers. To make the correct decisions, they must have good information about the research systems they manage: inputs to research (e.g., human, financial, and physical resources), research programs (experiments and studies), and research outputs (improved technologies and reports on experiments and their impact). This information must be easily accessible, up to date, and accurate. It should also be available in various forms so that it can be used to serve different management issues. This is the function of a management information system (MIS).

What is an MIS?

An MIS is a system using formalized procedures to provide management at all levels in all functions with appropriate information, based on data from internal and, if desired, external sources to enable to make timely and effective decisions for planning, directing, and controlling the activities for which they are responsible. It should be easy to use even by people with few or no prior computer skills—learning to use it should take no more than two days.

How can an MIS help managers?

An MIS for agricultural research provides easily accessible information for the various stages of the agricultural research management cycle and its major processes: matching the program with agreed priorities, planning, budgeting, monitoring, reporting, and evaluation. It can present any of a collection of standard reports on screen or in print.

At the national level, these reports include standard lists of all experiments and all scientists, sorted for example by ID number, station, main commodity, or scientific discipline. While national agricultural research organizations (NAROs) may already have the capability to produce such reports manually, it may take days or weeks to prepare just one, often at considerable expense of time by one or more individuals. An MIS can produce these reports in minutes at a few key strokes.

So greater efficiency is a major advantage of an MIS. Core data, for example on research activities and scientists, has to be entered into the system only once, serving a wide range of subsequent information needs that otherwise might be pursued with much duplication of effort, delay, and missing items. For example, figures 1 and 2 both use the same scientists' time information as shown in table 1, but each uses this information for different purposes.

With just a glance at figure 1, a manager can see the differences in resource use between the main research disciplines. The graph also compares this information for the current and previous years so that trends can be seen.

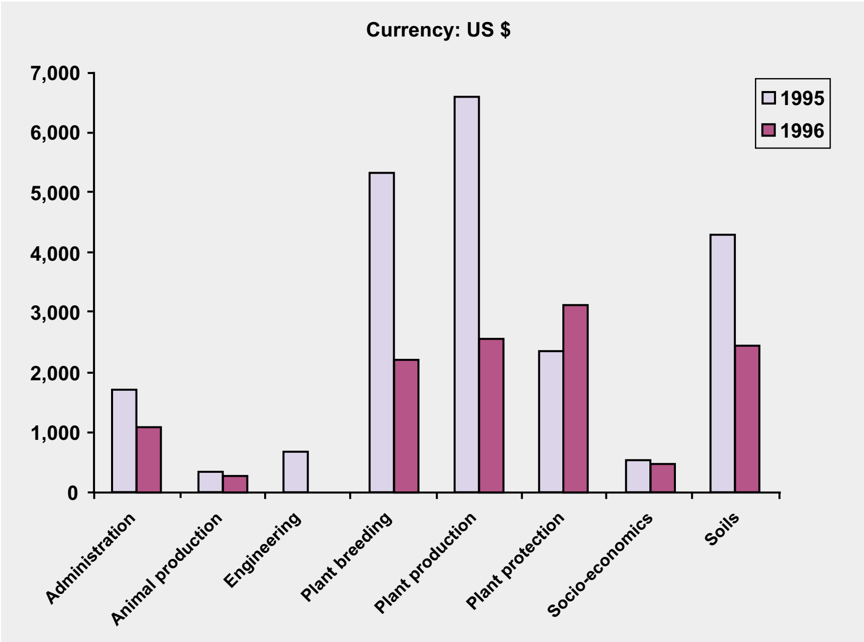


Figure 1: MIS graphical output: resource allocation between the main scientific disciplines

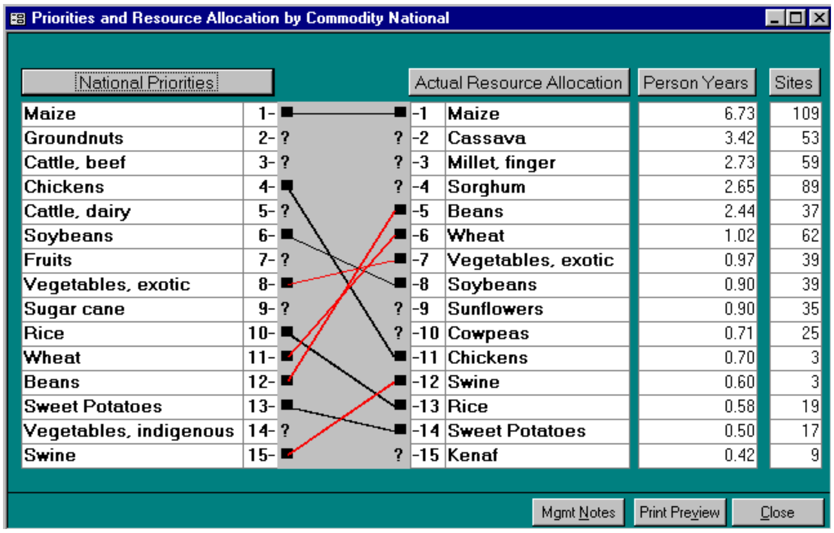


Figure 2: Comparison of national priorities and resource allocation to commodities

Managers may be able to access data on different levels of research activities through a hierarchy of experiments, projects, and programs. Details of research stations, their mandate, soils, and size are all readily available, as is information on external factors that could influence the research program.

The range of reports available from an MIS vary and depend on the expressed needs of management in each country or NARO. But the standard information it might contain should enable a wide range of questions to be directly answered:

- What proportion of the research program is allocated to each commodity? What proportion is allocated to each station and to each agroecological region? What proportion has nutrition (or pathology, or soils, etc.) as its main discipline?
- What proportion of a scientist's time is spent on research, management, extension, and training? What proportion is spent on specific research activities?
- How many researchers have BSc, MSc, or PhD degrees? How many fall into each age group, gender, or grade level?
- What proportion of the research budget is assigned to scientists' salaries, labor, or transport? How much is from national funds and from donor funds? How does this vary between forestry, crops, and livestock research?

Without an MIS, some of the information necessary to answer such questions may exist, but is often in several different places and formats. With an MIS, answers are easily depicted in graphical, tabular, or pie-chart form. The MIS aggregates the information, together with new data collected for the purpose, and then generates information in an easily usable form.

For users with a little further training, the system can also provide answers to a wide range of unforeseen, ad hoc questions that may arise in the day-to-day running of a research program and that cannot always be anticipated. If any such question is likely to be asked again, then the report can be stored in the system so that it is available when needed but always updated with the latest information. An MIS stores a report in its "logical" form only, that is with just the report structure; each time the report is called, a fresh copy is generated using the latest information that has been keyed into the system.

This wide range of facilities, coupled with ease of use, requires the use of modern database software and skilled system implementation, but once this investment has been made, managers have a rich support tool at their disposal.

Further MIS outputs used in research management

An MIS should be able to display resource allocation to main commodities and a comparison of such allocation with previously set priorities. The left column of figure 2 shows the relative national priority accorded to the highest priority commodities. On the right is MIS-produced data on actual allocation of research resources to these commodities. In the above case, resource allocations diverge considerably from national priorities. The research manager's task is to explore the reasons for such differences and, if necessary, to adjust the research program, usually over some years, to bring it closer in line with the agreed priorities.

Table 1 shows the main types of personnel information likely to be covered by a research program MIS. Table 1b shows on the left a scientist's estimates of time allocated to the main functions of research, management, training, study, and others. On the right of table 1b, the time attributed to research (80% of one person year in this example) is disaggregated further into the time spent on individual experiments. Using the formula of 1 week = 2% of a person year, such attribution can be done easily and with sufficient accuracy for management purposes. Salary information from the payroll may then be used to convert scientists' time to costs.

Managers can use the personal details as in table 1a to compare scientists' disciplines with the disciplinary requirements of the research program. One step in this process could be to generate a table, such as table 2, which shows the current and projected complement of scientists by discipline.

The deployment of research personnel can be assessed across commodities and disciplines. Date-of-birth data can be used to show the future effects of retirement, and recruitment and training schemes can be adjusted accordingly.

Finally, the system may address project budgeting in which budget figures can be entered for each project. These can be aggregated by the system to show composite figures for each program, station, commodity, science discipline, or agroecological zone.

Table 1: Examples of Scientists' Data Stored in an MIS for the Research Program

a. General Data

Personal	Admin	Scientific
Name	ID number	Highest degree
Date of birth	Rank	Scientific discipline
Gender	Job title	Main species (crop or animal) or noncommodity factor

b. Time Allocation Data, and Links to Research Activities

Percentage time allocations	%*	Research time breakdown	%*
Research	80	Experiment / study A	40
Management / Admin.	10	Experiment / study B	20
Extension	10	Experiment / study C	10
Training		Experiment / study D	10
Degree studies		Experiment / study E	
Other		Experiment / study F	
		Experiment / study G	
Total time	**100**		**80**

* % person year, where 1 week = c. 2%

Table 2: Using an MIS for Human Resources Planning: Current and Projected Numbers of PhD Holders in Each Discipline

Discipline	Current no. of PhDs	Retiring by year 2002	Currently in training	Potential no. in year 2002	% change
Plant production	12	3	4	13	8
Plant protection	8	1	3	10	25
Soils	7	2	4	9	29
Animal health	5	1	0	4	-20
Plant breeding	5	2	2	5	0
Socioeconomics	3	0	3	6	100
Post harvest	2	0	0	2	0
Animal production	2	1	3	4	100
All	**44**	**10**	**19**	**53**	**20**

The project life cycle

An MIS for agricultural research may cater for the life cycle of research activities, storing data related to stages before, during, and after the implementation of a research activity. It stores the research protocol or proposal, produces reports for research planning meetings, and records the appraisal of such a proposal. Approved proposals are then "rolled over" into research activities. The system can support the planning process by listing the proposals in various orders—e.g., by commodity, research station, or agroecological zone—and by displaying features of the existing research program.

During the active life of a research activity, the system may produce monitoring reports using a range of monitoring indicators. It should supply a base set of indicators and a facility for the organization to insert its own. More sophisticated systems can also store actual research results data. These may even be linked to a facility for the production of annual reports in a partially automated process.

On completion of an activity, the main results of the research, aggregated over all its years, can be entered into the system, providing a permanent record of research that can easily be stored, copied, and transmitted to all stations and to other interested researchers.

Implementation

The successful introduction of an MIS into a research organization depends on the following:

1. The quality of the MIS

The quality of an institution's MIS depends on a number of things: the choice of system and software, the system's development team, and the allocation of financial and human resources to the project that is to develop the system. There are several options for the choice of system and software. In order of decreasing cost, an MIS can be

- developed from scratch within the organization using a programming language
- developed from scratch using an off-the-shelf database management software
- adapted from a standard project management software
- adapted from a generalized existing MIS for agricultural research

2. The implementation process

An MIS is implemented in a sequence of activities. There are a number of decision points in this process at which it will be decided whether to proceed, change direction, or abort the project. Initial meetings with top management and institute directors are required, followed by seminars with researchers and their directors. At these meetings, the system development team demonstrates typical features of an MIS and explores the information needs of managers. It also explores and documents the current state of data on researchers, research activities, research outputs (e.g., annual reports, extension advice, research papers), budgets, and physical resources, and of any existing information systems.

A top management patron for the whole exercise has to be identified. He or she will ensure that the project and its role in supporting research management are given recognition throughout the organization, and that the necessary resources to develop and operate it are made available. Such a patron need not necessarily use the system directly but should request outputs from the system for his own use and require that these are of good quality and issued on time. In addition, it is useful to have a operational sponsor from senior management, who is an active user of the system and who monitors operational activities.

Two or three national information coordinators need to be appointed, typically members of a monitoring and planning unit at national headquarters. At each station or institute, at least two MIS practitioners will need to be appointed who will be responsible for operating the MIS. They must be selected with care and should be staff members likely to remain on post for at least the coming two years. It is best if one MIS practitioner is a senior researcher. The value of computer training declines rapidly if the newly learned skills are not used immediately, so the practitioners must have good access to a computer at the time of training.

Any additional data on research activities and scientists needed for management purposes is collected. Training is provided for the MIS practitioners at research stations in data capture, processing, and production of management reports. Senior managers and station directors are introduced to skills in using the system for their own management needs.

Follow-up visits by the trainers in the subsequent months will be needed to assist in solving problems that inevitably arise as staff take on the new system, its associated technology, and often new ways of doing things.

3. Institutionalization

Integrating an MIS into the regular research management cycle is often the most difficult and neglected stage. It calls for widespread acceptance of the system and its requirements—led from the top—and collaboration in its application. Various management initiatives can assist these behavioral practices.

The whole implementation should be monitored by agreed indicators and criteria set at the start. Table 3 provides some suggested targets in the successive stages of the implementation.

Table 3: Suggested Criteria for Assessing an MIS Implementation

Stage	Criteria
1	Existence of MIS database files for agreed subjects, e.g., scientists and experiments, for each research center and at the national level. Evidence can be provided by reports generated by the MIS itself of the numbers of records held for each subject and may be checked against existing figures. (For example, a total figure of 224 scientists can be checked against the payroll figure.)
2	Issue of some agreed standard MIS outputs by the MIS practitioners at each research center, for example: • directory of scientists (and their main discipline and commodity) • list of experiments (and the participating scientists) • resource allocation to commodities These outputs to be provided to all research managers and accessible to all scientists.
3	Institutionalization: • use of MIS outputs in research institutes and nationally, in research planning, monitoring and evaluation. • number of cycles (years or seasons) for which this has occurred

Part 1. For the Research Manager

Chapter

1 Information Management and Information Strategy

The main purpose of this chapter is to provide research managers at all levels with an overview of the principles of and options for managing information in agricultural research management.

1.1 Information management

1.1.1 Why computers?

Information can be managed entirely without the tools of pen and paper or electronic technologies. Hunter-gatherer communities manage much information about their environment in this way. Managers of a research organization, however, have to address large amounts of information in such a way that it can be shared between many staff in a short time. This is made possible by formal methods of information management. Paper has been the main tool for this purpose, but it is being increasingly challenged today by electronic media.

A legitimate question of many research managers today is why make the transition to a computerized system when good research has been conducted for many decades without it? There are clear reasons against such adoption: it brings significant new complexities, costs, and dependency on skills that most NAROs have little of and will find hard to maintain and keep once attained. And the technologies are changing so fast that the "write-off" time of investment is much shorter than those of traditional tools such as agricultural machinery and laboratory equipment.

There are a number of compelling reasons for adopting information and communication technologies (ICT). First, so many others are adopting it that those who do not, run the risk of becoming isolated. Second, there is widespread failure in information management where it relies on traditional paper-based information. There is a hope that the adoption of electronic technologies will help resolve this problem. Third, ICT can process and transmit large volumes of data and at high speed. Fourth, ICT can make and store copies of very large data sets. And finally, ICT can replicate such large volumes of data so that, for example, each research station can have a copy.

Electronic systems are proving much more powerful than paper-based systems. In his early days as a researcher, the writer progressed from a slide rule to an electronic calculator for numerical work such as statistical analyses of variation. A mechanical typewriter and carbon paper allowed only about six copies, with little allowance for error correction. Now draft copies can be e-mailed in seconds to colleagues thousands of miles away, and the comments received back can be incorporated into the text without the use of any paper. And the World Wide Web (see section 6.1.5) offers rapid access to enormous and increasing amounts of information located around the world.

However, these potential benefits remain to be realized in many developing countries. One of the goals of this book is to assist those in agricultural research management in harnessing ICT in order to overcome these difficulties in information management and achieve the benefits.

1.1.2 Managerial, policy, and technological factors

Senior managers need to address and agree on a number of policy issues before implementing a management information system (MIS). Developments particularly in hardware and communication technologies have led to dramatic rates of progress in the technologies for processing, storing, and transmitting information. And while the performance of hard- and software increases spectacularly, costs are falling. This has contributed to the spread of information technology to a large proportion of the workforce in the North. In the South, computing technology is also rapidly being introduced, but the related communication technologies lag behind, in part due to the slower pace of liberalization of the state-owned post and telephone services.

The ability to connect computers in a network throughout the organization and beyond is offering unprecedented potential for the sharing of information. This is placing much greater information processing power into the hands of each individual. A consequence is the decentralization of information management.

On the other hand, there are numerous examples where the expected benefits from ICT were not realized. In the 1990's, several high-profile and costly projects failed dramatically, and part of the blame was assigned to the technology. For example, the failure of the baggage handling system at Denver International Airport held up developments at the airport for months. A glitch in London's automated ambulance transportation system left patients waiting for about half an hour. And both the California Department of Motor Vehicles and the London Stock Exchange spent large sums of money on software that never became operational. These are examples of failures that were very expensive or even harmed people (Grinter 1996) and contrast starkly with the expectations from information technology of greater efficiency, better decision making, and increased service to customers.

Why is there such a great contrast between promise and delivery? Much of the attention on information management today focuses on the technological aspect. But the full benefits of these technological developments will be realized only by those organizations that also focus on the managerial and policy issues relating to the way information is managed. Brown and Grudney (1994) point out:

> "The benefits that advances in computer and information system technology offer to organizations in the twenty-first century seem indisputable, but whether organizations will actually achieve them seems very much in question. . . Technical ability alone cannot overcome the information systems challenges facing the public and private sectors. Organizations must surmount the managerial and policy requirements to reap the full benefits of the technological advances."

They advocate that this be addressed through a coherent philosophy for information resource management.

Information systems staff have to be knowledgeable not only about information technologies but also about the core business of the organization. In NAROs, particularly the smaller ones, information systems staff often have converted from the agricultural sciences to take up information science as a second career. While this ensures much-needed knowledge of the core business of agricultural research, they often lack higher-degree studies in information science. More seriously, this move to a nontraditional science often excludes them from a conventional career path. Management, personnel, and training officers, therefore, need to recognize that the agricultural / information science "hybrid," after some years of experience in such a position, may make him or her just as good a research manager as would a veterinarian or plant breeder.

Management also need to take decisions on who has access to which data and where a particular information system is housed and managed. In the past, these systems were usually managed centrally by specialized computing staff. In large organizations this may still be necessary. But there is evidence that in NAROs, a centralized system is too remote from the intended users (the research managers and the researchers) and consequently neglected. Instead, and in line with dissemination of information technology to users, systems should be established as close to the users as possible. For example, a personnel system, once it is set up (with assistance from staff from the information system of course), may well be kept within the personnel department. Similarly it is now quite feasible and indeed preferable to locate the operation and routine management of the research program information system near to the offices of the program leader, and the station and institute directors.

An important question that often arises in this area is whether to build a single system or a number of systems, e.g., one for research program management, one for personnel management, and one for financial accounting. While those from a technical computing background tend to advocate a single, all-embracing system, there are good reasons to establish individual systems for the major functions, while adhering to common standards, names, and software to facilitate links between them.

These and other management and technical issues are discussed in greater detail in section 4.1 "Issues in the implementation of an MIS" in this volume.

1.1.3 Information needs of managers

The following is a list of general information needs by managers, based on conclusions from a number of studies (see for example McNurlin et al. 1989; Aris 1992):

- **Summaries from the routine processes of the enterprise.** Operational staff need continuous and detailed information on the operations they are controlling. Managers need only a summary of operational data, e.g., the total output figure for the day, the week, or the department. It may also be qualitative, e.g., in monitoring quality control. An MIS must therefore accept the detailed information and aggregate it. The degree of aggregation is proportional to the level of the manager.

Managers of different areas of the enterprise need different selections of aggregated data.

- **Information on exceptional events.** For routine enterprise operations, aggregated data indicating quantities and qualities of outputs need to be supplemented by details of any exceptional occurrences. Within an MIS the boundaries of "normal" are specified, and the MIS is set to report occurrences outside these limits. (For example, an agricultural research MIS might have a report listing all those research activities for which the planned end date has passed but which are still active.)
- **Facility to find ad hoc information.** Delivery of the two sorts of information described above can and normally should be largely automated. But the MIS also needs to be able to meet specific questions that cannot be anticipated.
- **Time series information.** An MIS is largely concerned with the present and the future. But often there is value in comparing the present with the past. The MIS therefore needs to be able to store previous data and readily compare today's performance with selected periods from the past.
- **Comparative external information.** While an MIS is focused primarily on the processes within the enterprise, valuable insight can often be gained from data from other institutions in the same type of business in the same country or from other countries. This data may be of direct interest in itself, adding to the corporate knowledge of the institution, or it may be used in comparison with information from within the organization to judge performance. Such external information may help to establish "norms" for use in planning and performance assessment.
- **Contextual or environmental information.** The context or environment in which an organization operates usually has a number of effects on the organization's performance. Information from an organization's environment that can usefully influence management decision making and that is available needs to be identified, captured, and presented on a regular basis. It is important to be selective in deciding what to include in an MIS.

1.1.4 The importance of data sources

The cooperation of researchers in capturing data, for example on experiments and studies, is essential. It is best achieved by clarifying their role within an MIS and ensuring that they too will be beneficiaries of the system. At its simplest this may be by ensuring that they too receive output from the system, which is most effective if the outputs follow fairly closely the data-capture exercise—normally within a few weeks. Section 4.1.4 "Critical success factors," recognizes this as one of the key factors for successfully implementing an MIS.

1.1.5 Information systems

Most national agricultural research organizations have several information "systems," such as a library, a set of personnel files, and an accounting system, each usually under the responsibility of different people who manage it for their own needs. At a higher level the managers of an organization need a regular flow of information from the enterprise to effectively manage it. All these activities can be collectively described as information management, and many of these procedures are predictable and routine and can readily be set up in advance.

Today's computers and information technology greatly assist a manager in managing information. Computers not only take out the drudgery of repetitive tasks but can prompt organizations to adopt new management processes. An MIS generates information by collecting underlying data from various parts of the organization, processing it, and delivering meaningful information that managers can use to make decisions. But inevitably the need will arise to answer new questions and solve new problems. An MIS, therefore, needs also to be flexible and have a facility to provide ad hoc reports in which the manager selects the information provided. Today, MIS systems are so flexible that they can be easily customized to suit the user's needs. Easy-to-follow menu systems enable busy managers to learn how to use the system and get useful information from it without needing to learn a new language of special commands and procedures.

A generalized definition of an MIS is a "system using formalized procedures to provide management at all levels in all functions with appropriate information, based on data from both internal and external sources, to enable them to make timely and effective decisions for planning, directing and controlling the activities for which they are responsible" (Lucey 1991). For the purpose of this volume we narrow this down and define an MIS as "an ongoing data-collection and -analysis system, usually computerized, that provides managers with timely access to information on research inputs, activities, and outputs."

MIS have evolved into a number of models, some of which are briefly described below. While we will focus primarily on the generalized MIS in this book, it helps to be aware of the existence of other possibilities in case the need of such a system becomes apparent.

Decision support systems

McNurlin et al. (1989, p.383) have defined decision support systems (DSS) as "computer based systems that help decision makers confront ill-structured problems through direct interaction with data and analysis models." These systems are usually highly quantitative and enable users to explore the consequences of different decisions based on different incoming data. With DSS various models can be created and analyzed, for example in the form of "What if" questions. Spreadsheets provide a simple method of such modeling and analysis. Farrell et al. (1992) show how databases, deterministic simulation models, and rule-based (expert) systems may all contribute to DSS.

Executive information systems

Executive information systems (EIS) may be regarded as a specialized kind of DSS to meet the main information needs of senior management, listed by McNurlin and Sprague (1989) as follows:

- performance measures of critical factors
- descriptions of current key problems

- highlights of the things in which senior management is most interested
- detailed reports of subordinates' performance

Features typical of a good EIS include aggregates of a wide range of organization data, reports of "actual versus planned" actions, clear graphics, easy access to key internal and external information, and tailored targeting of the critical success factors of the individual executive.

Leidner et al. (1993) examined the effects of EIS use on aspects of the decision-making process by surveying 46 executive users of EIS. They found that the frequency and duration of using EIS increased the speed of problem identification and decision making and the extent of analysis in decision making.

Group decision support systems

Group decision support systems (GDSS) assist collective decision making through enhanced group collaboration and creativity, time saving, improved solutions, and social leveling. Lewis and Keleman (1990) reported several benefits from using GDSS, including a high degree of commitment for project implementation and better management of conflict. Users reported improved ability to address planning problems and greater productivity than in ordinary team working, partly through greater opportunity to voice opinions.

They also make some recommendations for developing such systems. One very important component is the user interface, which has to be simple, so that even computer novices can use the system, and it has to be transparent, so that attention is soon focused on the task in hand and not on the system. This applies equally to the generalized MIS. Built-in software timekeepers can protect participants from the danger of spending too much time in discussion and editing phases, analogous to an agenda in a well-managed meeting.

They also recommended providing training at two levels: first the GDSS itself and then the structured approaches to decision making. Groups should not be expected to learn a new tool at the same time as they are trying to resolve a difficult problem. Davenport (1994) too stresses the importance of training. He gives an example of how such a system with links to several external databases led to a progressive increase in the use of information technology by executives and in achieving the organization's goals with more and better deals.

1.2 Information strategies

by Byron Mook

Many NAROs currently face problems of "second generation" development. Over the past 20–30 years, most have dealt with "first generation" issues as they and the international community have invested in human and physical resources. Once these investment programs have paid off and buildings and scientists are in place, the urgent management challenge for the coming years is to mold these resources into efficient and effective research programs.

One of the most effective tools for doing so can be improved information. New computer and telecommunications technologies are revolutionizing agricultural research. Managers as well as scientists can have access today to more information than ever before, and PCs are becoming standard equipment in offices and laboratories. But the most critical question about all this information often remains unanswered: are managers and scientists able to use it to improve the efficiency and relevance of what they are doing? This section argues that an information strategy is a powerful aid to ensuring that the answer is yes. An information strategy provides guidelines as to the purposes for which information within an organization is to be used, as well as on how it is to be managed.

1.2.1 The case for an information strategy

The NARO hires architects to design research buildings, it asks social scientists for human resources and training projections, and it requires scientists to plan their experiments. But very often it acquires its information technology (IT) on an ad hoc basis and organizes its information management (IM) through existing structures. The costs of such an unsystematic approach to information development can be measured both in wasted money and in lost opportunities.

A NARO's decision to undertake an IM/IT development strategy is usually based on three assumptions:

1. The lack of systematic access to information is (and will continue to be) a serious constraint on both management and scientific decision making.[1]
2. New IT and new IM have the potential for improving use of information.
3. Early action is required if information use is to be improved. The challenge is to strike a balance between three interrelated considerations:
   - the need for information
   - the willingness of NARO managers and scientists to use it
   - the capacity of the NARO to manage both information content and information technology

But how can agricultural research managers justify to policymakers the need for an IM/IT strategy? The case must rest on information use rather than on information technology. In 1993, ISNAR designed a simple sequence of means and ends, which a number of NAROs have used since (figure 1.1).

1.2.2 The strategy development process

The first step in developing an IM/IT strategy is assessing what information users actually require. Such a "needs assessment" exercise will be somewhat labor intensive and time consuming as it must address four main challenges.

1. The important word here is "access." A large amount of information already exists in most NARS but in forms and places that make it difficult for managers and scientists to use.

Appropriate information technology

↓

Improved information management

↓

Stronger research planning, monitoring, and evaluation

↓

Better research

↓

More benefits to farmers

Figure 1.1: Pathway of means and ends in agricultural research management

The first is that it needs to take account of different types of users of information, including farmers, extension workers, research scientists, research managers, policymakers, and members of the international community (including donors). Second, an IM/IT needs to deal with different subjects, such as background agroecological and socioeconomic data, completed research, current research, finance, and personnel. Third, the strategy needs to look at different processes, including research project planning, implementation, monitoring, reporting, and evaluation. Last, it draws on various sources, such as databases, books, articles, abstracts, and reports.

Because "information" is such a big subject, users in NAROs often find it difficult to set priorities. Many are not fully aware of what information is actually available and in what formats. New computer and telecommunications technologies continually change the rules of the information environment. This last fact alone means that a formal questionnaire on information needs will not usually provide a complete picture. There is no substitute for field visits and face-to-face conversations. A sample of NARO scientists and managers may be asked needs-related questions like the following:

- What types of information do you use now?
- Where do you get it? How?
- What do you see as the current strengths and weaknesses of IM in your organization? Are you satisfied with the content that you get, as well as with the ways in which you are able to identify and use it?
- If you could have easier access to better information, how (specifically) would it help you to do your job better?

Only after such an assessment of needs has been completed can the strategy developers begin to think systematically about the technology component. The question at that point is clear: what kinds of technology are required to respond to the needs that have been identified?

1.2.3 Subjects for a strategy

Five "core" subjects are likely to form the basis of any NARO information strategy. They are listed here, not in order of importance, but rather in the sequence that a strategy development team will probably consider them: (a) scientific and technical information, (b) management information, (c) network development, (d) human resources, and (e) organization and management.

(a) Scientific and technical information

Research obviously depends on information. NARO scientists need to keep up to date if their work is to be both relevant and effective. They must be able to find out as quickly as possible what farmers need, what past research has done to address those needs, and what other scientists are currently doing.

The traditional way in which scientists have met such information requirements has been to visit libraries. NARO researchers have used libraries to identify what information they should consult and then to access that information. A major resource for information identification has been abstracts, and one of the main points of access has been journals.[2] A NARO that is committed to strengthening its scientific and technical information will need to give attention to the following action areas: organization; coordinated book, journal, and CD-ROM acquisition; national bibliographic databases; current awareness services; search services; document delivery services; and publications (both hardcopy and electronic). Figure 1.2 shows how these seven areas are related.

(b) Management information

Just as good research requires information, so does good management. For program planning, for instance, data is needed on research completed and research in progress. Well documented information is needed about the rationales, objectives, loca-

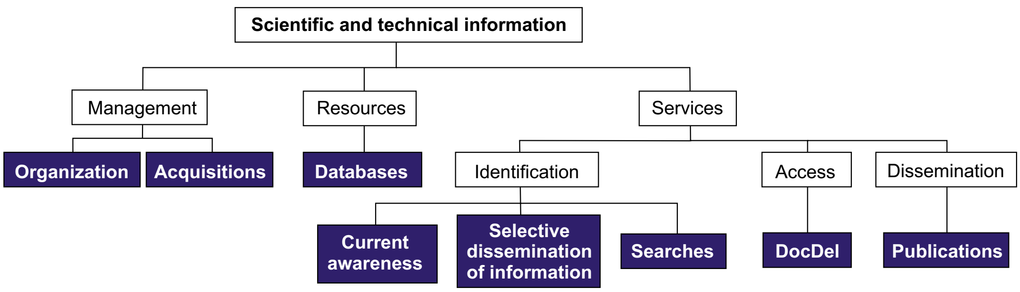


Figure 1.2: Seven action areas for providing scientific and technical information

2. In the past decade, even though new IT is rapidly changing the structure of libraries, the basic functions of IM remain the same. IM in any scientific field really consists of three tasks: finding information, handling information, and disseminating information.

tions, dates, resources, milestones, and outputs of programs. For program monitoring and evaluation, managers need data to support basic cost-benefit analyses. Why have some projects been chosen for implementation rather than others? How much are these projects costing, and what impacts are expected? For human resource management, data on qualifications and experience is required. What scientific fields need to be strengthened? What will the aggregate profile of scientists look like in five, 10, or 15 years? Finally, financial managers require data on projected and actual expenditures.

The matrix in table 1.1 provides a framework for a basic MIS. The four rows show the four categories noted in the above paragraph, while the two columns show the uses to which good data can ideally be put. Most NAROs face three challenges in designing and implementing an MIS that takes account of current organizational realities: (1) to develop further the databases that they already have, (2) to establish procedures by which these databases can be shared, and (3) to acquire the information technology that will make such sharing possible.

(c) Network development

Contemporary organizations with PCs and networks, public or private, anywhere in the world, can usually place themselves at one of the following five stages of IT management.

1. **Introduction.** A few PCs, mainly used for word processing.
2. **Proliferation.** More PCs and a few small local area networks (LANs)
3. **Confusion.**
4. **Rationalization.** Some centralization. Adoption of standards.
5. **Managed change.** Clear IM/IT policies and processes.

Many contemporary NAROs are somewhere between stages 3 and 4. Some of the most common symptoms of stage 3 are listed in table 1.2.

A NARO can deal with each of these issues separately. But a much more desirable approach is to have a strategy that encourages management to adopt a new set of structures and procedures specifically designed for IM/IT. The current argument that

Table 1.1: Matrix for a Basic MIS

Subjects for information/ uses for information	Planning/programming	Monitoring/evaluation
Research	annual workplans	impact assessment
Human resources	recruitment planning; career planning	performance appraisal
Financial resources	budgeting	accounting; auditing
Physical resources	procurement planning	inventory; stock control

Table 1.2: Common Problems Arising in Introducing Information and Communication Technology

Area	Issue
Policy	No IM/IT policy framework
	No hardware and software standards
	Insufficient contact with external sources of support
Management	No clear decision-making authority over IM/IT
	No clear rules regarding PC use
	No regular support services
	No regular acquisitions procedures
	No regular maintenance procedures
	Few resources
	Few trained users
	Variable annual expenditure
Hardware	Few standard PC configurations
	Widespread virus infestations
	No regular backup procedures
Software	Mostly unlicensed copies
	No regular upgrades
	Few manuals

a NARO should have a unit responsible for IM/IT must be just as compelling as one that it must have units responsible for financial management.

The focus of a network analysis will be determined by the points at which scientists and managers want to access information and between which points they want to move it. There are likely to be four different levels: (1) the research subcenter or field station, (2) the research center itself, (3) the national research organization, and (4) the international scientific community. The second level is the usual starting point and involves the development of a LAN. The first and third require a wide area network (WAN). And the fourth is dependent on access to the World Wide Web via the Internet.

(d) Human resources

By far the most important NARS resource in implementing new IM/IT programs is its people. Most scientific and technical information, MIS, and network development initiatives that fail, do so not because the technology does not work but because of inadequate human resources. Either an organization simply does not have enough people, or there is too much staff turnover, or the people who are there are not well trained.

"Information management" is a new concept to many NARS managers and scientists, and much new "information technology" is intimidating. A major goal therefore must be to create an environment in which information is regarded as important and in which IM/IT skills are seen as part of the toolkit of each member of staff.

The human resource development component of a successful strategy needs to address the priorities with which categories of staff require IM/IT training. The data shown in table 1.3 represents a first attempt at priority setting by one NARS with which ISNAR has worked.

(e) Organization and management

How can a NARO best manage its involvement in "the information revolution"? An IM/IT strategy will describe the most appropriate structures and procedures for dealing with (at least) the following organization and management issues:

- provision of continuing advice to NARS management on responses to new developments in IT
- ongoing formulation of new plans and programs for IM/IT
- organization of IM/IT training for NARS staff
- implementation of NARS-wide guidelines on IT acquisition and maintenance
- promulgation of rules regarding IT use
- development of NARS-wide data compatibility standards
- enforcement of software and hardware standards

1.2.4 Next steps

Finally, a good IM/IT strategy gives some indication of next steps. How do we get from here to there? Three NARS with which ISNAR has worked intensively on strategy development are doing some of the following:

- creating a national "information action team" to advise NARS management

Table 1.3: Priorities for Training in IM/IT Awareness and Skills

Target groups for training	Awareness	Skills
Managers	****	*
Scientists	*	**
Library/documentation specialists	***	****
Information technology specialists	**	***

**** = highest priority; * = lowest priority
Note: Data are taken from one client case.

- establishing a well-staffed and well-funded "computer services unit" at the national headquarters
- developing a research-oriented LAN at headquarters, both to serve headquarters itself and to become a "demonstration site" for other parts of the NARS
- agreeing on national hardware and software standards
- training professionals in science and technology information and MIS
- creating a Website as a means of educating policymakers about NARS plans and programs

A NARS with a good IM/IT strategy is in a position to turn the usual IM/IT development paradigm on its head. In most organizations, the sequence of events presented in figure 1.3 is the practice: hardware acquisition comes first. Growth then becomes unplanned and ad hoc. PC and network "specialists" are left on their own because management feels that it does not understand the technical issues involved. Problems appear. Hardware standards are difficult to bring about without a policy on software standards, and software standards are difficult to enforce without a policy on data compatibility.

In sharp contrast to this model, an IM/IT strategy provides a basis for starting at the bottom of figure 1.3. The first steps in this (more challenging) sequence are articulat-

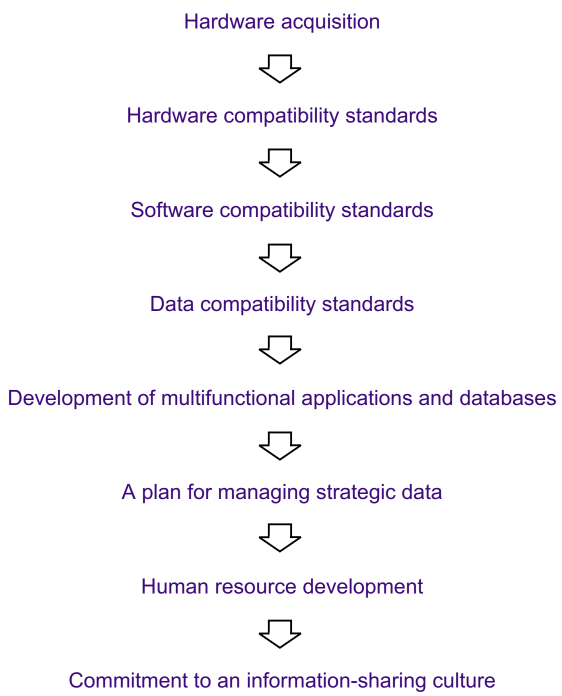


Figure 1.3: Typical sequence of information technology introduction

ing an IM/IT policy congruent with NARS goals, developing necessary human resources, and establishing clear management structures and processes.

The most obvious result of starting at the bottom will be cost savings. If appropriate policies and commitments can be put in place, the successive steps can then be taken more or less automatically.

References

Aris, J. 1992. Best practice in information systems management: Lessons from the IMPACT Programme. In Proceedings of the Developing Countries Specialist Group, British Computer Society, 1992 Summer Colloquium: Managing IT-Enabled Change in Developing Countries, 22 July 1992. Swindon, UK: British Computer Society.

Brown, M.M. and J. Grudney. 1994. Information systems and organizational development: Meeting the demands of information resource management in the twenty-first century. *International Journal of Public Administration*, 17, vol 3, 637-658.

Davenport, T.H. 1994. Saving IT's soul: Human-centered information management. *Harvard Business Review*. March-April 1994. 119-131.

Farrell, R.A., C.J. Pearson, and L.C. Campbell. 1992. Relational databases for the design and construction of maintainable systems in agriculture. *Agricultural Systems*, 38, 411-423.

Grinter, R.E. 1996. Understanding dependencies: A study of the coordination challenges in software development. A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Information and Computer Science. Irvine, US: University of California.

Leidner, D.E. and J.J. Elam. 1993. Executive information systems: Their impact on executive decision making. *Journal of Management Information Systems*. Winter 1993-94, vol 10, No. 3, 139-155.

Lewis, B. 1997. IS survival guide: Quality and service. *InfoWorld* 4 August 1997.

Lewis, F.L. and K.S. Keleman. 1990. Experiences with GDSS development: Laboratory and field studies. *Journal of Information Science* 16, 195-205.

Lucey, T. 1991. Management information systems (6th edition). London: DP Publications.

McNurlin, B.C. and R.H. Sprague (Eds). 1989. Information systems in practice. 2nd Edition. Upper Saddle Miner, US: Prentice-Hall International.

Chapter

2 The Agricultural Research Management Information System

This chapter is of interest to three main groups of staff who should be consulted on their information needs and on the desired features of the MIS: (1) a wide cross section of the intended users of the system including project and program leaders, station managers and headquarters-based research managers, and researchers, (2) the research managers who have responsibility for supervising information management in the organization, and (3) the information systems developers or managers.

The purpose of this chapter is to give all those who are to contribute to the "requirements specification" and contents of the intended MIS, a broad view of how such a system can serve the management of research.[1]

Operating a management information system consists of three processes: (1) collecting data and entering it into the system, (2) storing and processing it, and (3) delivering outputs such as reports, graphs, and tables that are useful for management purposes. This chapter focuses on the last process because the research manager and other users are most interested in these outputs.

The chapter describes how the outputs of an MIS serve research management processes, such as research program planning, project appraisal, monitoring and reporting. It suggests MIS outputs that relate to the main components of a research program, notably researchers and their experiments and projects.

2.1 Introduction

Agricultural research organizations everywhere have much in common. They have research scientists, experiments or other research activities, budgets and expenditure, and physical assets such as research stations with fields, laboratories, and offices. And each of these components has attributes that are much the same in every case: a scientist has a name, a date of birth, and usually a recognized scientific disci-

1. This chapter presents many examples from management reports from ISNAR's INFORM-R management information system. The concepts underlying these reports derive from many sources, including suggestions from research managers in several countries. In INFORM-R, "management notes" can be displayed on the screen with each report, providing ideas on how the report may be used. Where a report is illustrated in this book, the associated management note is reproduced beneath the illustration. The author would like to stress that the use of INFORM-R as an example MIS in this chapter does not imply that this system is superior to any other.

pline. A research activity has a title, objectives, and start and end dates; it focuses on one or a few commodities or noncommodity factors such as soils or socio-economics, requires a budget, and leads to a report.

Similarly, research management everywhere follows the same broad processes of setting priorities, planning, budgeting, and appraising research proposals at the start of a program, and reporting after and sometime during a research activity. Further management activities may monitor ongoing research, evaluate completed work and, for the small proportion that reaches the farmer, assess the impact on farm production. These common management processes rely on information about the major components and their attributes as described in the previous paragraph.

A consequence of this commonality is that it should be possible to identify the information needed by agricultural research management to a considerable extent wherever it is undertaken. This chapter explores this prospect and how a generalized MIS may provide such information. Conversely, in any particular research organization there may be details, particularly in the management processes, that are specific to that organization. It will therefore be necessary to adapt a generalized MIS model and its outputs (e.g., reports, tables, and graphs) to local needs.

When an MIS is implemented it is important to carefully question representatives of all the groups of likely users of the system as to their information needs. It is often helpful to show such users examples of outputs that are found useful elsewhere. Nowadays a successful information system may be expected to improve the way a business is conducted—it is no longer a passive tool to merely deliver faster what has been needed up till now.

This chapter describes many information system outputs that have been found useful elsewhere. Readers are invited to choose those that are appropriate to their situation. If too much is added to the list of requirements, then it will be difficult and costly to develop and manage the system. Each additional information item incurs a cost both in the complexity of the system and in data capture. Once a core of essential data is defined, a useful rule is to decline any additional items unless both a use and a user can be identified.

2.1.1 Information needs for managing agricultural research

The main information needed to manage agricultural research derives from the details of the research program and the underlying problems it addresses, from the financial and other resources allocated to it, from the researchers implementing it, and from its outputs. But we need a much more detailed specification of requirements to implement an MIS. Corbett (1989) suggests that an agricultural research MIS should do at least the following:

- describe the research in manageable units
- apportion costs of the research
- list the principal researchers involved
- report progress in research and in spending

- outline briefly future research intentions
- allow retrieval of all the information aggregated in different fashions
- be revised regularly so that it is current
- permit continuity from year to year
- provide an estimate of benefit or other reason for doing the research
- provide some measure of output of the research system
- provide the information needed for management purposes with minimum effort for the researchers

We pursue this analysis further in chapter 4 "Building a Management Information System."

Research on the development of commercial information systems[2] has revealed a significant difference between the information that managers say they want (i.e., on quantity and costs of production) and information on the factors that related most closely to the success of the company (i.e., product quality and speed of delivery). In the context of the NARO, this may suggest that while it will be very important to ask research managers what information they need, it may be legitimate to also have an eye on key indicators of the NARO's successes.

Information needs to be widely available, in different forms, and to many people: researchers, research managers, extension services, farmers' organizations, specialists, and senior administrators at the ministry of agriculture and donor representatives. Modern information technology, particularly database technology, enables such manifold requirements for information to be met very efficiently from a single repository of data. Providing useful information, at minimum cost and effort, is the primary goal of an MIS.

In this book we focus on the specialist information needed for managing the agricultural research program (e.g., experiments, surveys, projects) and its outputs (e.g., new technologies, annual reports, and scientific papers).[3] For administration of the research service, additional information on accounting and personnel will be needed. Detailed information on these resources, in the present state of information systems technology in many agricultural research organizations, particularly in developing countries, is usually best handled by separate systems for each. Information requirements for accounting and personnel management are common to many sectors, and there are many information systems on the market providing for such requirements. (Chapter 5 "Information Systems for Administration" gives an overview of selected topics.)

2. By Brian Houlden of Warwick University and David Woodcock of Nottingham University, UK. 1989.
3. This model has been adopted by ISNAR in its INFORM and INFORM-R management information systems. They focus on research program management but extend the scope to include limited data on finance (such as research project budgets) and some personnel data: the "scientific" aspects of the scientists viz. their qualifications and their commodity and discipline foci.

2.1.2 Key components of an agricultural research MIS

Figure 2.1 shows three dimensions of managing agricultural research, each with its own particular class of information. The following are the main resources in agricultural research:

- **people:** researchers, research managers, and the supporting staff
- **scientific information:** as found in annual research reports, scientific journals, text books, indigenous knowledge of farmers, and in electronic form (e.g., in on-line databases and on CD-ROM).
- **finance**
- **physical resources:** land, laboratories, machinery, etc.

Management will need information on each of these four and on their immediate products, the research activities and their outputs (new technologies, research reports). An MIS records data on these components of the organization, and it processes the data into meaningful information that can assist managers in making better decisions. An MIS in an agricultural research organization may focus on just a selection of these core components or it may extend to a broader set (see for example section 2.3.5 "Internal and external context.")

It is useful at this point to clarify some terms used in this text. The case for an MIS rests on the need for managers to be well informed on each area in their jurisdiction: the MIS provides information on each such area. Such information is called management information, which is quite different from scientific information. Research activity is used generically to comprise the following:

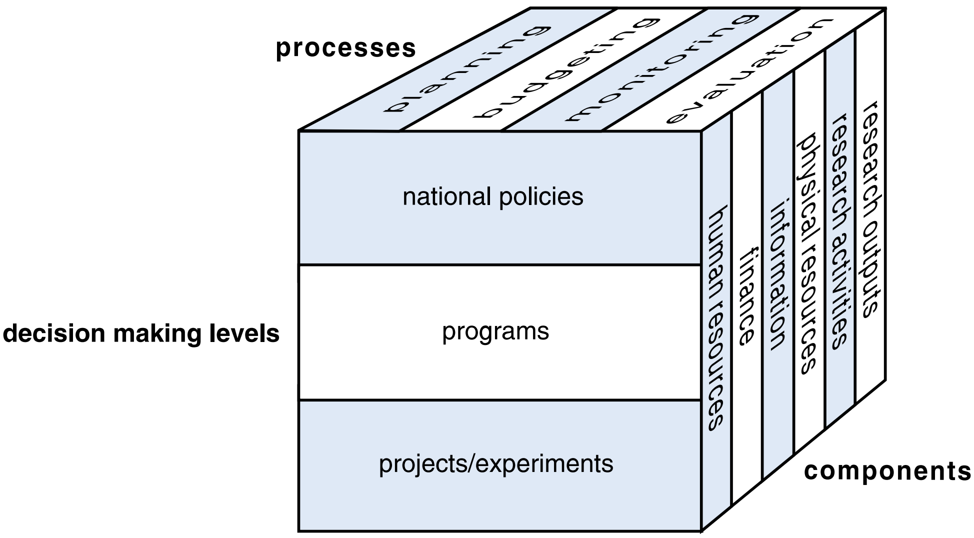


Figure 2.1: Three interacting dimensions of information needs

- **experiments and studies:** the lowest level of research activity for which a research protocol is prepared. It typically has a title, objectives, and treatments or hypotheses. The term is used here to include surveys and studies.[4]
- **projects:** a group of related experiments sharing a common objective, usually with a finite overall budget and time allocation.
- **programs:** a group of related projects, often within a single commodity group or discipline, and usually of sufficient national importance to be expected to continue indefinitely, i.e., without a specific end date.

To distinguish ex ante and ex post examination of research activities, we use appraisal for the former and evaluation for the latter, reserving monitoring for the examination of ongoing activities.

2.1.3 A hierarchy of levels

In any organization there is a need for detailed information at the "production area" level and more aggregated or summarized information at management and executive levels. This also applies to agricultural research (see also figure 2.1)

Researchers and their experiments and projects

Researchers should have good access to information, for example on what other research is going on in their country to ensure that there is collaboration rather than duplication. This may sound obvious, but experience shows that scientists are often oblivious of very similar work done in the same country. An MIS is a very suitable instrument to provide the necessary communication. Researchers should be able to explore what has been done locally in their own discipline and commodity in the past. This should be available in the form of annual reports, but good sets covering the past many years are often not available.

Today an MIS offers the prospect first of assisting in generating annual reports and second of providing an electronic archive that is easily copied and replicated for each research station. Such an electronic archive is also much more easily interrogated than a paper record. For example, it should be easy to identify all occurrences over many years of research in a particular scientific discipline in a particular commodity, e.g., all experiments examining the nutritional requirements of cotton, or diseases in groundnuts. This might take minutes with an electronic database compared with an hour or some hours going back through 10 or 20 years of annual reports or a local agricultural journal. (Scientific information is considered in more detail in chapter 6.)

Program leaders

Program leaders may be interested in external scientific information, but a more immediate need is for information on the program that they are leading. They will need

4. The use of the term "experiment" for activities such as observations, to which it should not strictly apply, is unfortunate, but it has acquired a legitimacy of usage. Perhaps "investigation" would be a better term.

details of every experiment as in a research protocol and feedback on the progress of its implementation, culminating in a final report on its completion. Program leaders need such information to oversee progress within the program, to ensure that relevant results are passed on to the stakeholders, and for planning the next research year's program.

Station and institute directors

Station and institute directors inevitably have to split their time between research management and administration. The latter requires much information on budgets and expenditure; personnel issues such as performance assessments, promotions, recruitment, and training; equipment procurement and maintenance; and many other issues. In this book, we focus on the information needed to manage research (some administration issues are discussed in chapter 5). Station and institute directors will have an interest in both detailed and aggregated information—-experiment reports provide the detail, while summary reports such as lists of projects and their scientists, graphical illustration of resource allocation across commodities, etc., provide aggregated information.

National research leaders

National research leaders are responsible for keeping three factors broadly in line with each other: research priorities, derived from national strategies, the research program, and the cadre of research personnel. The priorities are used through program planning to guide composition (e.g., at commodity level) of the research program, which in turn indicates the need for research personnel. Information on these three areas is therefore needed to facilitate comparison and check that there is a reasonable match between them. Managers will need to have access to aggregated data of the country's current research program that will show, for example, allocation across commodities and across agroecological zones. This picture can then be compared with agreed priorities. Similarly, information on the current research cadre broken down by commodity and noncommodity factors is needed for comparison with priorities and with the current research program.

With little further effort a well-maintained MIS can provide graphical and tabular outputs showing the balances between these three factors.

2.2 Research management processes and their information needs

2.2.1 The agricultural research management cycle

Every agricultural research organization goes through a (usually annual) cycle with stages such as the following:

- **research planning:** identifying significant changes that are needed in the coming season
- **appraisal:** judging the merits of research proposals
- **monitoring:** the progress of ongoing research activities

- **reporting:** interim results for ongoing work and final results of completed work

To learn from the experience of each research activity after its completion, ideally there should also be

- **review or evaluation:** of the work done and its outputs

This may be taken further by examining, sometime later

- **impact:** of research outputs on farming practices

On a longer time frame, e.g., five years, there may be a review of

- **research priorities**
- **long-term research planning:** deciding the main thrusts for future research

The best decisions made at each stage will be based on good information. Two core reports that managers need in order to make good decisions at each stage are (1) a list of experiments and (2) a directory of researchers. It should be easy to have either sorted respectively by experiment ID and surname, as well as by commodity, research station, and scientific discipline.

Each stage of the management cycle has its own need for information. Examples are the resources available (e.g., the numbers of scientists in each discipline and commodity, the budget allocation for the coming year, and the current distribution of funds across commodities, stations, and disciplines), and the details of the current research program and its match or mismatch with agreed priorities. An MIS can provide this information (in detail or summarized) for each management process.

The following sections describe the stages and processes of the agricultural research management cycle in more detail, using examples of MIS outputs that can support them.

2.2.2 Priority setting and implementation

The complementary processes of priority setting and implementation exhibit a remarkable difference. The former has attracted far greater attention than the latter, with an extensive literature, much of it sophisticated and often theoretical.

Priority setting

Today the resources to do research seldom allow a NARO to pursue all the suggested lines of research. The need arises to select those lines that are most important. Defining the criteria for "importance" is a sequential process that starts at the level of policy. Examples of such criteria at the national level include food security, import substitution, export earnings, poverty alleviation, improved management of the natural resource base.

When these policy factors are set (or changed), they are applied usually through analyzing their effects on priorities across commodities (crop and livestock species) and noncommodity factors such as soils, irrigation, post-harvest technology, socio-economics and marketing. Within each commodity or factor it is then necessary for scientists and research stakeholders (e.g., extension agents, farmer representatives, and food processors) to identify the key problems, their relative cost, and chance of success. This ranked order of problems advises which scientific disciplines[5] are needed to undertake the research.

Implementation of agreed priorities

Responsibility for developing and publishing research policy usually rests with senior management in the ministry of agriculture, supported by or in close collaboration with senior research managers. The latter are then responsible for implementing this policy, ensuring that it is broadly reflected in the research agenda. It is very useful in this regard to compare resource allocation across commodities with agreed priorities. A good MIS may help; it monitors resource allocation across commodities and noncommodity factors over the entirety of national agricultural research, which may be dispersed over numerous stations and supported by a range of donor funds (see figure 2.2). In practice, however, resource allocation and agreed priorities differ to a surprising degree, which may be due to the absence of an MIS.

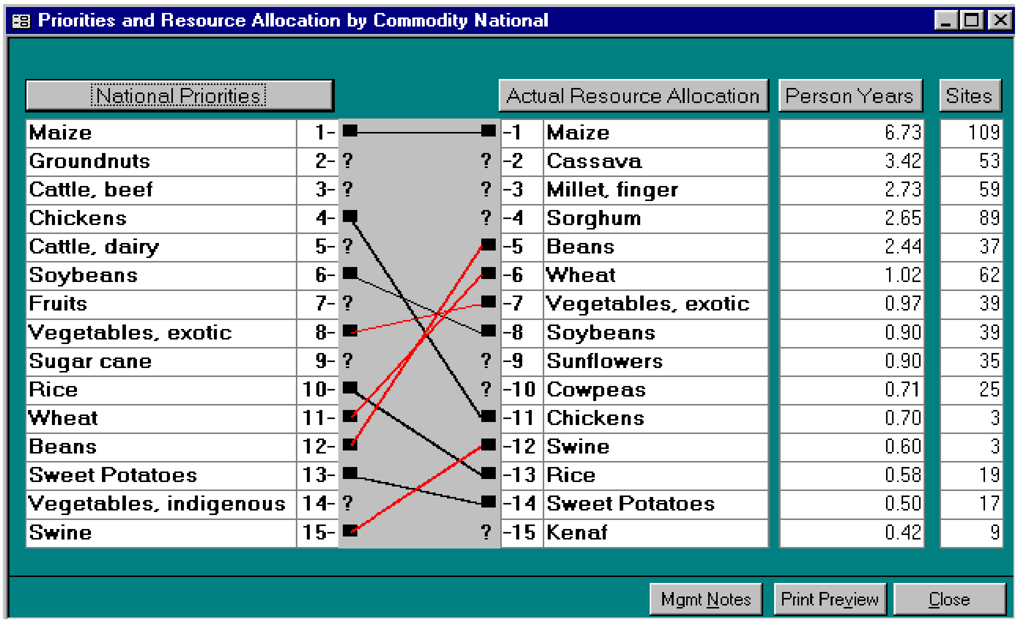
Implementing agreed priorities probably requires much more practical management and less sophisticated theory than does priority setting. Maybe for these reasons it has attracted less attention by development specialists and research managements, which is borne out by figure 3.1 in chapter 3 (page 87), in which resource allocation across commodities is tracked over several years and compared with priorities. There is little evidence of the two being brought more closely into line. ISNAR has similar data from other countries. This emphasizes that, if used by senior management, an MIS can be a valuable tool in demonstrating trends and identifying such mismatches.

Another reason that priorities may not match the research agenda can be a lack of confidence in the soundness of the priorities—they may have been drawn up several years ago or the method employed in developing them may be unconvincing.

A third cause of mismatch between agenda and priorities may be lax program-planning procedures, which leave researchers to find their own research thrusts independently of policy. In recent years this has been encouraged by donor programs that are not or poorly coordinated with the NARO and that link directly with research stations, research projects, or even scientists, bypassing and hence undermining the central research program planning processes. Where donors provide the greater part of the research current budget, this can easily skew the research program in ways not necessarily in the best interests of the national agricultural production. An MIS report, such the output shown in figure 2.2, may be presented to donors to encourage them to pay greater attention to national needs. In the case illustrated, a donor proposing to support bean research could be shown that beans are already being researched out of proportion to their priority, and it can be asked to consider one of the underresourced commodities such as fruits.

Implicit in these considerations is the need for information: on agreed priorities across commodities and noncommodity factors, on the commodity or factor and discipline of all scientists, and of all research activities. Priorities across commodities

5. See section 2.3.1 for a discussion of the term 'discipline' and table 2.8 for examples.



Management Notes[6]

Three issues often arise from this output.

1. Does the priority list need revision?

2. How should resource allocation be adjusted to be closer to priorities? Usually this is a gradual process, done through annual research planning meetings, at which this output can be studied.

3. How have any mismatches come about? And what can be done to minimize this in future?

When using this chart, the List of Commodities and their Resource Allocation* which provides data on all commodities worked on, rather than just the top 15 as here, may be useful.

* See figure 2.3

Figure 2.2: Comparison of agreed priorities and actual resource allocation across commodities

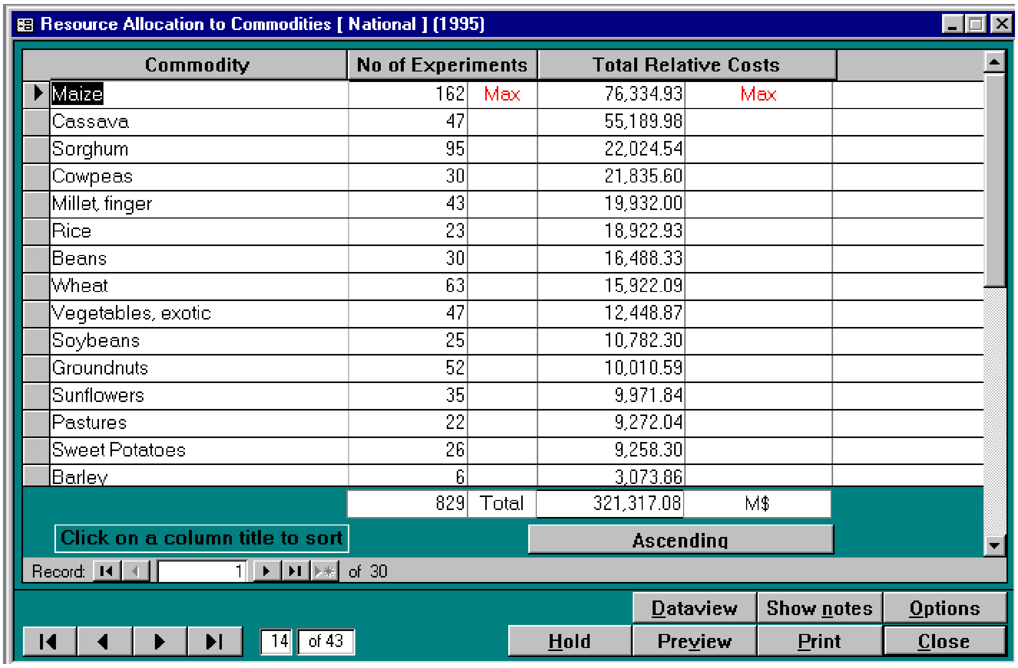
and factors are usually set or fully revised at intervals of some years, and there may be minor adjustments more frequently to reflect changes in the external environment (see section 2.3.5 "Internal and external context"). An MIS may therefore store a single, current set of priorities. This set may include subsets for each agroecological zone for regional program planning, while using the national set for institution-wide issues such as personnel planning.

6. As explained in the introduction to this chapter, management notes can be displayed on-screen with the graphic.

2.2.3 Planning

Research planning seeks to turn agricultural research priorities into practice within the constraints of available resources. Three sets of information are needed: on agreed research priorities, on available resources (human, financial, and physical), and on current research.

Long-term program planning will be conducted only occasionally (say every five years), when the national research priorities have been adjusted following priority analysis and setting. A list of the commodities and factors in the existing research program, displaying resources (human, financial) allocated to them, can be prepared from the data available in the MIS along the lines of figure 2.3. Such a list, in conjunction with the output shown in figure 2.2, facilitates continuity of research



Commodity	No of Experiments		Total Relative Costs	
Maize	162	Max	76,334.93	Max
Cassava	47		55,189.98	
Sorghum	95		22,024.54	
Cowpeas	30		21,835.60	
Millet, finger	43		19,932.00	
Rice	23		18,922.93	
Beans	30		16,488.33	
Wheat	63		15,922.09	
Vegetables, exotic	47		12,448.87	
Soybeans	25		10,782.30	
Groundnuts	52		10,010.59	
Sunflowers	35		9,971.84	
Pastures	22		9,272.04	
Sweet Potatoes	26		9,258.30	
Barley	6		3,073.86	
	829	Total	321,317.08	M$

Management Notes

This output can be used in conjunction with the Priorities and Resource Allocation by Commodity chart* in periodic major reviews of the research program.

Use this output also, by scrolling down to the bottom of the list, to check the smallest commodity programs. Are they all justified? Some argue there is a minimum critical mass of research effort beneath which the chance of realizing new technology is very small. A small program may be justified for initial exploratory work of a new area, but if this small input has persisted for several years the limited resource may be better invested in maintaining a literature survey of what is being done elsewhere.

* See figure 2.2.

Figure 2.3: Resource allocation to commodities

and realistic decision making for an efficient roll-over from the past to the future research agenda. By this we mean that existing research projects identified for reduced focus may be allowed to reach a milestone at which results can be reported. The reduced focus is achieved over some seasons by allowing current work to be sensibly wound up and not starting new initiatives.

Individual projects are defined from the key problems within each (commodity or factor) program, and ranked in their order of priority. Examples of criteria for priority ranking are the severity of the problem (number of farmers affected, size of area infested by a disease), and the chance of finding a solution. Objectives for each project have to be identified that contribute to the achievement of the program objectives. An MIS may issue a list of projects in ranked order of priority according to the mean of several scored factors.

The budget and research personnel required to continue ongoing experiments need to be known before resources can be committed to new experiments. A useful MIS output is one showing which projects or experiments have terminated or will terminate in the near future (see figure 2.4). The freed resources, e.g., research staff, can then be redeployed elsewhere. Individual data sheets for projects and experiments are useful if changes to the scope and direction of research have to be discussed and the research outline has to be changed. This may be necessary following the recommendations of monitoring or interim evaluation from the previous season.

Data on research staff is important for detailed project planning and decisions on the assignment of responsibilities. The MIS can provide information on allocation of researchers' time to research activities, to commodities and factors, and to disciplines. Other personnel information, specifically the disciplinary and commodity focus of

Page 1/1

List of Experiments due to finish next year
(Sorted by Commodity)

Printed: 18/8/1998
Research: 1996-1996

Commodity	ID	Title	Start Date	Planned End Date
Maize	6029	On-farm Evaluation of New Maize Hybrids	01/12/1994	01/06/1996
Many	6443	On Farm and On-station Field Days	01/01/1995	30/09/1996
Sweet Potatoes	5820	Germplasm Collection, Introduction and Evaluation of Sweet Potato	01/12/1994	01/12/1995
Wheat	6097	Date of Seeding of Plant Density X Variety Trail For Irrigated Wheat	01/10/1995	31/12/1995

Total experiments due to finish next year: 4

Management Notes

This gives a view of all active experiments scheduled to stop by next year. The list is sorted by commodity so some idea of the effect on each such commodity can be seen.

Figure 2.4: MIS report showing experiments due to close

scientists, helps identify staff with the required experiences and expertise to achieve the project's objectives successfully. If the right scientific mix is not available among in-house staff, research managers have to find alternative solutions such as training, retraining, or recruitment (see box 2.1).

An MIS can produce a range of reports, with aggregation of data needed at different levels of management. At higher levels, managers should not need detailed project proposals that would in any case merely duplicate the work already done at the correct project management level. Outputs showing resource allocation to projects, disciplinary or commodity focus, as well as outputs showing the relation of project objectives to program goals and national agricultural-sector priorities are useful at this higher level. Various MIS outputs useful for research planning are listed in table 2.1, each with its set of management notes that can be displayed on-screen over the report.

2.2.4 Appraising the research proposal

Each new proposal requires the preparation of a detailed protocol. It should describe the background, objective, expected outputs, review of past related research locally and (through a literature search) elsewhere, and the proposed methodology. It may include a work plan and budget, which may be broken down by year. From this data, the MIS can print individual data sheets for proposed experiments and a summary of all proposed experiments. These outputs can be used in the peer review to assess the scientific and technical soundness of the research.

A research proposal is normally assessed against criteria such as estimated benefits and costs, its relevance (i.e. a comparison with current research priorities), and the scientific soundness of the proposed study and likelihood of success. Table 2.2 gives examples of proposal assessment indicators. This table can be a form in an MIS,

Box 2.1: Redeployment Between the Sciences

Applying a new research agenda to an existing agenda may result in a mismatch between the new and the existing cadre of scientists. Figure 2.5 shows how the distribution of scientific disciplines across the research program and among the scientists can be compared by MIS pie charts, with good visual impact. This may even call for recruitment of the new specialists needed or training and reassignment of existing staff. Generally it is easier to reassign staff to a new commodity than to a new discipline. Plant protection specialists often apply their respective sciences to several crops as part of their normal duties. A maize breeder can be reassigned without too much relearning to breeding in another crop (particularly if it has similar characteristics), as can an agronomist. The main principles within each science are much the same from crop to crop.

Changing discipline is a much greater transformation and is not usually undertaken. A plant pathologist and an entomologist cannot easily swap roles, nor assume those of breeder or agronomist. Changing from a crop to a livestock commodity also involves a change of discipline.

Table 2.1: Examples of Information System Outputs Used for Research Planning

Report Title	Contents	Management Notes
National list of experiments, sorted, for example, by: • experiment ID • research station, commodity • commodity, discipline • main discipline (see figure 2.5).	Each of these lists could give experiment IDs, titles, and objectives of all current experiments.	• The primary catalogue or directory of research activities, useful to national research managers and the biometrics office. • Shows the current experiments at each research station. • Shows the distribution of research on a particular commodity across all stations. National commodity specialists and program leaders can easily see how their commodity is being addressed across the country. • Shows the experiments making prime use of each of the standard fifteen scientific disciplines. Specialists in a particular discipline, whether researchers or managers, can examine what is being done in their field throughout the country.
Research priorities and resource allocations (see figure 2.2).	Research priorities are often applied to commodities. This output comprises a ranked order of high-priority commodities and, for comparison, the commodities currently allocated the most resources. Ideally the two lists should be the same but in practice this is rarely even approached.	Three issues often arise from this output: • Does the priority list need revision? • How should resource allocation be adjusted to be closer to priorities? Usually this is a gradual process, done through annual research planning meetings, at which this output can be studied. • How have any mismatches come about? And what can be done to minimize this in future? Can be used in conjunction with the Resource Allocation to Commodities output (fig. 2.3), which provides data on all commodities worked on, rather than just the top 15 as here (see next item in this table).
Resource allocation to commodities (see figure 2.3).	A list of all commodities being researched at a station, or nationally, along with a resource allocation indicator e.g., total researcher FTEs,* or budgets, or number of experiments, allocated to each.	Check that the ranked order is in line with the mandate of the research station. A commodity with very low total researcher input, e.g., perhaps < 10% person year, raises the question of how likely it is that this work will yield useful outputs. It may be permitted for initial exploratory work of a new area, but if this small input has persisted for several years the limited resource may be better invested in maintaining a literature survey of what is being done elsewhere.
Researchers' time allocation to commodity groups	A list of the commodity groups (e.g., cereals, oil crops, livestock) being researched at a station, with the total researcher FTEs allocated to each.	Check that the ranked order is in line with the mandate of the research station.

(continued on next page)

(Table 2.1 cont.)

List of experiments due to finish next year (see figure 2.4).	List of experiments where "Expected End Date" = current year +1.	This gives a view of what is scheduled to stop next year and the impact on the size of commodity thrusts, e.g., if 5 out of 8 groundnut trials are due to end.

* full (person) time equivalent

Table 2.2: Appraisal Indicators for a Research Proposal

Indicator	Appraisal
A. Benefits	
1. Reasonable chance of achieving the objectives, within the proposed time: 4-point scale, e.g. 1. excellent, 2. good, 3. moderate, 4. small	___
2. Estimated value of success: e.g., productivity gain per ha x applicable crop area x estimated adoption rate =	$ ____
B. Relevance	
The first four items below address the concern that research is often initiated without checking what *has been done before* on the same issue, or is *being done elsewhere*.	
1. national literature search completed?	y / n
2. no. of related national references cited	___
3. international literature search completed?	y / n
4. no. of relevant international references cited	___
5. experiment objectives: compatible with project / program objectives?	y / n
6. expected outputs: (list) : ___________________________________	
7. are these outputs measurable?	y /n
8. if so, list measurable parameters: ___________________________________	
9. expected main beneficiary: ___________________________________	
C. Quality	
1. is it "good science"?	y / n
2. peer review held?	y / n
3. if peer review held, date:	__/__/__
4. peer review result: 1. approved, 2. approved with modifications, 3. withdrawn	__
5. design approved by biometrician?	y / n
D. Cost	
1. estimated time to achieve objectives:	__ years
2. estimated scientist time needed (person years) :	__ FTEs
3. estimated cost of experiment:	$ ____
4. resources needed specified?	y / n
5. resources needed approved?	y / n

Notes on the indicators listed in table 2.2

An important feature of these indicators is that as they are known to the researcher, they act as a check list that he or she needs to address carefully before submitting a research proposal. They constitute a self-evaluation that saves the review committee time, and the researcher the embarrassment of an unexpected rejection. This applies particularly for example to indicators B1–B4. The peer review then merely has to assure itself that the information provided is true.

A. Benefits

Chance of achieving the objectives: This is always subjective and best judged by a small number of experienced researchers. A proposal by an irrigation group to embark on cloud seeding experiments in a low-rainfall area prompted a very interesting debate but was eventually turned down as there was agreement that the chance of achieving demonstrable effectiveness was small.

Estimated value of success: Can be estimated by multiplying (estimated) productivity gain per ha x applicable crop area x estimated adoption rate. As two of the three factors are estimated, the result cannot be an accurate prediction but may still provide good guidance on the general magnitude of the potential value of the work.

B. Relevance

The first four questions in this section seek to avoid wasteful duplication of previous work. This may be largely avoided by a good review of local and international literature. This is of course easier said than done today when many NAROs have largely ceased to provide their libraries with operating budgets. The new information technologies promise to offer help here, yet in many countries this promise awaits fulfillment.

Questions 1 and 3 ask the researcher whether literature searches have been made. This is more a reminder to do so before submitting a proposal for consideration. Questions 2 and 4 merely reinforce the point by asking how many references were found.

Question 5: The objectives of a research proposal are very important. Asking the proposer to confirm congruence with those of the parent project or program ensures both that he or she gives it due thought and that the proposal is indeed likely to be in line with its parent project or program, and supporting the integrity of the latter.

Questions 6–8 prompt the proposer to think carefully about the enterprise. They provide a guide to those assessing the proposal as to its merits.

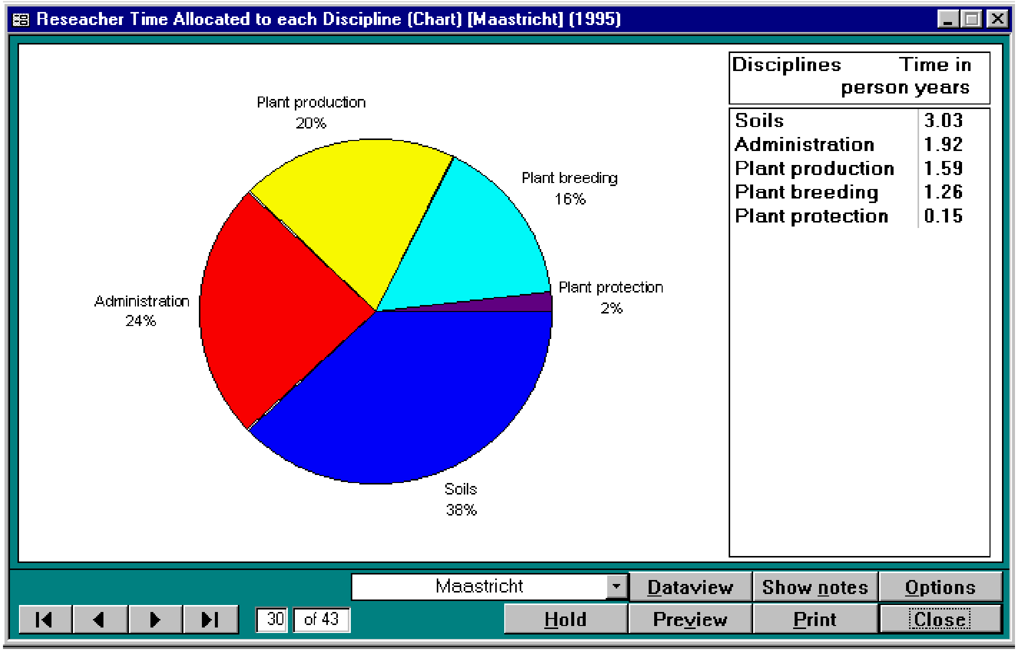
Question 9: Busch and Bingen (1994) stress the importance of declaring the beneficiary: "it appears that most annual research program reviews rarely question the adequacy or responsiveness of the research to specific societal interests or needs .... In short, the historical evidence is clear: linkages to client groups are an essential and vital feature of effective agricultural research."

C. Quality

Question 1: The object here is to focus the mind, first of the researcher and then of the reviewer, on this point. It is argued that most experienced researchers would be able to offer useful judgment on this issue.

D. Cost

Question 4: The MIS merely records whether or not they have been (adequately) specified. It may or may not be thought worthwhile to record them within the MIS.



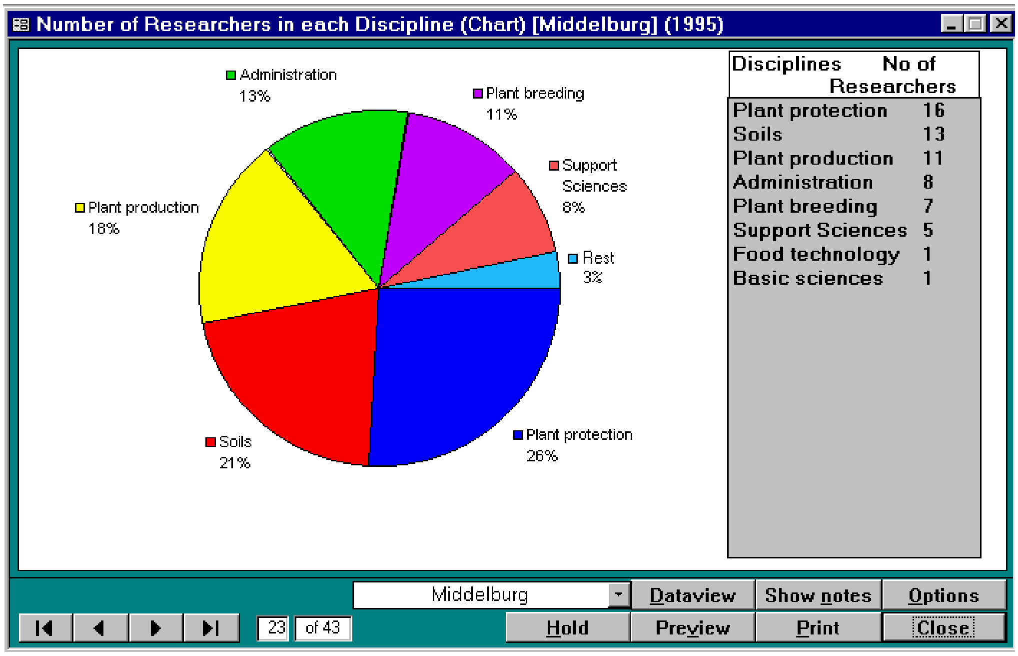
Management Notes

Do these reported time allocations match your perception of the relative seriousness of the problems they address and are they in accordance with the station's mandate? And if not, is there scope for, over time and perhaps through selective recruitment, changing the mix of disciplines to better match the problems facing the local farming community?

Figure 2.5a: Comparison of discipline focus in the research agenda

which can be printed out for a scientist to complete and attach to a research proposal.

When there are many proposals, it is helpful to convert such judgments to some kind of numeric value to facilitate ranking. After all proposals have been peer-reviewed, the resulting data is keyed into the computer. The factors may be previously weighted to reflect management's view of their relative importance. The MIS can convert each keyed-in response into a score that is multiplied automatically by the appropriate weight. A list of all proposals, ranked in the order of their total score, can then be produced. This provides a rapid, efficient, and relatively impartial method for ranking a large number of proposals. As with all such scoring systems, the researcher peer group or research managers should carefully inspect the result for any anomalies. Some comments on the indicators in table 2.2 are given in the "Notes on the indicators listed in table 2.2" (previous page).



Management Notes

This output “Numbers of Researchers Working in Each Discipline” can be compared with the similar pie chart “Researcher Time Allocated to Disciplines.” Large mismatches may suggest the need to increase the human resource available in “over-worked” disciplines through transfer, higher degree training and recruitment.

Figure 2.5b: Comparison of discipline focus in researchers

The ex ante and ex post economic assessment of research activities

Some of the research management stages examined here (e.g., priority setting and implementation, appraising research proposals, evaluating a research activity, and assessing impact) have been the focus of much attention, in particular from economists. Sophisticated econometric methods have been developed to assess their expected and realized values. These methods can be divided into two groups: ex post, which attempt to assess the effect of completed research on agricultural output, and ex ante, which assess the prospects of research before it starts.

The advantage of ex post methods is that they draw on existing information on research outputs, agricultural production, and the general environment. They provide lessons that may guide management in the future. The advantage of ex ante methods is that they look to the future and have consequently a more direct appeal to decision makers.

Problems of these methods include a requirement of much data that is expensive to collect and may be only partially available, and the need to separate out effects of other factors. Some of the methods are very complicated and require scarce, skilled personnel. Their development and sophistication appear to have outpaced both the availability of data and their actual use for decision making at least in publicly supported agricultural research. Their potential value may be curtailed by a change of minister or even government that brings with it a change of priorities before they are fully implemented. This argues against large expensive priority-setting exercises that may become irrelevant before they are implemented. These methods are not described further here but see Schuh and Tollini (1979) for a critical appraisal in general and a description of several models, and Alston et al. (1995) as advocates for, and for a more detailed review of, such techniques. Prinsley (1992) also describes and gives examples of different ex ante and ex post evaluation techniques in research.

2.2.5 Monitoring

Monitoring is here taken to mean the examination of progress of an ongoing research activity to identify and warn of problems, deviations from declared plans and objectives, or any other reason to intervene at a time in the research activity's life when corrective action is still possible. This contrasts with evaluation (see below), which seeks to learn from the experience for the benefit of future research. We are here concerned with formal monitoring supported by outputs from an MIS. A good manager will additionally monitor research activities informally by visiting research sites and talking to scientists. Monitoring may also have a role in providing accountability, particularly in donor-funded projects.

A research activity has the following components, each of which is important and therefore justifies monitoring:[7]

- **inputs:** scientists' time, budget, supplies (seeds, fertilizers, etc.), and the timeliness of each of these
- **implementation:** literature survey, planting, treatment applications, results measurements, harvesting
- **outputs**: experimental results records, reports, new technologies, local and international published articles, presentation to farmer groups
- **context factors:** the needs and situation of clients, outbreaks of disease, markets and prices (local and world), agricultural and price policies, national and donor policies, and advances in science and technology. These are addressed usually in the planning stage. (Context factors affecting the research program and institute are addressed in section 2.3.5 "Internal and external context.")

Outputs are arguably the most important of these components and therefore the first focus for monitoring. A key output will typically be an annual report that includes actual research results for the last season. Monitoring output is more important than monitoring input, not only because outputs are closer to the ultimate goal of re-

7. These four also constitute the elements of the context, inputs, process, and product (CIPP) model used mostly in the USA to monitor and evaluate educational projects.

search and have impact on the target beneficiaries, but also because output monitoring has both short-term research management information and long-term scientific information. In a well-managed organization, the monitoring of the input of supplies can be delegated to the stores department for example, and expenditure input data should come from accounts staff.

For both external and internal purposes it is useful to record—a month or six weeks after the main planting season—all experiments that have actually started. Socio-economic studies and livestock and perennial crop experiments can usually accommodate to this timetable. With a little processing this can provide a list of current research that is useful to researchers, research managers, visitors, donors, and any others who need to know, see for example figure 2.6. The MIS can present this report with research activities sorted by commodity (as in this case), by scientific discipline, research station, or activity identity number (ID).

The intensity of monitoring is a contentious issue. In a well-funded research organization, researchers tend to be well motivated and therefore do not require more intensive monitoring than provided for by their annual reports. This ideal situation is very efficient. Ironically, as resources decline, monitoring becomes more intensive in the form of more frequent reports and more indicators. These are often focused on inputs and process or on immediate outputs such as physical target achievements, rather than reports with actual research results. In a financially stressed environment

Page 5/30 — National List of Experiments — Printed: 25-May-01
National — (Sorted by Commodity) — Research Year: #Name?

Commodity	Research Station	ID	Experiment Title	Main Researche	Start Date	Planned End Date	Actual End Date
Chickens			No of experiments on Chickens: 1				
	Madurodam	6330	Feeding trial to evaluate the effect of Granding a sorghum supplementary ration on theperformance of village chickens.	Oosten, Simon, S.	01-Mar-96	01-Mar-98	
Citrus			No of experiments on Citrus: 5				
	Ameland	6319	Tree Crops Nursery	Hiensch, Derek, D.	01-Aug-95		
		6320	Tree Crops Orchard / Variety Maintenance	Hiensch, Derek, D.	01-Aug-95		
	Middelburg	6234	NATIONAL SURVEY OF CITRUS WOOLLY WHITEFLY (CWWF)	Garder, Ephraim, E. M.	10-Jun-95	10-Jul-95	
		6367	Development of IPM for control of Cercospora Angolensis of citrus.	Fransisca, George, C.	01-Dec-95	01-Jun-97	
		5911	Distribution and severity of cercospora angolensis and Oidium mangefera in Stapvorstap	Fransisca, George, C.	01-Dec-94	01-Jun-96	31-Dec-95
Coffee			No of experiments on Coffee: 11				
	Maastricht	3227	Liming Trial Under Rainfed Conditions (Tweelingloo)	Klein, N, N.	01-Jan-81	01-Dec-91	
		6291	Cattura Progeny Trial / Latin American Progeny Trial	Klein, N, N.	01-Jan-84	31-Dec-99	
		5909	Coffee berry disease chemical control experiment	Klein, N, N.	12-Dec-94		
		3228	Spacing Trials under rainfed Conditions (Tweelingloo)	Klein, N, N.	01-Dec-91	01-Dec-91	
		5946	Evaluation of indigenous phosphate fertilizer for Agronomic Effectiveness	Candan, Humphrey, H. C.	15-Nov-93	01-Jul-00	
		5787	Planting systems in coffee (COVA)/Fertilizer rates trial	Klein, N, N.	12-Dec-94	12-Dec-01	
		2949	Portugease Variety trial	Klein, N, N.	01-Jun-77	30-Dec-97	
		6290	Coffee National Variety Trial	Klein, N, N.	01-Aug-95	31-Dec-05	
		6288	Intercropping Coffee With Leguminous plants / fertilizers Trial	Klein, N, N.	12-Dec-95	12-Dec-01	
		6282	Catimor Progeny Observation	Klein, N, N.	01-Jan-82		
		6285	Meseum variety collection	Klein, N, N.	01-Oct-64		
Cotton			No of experiments on Cotton: 13				
	Den Bosch	6203	AN ANALYSIS OF FUTURE PROSPECTS OF COTTON PRODUCTION IN THE LIBERALISED MARKETING SYSTEM - A CASE FOR Jezuseik Goupil	Maarschalk, M, M.	01-Jul-95	01-Aug-95	
		9005	Cotton spacing trial	Eck, Christopher, C.	01-Dec-92	01-Jul-96	
	Lelystad	8005	Technical coordination of fibre research programme at Roerzuil	Besten, West, W.	01-Jan-95	31-Dec-95	
		5921	Evaluation of active ingredients (insecticides) in bollworm control	Betten, Kangwa, K.	01-Nov-94	01-Jan-96	
		5923	Effect of the programme of spray on the control of bollworms	Betten, Kangwa, K.	01-Nov-94	01-Jan-96	
		5811	Response of Variety F135 to different planting dates and plant densities.	Blankenstein, David, D. K.	01-Dec-94	01-Dec-97	

Figure 2.6: List of current research activities, sorted by commodity

such micromanagement adds to the management overhead of the research program and to the task of managing information.

Monitoring also applies to the institutional level. Some examples of institutional monitoring indicators, mostly taken from Peterson (1994), are shown in table 2.3. Peterson also explores the prospect of quantifying output indicators and then, by dividing by a measure of input such as cost or researcher time, developing an efficiency indicator.

Indicators

To help make monitoring more consistent it is common to agree on a set of indicators: predesigned data items known to reflect the quality of research performance. They comprise input, processing, and output factors against which an organization, project, or process can be measured. They may demonstrate conformance with or deviation from agreed features of "good" research.

Good indicators should help research managers make decisions that enhance the performance of the research program. Indicators work best if they are objectively verifiable. Typically this implies a quantitative indicator, often with a supporting "norm" within which the feature should lie, or a logical "yes/no" answer. They should be simple and easy to measure, and any data they depend on should be either already available or easily and cheaply collected. Finally, the number of indicators should be limited to the minimum, as each adds an overhead cost to the research. Selection of indicators is therefore important. A useful test of the value of an indica-

Table 2.3: Examples of Factors for Monitoring the Research Institution

Factor	Indicators
Inputs	
Scientists' access to external scientific information	Number of journals subscribed to by library
Processes	
Institutional sustainability (World Bank 1997) (see also section 2.3.5 "Internal and external context")	Proportion of budget supported by sales of products and cost recovery
	Operating cost: salary cost ratio
	Donor dependency: proportion of budget supported by donors
Outputs	
Improved varieties of plants and animals	The number produced per program over, say, five years
Improved cultural practices	As above
Training for farmers and extension personnel	Number of events; number trained
Research publications	Number in (a) refereed international journals, (b) local farming press, and (c) extension literature
Scientific services (soil and plant analyses, vaccine preparation, seed quality tests, etc.)	Number per year; revenue against cost

tor is to ask what conclusion may be drawn from it and what management action can be taken as a result of its information. Table 2.4 illustrates the application of some of these requirements of indicators. A more general approach is shown in box 2.2.

Figure 2.7 shows an MIS-generated monitoring report, which allows scores on a small number of selected indicators to be aggregated into a single figure, helping a manager to quickly focus on any projects facing difficulty.

For several indicators it will be possible and useful to compare actual with planned performance. This variance invites an explanation and corrective action. An MIS should be able to present trends over recent years. In a relational database model this is facilitated by including a "year" field in the relevant database tables.

2.2.6 Research results reports

A distinction can be made between reports that provide information about the research to assist managers in managing the research program (see 2.2.5) and reports that contain actual research results. The latter are very important as they present the outputs—the justification for doing the research in the first place. They provide a permanent record of what has been done and provide the necessary information for that part of research that is suitable to be transmitted to clients. As an indicator of research output they provide research managers and donors with a significant monitoring indicator.

Table 2.4: Features of Monitoring Indicators to Identify Problems in Research

Common problems faced in res. mgmt.	**Indicator**	**Ease of objective verification**	**Ease of data collection**	**Amenability to management correction**
Research reports are lacking.	Presence of a research report: yes / no	*High*	*High*	*Moderate:* require report prior to approval of further research.
Research continues longer than necessary.	List of projects still active after their planned end date.	*High:* it is either past its end date or not.	*Moderate:* needs a project management system to be in place.	*High:* cease funding unless continuation can be justified.
Research and agreed research priorities are misallocated.	Comparison of agreed priorities with resource allocation across commodities, agroecological zones.	*Moderate:* resource allocation data usually only relative. Priorities usually based on subjective multifactor judgments.	*Low:* priorities may not be available; resource allocation data requires comprehensive project management system.	*Low:* requires firm management action over long period. ISNAR has found success here to be rare.

Box 2.2: Criteria for Selecting Performance Indicators

The following are criteria for selecting performance indicators and measures in developing public-sector performance measurement systems:

Data criteria

- Availability: Do the data exist, can they be collected?
- Accuracy: Are the data reliable? Check for errors, biases.
- Timeliness: Are data timely enough for evaluating performance?
- Security: Are there privacy or confidentiality concerns?
- Costs of data collection: Are sufficient resources available?

Measurement criteria

- Validity: Does the indicator or measure address financial or program results? Can changes in the indicator be clearly identified as desirable or undesirable?
- Uniqueness: Does duplication between different indicators exist?
- Evaluation: Are reliable benchmark data, standards available to interpret the indicators?

Measurement system criteria

- Balance: Is there a balance between input, output, and outcome indicators?
- Completeness: Are all major programs and their components covered?
- Usefulness: Will management use the system to effect change based on the analysis of the data? Are there incentives for management to use the data after they are collected? Are management reports "user-friendly," i.e., clear and concise?

Based on Tuck and Zaleski (1996)

Managers need to study the results of research to assess its quality and relevance to farmers. Extension staff will be interested in the results of the research that has successfully identified appropriate technologies. Fellow scientists will be interested in details of research done previously and research in progress to guide them in further research and in avoiding wasteful repetition of similar studies.

Table 2.5 lists some types of research report that meet these different requirements. No organization would be expected to use all these models, however; that would create an excessive overhead. The table helps the information systems designer to agree with research managers and scientists on the models that the MIS should support. The models support three levels of research activities: experiment, project, and program, but reports may not be needed at all these levels in a particular research organization.

Page 1/1 **Annual Monitoring of Experiments (sorted by Commodity)** **Printed:** 18/8/1998

Research: 1996-1996

Scores:

Quality of Experiments Annual Report	Attainment of Experiment Objectives	Continued Relevance	Future Status
3: Full report 2: Adequate report 1: Partial report 0: No report yet	3: As planned 2: Behind schedule 0: Nil	3: High 2: Medium 1: Low	C: Continue unchanged M: Modify & Continue T: Terminate

Exp ID	Experiment Title	Annual Report Quality	Attainment of Objectives	Continued Relevance	Total Score	Future Status
	Commodity: Cowpeas					
5896	Agriland National Cowpea Variety Trial - Dual Purpose	0	0	0	**0**	
	Commodity: Engineering					
6056	Farm Power Machinery Testing and Development	2	3	3	**8**	C
	Commodity: Groundnuts					
6006	Advanced Groundnut Variety Trail (Spanish/valencia)	0	0	0	**0**	
6063	Groundnut Variety Verification Trial 1994	2	2	3	**7**	C
	Commodity: Millet, finger					
5899	Crop Rotation In Finger Millet	2	3	3	**8**	
	Commodity: Sorghum					
5784	Sorghum Preliminary Variety Trial	0	0	0	**0**	

Management Notes

Research activities should normally be monitored annually. This report may be used in the annual research program planning meeting. The ranked scores can be used as a first ordering of experiments in importance. Such an automated ranking is valuable particularly with a large number of experiments but should always be reviewed by experienced researchers/research managers to include from their experience a wider set of criteria than an automated system can handle.

Figure 2.7: A monitoring report showing aggregation of scores across three indicators

A problem in some NAROs is that even though annual reports of work done are very important for the organization[8], there is often much delay in producing them, and the quality is often disappointing. In part this is due to institutional and management issues, but in part also to the sheer difficulty and labor involved in assembling such reports. An MIS has a great deal to offer here. A well-maintained MIS can greatly ease the burden and enhance the value of the annual research report at many levels: research experiment, project, program, station, institute, and country. A single set of carefully selected data items can be collected and stored in one operation and then

8. In section 6.1 of this volume, Allmand et al. observe that "the marketing flagship of a NARO may be its annual report."

Table 2.5: Examples of Research Reports

Level/type	Description of report
1. experiment annual report	Contains the detailed results of a research experiment from a single year. It provides a record that can be used (1) for research review and planning processes, (2) to compile an experiment completion report from the annual reports through the life of an experiment, and (3) to compile program, station, and department annual reports.
2. experiment completion report	Contains the main details and results from the full life (i.e., all years) of an experiment, along with conclusions, any extension message, and identification of any further research still needed to address the main objectives of the experiment.
3. project annual report	A project usually comprises one or more experiments or studies that share a common overall objective to meet a particular problem. Typically there is a finite time period and funding and sometimes a sole donor. The report provides an overview of the work done and any significant steps towards the realization of the project objectives, the funds spent, and the next steps. Background information may list the objectives of the project, its starting and ending date, and who is providing either technical or financial support. There will be an overview of the progress made to date and in particular in the year being reported, written by the project leader. It should include a forward look to what is planned in the next and following seasons and estimate the time still needed to reach a conclusion. There should be a report of the project's finances, including allocated budget, funds actually released, and funds spent, to date. Optionally, there may be a set of attached copies of the individual reports of each experiment.
4. donor project annual report	As for the project annual report above, but drawing from experiments supported by a particular donor.
5. station annual report	Provides both an overview and a detailed record of the work done by the station staff in the past year for the following purposes: It enables senior research management at headquarters to monitor the work at the station technically and administratively. The assembled station annual reports from all stations can provide most of the contents needed for the national organization's annual report. It provides those with a special interest in the work of the station with the information they need. The station annual report should contain the following background information: the mandate of the station, a list of its scientists with their commodity interests and scientific discipline, and a summary inventory of facilities (land area, arable area, altitude, average climate conditions, etc.). An MIS may be able to provide much of this automatically. There may be an overview, by the station director, of the highlights of the year, both technical (research) and administrative. It may be followed by a report of the station's finances, including allocated budget, funds actually released, and funds spent, to date. This can be entirely written by the station director or head of accounts, or can be an attached copy of the report generated by a budgeting and expenditure system with a manuscript summary.

(continued on next page)

(Table 2.5 cont.)

Level/type	Description of report
	The report may also contain summary reports of program leaders and the individual reports of each project and experiment, which an MIS can generate automatically. A station typically is a site dedicated to research of particular relevance to a local area or region. Thus it may address agricultural production appropriate to the local soils, rainfall, and temperatures and to the local farming culture and socioeconomic situation. To fulfill this it usually has land, sometimes at more than one site to address different representative soil types or other conditions, farm machinery and workshops, laboratories, offices, and a small library. It may include residential accommodation for its staff. The station annual report may report on the status of any or all of these entities.
6. program annual report	A program usually comprises one or more projects that share a common overall objective to meet a particular component of the national research strategy. Usually there is no set time limit on a program as it is expected to last many years or indefinitely. This report may often not appear as a separate report, but rather as a constituent of the research department annual report. It provides an overview of the work done within the program in the last year, the funds spent, and any significant achievements contributing to the realization of the program's objectives.
7. research department annual report	Ideally two types of annual report of a research department are prepared, reflecting the different needs of different audiences: Part 1: summary. A high-level review is required for circulation to other ministries, especially finance and those ministries concerned with natural resources, and to parliament as may be required by law and as the source of government funds. Part 2: detail. A more detailed and technical report is needed both as a permanent record of information for the home NARO and for other research institutions at home and abroad. If their needs are to be fully met, much more data is needed than is required for the first audience mentioned above. This second report is often missed and details then have to be obtained from the biometrics office or the individual scientists.

used flexibly for a range of different purposes as and when needed. Outputs can be generated to meet different needs: those of management at different levels, those of other NAROs and NARS and specialist networks, donors, and international information systems, such as the Current Agricultural Research Information System (CARIS) of the Food and Agriculture Organization of the United Nations (FAO), the Current Research Information System (CRIS) of the US Department of Agriculture (USDA), and the List of Research Workers in the Agricultural Sciences of CAB International.

Donors sometimes specify the format of the report that they wish to receive. This may be when they previously received either no report or one with inadequate information. Their specification is largely a device to get the essential information they need. It is very likely that most donors would accept a standard format if a research organization provided a good factual report in a timely manner. A good MIS should enable this standard format to be repeated for each donor with just the experiments and projects supported by that donor, dispensing with the need for freshly tailored reports for each. Donor representatives are sometimes more interested in the financial records of a project than the science or the agriculture. An MIS may need to track budgets, released funds, and expenditure at project level and generate an appropriate report for each donor.

Researchers should be encouraged to publish in international peer-reviewed journals. This tests the quality of the research and its reporting and contributes to the general pool of knowledge. But they should also be encouraged to first focus on publishing where their own clients, the country's farmers and extension service, may benefit, namely, in local farming publications and extension literature. This point is considered further in section 2.2.7 and in the evaluation indicators shown in table 2.6.

Importance of recording "failed" experiments

It is in the nature of research that many activities do not yield results that are suitable to be transmitted to farmers. Often, only a limited proportion of all projects represent the major research output of an institute or program. However, a successful agricultural practice recommendation is usually based on a large number of trials, many of which have not in themselves yielded a result directly applicable to farming. Tuck and Zaleski (1996) observe that "one of the major obstacles to measuring R&D is the measurement of failed experiments. However, they can be measured by ensuring that the failed experiment results in an output, such as a report on the results of the experiment." In other words, an unfruitful experiment is not necessarily a failure unless the organization fails to learn from it. To enable an organization to learn from unfruitful experiments it is essential that the experimental design, implementation details, and results are properly documented and made available to other researchers. The MIS, with its associated archive, is an obvious repository for such records.

2.2.7 Evaluation

Evaluation, as used here, is the systematic assessment of the outputs of completed research for a number of purposes:

- The primary role of evaluation is to enable managers and researchers to learn from experience and feed the lessons back into the planning and implementation of more effective future research activities (Nestel 1989).
- The knowledge that their work is to be evaluated, against known criteria, provides researchers with encouragement to address those criteria. For this to be effective the criteria need to be clear, well chosen, and applied consistently, year after year.
- The evaluation process provides management with material (in addition to annual reports, which remain potentially the best evidence, prior to impact studies some years later) to demonstrate the outputs of research and encourage budgetary support from finance and agricultural ministries and from donors.

Table 2.6: Examples of Research Evaluation Indicators

Indicator	Response
Benefits:	
1. Degree to which the objectives in the research protocol have been realized: 1. fully realized , 2. largely realized, 3. significant problems, 4. activity abandoned:	___
2. Were objectives changed?	y / n
3. If 1.3 or 1.4 above apply, or if objectives were changed, describe, with advice on future avoidance:	___
4. Contribution to knowledge: 1. significant, 2. some, 3. insignificant:	___
5. If 4.1 or 4.2 apply, briefly describe. If 4.3, explain why and whether in future similar work this could be foreseen:	___
Publications:	
6. reported in station / institute annual report?	y / n
7. full report prepared?	y / n
8. in local journal?	y / n
9. in international journal?	y / n
10. in extension literature?	y / n
11. Other outputs achieved as listed in research protocol?	y / n
12. Results verified in on-farm research?	y / n
13. Conclusion reached?	y / n
If yes and if useful to farmers (or other targeted beneficiaries):	
14. Introduced to extension service (or used by other target beneficiary)?	y / n
Costs: for comparison with *estimates* given in research protocol	
Time to achieve objectives (years):	___
Scientist time needed (person years):	___
Cost:	___

Evaluation needs to include at least a statement of what research outputs have been produced, e.g., reports prepared, what results were worth passing on to the extension service and whether this is in hand, and which results need further investigation. It may also contain an assessment of these outputs. Such assessment should first be made in the light of the objectives of the research; these are the focus of the first three indicators in table 2.6. The comments made on indicators in 2.2.5 apply equally to evaluation indicators.

Effectiveness and efficiency

Evaluation may attempt to measure effectiveness and efficiency. Effectiveness is demonstrated by outputs such as new agricultural technologies, processes, products, and knowledge potentially valuable to the client, usually the farmer. Earlier we suggested that outputs were the most important source of indicators for monitoring and evaluation. They are the focus of indicators 4–14 in table 2.6.

Efficiency implies cost per output and invites economic assessment, but this is not often done (World Bank 1996), partly because economic expertise in NAROs is scarce. On the other hand it is not essential to apply detailed economic evaluation to all research projects; targeting a selected subset may prove more efficient. Economic evaluation of a research program may be applied at periodic intervals of, for example, five years.

An MIS can routinely apply a simpler evaluation to all projects and experiments with modest input from researchers and program leaders. The indicators of table 2.6 can be applied in this way. The table represents a form that a researcher can quickly complete. The indicators help researchers and management to revisit the intentions of the original project and determine the degree to which objectives were realized and, where a shortfall is seen, ask for an explanation. This facilitates learning from the experience. Other indicators probe the outputs of the activity and inevitably focus the researchers' attention on the need to plan for these. The distinction in the questionnaire between international publications and the local farming press can send a signal to scientists that both are important.

The inputs into such a simple evaluation form can be entered into an MIS to produce a summary record of every experiment. This data can be converted to a numeric score for rapid comparison. An example of the result is shown in figure 2.8. Such an automated system allows management to apply weights to differentiate between the relative importance of indicators.

Impact assessment

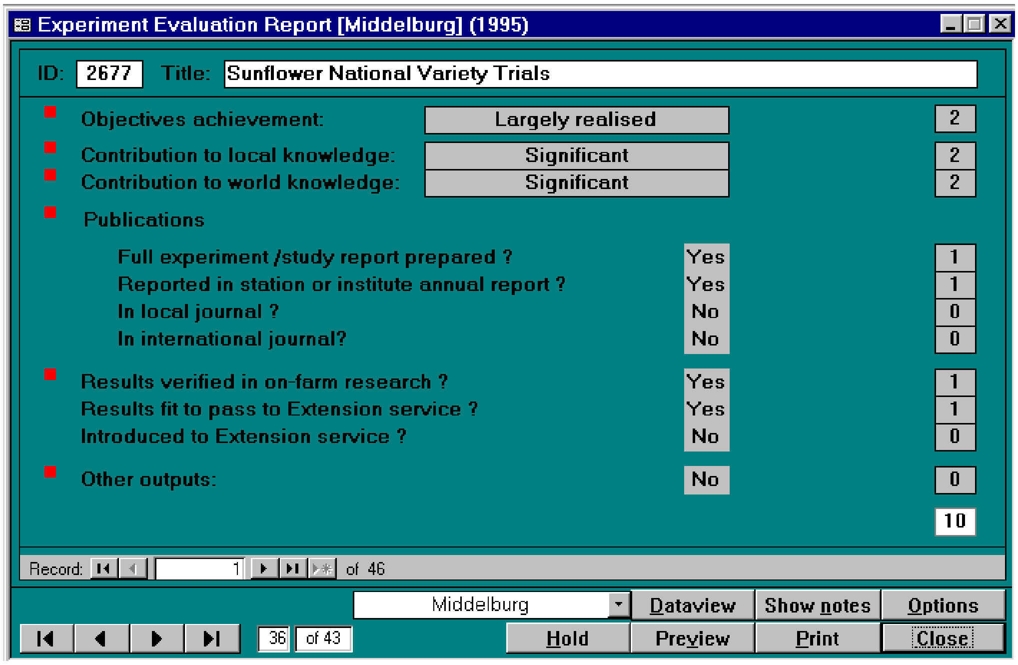
Impact is the effect of research on the farmer's practices and production. It is the ultimate test of the effectiveness of research. Its value depends on lessons accruing from impact assessment being fed back into the research process.

In practice these assessments and the value of such assessment are beset with difficulties. Impact usually cannot be assessed until some years after the research has been completed and relevant outputs transferred to the farming community. Changes in farmer practice and production may be a function of several factors apart from research recommendations, such as other extension messages, price changes in inputs and in farm produce, and credit services, the effects of which may be difficult to isolate.

Some estimates of impact derive from comparisons with estimates made in the research proposal, for example:

- hectares to which the new technology applies
  - potential adopters as % of farming community
  - actual adopters as % of farming community
  - number of actual adopters
  - hectares on which adopted

- value
  - mean productivity gain per ha x hectares on which adopted



Management Notes

An experiment or study is evaluated after completion. First the results over the life of the study are compared with the objectives. Additionally a set of general evaluation indicators, which can be modified from country to country, is recorded. These subjective assessments for a dozen or so indicators are each converted to a numerical score and these are automatically added together to give the Total Score for each study.

Use the record selector to move to another experiment.

The Experiments Evaluation List enables the same data to be compared across all experiments, and ranks these in Total Score order.

Figure 2.8: Experiment evaluation scores processed by an MIS

Impact assessment is best done by an independent party other than the researcher. It may be the work of a separate socioeconomic unit within the NARO or in a university. The report is therefore usually not part of a NARO's MIS, although MIS outputs may be a useful starting point for the impact assessment. It is important that such completed impact studies are readily available to researchers and research managers. We know of one organization that requires all new proposals to include a statement that the project evaluation and impact records have been consulted for any previous projects of the same commodity and discipline.

2.3 The main components of an agricultural MIS

This section discusses information relating to the following research elements: researchers, research activities, the context or environment in which research is conducted, and some special areas. (Finance, scientific information, and physical resources are usually the subject of separate information systems. See chapter 5 "Information Systems for Administration.")

2.3.1 Researchers

Definition

An MIS for management of the research program will be concerned largely with personnel that are directly linked to the research program—the research scientists. A research scientist is usually a graduate, but large NAROs in particular give first-degree graduates the standing of an apprentice and recognize only higher-degree staff as scientists. For MIS purposes a useful rule is to include all personnel who are actually responsible for one or more experiments. In small NAROs this often includes a small number of experienced technicians who, though without a university degree, have over the years gained sufficient experience and knowledge to be given responsibility for managing research activities. The generic term "researcher" can be applied to cover all these variants.

Some MIS outputs may extend to using data on other staff too, particularly technicians who are attached to specific projects or programs. Data on other categories of staff are usually left to a personnel system (see section 5.2 "Personnel information for research management").

What information to include

Faithful to the dictum that an MIS should hold only the minimum data required for the purpose, the data held on researchers in an agricultural research MIS may be limited to that having an impact on the research program. Any additional data is properly located in a separate personnel system. In more advanced systems the two will be linked and all, or most, of the personnel data can then be kept in the personnel system and retrieved as necessary for MIS outputs.

The scientists' attributes important to research management will include the person's qualifications, their scientific discipline and the commodity that they specialize in, and their research post. All such data is valuable to research planning and personnel planning. It ensures that decisions are based on good information and that the major directions of the research programs are supported by sufficient research staff of the appropriate skills.

Reports

Table 2.7 shows examples of reports used for research personnel management. Perhaps the most basic report relating to scientists is a directory, showing their names, home station, the commodity they primarily work with (if appropriate), and their main discipline. Figure 2.9 shows an example from an MIS that can be displayed on

Table 2.7: Examples of Information System Outputs Used for Research Personnel Management

Report title	Contents	Management notes
National directory of researchers, sorted in various ways e.g., • nationally • by research station • by commodity • main discipline (see figure 2.9)	List of all researchers and some of their important features, with a choice of how the list is sorted.	• Allows any researcher to be looked up along with whatever base data is selected to be shown e.g., their station, commodity, main discipline. • Enables the research staff at each station to be seen, along with their main professional features. • Shows the scientists working in each commodity, along with their main professional features. • Lists the scientists within each of the (15) standard disciplines along with their main professional features.
Number of researchers and research activities at each station	Table of stations, with columns for the numbers of researchers and experiments at each.	A small number of researchers at a station suggests relatively high administrative overhead costs, and less opportunity for peer consultation. Justification would usually rest on the station being the only representative of an agriculturally important agroecological zone. [Such a table also allows the numbers of researcher records entered into the system to be compared with the full number as reflected in the payroll. The two should be the same. Differences, usually due to incomplete data capture, call for attention.]
Researchers' main disciplines and highest degrees (see figure 2.11)	Chart of the 15 standard main scientific disciplines and the numbers of scientists in each at BSc, MSc, and PhD degree levels.	Human resource allocation to scientific disciplines should reflect the problems found across all commodities and farming systems. Look for any surprises: disciplines with more, or with less, scientists than you would have expected. Then look for disciplines with a particularly high proportion of BSc degrees and for those with a low proportion of MSc and/or PhD degrees. These may be targets for further higher degree training. See also the Age and Discipline for Staff Forecasts output below.
Age x gender distribution of researchers (see figure 2.13)	Can be presented in the form of a histogram with age along the X axis, where one unit may be 1 year or 5 years.	The top end of the age scale shows any researchers near to the retirement age. This can assist personnel planning through management decisions on recruitment and on higher degree training to fill anticipated gaps. Gender issues can also be addressed here. See the two following outputs.

(continued on next page)

(Table 2.7 cont.)

Report title	Contents	Management notes
Researcher age and discipline for staff forecasts	A table of disciplines with columns for the number of researchers at present, and expected to retire in the next *n* years, separately for MSc and PhD degrees.	This table is used to further analyze the research staff position in relation to disciplines, now and in the future. It shows within each science discipline the expected *losses* due to retirement and *increments* to be expected from those currently studying for higher degrees, and the resulting net number for each discipline. This can advise training and recruitment plans.
List of researchers likely to retire (see figure 2.12)	Provides a listing of named scientists likely to retire.	The result of the analysis shown in the Staff Age & Discipline Forecasts table. As in that output, only those likely to retire on the basis of age are shown.

screen or printed to paper, sorted in any order: by Person ID, Last Name, Main Discipline, Commodity (as in the case illustrated), or Station.

Main discipline

Scientific discipline is an essential attribute of researcher scientists. It is much used in personnel planning and can be useful to designate a standard (controlled) list such as table 2.8, which has been used in several countries. When applied to scientists and research activities, such a list facilitates useful correlations between the scientist pool and the research program. A significant mismatch may indicate that research activities are being pursued by scientists not specialized in that discipline or that scientists of one discipline are having to spread themselves over more research than those of another (see figure 2.5 above). In such a case management would be prompted to see whether the skill base of researchers could be altered to bring it more in line with the needs of the research program, perhaps through in-service or higher degree training and/or recruitment.

This standard list of scientific disciplines has proved useful in many countries, but experience in the field has shown that scientists sometimes feel restricted using just this short list. The solution is to have a second field in the list, for example for subdiscipline or specialism, for station or institute purposes. The contents of this field can be agreed within the station or institute, or better, between all those of that discipline in the country. This field can be useful in the search for a particular specialist in large institutes or in regionally shared data sets.

Highest degree

Data on scientists' main discipline can be aggregated and combined with data on their highest degrees to provide an analysis as shown in figure 2.10.

The distribution of scientists' degree levels in a NARO should be a key concern of management. There is a general view that the bulk of research is conducted by MSc degree holders. A relatively small number of PhD holders is needed to provide the

Page 1/6		Directory of Researchers (Sorted by Commodity)		Printed: 15/07/1998
				Research: 1990-1998
Commodity	Main Discipline	Research Station	Name	Specialism
	No of Researchers: Beans		5	
Beans	Plant breeding	Maastricht	Dijk, Edward, K.	
	Plant breeding	Maastricht	Douane, David, P.A.L.	
	Plant production	Den Haag	Deacon, Howard, H.	Agronomy
	Plant production	Maastricht	Cnand, John, J. C.	Agronomy
	Plant protection	Madurodam	Engels, Kees-Jan, H.	Entomology
	No of Researchers: Cassava		13	
Cassava	Food technology	Maastricht	Halesbury, Robert, R.L.M.	
	Plant breeding	Arnhem	Bogaard, Davies, D. C	
	Plant breeding	Arnhem	Brown, John, A.	
	Plant breeding	Arnhem	Flowers, Roberto, R.	
	Plant production	Arnhem	Brown, Abel, M. S. C.	Agronomy
	Plant production	Groningen	Bloomsbury, Martin	Agronomy
	Plant protection	Arnhem	Bressers, Codrin, C.	Entomology
	Plant protection	Arnhem	Gavel, Kalebe, M.	Biological control
	Plant protection	Arnhem	KattenburKayang, M, A	Pathology
	Plant protection	Arnhem	Short, Charles, C.	Plant pathology
	Plant protection	Den Haag	Brunt, Alfred, J.	Entomology
	Plant protection	Den Haag	Valverde, Juan	Entomology
	Plant protection	Groningen	Breman, Barend	Entomology
	No of Researchers: Cattle		6	
Cattle	Animal health	Maastricht	Krabben, Martin, M.	
	Animal production	Haarlem	Janssen, John, C.	Animal husbandry
	Animal production	Madurodam	Danun, John, S.	Animal nutrition
	Animal production	Madurodam	Davids, Joseph, J.	Animal nutrition
	Animal production	Madurodam	James, T, F.F.	
	Animal production	Madurodam	Welzen, Theresa	Animal nutrition
	No of Researchers: Cattle, beef		1	
Cattle, beef	Animal production	Haarlem	Kraan, Robert, R.	
	No of Researchers: Chickens		1	
Chickens	Animal breeding	Gouda	Janssen, Jim, Y.M	
	No of Researchers: Coffee		1	
Coffee	Plant production	Maastricht	Klein, Fred, N.	Agronomy
	No of Researchers: Cotton		3	
Cotton	Plant breeding	Lelystad	Besten, Miep, W.	
	Plant production	Lelystad	Blanksten, David, D. K.	Agronomy
	Plant protection	Lelystad	Janssen, Ed, K.	Entomology

Note: The MIS from which this was printed can present the list sorted by commodity (as shown here), by research station, by scientific discipline or by person's surname. (All names and research stations have been changed to fictitious names.)

Figure 2.9: Directory of researchers

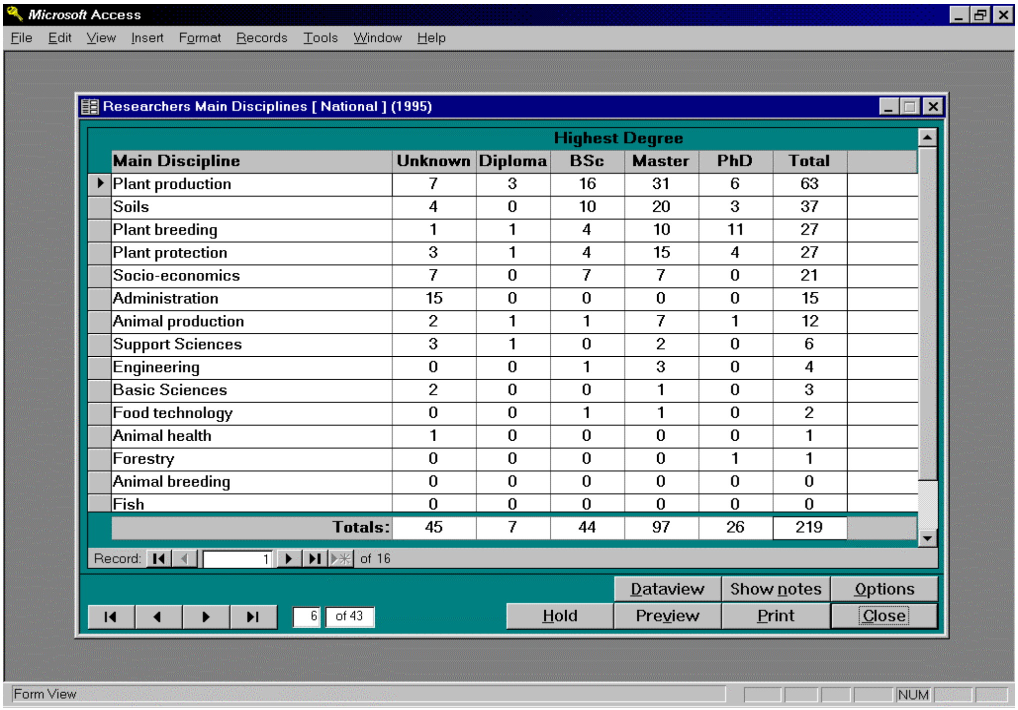
Table 2.8: Scientific Disciplines and Their Specialisms for Research Activities and Scientists

Discipline (controlled terms)	Examples of included specialisms (uncontrolled)
Animal breeding	reproductive physiology, hybrid production
Animal health	veterinary, immunology, toxicology, animal virology, animal parasitology
Animal production	animal husbandry, animal nutrition, animal feeds, animal physiology
Engineering	agricultural buildings, machinery (tillage, harvesting, etc.), postharvest technology (drying, shelling, storage)
Fish	aquaculture, fisheries, limnology
Forestry	
Plant breeding	germplasm collection
Plant production	agroforestry, agronomy, horticulture, irrigation, pastures, plant physiology, tissue culture
Plant protection	biological control, entomology, integrated pest management, nematology, plant pathology, weed science, virology
Food technology	human nutrition
Socio-economics	economics, farming systems, production statistics, sociology
Soils	soil sciences, soil surveys, soil conservation, soil testing
Basic sciences	physics, chemistry, biology, ecology
Support sciences	biometry, computing, extension, information, library, meteorology, microbiology, hydrology, chemical analysis, production (e.g., seeds, vaccines)

inspiration for new ideas. New graduates with BSc degrees provide a source of future scientists through an apprenticeship of field experience and, for those showing promise, MSc degree training. Figure 2.10 presents the manager with a view of the current position. Figure 2.11 adds current degree-training data and date-of-birth data to provide a forecast of the staff position at any selected year ahead. It forecasts future numbers of scientists in each discipline and for each degree level; to any specified number of years ahead, based on the current level; those currently in degree training; and attrition from retirement. Figure 2.12 then identifies the scientists scheduled to retire within the time interval specified. For each retiree it shows the station at which they are located and the commodity and discipline of their work. A manager can see at a glance gaps that are likely to arise in the coming years, and can take the necessary actions such as recruitment, redeployment, or training, as appropriate.

Measuring researchers' input as a resource

Researchers are a major part of the cost of research and their cost is therefore a key factor in research management. Information on scientists' time allocation to research activities, as a proxy for their cost, is therefore a valuable contribution to mon-



	Highest Degree					
Main Discipline	Unknown	Diploma	BSc	Master	PhD	Total
Plant production	7	3	16	31	6	63
Soils	4	0	10	20	3	37
Plant breeding	1	1	4	10	11	27
Plant protection	3	1	4	15	4	27
Socio-economics	7	0	7	7	0	21
Administration	15	0	0	0	0	15
Animal production	2	1	1	7	1	12
Support Sciences	3	1	0	2	0	6
Engineering	0	0	1	3	0	4
Basic Sciences	2	0	0	1	0	3
Food technology	0	0	1	1	0	2
Animal health	1	0	0	0	0	1
Forestry	0	0	0	0	1	1
Animal breeding	0	0	0	0	0	0
Fish	0	0	0	0	0	0
Totals:	45	7	44	97	26	219

Management Notes

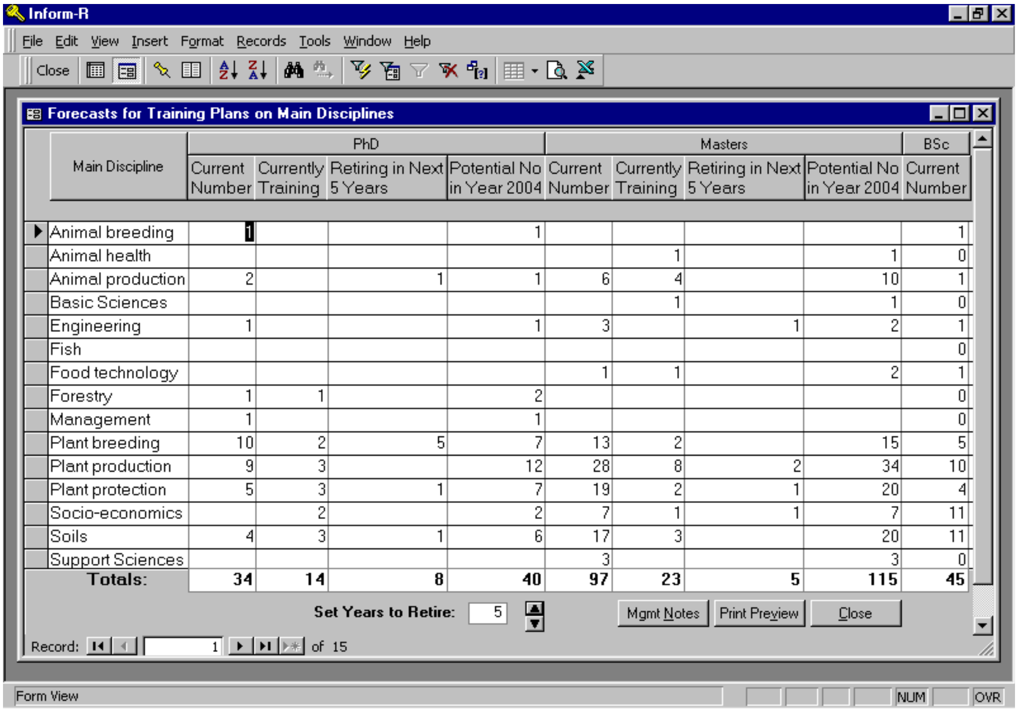
Human resource allocation to scientific disciplines should reflect the problems found across all commodities and farming systems. Look for any surprises: disciplines with more, or with less, scientists than you would have expected. Then look for disciplines with a particularly high proportion of BSc degrees, and for those with a low proportion of MSc and/or PhD degrees. These may be targets for further higher degree training. See also the outputs Forecasts for Training Plans in Main Discipline* and List of Researchers Likely to Retire**.

* figure 2.11; ** figure 2.12

Figure 2.10: MIS output of researchers' main disciplines and highest degrees

itoring a research program. The simplest measure is the number of weeks worked on each activity in a given time period, e.g., the year, cropping season, or month.

In one method each researcher is asked first to divide their total working time into the major time slots: research, management, extension, full-time/long-term study, short-term training, teaching, and extension, using percentage figures (see table 2.9). Then each is asked to split the percentage time allocated to research among the several research activities (projects or experiments). In this exercise, a week is taken as 2% of a year. The collection of scientist time data is fairly straightforward. The use of time sheets need not be very demanding on staff. The subject is discussed in greater detail by Gijsbers and Fraser in section 5.1.



Main Discipline	PhD				Masters				BSc
	Current Number	Currently Training	Retiring in Next 5 Years	Potential No in Year 2004	Current Number	Currently Training	Retiring in Next 5 Years	Potential No in Year 2004	Current Number
Animal breeding	1			1					1
Animal health						1		1	0
Animal production	2		1	1	6	4		10	1
Basic Sciences						1		1	0
Engineering	1			1	3		1	2	1
Fish									0
Food technology					1	1		2	1
Forestry	1	1		2					0
Management	1			1					0
Plant breeding	10	2	5	7	13	2		15	5
Plant production	9	3		12	28	8	2	34	10
Plant protection	5	3	1	7	19	2	1	20	4
Socio-economics		2		2	7	1	1	7	11
Soils	4	3	1	6	17	3		20	11
Support Sciences					3			3	0
Totals:	**34**	**14**	**8**	**40**	**97**	**23**	**5**	**115**	**45**

Management Notes

This table is used to analyze the research staff position in relation to the main scientific disciplines, now and in the future.

It shows expected losses due to retirement, and the expected gains from those currently studying for higher degrees. The "Years to Retire" can be set to look ahead for any desired period.

Figure 2.11: MIS output showing forecasts of scientists by main discipline and highest degree

The number of person-weeks allocated to all projects in a certain discipline, program, or region can provide a useful approximation of resource allocation. This indicator can easily be converted into monetary cost if salaries, benefits, and allowances are known. However, this deduced monetary value is not the true monetary cost—it is a proxy for the true cost and while proportionally sufficiently in line with true costs for research program management, it is not the real thing as needed for financial accounting. It may still be the most accurate indicator of resource allocation available to managers and has proved valuable in some NARS.

Data collection frequency

How frequently scientist time data needs to be collected has to be considered carefully. Points of view to consider are the researchers who have to provide the data, the

List of Researchers likely to Retire

Name	Age	Retire	Research Station	Commodity	Main Discipline	Specialism
Groen, James, J.	55	0	Den Haag	many	Plant protection	Entomology
Baumen, Sjaak, J. A.	54	1	Amsterdam	Kenaf	Plant production	Agronomy
Gardingen, Toon, K.	54	1	Den Haag	Soils	Soils	Soil chemistry
Jasper, Jules, F	53	2	Almelo	Soybeans	Plant production	Agronomy
Douane, Nigel, P.	52	3	Haarlem	Many	Plant breeding	
Hazebroek, Josy, B	51	4	Zwolle	Irrigation	Engineering	
Burgh, Eduard, B. L.	50	5	Maastricht	Millet, finger	Plant breeding	
Deelder, Donald, M.	50	5	Madurodam	Many	Plant production	Agronomy
Ent, Bruce, B.M.	49	6	Den Bosch	Groundnuts	Plant breeding	
Bruyn, Abel, M. S. C.	49	6	Arnhem	Cassava	Plant production	Agronomy
Hento, Ian, S K	49	6	Zwolle	Many	Plant production	Agronomy
Janssen, Jan, M.	49	6	Den Haag	Soils	Soils	Soil science
Captein, Pim, M.	47	8	Maastricht	Soils	Soils	Soil science
Geus, Theo, W. M.	47	8	Den Haag	Many	Management	
KattenburKayang, M, A	47	8	Arnhem	Cassava	Plant protection	Pathology
Janssen, T, F.F.	47	8	Madurodam	Cattle	Animal Production	Animal Nutrition
Groeneberg, Edward, E	47	8	Den Haag	Many	Plant protection	Entomology
Henkes, Percy, P.C	47	8	Den Bosch	Multipurpose trees	Plant production	Agronomy
Gatsby, George, C.	47	8	Den Haag	Many	Plant protection	Plant pathology

Set Years to Retire: 8 Mgmt Notes Select Print Preview Close

Record: 1 of 19

Note: The number of years that the system looks ahead can be selected. It is currently set to eight years. Data is from an actual case but person and station names are fictitious.

Management Notes

The result, in terms of actual names of expected retirees, of the analysis shown in the Forecasts for Training* output. As in that output, only those likely to retire on the basis of age are shown.

* figure 2.11

Figure 2.12: Table of researchers due to retire and their commodity group and highest degree

information practitioner involved in processing the data, and the manager using the data. Experience shows that collecting data once every season is sufficient and efficient. More frequent collection needs to justify the extra work involved.

Other important attributes of researchers

The "Persons" data tables in annex 2 "Database Tables and Fields in INFORM-R" give examples of researcher attributes that are useful in research management.

2.3.2 Research activities

The term research activity is here used as a generic term for the various levels in the hierarchy of experiments, projects, and programs. (See the glossary in annex 1 for general definitions of these categories.)

Table 2.9: Examples of Scientists' Data Stored in an MIS for the Research Program

a) General Data

Personal	**Admin**	**Scientific**
Name	ID number	Highest degree
Date of birth	Rank	Scientific discipline
Gender	Job title	Main species (crop or animal) or noncommodity factor

b) Time Allocation Data, and Links to Research Activities

Percentage time allocations	**%***	**Research time breakdown**	**%***
Research	80	Experiment / study A	40
Management / Admin	10	Experiment / study B	20
Extension	10	Experiment / study C	10
Training		Experiment / study D	10
Degree studies		Experiment / study E	
Other		Experiment / study F	
		Experiment / study G	
Total time	100		80

* % person year, where 1 week = approx. 2%

The term experiment is here used somewhat loosely. Strictly it refers to a research activity where conditions are deliberately altered in a controlled way, and the effects of such imposed change on one or more other factors are measured. Observations are research activities where, typically, no such change is induced; one measures the undisturbed situation. Surveys, such as soil, botanical, socioeconomic, and pest incidence surveys, are examples of observations. Here we use the term experiments to mean the lowest level research activity for which a protocol is prepared and to include such observational research.

The actual number of experiments is the most basic measure of research activity (see figure 2.3 on page 26), but it is a crude measure due to the large variability in size. A maize agronomy trial, for example, may conclude in three seasons and have a small number of measured parameters. A perennial tree crop or cattle breeding trial, on the other hand, will take many years, and a socioeconomic survey of many villages may embrace a wide geographic and cultural diversity and a large number of factors. So while the number of experiments in each commodity, station, or discipline does provide some guidance to resource allocation, it is improved if augmented with researcher time or financial expenditure before aggregating up from each experiment's use of these resources.

MIS fields for research activities

The design of an MIS is addressed in part 2 of this volume, but we can already state here that the careful selection of fields in an MIS can contribute much to its value and ease of use. The contents of some fields, such as Research Station Name, will be drawn from a known list, which can be built into the system. The provision of such lists makes for much more consistent data, which in turn leads to more accurate reports.

A **title** is needed to identify each experiment. The title is an important device to (1) inform a reader of the essence of the research activity and (2) enable the study to be referred to, easily and meaningfully, in conversation and reports. For the first purpose it should probably include the main commodity or commodity group or factor under study and the nature of the study. This is often shown by the main discipline used. And for the second, it should be concise, but not too concise. Table 2.10 shows some examples taken from an MIS implementation, with titles that are too terse or verbose, and enhanced experiment titles. The enhancements have been made after looking at other fields, particularly that for the objectives.

An **identification code** (ID) is a short and unique identification term for each research experiment or project. There is merit, from the view of good database design, in ensuring that no meaning is imbedded in such an ID. The ID often carries addi-

Table 2.10: Examples of Titles That Are Too Terse or Verbose, and Suggested Improvements

Original experiment title	Objectives	Improved experiment title
Sunflower Preliminary Trial	To determine the performance of newly developed hybrids in terms of seed, oil and disease resistance. To compare the performance of different sources of germplasm for locally important traits. To screen all elite material for adaptability and combining ability in all agroecological zones.	Sunflower Preliminary Germplasm Screening Trial
National Variety Trial	To evaluate overall performance of advanced varieties across diverse environments. To identify promising varieties for future production.	National Maize Variety Trial
The Effects of Establishing Two Year Fallows by Drilling or Potted Seedlings in the Presence and Absence of Crops on Fallow Performance, Maize, and Beans	To examine the potential of establishing direct drilled and raised seedlings of Sesbania to increase economic return during the year of improved fallow establishment seeding. *Justification:* Mitigation of deforestation, provision of fuel (fire wood) and increased crop yields; under reduced or no fertilizer application are three important attributes of improved fallow technology. The demerit is that labor is invested without any crop yield benefits in the year of fallow establishment. The aim of this experiment is to make this technology more economically viable by intercropping crops with multipurpose tree species in the fallow period.	Sesbania for Enhancing Fallows

tional information, such as year or station, but these are better dealt with by specific fields for that purpose.[9]

Commodity, together with noncommodity factors such as soils and irrigation, is an essential attribute of most research activities. Annex 3 lists some 200 commodities, mostly animal and plant species; each is attributed to one of about 20 commodity groups. This may provide a useful reference list for system designers. Some NAROs will still need to add some extra names such as those of local vegetables.

Scientific **discipline** is also an essential attribute of most research activities and is much used in research planning and monitoring. It applies also to researchers (see section 2.3.1). Table 2.8 above gives a suggested standard list of disciplines, along with examples of their narrower "specialism" terms.

A **key word** field is useful for attributes not covered by specific fields for the more commonly used attributes such as commodity, discipline, main pest, station, and AEZ. Examples of such exceptional search terms include varieties of plant commodities, breeds of livestock, and strains of pathogens (the last one may actually be found in a "Main Pest" field, which usually is uncontrolled due to the large number of terms).

The careful appraisal of an experiment is important prior to its approval. Additional information is needed for this purpose: **objectives**, the intended **treatments**, the **statistical design**, the **site(s)** it is to be effected at, and the **number of years** it is to continue, and the **responsible scientist** and names of other scientists who may be involved. (See the "Experiments" database tables in annex 2 for a fuller list of experiment descriptors.)

On-farm research, soil surveys, and socioeconomic surveys call for considerable extra information in addition to that of the standard station experiment. Mutsaers et al. (1997) provide a more detailed discussion. Much of this information is variable and may not be stored in the regular MIS. The system will however contain for each such study the same core data as for station experiments and may include some additional fields as shown in table 2.11.

Experiment results

There is a good case nowadays for storing experimental results electronically. Electronic storage continues to become cheaper and easier as the cost of computer hardware declines and the power of software increases. It enables the data to be processed, stored, copied, and transmitted to others more easily, securely, and cheaply than in the past.

Recording the data from experiments in an MIS database presents some difficulties. It requires a match between the MIS structure and the data, and research results may not occur in the same form. It will be difficult to plan for all possible data structures

9. Readers wishing to explore the reasoning behind this advice can find a good account of it in Date, C.J. 1992. Relational Database Writings, 1989-1991 pp. 461-466. Reading, USA: Addison-Wesley.

Table 2.11: Examples of Extra Research Activity Fields Appropriate for Surveys

Field	Description or examples
Experiment ID	a unique code for all research activities
Survey type	e.g., adoption rate; botanical: taxonomic; botanical: genetic resources; case study; diagnostic; ecological; exploratory; farming systems; impact; monitoring; participatory rural appraisal; zoning
Target population	the group of people who are the object of the survey
Sampling frame	a description of how the sample of the survey is to be defined
Sample unit	e.g., household, village, cooperative, agroecological zone, farmer's field
Sample size	e.g., four agroecological zones
Data collection method	e.g., questionnaire, field survey

that may ever arise or even to cater for a large range. The best option may be to provide for the most common data types and design a flexible database.

Four levels of data coverage can be recognized. The base level is restricted to "metadata:" a record is provided for each experiment of what results data exists. This will enable researchers embarking on new research to find out what has been done before and whether and where the results data are to be found.

The next level provides a text field for a short account summarizing the results and conclusions of the work. Many database software products have a "memo" field type, which can accommodate an account of up to several hundred or a few thousand words. It will usually include some numeric data within the text.

The third level of results coverage can store actual results data stored in separate fields for the purpose, so that such data can be manipulated in different ways in system-generated reports. At this level only the means of each treatment and of the whole experiment are stored rather than data for each plot or sample. It also holds summary data indicating its quality, such as the coefficient of variation, the statistical design used, and the number of replicates. There may also be some site data, such as soil pH and previous year's land use, and nontreatment management factors such as fertilizer applications and crop variety used, and there may be a place for comments. These all assist the reader in interpreting the results and are particularly useful if subsequent comparisons are to be made of time series data or data across several sites.[10]

The fourth and most detailed level stores additionally data from every plot. This adds very considerably to the data volume and complexity and takes the system from a generalized agricultural research MIS to a specialist system of particular interest to the biometrician. Some statistical analysis packages provide this facility and require all plot data to be entered in order to run the analysis. Its main purpose is to facilitate

10. ISNAR's INFORM-R operates at this level and has four database tables applied to research results. Details are given in annex 2.

statistical analysis. Storing data in this way also allows for re-analysis later, for example using new statistical techniques.[11]

2.3.3 Completed research: Data archive

In an ideal situation completed research is recorded more or less fully in annual reports, and the more significant results are published in scientific journals. This helps researchers presented with a farming problem to see whether similar problems have been encountered before and what research has been done locally on the topic (see box 2.3 for an illustration of the value of traditional records). Unfortunately, complete runs of local journals and annual reports going back several decades are no longer available in many countries. Electronic information systems may offer an alternative.

An agricultural research MIS is stocked initially with the data of current (ongoing) research projects. The system needs to distinguish the data from different years for projects in their second or subsequent year and enable these data sets to be aggregated, so that when a project is completed, its final report can be prepared, summarizing the records of all years.

Records of completed research may be held in a separate "archive" system in a simple MIS. More sophisticated systems will link current and past research records to facilitate a search for what has been done before locally on any particular topic. This is important in the planning of new research projects. The amount of detail may vary from just a title, ID, and year, pointing to more detail in paper records, to a record of a summary report complete with results data as described in section 2.3.2 above.

Box 2.3: The Value of Good Archives of Past Research

In the 1970s Witchweed (*Striga asiatica*) became a significant problem in maize production for resource-poor farmers in parts of Southern Africa growing maize continuously on the same ground with low inputs. In severe cases crops yields were reduced to almost zero. Studies of research elsewhere identified some herbicides with reported activity against this weed, but a series of experiments over several years failed to find a treatment that farmers could successfully apply.

Solutions were eventually identified in research done four decades previously. In the mid-1930s Witchweed was also reported to be the most serious problem in maize production in parts of Southern Africa. Intensive research at that time identified a series of agronomic measures that provided effective control of the weed.

These solutions came to light only because they had been thoroughly described in the local agricultural journals of that time and because a complete run of that journal was held in the local research station library.

11. Examples of such packages include Genstat, Minitab, and SAS/STAT.

A third function of archived data is to support the preparation of subject reviews. An MIS with many years of archive research data can offer an efficient source of very relevant information for a review article or extension literature. In a pilot survey of the information needs of research managers such archive data ranked one of the most important (see box 4.3 on page 145).

Many countries have half a century or more of previous research recorded in paper files. This material can be transferred to electronic media (Potts et al. 1996). A number of international aid projects provide support for such initiatives (Chitala Zimba and David Parker, personal communication, 1997; also see the case study from Zambia in section 3.1).

> "The Electronic Rothamsted Archive (ERA) provides a permanent managed database for the secure storage of disparate data relating to Rothamsted's long-term experiments. . . . The famous Rothamsted Classical Experiments were started during the last century and have continued to this day. Such long runs of data are increasingly in demand for research into issues such as the impact of climate change and of pollution" (Potts et al. 1996).

2.3.4 Areas of special interest

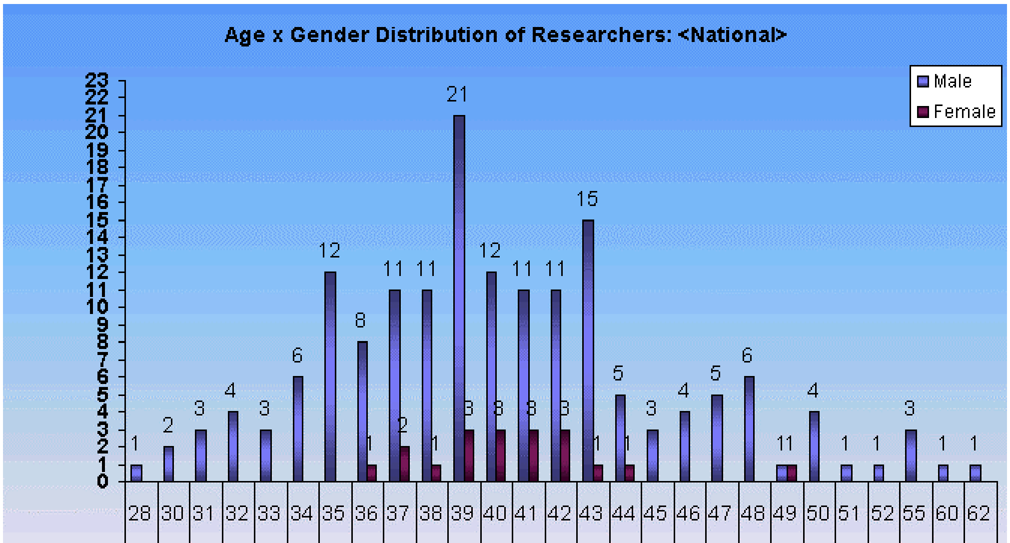
An MIS can be an efficient tool to track the degree to which a research program is addressing particular issues. The selection of issues will vary with the country and the research organization and with time as new issues gain importance. It can be helpful for a NARO to be able to demonstrate to donors with an interest in a particular issue the degree to which a program supports the issue. It can provide detailed information about its efforts in a particular issue by an ad hoc search of selected fields, such as project title and objectives. Some examples of issues and suggested indicators follow.

Sustainable agriculture

The term sustainable agriculture is generally used to indicate a broadening of focus from traditional commodity-based research for maximum production to include research on the environmental side effects of plant and animal production. It applies at three levels: national, farm, and research policy and program. We will focus on the last level here. Indicative terms include soil erosion, water and air pollution, nonpoint pollution, soil salination, zero tillage, pesticide contamination (of produce), (emergence of) pesticide-resistant pests, and "high precision" farming techniques, such as integrated pest management and variable fertilizer application. For a more detailed consideration see Runge (1992).

Gender

An MIS can assist in determining whether a project has a significant gender component and in providing information on the share of the budget devoted to gender-related activities. It can also demonstrate very clearly the status of women in a research service. Figure 2.13 shows not only the proportion of female and male scientists but also their age distribution, which in turn provides further leads to explore why none are older than 44; was this due to a late start in recruitment or are they



	28	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	55	60	62
Male	1	2	3	4	3	6	12	8	11	11	21	12	11	11	15	5	3	4	5	6	1	4	1	1	3	1	1
Female								1	2	1	3	3	3	3	1	1					1						

Management Notes

The top end of the age scale shows any researchers near to the retirement age. This can assist personnel planning through management decisions on recruitment and on higher degree training to fill anticipated gaps. Gender issues can also be addressed here.

Figure 2.13: MIS output showing the age and gender distribution of scientists

leaving the service at this age and is this due to family issues, frustration at the lack of promotion opportunities, or some other factor(s)? In the case illustrated, the exceptionally high number of 39-year-old males may be due to the ministry of agriculture having been under pressure some years earlier to employ the country's first graduating class of a new agricultural faculty at the local university.

Client-oriented research

Information on farmer involvement in research and linkages to farmers' organizations or the private sector may be built into the MIS. An MIS could readily report on the number of on-farm experiments, and the number of experiments with outputs that have been passed to the extension service or published in local farming literature.

2.3.5 Internal and external context

A research organization does not exist in isolation. It serves its clients and it is accountable to its funding sources. External political, economic, and social forces can have a major effect on the opportunities and constraints of a research organization— they set limits on what is possible and influence actively or passively the direction of research, style of management, and degree of autonomy, recruitment, and reward practices. Information on such parameters is important to senior management, charged with steering the organization through the constraints to the opportunities. As such, the information will be mainly of use in setting policies at the NARO's headquarters level and may have little value in the day-to-day management of the research program.[12]

We divide such information into internal and external information, i.e., that coming from within the research organization and outside it. The MIS may not be the only or the best medium for storing and distributing context information. Decisions are therefore needed on which information is required, where it is to be collected from, where it is to be stored, and how frequently it is to be updated. Table 2.12 offers examples for the first of these decisions. Sources have to be locally determined. The same applies to the choice of storage site; one could simply record the titles of data items agreed to be useful, along with their location to enable them to be derived as and when needed. With the rapid development of telecommunications and on-line data services these may provide an efficient repository of information needed by different sectors. But most NAROs do not yet have such good connectivity to information repositories. In such cases the manual collection and entry into a NARO-wide MIS could bring valuable information to all research managers. Steps will be needed to ensure its regular updating.

The internal environment

The environment in which researchers (arguably the most important entity in a research institution) operate may profoundly influence their motivation and hence performance and outputs. The MIS can provide the research manager with current information on the factors involved, such as information on researchers' salary levels and scales, promotion records, prospects, and turnover.

One important indicator is the level of operational funding per researcher. This is the total cost of research less that of salaries and wages and capital expenditure, divided by the number of researchers. While salary and wage costs vary greatly between countries, there is evidence that operational costs tend to be much the same worldwide. Costs of vehicles, fuel, chemicals, cement, electricity, and laboratory equipment are generally more expensive in the South while subsistence allowances for off-station work etc. will usually be lower. Some work by ISNAR suggests a figure of US$28,000 for operational cost per researcher (Dagg, personal communication). Most NAROs operate with a much lower figure however.

---

12. It may be useful, however, to make the information accessible to station directors, encouraging their interest in policy issues.

Table 2.12: Some Indicators of Agricultural Research Context

Item	Reference figure	Use in research management
A. Budgetary		
Research expenditure (annual research budget)	Not applicable.	Can be compared graphically with trends in (a) the AgDP and (b) agricultural export value. If both are declining, then it provides a strong case for increase in research expenditure.
Operational expenditure per researcher (Expenditure per researcher net of salary)	While researcher salary costs vary greatly between countries, operational costs tend to be much the same worldwide. Initial work by ISNAR suggests a figure of US$28,000 as an approximate target.	Costs of vehicles, fuel, chemicals, cement, electricity, laboratory equipment are more expensive in the South, while per diems for off-station work etc. will usually be lower. A low figure indicates a reduced chance of productive research.
Research expenditure as % of AGDP (also known as the agricultural research intensity ratio)	A 1980 World Bank Policy Paper suggested 2.0% as target for 1990. Mean for developed countries 1980-85: 2.03. South Africa 1991: 2.59. USA 1992: 2.22. Most DCs are < 1.0%.	The 2% value has no fundamental basis. But using it as a target value is justified by the high rates of return (ROR) from investment in agricultural research. Despite some ambiguity in ROR studies it is clear that ROR from research is much higher than government borrowing rates, suggesting this to be a valuable investment by governments.
AGDP as % of GDP	Mean for Africa, 1990 = 31 % (World Bank report).	Relates the agricultural sector to the national economy. If the trend is downward, as expected in economies moving to an industrial base, it will be harder to justify an increase in investment in agricultural research. But there may still be a case for increase in research on commodities with industrial use or for export. If the trend is upward, it supports a corresponding increase in allocations to agricultural research, particularly if the research expenditure as %AGDP is lower than the national target figure or, if this has not been agreed, the mean for the region is low.
Researchers per million farmers	Mean for 23 SSA* countries in 1991 = 70.	Has to be used in conjunction with measures of personnel quality. An MIS should have a standard output of numbers of PhD, MSc, and BSc staff and preferably a predictive facility based on numbers currently training and expected attrition on retirement.

* SSA = sub-Saharan Africa

(continued on next page)

(Table 2.12 cont'd)

Item	Reference figure	Use in research management
% of research expenditure provided by donors	Mean for 11 SSA nations was 35% in 1986, rising to 43% in 1991. Some exceed 60%.	A high figure represents gain to the country but may also carry risk. The most immediate is the danger of influencing resource allocation in ways that may not be in accord with agreed priorities. It then becomes important to regularly generate a report of actual resource allocation compared with agreed priorities. This can be used to encourage donors to target needed areas. In the longer term a high figure may create dependency with risk of dislocation of the research program should donor funding end.
Agricultural gross domestic product (AGDP)	Not applicable.	Provides tangible evidence of the state of the agriculture sector. Compare with investment in agricultural research. A decline in the former is an argument for increase in the latter particularly if the research expenditure as %AGDP is low.
External debt	25 of 32 "severely indebted"*countries are in SSA. SSA has debts of US$235 billion = 76% GDP, or 236% exports. (1997)	A high figure may recommend more active pursuit of grant aid or aid on favorable terms.
B. Land Factors		
Irrigation as % agricultural land	Average for Africa is 2%, Latin America 7%, and Asia 27%.	The green revolution of Asia was based on plant breeding for moist conditions. Africa's much drier environment is likely to depend more on plant production/agronomy, plant protection, and post-harvest work, except in places where this figure is high relative to the region.
Fertilizer use: kg/ ha of arable land	Mean for low income economies, excluding China and India, 1991-92 = 40.3.	A figure substantially lower than the mean for the region may be due to economic factors, for example the high costs in a land-locked country. But it may also indicate more work is needed into optimal fertilizer rates, calculations on the returns to the farmer of using these rates, and dissemination of this information.
Technical efficiency gap	Has been shown to be as high as 30%.	This is the gap between farmer yield and that achieved *at the same input level* with best knowledge and skills applied. Indicates the prospects for research to raise farm productivity without raising input costs, e.g., fertilizer timing and placement, water use, and pesticide application methods.

* Severely indebted low-income countries are defined as those with debt-to-GNP and debt-to-export ratios greater than 80% and 220 % respectively, and GNP per capita less than US $675.

(continued on next page)

(Table 2.12 cont'd)

Item	Reference figure	Use in research management
C. Population Factors		
Economically active agricultural population as % of economically active population	Mean for Africa, 1990 = 70% (World Bank report).	A high figure demonstrates that agriculture serves not only food security but also as a source of employment. This is a legitimate argument for a comparably high research allocation as % of AGDP.
Extension staff per million farmers		A research facility is very dependent on an effective extension service.
Current population and annual growth rate		Various estimates of food needs in comparison with food production and imports.
Per capita food availability	3,300 cal/day in the North, 2,800 in developing countries, 2,100 in SSA (1990).	Various estimates of food needs in comparison with food production and imports.
Urban proportion of the population	Urban population in developing countries is growing much faster than the rural and may exceed it by 2020.	Some argue urban food production to be included in the research program.
D. Import and Export Factors		
Agricultural imports value		A high figure prompts consideration of research aimed at import substitution through local production. The absolute figure of each commodity may advise the scale of research justified in each case.
Agricultural imports as % total imports		A high figure prompts consideration of research aimed at import substitution through local production.
Main agricultural imports	Not applicable.	These are candidates for consideration for research to support local production.
Agricultural exports value	Not applicable.	In research program priority setting.
Agricultural exports as % total exports		A high figure can be used to demonstrate the importance of agriculture to the national economy and the need to support it through research.
Main agricultural exports	Not applicable.	A check should be made that these commodities are adequately served by research, particularly if their yields are low in comparison with known potential yields.

(continued on next page)

(Table 2.12: cont'd)

Item	Reference figure	Use in research management
World prices of imported / exported commodities	Beef: while per capita consumption in emerging economies in the South is increasing, there has been a steady decline in the North over the past 20 years, due largely to health considerations.	This kind of information may advise the proportions of resource allocations allocated between commodities.

With budgets falling and the number of researchers increasing, there has been a serious decline in the total expenditure per researcher in many countries. Operational costs per researcher have often sunk to levels seriously undermining the research program. An early and serious casualty of this is a decline in on-farm adaptive research, as this type of research depends on adequate travel and subsistence funds. Ironically, where some donor programs have directed support exclusively to this use, scientists spent a lot of time traveling and attending workshops, devoting less time to research. An MIS can easily identify and quantify the effects of this, but there is a risk of shooting the messenger rather than resolving the cause—there is a temptation for those providing the necessary data to withhold it or supply false data. This underlines the importance of an information system patron in top management to ensure that appropriate corrective actions are taken at the highest level.

The external environment

The external environment includes the administrative and legal, technological, political, economic, and social and cultural regimes of the country (Lusthaus et al. 1995). All these are relevant to understanding an organization's performance and attempting the management of change. Here we are focusing on those issues, often quantitative, that can be more precisely defined. Managers with a knowledge of these issues can take actions that enable the organization to respond to conditions in the external environment.

Figure 2.14 shows an example, using fertilizer use on the farm and its variation over years. Two conclusions can be drawn: first, a steady increase occurred from 1971 to 1988, followed by a sudden 33% drop in 1990. This coincided with the removal of state fertilizer subsidies. Second, these rates are well below agronomic optima, ignoring in this instance economic data such as fertilizer and commodity prices.

This contextual information necessarily comes from many sources. It is often incomplete, inconsistent or plainly unreliable, perhaps even manipulated (Schutzelaars et al. 1995). An MIS can greatly enhance its quality, quantity, and timeliness. The first step is to include in the MIS design process an analysis of what information is required. This will identify those information items that are available, their most appro-

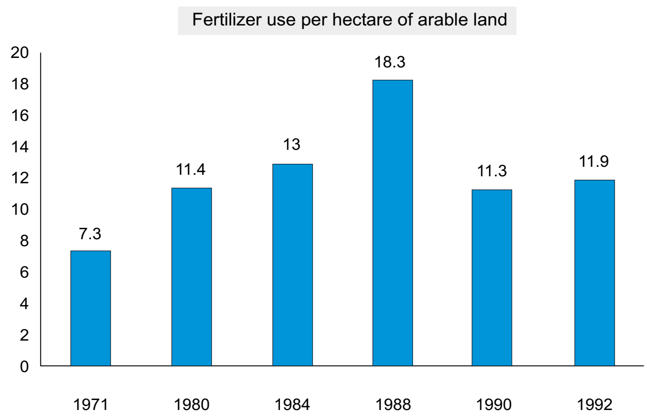


Figure 2.14: An example of context information

priate sources, and mechanisms to tap them in a reliable manner and at intervals ensuring sufficient but not excessive accuracy and timeliness.

For much contextual data it will be necessary to have some or all of the following:

- a recognizable title of the data
- the current value
- a record of the value over time (the trend)
- a reference for comparison (e.g., the mean for all countries in the region)
- a local goal or target figure that has been agreed at ministry level

The MIS may additionally offer guidance on how the item may be used to enhance the research program. Table 2.12 above gives examples from four areas: budgetary, land, human population, and imports and exports.

Factors should only be included if they are meaningful. For example, the number of researchers per million hectares of agricultural land is too imprecise a statistic for comparison with research services in other countries, because the differences between, say, intensive horticulture and pasture rangeland are huge. Researchers per million farmers or per million hectares of arable land would be more meaningful.

Caution is also needed in using international comparisons, particularly between different regions. Intraregional comparisons and trends over time within one country may be more useful. Some ratio indicators may be useful policy targets without assuming a cause-and-effect link. For example, between the 1960s and the 1980s, research expenditure as a percentage of AgDP doubled in both less developed and more developed countries. This may be encouraging, but output per dollar invested in research declined by about half during this period. Is this a case of diminishing returns? Probably not; there are other significant factors involved.

The World Wide Web is increasingly used by commercial businesses for relevant external information, and it should prove useful to NAROs too. The limiting factor at present is less one of content (see annex 4 "Worldwide Agricultural Research Information Systems and Sources.") than access to the Internet in developing regions and particularly rural areas. Where access is available, further problems are the unreliable quality of the information and the low proportion of relevant information resulting from a particular search.

Hackathorn (1997) suggests some steps to address the latter problem by establishing a facility to routinely derive information from the Web. First, it is necessary to agree on a provisional prioritized set of factors that are considered important to be informed on. Next a systematic plan for searching the Web for such information is formulated. Then processes are established to format identified material so that it can readily be presented to users, usually as memos, reports, spreadsheets, tables, notes or charts. These are distributed at once to senior managers likely to be interested and feedback is invited from users on the value of the information supplied. The steps so far might be regarded as a pilot to test the value of the concept. If feedback is sufficiently positive and where the size of the organization warrants it, the next step is to present management with a plan with costs for the additional resources. This will have to include an appraisal of the feedback from users. If approved such a project would integrate the structure of Web-derived data into the existing information system.

2.3.6 Regional information sharing

by Richard Vernon and Hope Webber

NAROs with similar environments that share information on, for example, commodities, soils, climate, and farming systems have much to gain: scientists can benefit from exchange of research plans and research results with fellow scientists working in the same commodity and discipline, and research managers can make best use of scarce resources by planning research programs in collaboration with their peers in neighboring countries. Some regional initiatives already recognize this:

> "It is unlikely that any of the countries in the sub-region can individually set up and finance a national agricultural research system which has the critical mass of scientists and facilities to adequately cover all these (...) 101 (...) commodities/factors of production. (...) The Team recommends that ASARECA should establish and maintain a database of all major research programmes/projects being undertaken by its member institutions and avail these to the other NARI's / partners to avoid duplication of efforts and optimise the utilisation of research in the subregion." (ASARECA 1997)

It is clearly more efficient to build on what has already been done before elsewhere or is currently being done than to unknowingly duplicate research. A NARO may lack the necessary scientific discipline for a specific, new agricultural problem, but it may exist in a neighboring country. And some topics need not be covered by every country. The costs involved in some kinds of research may be too high for a single NARO to bear and could be shared with other countries, e.g., the services of a plant virologist with an electronic microscope or a veterinary diagnostic laboratory for a particular disease requiring special and expensive equipment.

Needs of the scientists

Getting the results of completed research clearly helps scientists conduct research more efficiently and effectively. New research projects can start several steps ahead instead of repeating what has been done elsewhere. In the past, detailed annual reports, if produced and distributed in a timely fashion, and a wide range of international journals, provided just such information. But those days have gone for many regions. Knowing which other scientists are working on the same commodity or discipline in the region helps the researcher to establish contact and exchange on a personal level each other's plans, ideas, programs, and eventually research results. This is a key component of networking.

Needs of research managers and policymakers

Galante et al. (1995) identify the information needed to enhance research cooperation and coordination efforts, particularly at the policy level:

- a profile of a national agricultural research system, its agricultural research organizations and their sponsors
- national research programs and major coordination networks
- research institutions: not limited to contact information but also including a description of their research fields, expertise, and organization structure
- researchers and experts (including details about their fields of specialization)
- project reports
- research results
- potential users and application areas (to make it possible to estimate the impact and the application potential of research projects).

Facilitating regional information exchange

One of the simplest exchanges is that of a directory of scientists along with their commodity focus and scientific discipline and location (research station). An MIS can assist in (regularly) exchanging either printed directories or a diskette with the information on it. If the research organizations use similar software it should be possible with moderate information technology skills to merge the necessary data from their systems to provide a composite or regional directory. Webber (1999) has examined this, using data aggregated from 10 NAROs that were using ISNAR's INFORM MIS. Figure 2.15 shows a page from a (fictional) regional directory of scientists that was derived from this combined data set. Such information provides scientists and managers from neighboring countries with access to aggregated data from their combined management information systems. Similarly, with an MIS a list of research projects (a regional version of figure 2.6) can be produced annually with very little extra effort. Figure 2.16 shows how such an aggregated system can provide useful information for regional research planning. In this case the question posed was: "What types of biotechnology research are currently in progress in the region?"

18-Mar-99 **Global MIS Data** Page 1 of 46

Regional Directory of Researchers (Sorted by Main Discipline)

Main Discipline	*Commodity*	*Station*	*Country*	*Surname*	*First name*	*Initials*	*Specialism*
Animal breeding							
	cattle	Madurodam	Shangrila	Henning	Harold	S. C.	repr.phys
	Chickens	Gouda	Northland	Janssen	Jim	Y.M	
	Silkworms	Zwolle	Red Country	Oyen	Antony	A.C.	
Animal Health							
	Cattle	Maastricht	Blue Country	Krabben	Martin	M.	
Animal production							
	Cattle	Madurodam	Utopia	Danun	John	S.	Animal nutrition
	Cattle	Madurodam	Utopia	Davids	Joseph	J.	Animal nutrition
	Cattle	Madurodam	Utopia	James	T	F.F.	Animal Nutrition
	cattle	Haarlem	Northland	Janssen	John	C.	animal husbandry
	cattle	Madurodam	Shangrila	Welzen	Theresa		animal nutrition
	Cattle, beef	Haarlem	Blue Country	Kraan	Robert	R.	
	Farm. Syst.	Maastricht	Utopia	Deelen	Fred	W	animal physiology
	many	Madurodam	Shangrila	Thompson	Timothy	C.	animal nutrition
	Many	Madurodam	Utopia	Dillloe	Eric	E. M.	
	Many	Den Bosch	Northland	Drake	Daniel		animal nutrition
	Many	Haarlem	Eastland	Driel	John	G. S.	
	Pastures	Den Bosch	Eastland	Eck	Donald	D.	Animal nutrition
	socio-econ.	Madurodam	Yellow	Pijl	John	I.	animal husbandry
Basic sciences							
	Pastures	Den Haag	Westland	Foet	Margaret	M	Ecology
Engineering							
	Engineering	Lelystad	Northland	Blans	Jan	J.	
	Irrigation	Zwolle	Southland	Hair	Peter	S. D.	
	Irrigation	Zwolle	Southland	Hazel	Josy	B	
	Irrigation	Zwolle	Southland	Potts	Angel	A. E.	
	Many	Lelystad	Eastland	Elliot	Alan	M	machinery
	Many	Madurodam	Utopia	Johnson		M	agric. buildings

Note: All names have been changed.

Figure 2.15: Example of information aggregated from 10 NARO MIS datasets, offering a regional directory of scientists

Creating the regional data set need not be difficult. The information can be combined from diskettes sent by each collaborating institute or NARO to a central regional agency, which aggregates the data, generates appropriate reports for regional sharing, and distributes them on paper and/or (increasingly) electronically. The technology and information systems are already in place for some NAROs to implement such an arrangement.

The Internet promises a further means to the same end, and although only a few NAROs have the requisite access to the Internet today, many are likely to have good connections in the coming few years. ISNAR is exploring the mechanisms for mounting a selected subset of a NARO's MIS data on the Internet for this purpose. Figure 2.17 shows an output from the same data set aggregated by Webber from 10 NARO MIS, mounted on ISNAR's Website. The SINGER project operated by the In-

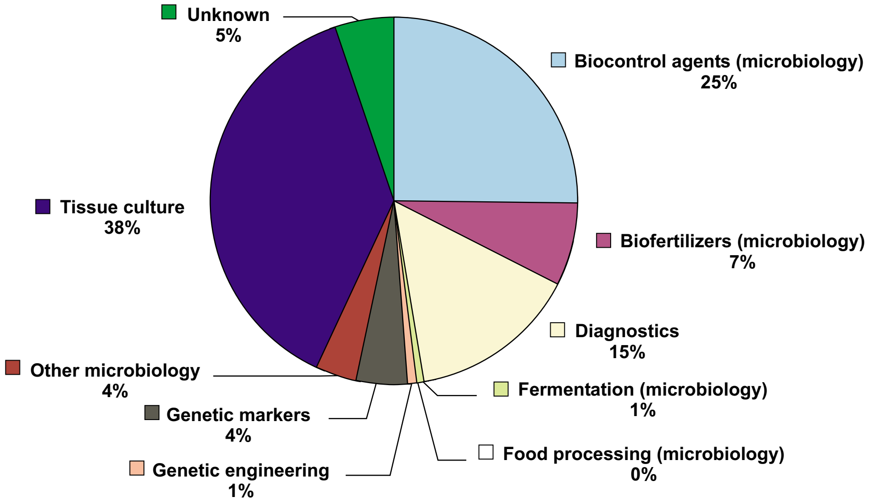


Figure 2.16: Example of information aggregated from 10 NARO MIS datasets to determine what types of biotechnology research are currently in progress in the region

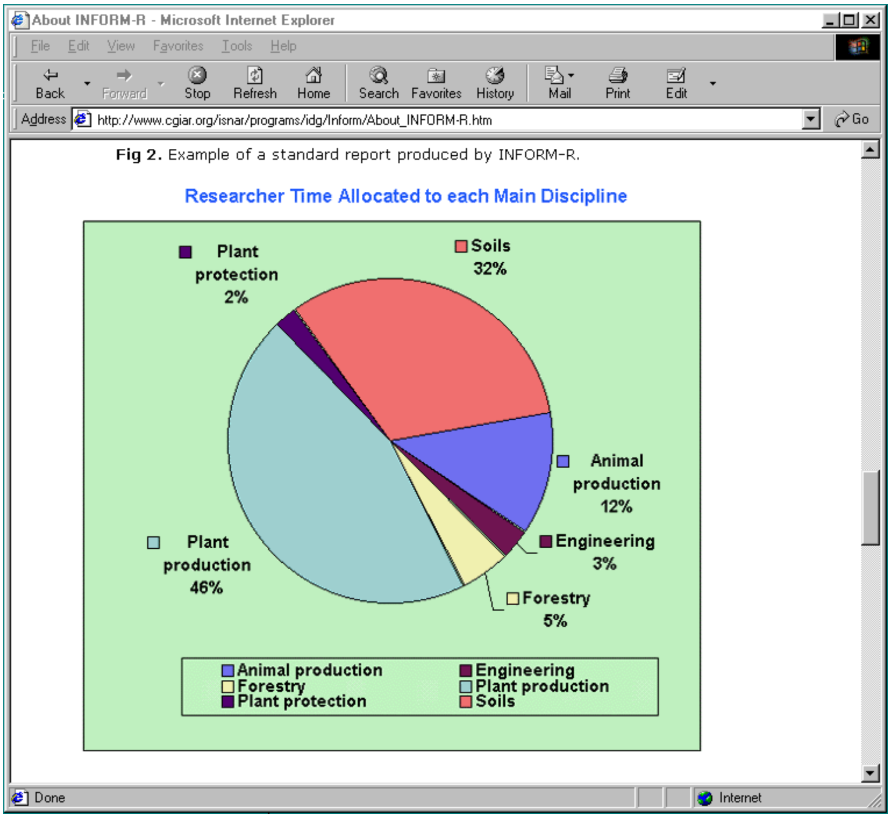


Figure 2.17: Example of information from an MIS, made available on ISNAR's Website

ternational Plant Genetic Resources Institute (IPGRI) in Rome (see annex 4 "Worldwide Agricultural Research Information Systems and Sources") is a good example of a similar initiative for sharing information on genetic resources held by the centers of the Consultative Group on International Agricultural Research (CGIAR).

Prior requirements

Some issues need to be addressed before setting up such regional information sharing. It is important for all parties to agree on which information is useful to share. Normally this will be only a small subset of the data typically held in a national MIS. Should a distinction be made between information for sharing between collaborating NAROs and that made available for public access on the Internet? The actual information likely to be offered in these shared systems, as shown in the above examples, is unlikely to be contentious, but there needs to be close consultation and agreement with the scientists, as well as compliance with any legislation, affecting the transfer of electronic information containing personal data. Some kind of memorandum of understanding between the collaborating NAROs, covering the content, the timing, and the mechanism would be needed.

References

Alston, J.M., G.W. Norton, and P.G Pardey. 1995. Science under scarcity: Principles and practice for agricultural research evaluation and priority setting. The Hague: International Service for National Agricultural Research.

Busch, L. and R.J. Bingen. 1994. Restructuring agricultural research: Some lessons from experience. ISNAR Briefing Paper No. 13. The Hague: International Service for National Agricultural Research.

Corbett, D.C.M. 1989. A current research information system for the management of agricultural research. *R&D Management*, 19, 3, 251-263.

Dagg, M. and F. Haworth. 1988. Program formulation in national agricultural research. ISNAR Working Paper No 17. The Hague: International Service for National Agricultural Research.

Date, C.J. 1992. Relational database writings, 1989-1991. pp. 461-466. Reading, Massachusetts: Addison-Wesley.

Galante, E., and C. Sala. 1995. Information needs for agricultural research managers and policy makers. Current research information systems in Europe: A step further. Proceedings of the CRIS 95 meeting in Milan, 11-13 May, 1995. Milan: National Research Council.

Hackathorn, R. 1997. Farming the Web. *Byte Magazine*. October 1997: 43-44.

Horton, D., P. Ballantyne, W. Peterson, B. Uribe, D. Gapasin, and K. Sheridan. 1993. Monitoring and evaluating agricultural research. A sourcebook. Wallingford: CAB International and The Hague: International Service for National Agricultural Research.

ISNAR. 1998. Information and Discussion Forum: Priority Setting in Agricultural Research. Internet Website: www.cgiar.org/isnar/Fora/Priority/index.htm. The Hague: International Service for National Agricultural Research.

Lusthaus, C., G. Anderson, and E. Murphy. 1995. Institutional assessment: A framework for strengthening organizational capacity for IDRC's research partners. Ottawa: International Development Research Centre.

Mutsaers, H.J.W., G.K. Weber, P. Walker, N.M. Fischer. 1997. A field guide for on-farm experimentation. Ibadan: International Institute of Tropical Agriculture.

Nestel, B. 1989. Some practical guidelines for evaluation within national agricultural research systems using the checklist approach. ISNAR Working Paper No. 22. The Hague: International Service for National Agricultural Research.

Peterson, W. 1994. Assessing organizational performance: Indicators and procedures. ISNAR Discussion Paper No. 93-18. The Hague: International Service for National Agricultural Research.

Potts, J.M., P.J. Verrier, R.W. Payne. 1996. The electronic Rothamsted archive and its use in agricultural research. Proc. 6th International Congress for Computer Technology in Agriculture. Wageningen, June 16-19, 1996. pp 171-176.

Prinsley, R.T. 1992. Research and development evaluation: A review. Barton, Australia: Rural Industries Research and Development Corporation.

Runge, C.F. 1992. A policy perspective on the sustainability of production environments: Toward a land theory of value. In Future Challenges for National Agricultural Research: A Policy Dialogue. The Hague: International Service for National Agricultural Research.

Schuh, G.E. and H. Tollini. 1979. Costs and benefits of agricultural research: The state of the art. Washington DC: World Bank.

Schutzelaars, A., G. Engelen, I. Uljee, and S. Wargnies. 1995. Computer systems that enhance the productivity of public-sector planners. *International Journal of Public Administration* 17:119-153.

Tuck, N. and G. Zaleski. 1996. Criteria for developing performance measurement systems in the public sector. *International Journal of Public Administration* 19 (11,12): 1945-1978.

Vernon, R. 1999. The information needs of agricultural research management. *Quarterly Bulletin of the International Association of Agricultural Information Specialists.* XLIV (1,2): 45-55.

Webber, H. 1999. Personal communication on his work with aggregated data sets from many NAROs' INFORM management information systems.

World Bank. 1997. Agricultural extension and research. Achievements and problems in national systems. A World Bank operations evaluation study. Operations Evaluation Department. Washington DC: World Bank.

Chapter

3 Developing a Management Information System: Five Case Studies

This chapter gives research managers and information systems developers a broader view by learning from the experience of others in developing an MIS. For the busy manager, a review of lessons learned from these studies is given at the end.

Five case studies of MIS development provide (future) implementers a view of what is involved in developing an MIS. The case studies provide a record of what was done, what was achieved, what were the main problems encountered, and what lessons were learned.

Three cases recount the experience over several years of implementing ISNAR's original INFORM MIS and how these systems are managed with little or no external intervention. Another covers the development of INFORM-R (the new relational model of INFORM) in its early prototype stage in the field. The case from Brazil is an account of an MIS being developed for a large NARO, where the demands are more complicated but the resources are also more sophisticated. These accounts have provided an additional source of knowledge on which this book has drawn.

3.1 Zambia: Establishing an MIS in the Ministry of Agriculture, Food and Fisheries, Research Branch

by David Parker[1] *and Richard Vernon*

Background

Zambia has a long history of agricultural research. Some is recorded in a collection of paper records of several thousand experiments dating from the 1950s to the present. The 1990s have seen the introduction of computers to support the management of the research program, culminating in the introduction and development of a formal MIS.

The Zambian Ministry of Agriculture, Food and Fisheries, Research Branch, today has eight research stations and seven technology assessment sites serving the three agroecological zones of Zambia. The research stations implement 13 research pro-

1. David Parker is former biometrician in the Soils and Crops Research Branch, Ministry of Agriculture, Fisheries and Food, Zambia.

grams through 27 research projects. Some of these projects are sponsored by international agencies in cooperation with the Government of Zambia, while others are run entirely by the Government. The Research Branch is led by the deputy director of agriculture (research) who is assisted by seven chief agricultural research officers: for crop improvement and agronomy, soils and water management, farming systems and social sciences, plant protection and quarantine, and for each of the three agroecological zones I, II, and III.

Before the introduction of the new MIS several forms of information were already available to research managers and researchers. These paper-based information sources included the accounting and personnel systems, a library, and an experiment records system.

3.1.1 Research management before computers

The research program was managed through a formal procedure of project approval of research protocols, established in the early 1960s. The project approval mechanism consisted of a scientist writing a project proposal, which was submitted for peer review before being approved for funding.

The main crop growing season (rainy season) is typically mid-November to late March or April. Most experiments were harvested by June, and results processed and written up in summary form for the annual Research Committee Meetings in August. These lasted two weeks, with a half day allocated to each commodity or commodity group, each of which had a commodity leader. The leader received all experiment reports for that commodity and aggregated them into those (a) to continue unchanged, (b) to continue but with changes, (c) to stop, and (d) new proposals in the form of the protocols mentioned above.

The Research Committee Meetings reviewed the past year and planned the next. Adjustments in priorities between and within programs were included. All scientists from all stations were present and all those working on a particular commodity attended that meeting.

Following these meetings, the scientists returned to their stations and expanded their summary reports of the previous season into detailed annual reports for each experiment. These reports were aggregated at headquarters into part II of the national annual Research Branch report. Part I contained summaries, which were more suitable for ministry staff, donors, etc.

Two other annual events, held in mid-season (February) were the following:

- Provincial tours: a mixed team of HQ, provincial, research, and extension staff made a tour of one week in each province, visiting farmers and the local research stations.
- Provincial experiment committee meetings, one-day meetings held at the end of a provincial tour: farmers, extension staff, and research staff scrutinized the local research program and the findings of the farm visits of the previous week and drew up recommendations for the national research committee meetings described above.

These three events and the annual reports were the main information vehicles of research management. In the 1960s and 1970s, this system was highly productive with an impressive flow of new technologies moving from the research stations to the farming community. Evidence can be seen in the Research Branch annual reports of that time and the numerous technical articles written by researchers for the local farming community and published in the government-run *Farming in Zambia* journal and the *Productive Farming* journal of the Commercial Farmers Bureau.

In the 1980s, shortage of funds severely curtailed the provincial tours and experiment committee meetings, and they were finally abandoned. The associated documentation of research protocols and research reports broke down around 1991 but has been partly revised with the development of the new MIS.

3.1.2 Establishing a computerized MIS

The project to establish a computerized research MIS was instigated by the government as an activity under the Zambian Agricultural Research and Extension Project (ZAREP)[2] in order to assist in monitoring and evaluating agricultural research projects within Zambia. This project also provided for the reestablishment of the Biometrics Unit and an expatriate biometrician in early 1993 was appointed to help reestablish the experiment records system. The MIS was to fulfill the dual needs of monitoring and evaluating past and present research and maintaining a database of current research work for the biometrics office. ISNAR was approached to assist in establishing the MIS.

In May 1993 an ISNAR staff member visited Zambia and went with the biometrician to a representative research station (the National Irrigation Research Station at Nanga) to demonstrate a data capture method. He gave a presentation to management and scientists at the station on the objectives of the MIS and how it could be implemented, what kinds of outputs it offers, and what would be asked of them. Scientists were then asked to complete a set of data capture forms. Questions about the system and on completing the forms were answered in plenary session so that all could learn simultaneously.

The data on the forms was keyed into a simple prototype database to demonstrate some example outputs, based on real data, to research management and ZAREP staff at the headquarters. On the basis of the presentation and ensuing discussions, management decided to implement an MIS.

ISNAR drafted a proposal in October 1993 to help develop and introduce an MIS for the Zambian Agricultural Research Branch. The proposal included data collection, data input, an implementation workshop, and two supervision visits by an ISNAR MIS expert. The Research Branch management accepted this and also identified some initial donor funding.

Ex ante data was again collected in September 1994 on research activities to be carried out during the coming season, and personnel information was updated. The in-

2. Operated from 1988 to 1996.

formation was transferred to Microsoft Access 2, the database software that ISNAR was using to develop its new INFORM-R MIS.

ZAREP was in the process of procuring new computer equipment to establish the MIS at each research station. The procurement, however, took a total of two years due to bureaucratic problems, leading to a major delay in the implementation of the project. The optimum specifications of the equipment changed significantly in the meantime, but could not be modified.

After the equipment arrived in September 1995, the Eastern and Southern African Management Institute (ESAMI) organized a workshop on database management principles in Lusaka in October. The workshop brought together several of the research managers who would ultimately be using the new MIS.

Phase 2

The second phase of the MIS introduction began in December 1995 with a workshop. Prior to the workshop, all participants were invited to a series of computer software courses, including an introduction to Access 2.

At the four-day workshop, a prototype relational[3] model MIS was introduced to some 20 research managers, researchers, and secretarial staff. The first purpose of the workshop was to allow different users to get acquainted with the prototype MIS and to record what they liked and disliked. The second was to assess whether the system was easy enough for managers to learn in no more than two days, to show them how to navigate through the system and find and manipulate any of the built-in reports. A survey on the fourth day suggested that these goals were achieved. On average, participants felt that they had reached a useful degree of competence on the third day.

But several issues were raised at this workshop and later through e-mail correspondence:

- One significant issue was the difference between ISNAR's initial understanding of the main requirements for such an MIS, that of senior research management, and that of a donor representative. ISNAR had taken a scientific view, assuming that the main purpose was to assist research planning, monitoring, report writing, and evaluation of research experiments, projects, and programs. The MIS emphasized resource allocation across commodities (crops and livestock species) and noncommodity factors (soils studies, socioeconomics, etc.), and across scientific disciplines and agroecological zones.[4] Research managers, however, were more concerned with administrative issues.
- Managers also wanted resource allocation reports to apply to "programs," but it proved difficult to have these defined in terms of their constituent projects and experiments, which was needed for the system to aggregate up data from the experiment level.

3. See glossary in annex 1 for an explanation of a relational database.
4. See chapter 2 for examples of commodities, noncommodity factors, and scientific disciplines.

- Managers were uncomfortable with the use of scientist time allocations as an indicator of resource allocation. They preferred to see actual financial data, but in the absence of an effective accounts system this proved impossible. A solution was later sought in the development of a budgeting and direct costs system, but this was never brought into use.[5]
- The prototype menu system proved less easy to use than had been expected.
- One donor representative wanted very detailed accounting figures for which he supplied a large and complex system of spreadsheets. He had no interest in any scientific aspects. While this presented the developers with a problem, meeting the information needs of donors was an important requirement of the system.
- There were also several technical points with the prototype software, e.g., data entry for some sections was not properly set up, though it was not expected to be a complete system at this stage.

System development

To address these issues, the developers completely redesigned the menu system. A separate budgeting and expenditure system was developed that allowed budgets and all experiment-related expenditures (i.e., excluding indirect costs) to be recorded. These could be used to produce monthly reports that aggregated experiment to projects and programs, showing budget allocated, budget released, expenditure, and remaining funds left unspent. These enhancements were shipped to Zambia in March 1996.

In parallel, ZAREP funds were used to procure for the accounts department at Mt Makulu Central Research Station a sophisticated accounting package (SUN Systems) to help solve its financial accounting problem and, eventually, to link the system with INFORM-R. But as the system was poorly supported, the accounts department did not adopt it. The experiment budgeting and expenditure system was also not brought into use beyond a brief trial period at Mt Makulu. First, it was designed for use at the research stations (where expenditure would be mostly incurred), but the system was never fully distributed to the stations due to shortage of funds for training workshops. Second, there was strong resistance among researchers, who, in a climate of considerable financial shortfalls, saw it as a threat to existing expenditure practices. Third, there were insufficient resources to properly test the system other than in actual use. Fourth, its introduction coincided with the removal of the accountant post at Mt Makulu, leaving a serious gap in accounting skills and management. Finally, the limited attempt that was made to use it indicated that accounting for expenditure at the experiment level was not practical; it was suggested that accounting at the project level might be more practical, but this needed to be tested.[6]

However, Zambia was fortunate in having a good Internet provider at the University of Zambia, which allowed large files to be attached to e-mail messages. This meant

5. Experience elsewhere suggests that the use of scientist time is a very efficient way of tracking approximate resource allocation and users do get used to it in time.
6. A similar conclusion was reached in Bolivia. See the case study reported in section 4 of this chapter.

that ISNAR could send new or modified versions of INFORM-R at no more than the cost of a local phone call. Nevertheless, this was insufficient to adequately address questions of systems analysis and design. The biometrics office set up a parallel copy of the INFORM data files from which administrative reports could be derived until the INFORM system was more complete. The system was thus kept a centralized one for the coming season.

Faced with its own internal problems, ISNAR was unable to commit sufficient resources to MIS development. An initial advisory team of four was lost, leaving development to just one staff member, later augmented with a research assistant. There was no specialist expertise to advise on how an MIS should address the finance area and develop a comprehensive budgeting and expenditure system. What was prepared was intended as a stop-gap measure to meet a request from Zambia. Funds for system implementation through an external programmer were intermittent. These shortages led to insufficient attention being given to aspects of system development, notably to meeting all the requests from users, especially in aspects of financial management, and to maintaining thorough system testing procedures.

Phase 3

One of the planned activities in the third stage of the implementation was as a follow-up visit by the ISNAR consultant to monitor how the system was running. With project funds almost exhausted, and the parent ZAREP project soon to close, this visit was converted to a visit to ISNAR by the two Zambian counterparts responsible for running the INFORM implementation. This visit enabled them to participate in a workshop involving other MIS practitioners from developing countries.

The database system was transferred to the INFORM-R system, but extra tables were needed to link to budgeting. Some extra reports needed locally were developed, e.g., a list of approved experiments sorted by the research programs within research stations, together with the appropriate budgets.

Extra reports

Because different entities sponsored agricultural research in Zambia, different reporting formats were in existence. Consequently, some researchers wrote two or more reports for each project each quarter. Invariably, researchers gave priority to the donor reports that were directly linked to funding, and the situation arose that reports of research carried out by Zambian researchers in Zambia were sent to Oslo, London, Stockholm, or a foreign embassy in Lusaka but were not available in the Mt Makulu central library.

To address this situation, a standard reporting format was incorporated into INFORM-R that consisted of two interim reports and an annual report for each research activity. This structure was shown to representatives of key donors but they wished to see the system fully operational before abandoning their requirement for separate reporting formats. These reports were implemented at the start of the 1997 season, and first interim reports were collected in January 1997.

The MIS annual report facility, however, was not used. Instead, the Research Branch annual report for 1997, the first annual report since 1992 and a major achievement, was completed manually. It took four senior scientists more than one month after obtaining information from a one-week annual report workshop, attended by nine senior scientists. This may reflect the missed steps in MIS implementation, such as sensitization of senior managers and training of MIS practitioners.

Decentralization

During the development of the MIS, it was operated in a largely centralized mode. Data was sent to Mt Makulu and entered into the system. There were two key problems with the centralized system: the heavy burden on the Biometrics Unit of entering the data from all the stations and the delays in updating the data. The latter led to complaints that the MIS was out of date particularly with regards to staff movements, so decentralization could no longer be delayed. During January and February 1997 all stations were visited and key personnel trained in entering data. Each station was also invited to send a secretary to Mt Makulu to participate in a week-long training in entering data.

During May 1997, when the second interim report was due to be input, the INFORM implementation in several stations was inspected. None of the stations visited had entered any data since the last visit in January. There were various problems. For example, out of the three stations visited in the Southern Province, one station had its computer stolen, at one station the secretary had passed away, and the secretary at the third station found the software difficult to use. It was therefore decided to abandon the decentralization for the time being.

It had always been recognized that once the system had developed to meet users' needs reasonably, there would have to be an integrated training program. But there were no funds for this within the government budgets. So plans were drawn up for a project to support training workshops in each of the three agroecological zones and to upgrade the software and hardware. At the time of writing no donor has agreed to accept this project, so the system remains centralized. It is not sustainable in this mode however.

Links to other systems

The post of accountant has been reinstated at Mt Makulu and an officer placed there. It is hoped that in due course a computerized finance system will be established to which INFORM-R can be linked.

In July 1995 the Ministry of Agriculture, Food and Fisheries (MAFF) published a systems requirement for an MIS for the Agricultural Sector Investment Programme (ASIP). This system was designed to take aggregated data from ASIP component MIS such as the Research Branch's INFORM-R and summarize it for use by senior MAFF personnel. The Research Branch's INFORM-R was the first ASIP component MIS to be developed. The only other to have been developed since is that for the Extension Service, though it too faces implementation problems.[7] The apex system in the Ministry includes a financial MIS using the SUN system and a personnel information sys-

7. Other ASIP components intended to have their own MIS include Fisheries, Animal Health and Livestock and perhaps the Seeds Certification and Control Institute.

tem based on Access. It is hoped that the personnel system will eventually be linked to INFORM-R to provide research managers with access to information on all Research Branch personnel, so that the combined system will require only one data entry process.

3.1.3 An archive system of past research

Many people have observed that current research is often started with little or no inquiry as to what has been done before in that topic. With the demise of libraries and of the habit to produce full and timely national annual reports, it is sometimes quite hard to find information on earlier research. An archive component of an MIS can fill this gap.

The large collection of paper records of experiments referred to in the introduction of this case study had great intrinsic value, but being a single copy, and on paper, it was always vulnerable (borrowed files can be lost) and it was not readily accessible to scientists on remote research stations. There was also no easy way of aggregating and processing experimental data for management review purposes. Computerization solved the first two of these constraints, and if taken further to an MIS, the third could also be addressed.

In the early 1990s the Research Branch asked the German Volunteer Service to assist in computerizing its extensive paper records of research experiments. Two volunteers established a simple database[8] and began the immense task of keying in selected fields for each experiment.[9] Due to the large size of this task only a small selection of the many data fields was keyed in. After the departure of the volunteers the system was brought to the attention of an ISNAR information specialist who, recognizing its potential value, had it converted to the INFORM-R structure in Access. The biometrician later arranged to have some further data keyed into it and prepared a simple "front-end," which enables even those without computer or database experience to search the system for what research has been done before.

This brings scientists at remote stations one step nearer to being able to check previous research. The system contains details of some 6,000 experiments, along with their replications on different sites (see table 3.1). The remaining data from the paper files still needs to be entered, additional data from the historical files at all the major stations needs to be collected and keyed in, and this system then needs to be copied to all stations. Finally the system needs to be linked to INFORM-R so that the records of completed experiments can be transferred to this archive system.

8. They used the database software D-Base IV, a leading database software at that time.
9. Optical scanning with optical character recognition might have been an aid, but the paper originals were not suited to this; they were mostly rather irregularly typed and sometimes stenciled copy, structured into "standard" experiment protocols and reports with various tables.

Table 3.1: Report from Zambian Archive of Experiments by Disciplines and Years

Start year	Total	Animal breed.	Animal prod.	Engin-eering	Food techn.	Plant breed.	Plant prod.	Plant prot.	Socio-econ.	Soils
1943	1						1			
1948	1						1			
1951	1					1				
1955	10						10			
1956	1		1							
1957	19						15	1		2
1958	81		1			2	56	1		6
1959	238			3		4	163	11	1	42
1960	178			2		8	112	8	8	30
1961	171		5			5	100	6		15
1962	203			12	1	30	77	13	3	60
1963	174		1	4		7	60	8		42
1964	78		3	1		4	34	12		16
1965	164		2	2		4	55	25	1	26
1966	148		1	1		17	65	31	1	15
1967	169		1				63	25	1	27
1968	89		1			16	47	12	1	6
1969	112		5	1		12	50	17	1	18
1970	141		11	2		21	61	20		14
1971	9		2			3	1			1
1972	96		5	2	4	33	32	12	4	1
1973	81		7	7	4	23	26	10		4
1974	119		1	1	11	33	40	22	1	8
1975	84		1		3	27	31	14		8
1976	148	3	3	3		55	45	33		4
1977	376	1	22	17	3	115	146	48	15	9
1978	81	1	6		1	21	23	25	1	1
1979	116		6	3		46	38	16	3	4
1980	99			1	5	58	16	11		7
1981	86		1		2	49	14	13	4	3
1982	245	1	4	1	2	133	66	19	2	17
1983	221		2	2	2	115	66	29		4
1984	208		3	2		100	81	14		8
1985	207				1	39	136	13		15
1986	270		1	3	3	63	159	19	1	19
1987	177		3		1	27	119	12	1	12
1988	177		1	4		36	86	17	1	32
1989	207		1	8	1	47	97	27	2	24
1990	169		1	4		32	95	15	1	21
1991	64			2		9	42	2		8
1992	34		1	4		4	11	3	2	8
1993	74			3		13	38	13		7
1994	244		11	2		66	94	37	7	20
1995	220		11	6	5	39	76	25	25	28
1996	154	1	9	5		29	45	16	20	19
1997	66	1	10		3	12	11	7	7	5
1998	29				3	4	3	11	3	3
Total	**6040**	**8**	**144**	**108**	**55**	**1389**	**2607**	**673**	**117**	**679**

3.1.4 Project review

Implementation difficulties

The Zambian MIS took off very slowly due to a number of factors, many of which have affected and aggravated each other:

- A comprehensive training program to enable the system to be decentralized to the stations was not found. The funds that were available for the implementation covered only an exploratory workshop to test the first prototype and two follow-up visits from the ISNAR consultant. All funding in Zambia stopped in June 1996, well before the implementation was strong enough to sustain itself and even before the software had been fully completed.
- A strong patron in senior management was lacking.
- There were problems in procuring computers.
- There was a lack of funds for traveling.
- The Access database system needed to be operated in parallel until the INFORM-R model could be adopted.
- The project budgeting system was not adopted, which may have led to the nonadoption of INFORM-R. In the 1997 and 1998 season the budget was unlinked from the production of research protocols. This led to a dramatic fall in the number of protocols submitted for data entry, and a similar problem may arise in the submission of reports.
- Development work on INFORM-R was stopped as funds at ISNAR were running out. This happened before all the requirements had been addressed, including those for donors, for the further development of the project budgeting system, and for the Zambian program structure.
- The original concept of an MIS for monitoring and evaluation has not been expanded to deal with the management information needs for information on personnel, finance, and inventory.

Although the economists in the Farming Systems Team in the Research Branch were originally assigned to take over the MIS, this never materialized. And the new Farming Systems and Social Science Division is reluctant to take on this responsibility as long as it remains a government-funded program with the consequent lack of funds.

Successes

Paradoxically, INFORM-R is perceived as a kind of success within the Research Branch, even though it is underused. One reason for this may be the rich store of information held in the system, which could be built upon if a stronger implementation were put in place. Many visitors from within Zambia and from other countries have come to see it. A subset of the system was even put on the Internet as a demonstration.

Personnel data on research time can be used to determine how much time has been spent on every research activity during the past three years. An important result has been the highlighting of the time spent on never-ending experiments. It is now quite difficult for researchers to propose the same research year after year as the accumulated research time would stand out as being very unusual, particularly if there are no tangible results.

For the past two years it has been possible to see which research activities were planned, which were started, and which were completed. This is also the second year that interim reports have been collected.

Zambian counterparts have built up considerable skills in using the database software; these skills enabled them to also undertake short-term consulting contracts, which supplement their modest government salary.

The new structure of the Ministry of Agriculture, Food and Fisheries has created new positions specifically for the MIS function, which recognizes the importance the government has given the MIS. It is no longer considered a part-time job for the biometrician.

The future

A core system is in place at headquarters, complete with fairly complete data sets for the past four years. However, this has been achieved largely through the commitment, beyond the bounds of duty, of just a small group of individuals, two of whom have recently moved to new positions. For the MIS to play its intended role in the support of the management of the research program, serious attention has to be given to, in particular, a senior-management patron; sufficient funding for ISNAR to arrange its thorough testing and completion of the outstanding requests; and sufficient funding for the Research Branch to implement the planned training workshops, procure further hardware, and distribute the system to all stations.

3.2 Ghana: Introducing an MIS in the National Agricultural Research Project

by Kwasi M. Setsoafia

Background

In 1987 the Council for Scientific and Industrial Research (CSIR) of Ghana and ISNAR conducted a review of Ghana's national agricultural research system. On the basis of the review's findings, the National Agricultural Research Project (NARP) was established in 1991. To implement the project successfully, a report by the World Bank in May 1991 recommended implementing an MIS that included an inventory of current research programs. ISNAR's INFORM MIS was chosen for this purpose.

3.2.1 System implementation

The first training workshop in INFORM was held in 1992 (see table 3.2). The first set of data on human resources and research activities was captured in 1992 and 1993 through questionnaires that were sent to the directors and deans of the participating research institutions for completion by the scientists. Guidelines and supporting information facilitated the process. Personnel of the NARP Technical Secretariat, consisting of consultants from a UK-based consultancy firm (ULG Consultants) and their Ghanaian counterparts, provided further explanation during follow-up visits. Completed forms were sent to the Technical Secretariat, and by June 1993 two INFORM databases on human resources (i.e., agricultural researchers) and research activities were established using the database software Reflex 2.0.

However, not long after a second training workshop in 1994, NARP lost its head of monitoring. As he had managed the system almost singlehandedly, two years of inactivity followed. Early in 1996 two new staff were appointed to the Monitoring and Evaluation Section of the Technical Secretariat. Following a two-week MIS workshop in Kenya, they conducted a new (second) round of collecting data in April and

Table 3.2: MIS Training Events in Ghana

Date and place	No. of partps	Topic	Types of participants
20 March – 10 April 1992, Accra, Ghana	25	• Program budgeting	• research officers • scientific secretaries • administrative officers
4–8 July 1994, Accra, Ghana	16	• Introduction to Computers • Introduction to MIS Principles and Practice of INFORM	• scientific secretaries • research officers
8 July 1994, Accra, Ghana	7 *	• MIS for Research Managers	• heads and deputies of participating institutes
Sept. 1995, ISNAR, The Hague (1 week)	1	• Introduction to INFORM	• agricultural head, Monitoring and Evaluation Section
March 1996, Egerton Univ., KARI (2 weeks)	2	• Introduction to Computers • Introduction to MIS Principles and Practice of INFORM	• agricultural head and senior assistant secretary, Monitoring and Evaluation Section
15–26 July 1996, Ghana	20	• Introduction to Computers • Introduction to MIS Principles and Practice of INFORM	• scientific secretaries • librarians • research officers
26 July 1996, Ghana	10*	• MIS for Research Managers	• heads and deputies of participating institutes

* out of a total of 15 heads of participating institutes

May 1996. Because the MIS had been inactive for two years, most scientists had once again to be taken through the necessary steps required in completing the forms. This was done in one-day workshops at each participating research institution, with two members of the Monitoring and Evaluation Section as resource persons.

The outputs of INFORM include data capture forms for scientists' personal data and for experiments, which can be printed empty or complete with existing data. During the 1996 one-day workshops at each participating research institute, scientists who had completed the 1993 human resource forms were asked to update their forms. Blank forms were given to those who had not provided data in 1993 and to newly employed scientists. To set up the research activities database, blank forms were given to leaders of projects that were approved after the NARP midterm review in September 1995.

The completed forms were collected at the end of the workshops and the data was entered centrally by the NARP Technical Secretariat or later on by the scientific secretary of the participating research institution. By the end of June 1996 new human resource and research activity databases had been set up.

The 1996 round of data collection was hampered not only because the MIS had been inactive for two years, but also because of the rush of research activities in the field at that time of the year: in southern Ghana, where most of the research institutions are located, the months of April and May normally mark the beginning of the rainy season and the start of field research.

The quality of the data on personnel—an exercise often undertaken for personnel departments—was generally good. On the other hand, one new requirement, the quality of the data on the time spent by scientists on activities and projects, was poor. This may be because the human resource forms were completed by individual scientists, while the research activity forms were completed by project leaders on behalf of the members of their team. Because of the multidisciplinary and inter-institutional nature of the research projects, project leaders sometimes gave estimates of time spent or to be spent by scientists without consulting them.

3.2.2 Problems encountered

The system could not be updated in 1994 and 1995 due to frequent changes in leadership of the Monitoring and Evaluation Section.

The Reflex ("flat file"[10]) software that INFORM was based on could not be used directly to show the relationships between research activities and scientists, the data of which are held in separate files. These items therefore had to be entered twice, making inconsistencies difficult to identify.

Practitioners use the Reflex software infrequently and since it is quite different from any other software they use, they quickly forget the skills they have learned in workshops.

10. For an explanation of the term flat file, see the glossary in annex 1.

It became increasingly difficult to collect data for the research activity database, due to the increasing multidisciplinary and interinstitutional nature of our research approach.

3.2.3 Impact

The system was used extensively in 1994 to generate information for priority setting during the preparation of the National Agricultural Research Strategic Plan and later for the preparation of a human resource development plan.

From December 1996 the system has been used to monitor progress (achievements) of research activities. In 1997 the system's data collected over a number of years was used to compare resource allocation between commodities with a previous set of priorities. Figure 3.1 shows the results for 1993, 1996, and 1997. It shows very clearly the gap between priorities and resource allocation. It also shows that in many cases the gap is widening instead of getting smaller, as one would expect if this data been made available to and used by the research planning process. It is hoped that this situation will improve in the future.

3.2.4 Lessons learned and recommendations for the future

There is a need to:

- involve users (directors, program coordinators, and scientists) more in the design and implementation of the system to improve its use.
- supplement knowledge and skills of participants acquired during workshops with on-the-job training through more follow-up visits; the NARP Technical Secretariat has benefitted immensely from such intensive on-the-job training for its INFORM practitioner.
- accord INFORM practitioners formal recognition through the formation of a formal network, such as the Ghana Agricultural Information Network System
- replace Reflex 2.0 software with a Windows-based software to make it easier for practitioners to implement INFORM in their institutes and for managers to be able to interrogate the databases themselves.[11]
- consult and establish consensus between project leaders and scientists to improve quality of data collected.
- collect data during the off-season months of December–March.
- promote institutionalization of INFORM. The INFORM practitioner at each institution should collect the data and set up the database. The role of the Technical Secretariat would be limited to providing technical support for practitioners and cleaning and collating institutional databases into a national database.

11. A Windows-based replacement is now available under the name of INFORM-R. (Ed.)

Medium Term Plans			1993 INFORM Records			1996 INFORM Records			1997 INFORM Records		
Commodity	Priority score	Rank	Commodity	Time %	Rank	Commodity	Time %	Rank	Commodity	Time %	Rank
Group 1											
Yams	239	1	Cocoa	18.1	1	Maize	8.9	1	Cassava	9.5	1
Cowpeas	227	2	Maize	8.9	2	Cassava	8.7	2	Maize	7.8	2
Maize	223	3	Oilpalm	8.9	3	Yams	7.5	3	Yams	7.3	3
Cassava	222	4	Rice	6.3	4	Fish freshwater	6.4	4	Soybean	6.8	4
Sorghum	218	5	Cassava	5.5	5	Soybean	6.4	4	Citrus	5.9	5
Cattle	213	6	Soybean	5.3	6	Rice	6	6	Cowpeas	5.5	6
Millet	212	7	Wattle	4.9	7	Oilpalm	5.5	7	Fish freshwater	5.1	7
Cocoyam	209	8	Pineapple	4.6	8	Plantain	4.4	8	Rice	4.7	8
Goat	208	9	Fish freshwater	4.5	9	Pineapple	3.8	9	Oil palm	4.3	9
Sheep	208	9	Plantain	4.2	10	Citrus	3.7	10	Plantain	3.5	10
Groundnut	207	11	Cowpeas	3.9	11	Cowpeas	3.4	11	Sweet Potato	3.1	11
Fish freshwater	205	12	Fish marine	3.3	12	Sweetpotato	2.8	12	Pineapples	3.0	12
Fish marine	204	13	Sheep	3.1	13	Cattle	2.6	13	Mangoes	2.9	13
Oilpalm	203	14	Coffee	2.6	14	Coconut	2.5	14	Groundnuts	2.1	14
Rice	201	15	Goat	2.4	15	Tomatoes	2.2	15	Cattle	2.0	15
Poultry	198	16	Coconut	2.0	16	Groundnut	2	16	Coconut	1.9	16
Group 2			Sorghum	1.7	17	Sheep	1.8	17	Tomatoes	1.7	17
Plantain	185	17	Cotton	1.3	18	Cotton	1.7	18	Sheep	1.4	18
Coconut	183	18	Yams	1.2	19	Sorghum	1.4	19	Cocoyam	1.2	19
Pig	182	19	Poultry	1.2	19	Millet	1.3	20	Sorghum	1.1	20
Pineapple	180	20	Pig	1.2	19	Mango	1.1	21	Millet	1.1	20
Soybean	179	21	Eggplant	0.9	22	Okra	0.9	22	Onion	1.0	22
Cocoa J	176	22	Citrus	0.9	22	Cocoyam	0.6	23	Okra	0.8	23
Sugar cane	173	23	Tomatoes	0.9	22	Onion	0.6	23	Shallot	0.6	24
Tomatoes	171	24	Groundnut	0.7	25	Cabbage	0.6	23	Cabbage	0.5	25

Figure 3.1: Comparison of medium term plan priorities and

Sources: NARP Medium Term Agricultural Research Strategic Plan Final Report, Sept 1994, Table 7.5, and INFORM databases for 1993, 1996 and 1997. With assistance from DFID, London, UK, and ULG Consultants Ltd, Warwick, UK.

→ example of a commodity given higher priority in the actual program than in the plan

→ example of a commodity given lower priority in the actual program than in the plan

- improve the management accounting role of INFORM by including fields that show time actually spent in the previous year and time projected on various activities and projects for the coming year.

References

Warr, K. and K.M. Setsoafia. 1996. Report on the INFORM 1996 database, National Agricultural Research Project, Ghana. Report submitted to the National Agricultural Research Project and CSIR (Ghana), the World Bank, Washington, and Department for International Development (DfID), London.

Setsoafia, K.M. 1996. Report on INFORM implementation at the Savanna Agricultural Research Institute, Ghana. Report submitted to the GTZ (German Technical Agency for Cooperation) and the Savanna Agricultural Research Institute, CSIR, Ghana.

3.3 Uganda: Institutionalizing an MIS in the National Agricultural Research Organization

by Dan N. Kisauzi and S. Dickson Baguma[12]

Introduction

Uganda's National Agricultural Research Organization (NARO), established in 1992, is a semi-autonomous body under the Ministry of Agriculture, Animal Industries, and Fisheries. It has eight research institutes (see table 3.3 for their mandates). The diversity of institutional cultures in the precursor units of the NARO institutes made it imperative to have as one of the first interventions the harmonization of the management systems, particularly in the areas of resource and research management. Using funds that the Government of Uganda had obtained from the World Bank, NARO contracted the services of ISNAR to assist in developing its MIS.

This assistance was encapsulated within a framework of a five-year comprehensive institutional development (CID) program, in which a major component was the establishment of a robust research MIS to backstop the evolving management systems. NARO management endorsed INFORM, an MIS developed by ISNAR (Nestel 1991), and adopted it in 1993.

3.3.1 The strategy for introducing an MIS in NARO

The overall objective of introducing an MIS in NARO was to provide research managers at all levels of the organization with accurate and up-to-date information on human, financial, and physical resources, as well as on research activities in order to expedite objective management decision making. The strategy focused on the following:

12. Dan N. Kisauzi was Leader of the Monitoring, Planning, and Evaluation Unit (MEPU) of NARO, Uganda, and is now an independent consultant. S. Dickson Baguma is Management Information Systems Specialist at NARO.

Table 3.3: Institute Research Mandates

Institute	**Research mandate**
Kawanda Agricultural Research Institute (KARI)	Perennial cash and food crops, soils, crop protection, farming systems, and plant introduction and plant quarantine.
Namulonge Agricultural and Animal Production Research Institute (NAARI)	Annual cash and food crops, crop-livestock management systems, and pastures for the humid and subhumid areas.
Serere Agricultural and Animal Production Research Institute (SAARI)	Cereals, roots and tubers, legumes and oil seeds for semi-arid areas; semi-arid production systems; seed research and production; pastures and range management; and crop livestock management systems for the semi-arid areas.
Livestock Research Institute (LIRI)	Animal health, animal breeding, theriogenology and animal diseases, including human *Trypanosomiasis*.
Fisheries Research Institute (FIRI)	Freshwater fisheries; fishing and fish technology; aquaculture and fish production systems.
Forestry Research Institute (FORI)	Natural forests and plantation forests management; forests products and utilization; and agroforestry.
Food Science and Technology Research Institute (FOSRI)	Food preservation, processing, storage, marketing, and dietetics.
Agricultural Engineering and Appropriate Technology Research Institute (AEATRI)	Farm mechanization, crop processing and storage, soil and water engineering, and technology.

1. Sensitizing research managers and scientists to INFORM

The sensitization of senior managers began in May 1993 with a one-day workshop. An ISNAR management information expert demonstrated the INFORM system to NARO's director general, his deputy, and the directors of the research institutes. The demonstration showed how the system could generate reports for different management requirements. It also emphasized the level of commitment that was required from senior management to successfully implement an MIS.

Unfortunately, at that time, the head of the Monitoring, Evaluation and Planning Unit (MEPU)—the unit designated to spearhead institutionalization of INFORM—had not yet been appointed. This had a negative impact on the subsequent speed and smoothness of the implementation process, which demonstrated at the outset the importance of effective patronage at a high level of management.[13]

In October 1993 one day was spent at each institute sensitizing scientists. The uses of INFORM were demonstrated and the tasks involved in creating and updating the underlying databases. The system was designed in such a way that data could be en-

13. When the head of MEPU was eventually appointed, he was familiarized with INFORM by participating in a two-week training workshop organized by ISNAR for the Kenya Agricultural Research Institute (KARI) in March 1996.

tered in two different flat file databases, i.e., human resource details and research activity, using Reflex software (Nestel 1991). Prior to the demonstration to management, a number of outputs (reports, graphs, cross-tabs) had been designed. More were designed later, during and after training.

2. Developing a cadre of MIS practitioners

The most cost-effective strategy for NARO to achieve a sustainable MIS was to create the position of a research management information officer (RMIO) with specialist skills in MIS and coordinating a group of scientists (MIS practitioners) based at each research institute and the associate Makerere University Faculties. A total of 17 MIS practitioners were selected and trained and an RMIO recruited.

Practitioners were selected on the basis of having a fair understanding of using computers. There were two scientists from each institute, one from each of the Makerere Faculties of Agriculture and Forestry, Veterinary Medicine and Science, and two from Agricultural Research Information Services (ARIS). This was seen as the most cost-effective and sustainable way of capturing and updating data annually, since most data was to be collected from the research institutes.

The objective of the 10-day training was to equip practitioners with skills to generate summaries from routine processes of the organization and to use the Reflex software to find the required ad hoc information. The initial INFORM training workshop was conducted in December 1993 using actual data that had been collected from the different research institutes and faculties beforehand. Thus, the practitioners could readily appreciate what the system offers by internalizing outputs from their own data. Also there was a need to have a system with actual and accurate data by the end of the training, which the practitioners could take back to their institutes in an operational MIS.

The training was done in the form of lectures, seminars, practicals, and case studies. Practitioners were introduced to Reflex, which runs INFORM, and were given an overview of MIS. They learnt how to enter and process program and research activities and personnel data. They obtained skills in generating useful information by learning how to design reports, filter, sort, and use cross-tabs to analyze research and personnel data. They also learnt how to design and produce pie charts and histograms and add keywords to individual experiments.

At the end of the training, practitioners suggested their own terms of reference:

1. Training: attend INFORM training workshops and subsequent follow-up seminars.
2. Data capture: annually collect and enter ex ante and ex post data on researchers, research activities, and research budget data for institutes.
3. Data processing: within an agreed time period—typically three weeks from when the data should be supplied by researchers—prepare and pass to directors an agreed standard set of reports, even though some data may not be complete. Discuss with the directors and program leaders the provisional outputs and agree on any additions and revisions.

4. Outputs: produce final reports:
   - on paper and distribute to director general and institute directors and researchers
   - on diskette: the original database files of researchers, research activities and budgets, and supply these to NARO headquarters and ARIS for compilation of national output
5. Advise research managers at all levels on the use of INFORM in decision making.

Tasks 3 and 4 would additionally be done at ARIS, NARO headquarters, but using national data sets.

At the request of individual institutes, MEPU organized more INFORM training workshops to acquaint as many scientists as possible with the MIS so that it can be efficiently used in decision making, while scientists execute their research. Currently MEPU is responsible for providing technical backstopping to the MIS practitioners.

One of the MIS practitioners was recruited as RMIO and posted in MEPU at NARO headquarters. His main duty was to provide technical backstopping and coordination of the institutionalization of INFORM. In 1996, he was sent to the United Kingdom for a one-year specialized degree training in MIS.

3. Data capture and processing system

The initial data capture started in September 1993 and lasted for four months. Two ISNAR staff and two NARO-ARIS staff visited the various research institutes and the Makerere University collaborating faculties. They worked with the selected MIS practitioners at each institute and faculty. At each place they first met the practitioners and then all the scientists to explain the way forms about research activities and human resources would be filled out by the individual scientists. Subsequent data capture is done twice a year: when institute research proposals are made (ex ante) and presented to the Board for approval and then after a fiscal year to ascertain what was actually implemented (ex post). Table 3.4 shows the annual progress of data collection in terms of records per institute or faculty.

4. Integration of MIS into planning, monitoring, and evaluation

Data is captured at the institute and faculty levels, with backstopping from MEPU. All data is merged to create national databases on human resources and research activities. The data is processed in both institutes and faculties and at the national level. The outputs of this processing are various standard reports that were designed to address the most frequent information requirements about human resource management and development and research in a holistic perspective. In addition, some outputs are generated to meet ad hoc information requirements for different managers for their decision making.

The system has been in place from December 1993, and its integration into the management processes of research planning, program formulation, and monitoring and evaluation has been largely successful. The way the MIS feeds into the program planning cycle is shown in figure 3.2.

Table 3.4: Records in the Human Resources and Research Activities Databases to Date

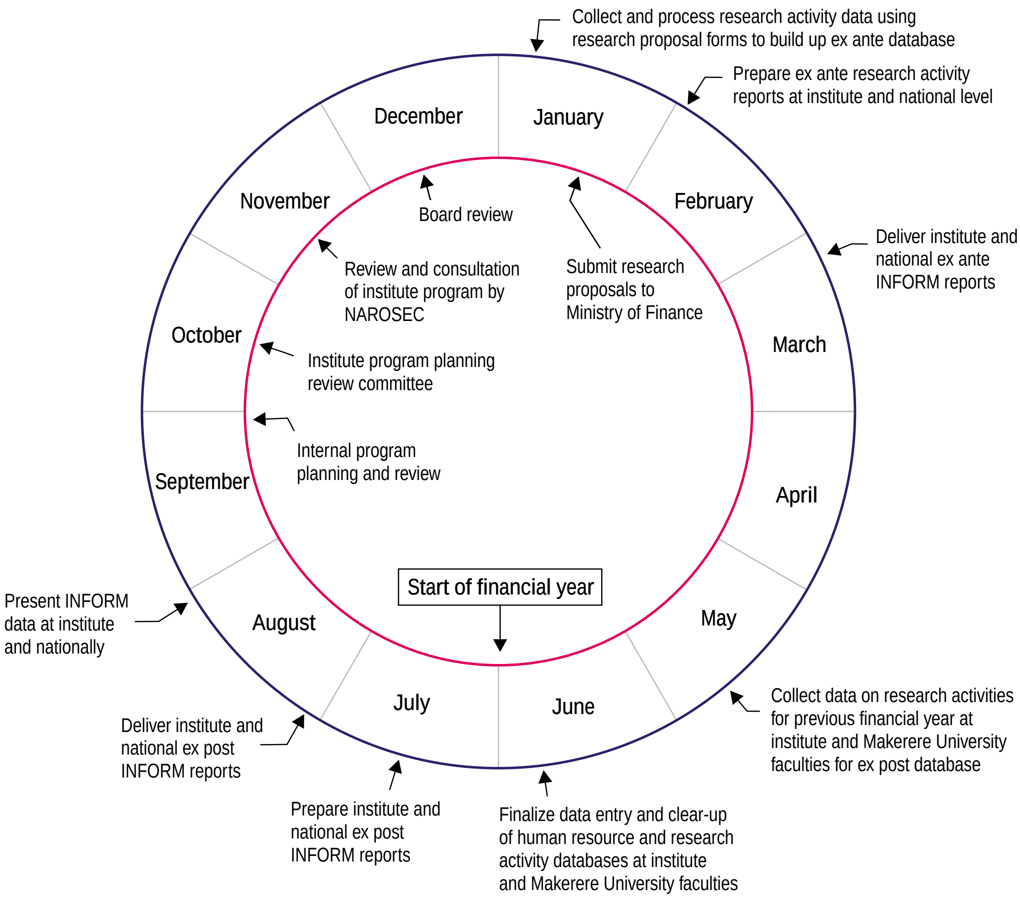
Institute and faculty	Number of records (ex post)									
	1993		1994		1995		1996		1997	
	HR*	RA**	HR	RA	HR	RA	HR	RA	HR	RA
NAARI	26	95	36	70	39	128	39	148	39	148
SAARI	13	85	28	90	29	103	28	105	28	105
KARI	48	187	45	206	48	158	48	158	48	158
FORI	14	16	18	11	18	22	17	43	17	43
FIRI	-	-	20	41	23	73	23	78	23	78
LIRI	26	53	25	44	25	41	25	46	25	46
MUFV	13	-	-	-	69	23	-	-	-	-
MUFAF	24	74	27	79	-	-	-	-	-	-
MUFSC	7	8	14	8	-	-	-	-	-	-
Total	**171**	**518**	**213**	**549**	**251**	**548**	**180**	**578**	**180**	**578**

* human resources; ** research activities

NARO's annual research program planning starts on 1 July and ends 30 June. In January practitioners at each institute collect and process research activity data using research proposal forms to build the ex ante database. (This is only for the proposals that are approved by the NARO Board in December.) This data is then aggregated into the national ex ante database. In February, institute and national ex ante MIS reports are prepared and distributed to different institutes and NARO headquarters in March.

In May, data on research activities and human resources actually implemented during the current fiscal year is collected from institutes and collaborating Makerere University faculties to form the ex post MIS database. Data entry and clean up of both human resource and research activities is finalized in June and the ex post reports at both institute and national level are prepared in July. These reports are distributed to all research institutes and to NARO headquarters at the beginning of August and presented by the practitioners and to the respective audiences towards the end of August. The outputs from the MIS at different stages provide input to the program planning process at an appropriate stage.

Individual scientists prepare their proposals and justification for the continuation of uncompleted research activities in July and August. During this period, the proposals are discussed, amended, or rejected, and prioritized at program level. In September, research program planning takes place at institutes at a meeting of scientists, stakeholders, and staff from MEPU. During this meeting, previous research work is reviewed, progress highlighted, constraints presented, and gaps discussed. This forms the justification for the continuation of the uncompleted research activities. In October, program planning and review committees sit to polish the individual programs presented. NARO headquarters consolidates them into institute research pro-



Inner circle = research program planning cycle
Outer circle = INFORM cycle

Figure 3.2: INFORM and the research program planning cycle in NARO (Uganda)

posals. These are presented to the NARO board in December for review and, where appropriate, approval, and the final product is NARO's research proposals.

3.3.2 Impact of the MIS

INFORM has been integrated into the day-to-day running of the research business of NARO. On the whole it has contributed to much more efficient research management by providing managers with (access to) accurate and timely information for decision making. Although hard to quantify, scientists' attitude and way of conducting research has tremendously improved. A series of training sessions for program leaders and other scientists at the request of institutes is a clear manifestation of the importance that the research managers attach to the MIS. The system is now recognized as a backbone for sustainable research.

MIS reports are presented first during research program planning to give a background of how research progressed in the previous year before research proposals are presented. The MIS has been used to answer ad hoc questions needed by different managers for decision making. Examples of such ad hoc requirements include:

- the current level of scientists' training
- approved research activities with their budgets
- genetic resources studies in NARO
- scientists' gender and distribution
- number of scientists per discipline per institute

Budget allocation is based on the facts provided by the ex ante databases. The MIS has also been used to match expected outputs of research activities with what has been achieved at each research stage for monitoring purposes and to filter out completed research activities for evaluation. It has also been used for human resources planning and development.

The management of physical resources is a prerequisite for a good research agenda. To achieve this it is important to know what you have, where it is, and what the status is. The first inventory of physical resources was created using INFORM, and this has been used by administrators and store managers to manage the assets.

3.3.3 Lessons learned and recommendations for the future

Before an MIS is introduced into an organization it is important to build confidence from the management in the system. Otherwise, it will be extremely difficult to obtain the necessary approval and support for implementing the MIS. ISNAR's demonstration to NARO's management serves as a good example to those who consider introducing and implementing an agricultural MIS in their NARS.

Uganda was the first country adopting INFORM where actual data collected from the research program was used in training MIS practitioners.[14] Much effort went into ensuring that the data collected was of acceptable quality. Despite this, the initial set of data was not very accurate, although it continually improved as more experience was gained and appreciation of the system grew. The main problem in the initial data capture was that some scientists were not used to monitoring how they spend their time. Thus time allocation to research, management, extension, training, studies, and other activities was mostly guesswork.

A few problems were encountered during the initial training. One requirement for participants was that they should have ample experience in using computers. How-

14. Previous INFORM implementations used a carefully constructed set of fictitious data for training purposes. (Ed.)

ever, several of them had very little experience, and it was therefore difficult for all to proceed at the same pace.

Some nontechnical problems that affected the speed of implementation include:

- lack of recognition by managers at all different levels of the efforts of practitioners in collecting, correcting, and processing data
- some practitioners were not provided with consumables, such as diskettes and printer ribbons, on time; some suggested that this could be resolved by a petty cash account for this purpose.

As the Reflex software that runs RMIS is a flat file database, it presents a great limitation in that you cannot produce an output from more than one database. This calls for a better software platform such as a relational database which can provide better flexibility. Other inherent weaknesses of Reflex include: failure to detect spelling mistakes, difficulty in exporting and importing data to and from other applications, and failure to cut contents in a field of one file and paste them in another file. In the near future, funds being available, we intend to convert to a relational model developed by ISNAR. However, this will require more powerful PCs, cost and effort of re-training and the cost of converting the current system to the relational one. In our view this would be more sustainable and friendly in meeting users' needs.

Reference

Nestel, B. 1991. An overview of INFORM, an information management system. INFORM Guidelines, Part 1. The Hague: International Service for National Agricultural Research.

3.4 Bolivia: Implementing INFORM and CIPFIS at PROINPA

by André Devaux and Govert Gijsbers, in consultation with Bertha Valverde, Alfonso Blanco, Antonio Gandarillas, and Carlos Niño Neira[15]

Background

The national potato research program of Bolivia (PROINPA) was initiated in 1989 through a project implemented by the Ministry of Agriculture and its Agricultural Research Institute (IBTA) and the International Potato Center (CIP). The Swiss Development Cooperation provided financial support. Implemented during 1989–1998, the project aimed to strengthen the human capacity of PROINPA, develop research and technology transfer programs, and contribute to increased potato production and productivity in Bolivia's production systems. PROINPA became the national center of technology development for potato and Andean tubers. It was recently converted into a private research foundation.

---

15. Bertha Valverde, Alfonso Blanco, and Antonio Gandarillas are staff members in the Programma de Investigacion de la Papa (PROINPA) in Bolivia. André Devaux and Carlos Niño Neira are from the International Potato Center in Peru. Govert Gijsbers is an ISNAR Research Officer.

As part of a strategy to increase the efficiency of the management of the program, the MIS of PROINPA would be improved. A review of CIP's collaboration with IBTA in the PROINPA project in 1990 suggested that the management of IBTA and PROINPA could benefit from the use of INFORM. A first visit by ISNAR took place in May 1991, during which plans for a pilot study and staff training were developed. In coordination with ISNAR and CIP, PROINPA presented a project proposal to a major donor of PROINPA, the Swiss Development Cooperation, to support the implementation of a management information and program budgeting system. In October 1991, the Swiss Development Cooperation agreed to fund the introduction and testing of INFORM at PROINPA.

One objective was to implement the system in PROINPA as a pilot exercise and explore the possibilities of expanding it to the other IBTA research programs. Another objective was to explore the possibility of combining INFORM and CIP's Financial Information System (CIPFIS), which PROINPA already used. CIPFIS is a CIP-developed project-based budgeting system, providing timely and accurate financial information for decision making. It is a tool for internal control and contributes to decentralizing responsibilities.

INFORM responded to several requirements for better management in PROINPA:

- It integrates and provides an overall view of research activities, budget, and personal data.
- It helps develop budgets by research project, including all the operative and personal costs.
- It improves monitoring of research activities by using data of project costs and staff time allocation by project and subproject.
- It decentralizes budget management to the project leaders.
- It provides PROINPA scientists with a tool to better manage their financial resources.

3.4.1 Implementing INFORM

In 1992 a two-person ISNAR mission to PROINPA held an introductory seminar to nearly all of PROINPA's 55 professional staff. During the next 10 days a small group received additional training. With the help of the ISNAR team, data for the 1991–92 season (160 experiments) were entered into the database. The database formats were translated into Spanish and adapted to PROINPA's needs (reflecting its particular structures, programs, terminology, etc.). The information was analyzed and a first INFORM report for PROINPA was produced.

At the end of the visit the results were presented to the PROINPA team, to staff from IBTA, and to technical assistance personnel.

The introduction of INFORM to PROINPA involved a number of firsts:

- The INFORM training course and analysis conducted in Cochabamba, Bolivia, was the first given in Spanish.

- It was the first time INFORM had been used for an activity of a CGIAR center.
- It was the first attempt to integrate INFORM's program budgeting with an institute's accounting system. (This involved CIP's financial controller in Lima.)
- It was the first attempt to use INFORM for monitoring and evaluation using a quantitative scoring approach.
- It was the first time that the INFORM key word system (thesaurus) had been extensively tested outside Sri Lanka, where it was developed.

The monitoring and evaluation work was particularly important as PROINPA had a specific interest in using INFORM for monitoring and evaluating its program implementation. This was done in a quantitative manner using a scoring system combining the progress of the activity, the quality of the work, and its relevance to national research priorities.

Following the first round of data collection and analysis, PROINPA appointed a coordinator in charge of INFORM activities. With good support provided by the two PROINPA directors from IBTA and CIP, the INFORM coordinator was able to produce comprehensive reports since 1993. ISNAR provided limited support in the form of follow-up visits for consultation and troubleshooting and a review of draft annual reports, but the program is essentially operated by national staff from PROINPA.

Two more follow-up visits from ISNAR, under the Swiss-funded project, took place in 1993. They focused on strengthening PROINPA's capacity to solve specific problems in financial analysis, using key words, using INFORM for evaluation purposes, generating reports for senior management, and introducing INFORM across IBTA.

Since 1993, PROINPA has made considerable advances. It involved a skilled documentalist in INFORM to enhance the quality and usefulness of the key words. More detail on geographic and agroecological coverage of the system was included. Efforts were made to increase the level of detail in the costing system so that it could include more direct costs. The monitoring of staff time on a monthly basis was introduced. Further work was done on monitoring and evaluation. Finally, a link between INFORM and CIPFIS was developed.

Following the completion of the PROINPA-ISNAR INFORM project, PROINPA has been largely maintaining the implementation of INFORM on its own, with very little support from ISNAR. It has done so quite successfully and PROINPA's INFORM work received very positive comments in CIP's last external review in 1994. Over the years CIP has consistently provided significant support to PROINPA's INFORM work, in particular through the development of a link to CIP's CIPFIS.

3.4.2 Data collection

To implement the system, the database formats provided by ISNAR were used. The first set of data on personnel and research activities was captured in 1992–93. The project database was adapted to the operative plan and each scientist was requested to complete a personal biodata questionnaire. The scientists were requested to estimate the time they would allocate to the research activities of the operative plan and to other types of activities related to their work plan. These classes of time allocation are shown in table 3.5, which shows changes over a five-year period. These changes were made to simplify the time allocation exercise.

3.4.3 Adapting INFORM to the requirements of PROINPA

Unit of analysis

The PROINPA research program was organized into projects, subprojects, and activities. Initially the unit of analysis used for INFORM was the activity (trial or survey), but experience showed that it was a too detailed level for both financial and information data management. Processing the information took too long. It was decided to use the subproject (combining three to four different activities on the same theme of research) as the unit of analysis.

Later on, other criteria were used to analyze PROINPA activities, such as the distinction between research and technology transfer, and, due to the decentralization of the research system in Bolivia, activities were organized by geographical and political regions. INFORM codification was useful for organizing and grouping the activities. This codification was also adapted to combine INFORM to the CIPFIS.

Collecting data on researcher time

Rather than collecting data on research time on an annual basis as proposed by ISNAR, PROINPA gathered this information on a monthly basis, using workdays instead of percentages. This use of time sheets has improved accuracy and helped estimate the personal time and cost requirement in planning activities. However, there were some inevitable problems:

- Instead of reporting on a forced 100% of time, staff now report actual days. Underreporting and overreporting (e.g., staff working weekends in the field) occur. A decision was made to adjust this once a year, using a simple formula.

Table 3.5: Classification of Actions for Allocation of Time Data Collection: Comparison between the 1st Year of INFORM Implementation and 5 Years Later

1991–92	**1996–97**
Research organized by activities	Research organized by subprojects
Administrative matters linked to research such as planning, evaluation, and budgeting	Administration and internal coordination
Technical cooperation given to other institutions	Institutional cooperation including training and technology transfer activities
Communication: Conferences, preparation of technical leaflets	
Training given by scientists	Training received by scientists
Training received by scientists	Services given to other institutions: sample analysis, consultancies

- It was difficult to get some people to complete the forms each month.
- In 1992, it was decided to include graduate students (working on a stipend) in the system. While they cost very little (hardly affecting resource allocation), they presented a problem in data collection as they are scattered over many activities. It was decided to remove students from the monthly data collection exercise and allocate 80% of time to their thesis subproject and the rest to "Other."
- The timesheets that were used were rather user-unfriendly (long lists of project numbers that were photocopied). One way to correct this was to use personalized and partially filled in forms to make it as easy as possible for researchers. These personalized timesheets had another major advantage: they could be used to present the researcher with the time allocation data provided earlier in the year. This feedback also helped researchers realize that the information they provided was actually being processed and used.

Quantitative scoring approach for monitoring and evaluation

As PROINPA had a special interest in developing a monitoring and evaluation methodology, a quantitative scoring system was introduced at the subproject level. The evaluation was based on the following criteria:

- **Operational progress of the subproject**—to evaluate how much of the work planned had been implemented. The score was between 1 (not initiated) and 5 (all the activities implemented and reported).
- **Efficiency**—to estimate the progress of the subproject based on the indicators established at planning time and their contribution to the main objectives of the project. Score between 1 (little progress) and 15 (more than expected).
- **Applicability**—to evaluate the potential application or diffusion of the results as a methodology or a technology. Score between 1 (little application) and 15 (outstanding).

The subprojects were evaluated at the time of the annual review of the program and scored by a team of three scientists. The total score obtained was used as an indicator of the performance of the subproject. This data combined with the operative cost and the time allocated to the subproject were used as an evaluation tool and were useful at the time of planning of the next operative plan. The quantitative scoring was not used systematically for all subprojects; it was useful when there was some discrepancy about the continuation of a subproject. The information processing required quite a lot of time for the INFORM coordinator and could not always be available at the planning time. It was then decided not to apply this evaluation systematically.

Integrating INFORM and CIPFIS

Attaching much importance to finance management, PROINPA, assisted by CIP, built a link module between INFORM and CIPFIS. This helped PROINPA build its budget and to monitor more accurately the direct as well as indirect costs of the subprojects. It was also useful to address the reporting requirement of the different donor agencies PROINPA had to deal with and estimate its research investment by geographical and political region.

The integration of CIPFIS and INFORM provide PROINPA's MIS with an on-line tool for financial management, budget preparation and monitoring, and information for decision making. The dynamic and the speed of processing information process in CIPFIS provided actual information to the INFORM structural analysis. The combination of the two systems enabled access to accurate information on the personnel and operational costs for the activities in which PROINPA technical staff was involved (see table 3.6). The differentiation of personnel and operative costs by research subproject was useful at planning time, and for project proposal preparation it helped the directors and project managers to better quantify the human and financial requirements.

Cost allocation

As many MIS users within PROINPA did not agree with the original INFORM approach of prorating all direct and indirect costs to all the research projects, PROINPA has developed a system where all direct costs are charged to subprojects and indirect costs to five other cost centers, as indicated below and shown in figure 3.3. The difference with the old INFORM approach is that the expenditures related to general costs centers are not redistributed to the subprojects. Only in the case of services expenditures, costs were divided between research and administration following a ratio of 70% for research and 30% for administration. These costs were redistributed to the research subprojects but in a transparent form, so that researchers knew how they were charged for services. The advantage of this system was that it gave project managers a more realistic and intuitively right feeling for the costs of their projects.

Table 3.6: Personnel and Operative Costs by Activities at PROINPA (1996–97 Season)

INFORM Code	Description	Person years*	Personnel costs (in US$)	Operative costs (in US$)	Total costs (in US$)
11111	Administration and internal coordination	7.94	72,101.27	14,676.04	86,777.31
22222	Institutional cooperation including training and technology transfer activities	5.96	57,568.58	16,572.31	74,140.89
33333	Training received by scientists including postgraduate students	2.85	20,572.02	50,267.20	70,839.22
44444	Services	2.30	10,280.39	4,618.33	14,898.72
	Research organized by subproject	80.96	350,555.78	275,087.97	625,643.75
	Total	100.00	511,078.04	361,221.85	872,299.89

* including scientists and students working in PROINPA

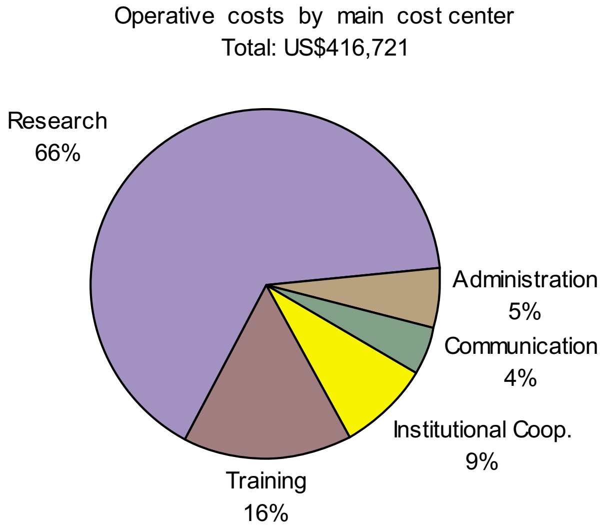


Figure 3.3: How nonresearch costs are attributed to separate cost centers

The INFORM structure was organized around the following cost centers:

- operative costs linked to the PROINPA management and administration such as institutional cooperation and services
- operative research costs linked to subprojects
- administrative costs
- training received by PROINPA scientists, including post-graduate training
- communication costs including publication and the documentation center
- non-INFORM, extraordinary expenses that covered nonbudgeted activities which were normally refunded later through external funds.

With the help of the CIPFIS, those costs were then sorted by agency or donor and by geographical and political region.

3.4.4 INFORM-R

INFORM-R is an upgraded version of INFORM that ISNAR developed in Access, a relational database software. As it was designed for a nationwide organization with national and institute levels, it could not be used in this way by PROINPA. Customizing INFORM-R to PROINPA's requirements proved to be difficult. Instead, PROINPA took the existing and well established INFORM and rebuilt it in Access, using its relational features to avoid double data entry, and using features such as look-up tables as ISNAR had done in INFORM-R. This adapted INFORM version was a basic system with most of the main functionality of INFORM.

3.4.5 Problems encountered

Training

To be run adequately, INFORM requires a well-trained coordinator who understands the system and has experience in using computers. The initial training at PROINPA on INFORM gave a good general basis to start implementing the system. But later on, more follow-up training support was required to enhance the efficiency of the coordinator and to improve users' understanding and appreciation of the system. Users were not always convinced of the need to complete the timesheets and of the usefulness of the system. Moreover there was a change of coordinator after three years, but there was no provision for training the new coordinator, who consequently had to learn by doing. This situation weakened the implementation of the system in the program.

Organization of INFORM structure

Initially INFORM was organized around project activities, but the output of the system did not justify the input in terms of time and work required for processing the information. It was then decided to change the unit of analysis from around 300 activities to 100 subprojects. The lesson learned was that when starting with a new management system, there is a tendency to be scrupulously detailed, which tends to complicate the implementation process.

Collecting researcher-time data

As mentioned above, evaluating researcher time presented a problem. But the exercise helped PROINPA's staff to better appreciate how they used their time, and it helped the program management to better estimate the personnel cost in the project planning process.

Cost allocation

The original INFORM redistributed all the indirect costs to the activities, which meant that each activity carried a load that represented different items and increased the cost of the activity a lot. The researchers, however, did not accept the figures. They did not recognize their experiments and intuitively felt that these were not really the costs of their activities. Even though the logic was explained to them and some then understood it, others still did not accept the figures.

An alternative approach was to include in the system other cost centers than the research activities in order to collect general cost items. Examples of such cost centers are explained above. When the expenditures related to these general cost centers are not redistributed, the costs of individual activities become more realistic from the point of view of the project managers. It also helps to figure out better the importance of the indirect costs, especially those related to administration. In a project-based management system it helps to determine the percentage of overhead by project.

3.4.6 Lessons learned

The application of INFORM and CIPFIS in PROINPA has contributed to a better organization of the program on a project-based structure. The project database was especially useful for the directors of the program. The graphical illustration of combined factors such as cost of projects, number of persons by project, personnel costs, etc., allowed the directors to obtain a global view of the situation of the project and could follow-up more efficiently, analyzing also specific information by subproject. This differentiated information by operative and personnel costs was also useful at the time of planning and for project proposal elaboration. The personnel database was less used and should have been updated every year, for example to monitor the short-term training activities by scientists.

INFORM and CIPFIS were also useful as an information tool to present data on budget, geographical distribution of activities, personnel, and other data to different audiences: donors, local authorities, evaluation missions, and annual report writers.

More emphasis should have been given to training and adequate follow-up of the coordinator of the MIS. This person should be able to communicate and obtain the required information from staff and provide the processed information to management.

The flexibility of INFORM and its logical concept has allowed PROINPA to adapt it without depending too much on ISNAR. With the help of CIP, CIPFIS was adapted to the evolving structure of INFORM in order to maintain the complementarity of the two systems.

3.4.7 The future

As PROINPA has recently changed its institutional status and become a foundation, it is in a transition period, adapting new management tools and a new accounting system. The concept of project databases developed with INFORM has been used to develop the new databases of the foundation. The foundation has not yet implemented an MIS. A follow-up and some training from ISNAR and CIP could be useful to support the foundation's management in this transition process.

3.5 Brazil: Implementing a state-of-the-art MIS in Embrapa

by Suzana Lima, Antônio Maria Gomes de Castro[16], and Moacia Pedroso

3.5.1 Embrapa's early experiences with MIS

Since it was established in 1973, Brazil's national agricultural research organization, Embrapa, has developed several MIS systems with various degrees of success. While all aimed to increase the organization's efficiency and effectiveness, some

16. Suzana Lima was the leader of the development project of SIGER, Embrapa's Management Information System. Antonio de Castro is SIGER's Implementation Coordinator.

systems developed in the 1980s had no clear linkage to Embrapa's planning system. Others were initiated almost at the same time as part of the strategic planning process and hence were an integral part of the process (Flores et al. 1994). In any case, all these systems may be characterized retrospectively as reflecting a general understanding of key management principles. These relate to planning, the importance of MIS in management, and technical conditions for their development and implementation.

Since the 1980s, Embrapa used at least three research and development MIS systems:

1. The Research Project Information System (SIPP) focused on Embrapa's research projects and experiments, requesting information on their execution every six months. Data was captured on paper forms, which were sent to an Embrapa sub-unit in charge of updating and storing this information in a central computer. SIPP was operational until the early 1990s.
2. Almost at the same time as Embrapa began its strategic planning process (Flores et al. 1994), a new MIS (SISPAT) was developed to facilitate the management of the research centers' annual working plans. This system, which is still operational, was first developed in Clipper software but was redesigned in Delphi.[17] Embrapa's centers enter the data on diskettes, which are sent to an Embrapa sub-unit in charge of the system's data updating and storage.
3. Just after the strategic planning process, and with the clear aim of making operational the planning system that emerged from that process, a third MIS was developed: the Planning System Information System (SINSEP). This was also first developed in Clipper but later made to be operated in Access.[18] It also used data captured on diskettes in a manner similar to that described above for SISPAT. SINSEP focuses on Embrapa's projects and subprojects, although it also has some basic information on Embrapa's programs.

However, these systems had several weaknesses:

- No systematic needs assessment was undertaken prior to their development. The systems thus did not meet the information needs of the clients.
- The information that they provided was fragmented because the systems were managed by different sub-units in Embrapa.
- The linkages with Embrapa's strategy were poor, as the systems focused on the operational decision level and did not relate to Embrapa's mission and objectives.
- Emphasis was placed on planning at the expense of monitoring and evaluation. The process of planning, monitoring, and evaluation as an integrated whole was not considered.

17. Clipper is a database application software for small- to medium-sized systems. Delphi is a database application development tool for medium-sized to large systems.
18. Access is a database application software for small- to medium-sized systems.

- The systems could not communicate with each other, although they represented different levels of aggregation of a unique information basis (subprojects and projects). The main reasons for this are (1) they were developed using different kinds of software and (2) they were conceptualized in isolation.
- There is little emphasis on feedback flows in these systems.

As a direct consequence, the output of these systems (information for decision making) was of poor quality, unreliable, and hardly useful for most clients. Users viewed the systems as merely bureaucratic rituals, with no connection to research reality and needs.

3.5.2 From SINSEP to SIAVE to SIGER

SINSEP was mainly developed to meet the demand for an innovative research and development MIS after Embrapa’s new planning system introduced several changes in the management of research and development. However, as the demand was rather urgent, SINSEP was developed in a very short time and initially showed several design errors, and users were reluctant to accept it.

The following were the main initial limitations of SINSEP:

- It was heavily oriented towards research and development planning and largely ignored monitoring. It paid almost no attention to ex post impact evaluation.
- It was oriented towards projects without any linkage to, e.g., research and development center planning, programs, and strategic plan.
- There were too many open-input information items, making the retrieval of aggregated output very difficult.
- The support in software and hardware during the design phase was weak, resulting in an unfriendly user interface.
- The system was loosely implemented, causing confusion among researchers of SINSEP characteristics.

These limitations led management to appoint a team to develop a new MIS to replace SINSEP. The team conducted an extensive survey of the information needs of a sample of 210 researchers (10% of the total number of Embrapa’s researchers). The findings indicated the needs of managerial information at the strategic, tactical, and operational levels. The group then proposed a system, SIAVE (Monitoring and Evaluation System of Embrapa) (Paez et al. 1994), which was based on SINSEP but with a revised conceptual framework, better computer facilities, and improved usefulness. The conceptual model of the system was constructed covering the monitoring and ex post impact evaluation of subprojects, projects, and research centers’ annual working plans.

Unlike SINSEP, which was first conceived as an auxiliary planning device, SIAVE took the format of a true MIS. At the heart of design was the notion that the system should be useful (and used) in the decision making process of Embrapa’s research managers. Its design was client oriented from the beginning.

Monitoring and evaluation formed the core of the system. Better decision making in the management of research and development was the system's main goal. User-friendliness, conceptual simplicity for increased understanding, and controlled input were all constraints. The target was to produce an MIS that was powerful enough to be used as a research and development MIS but that researchers would not consider a bureaucratic burden.

SIAVE was fully designed but not implemented. During the design phase, the team in charge of its development realized that SINSEP would have to be heavily reshaped in order to accommodate the conceptual and instrumental changes demanded. Therefore, instead of upgrading SINSEP, it was decided to construct an entirely new MIS, SIGER (Embrapa Management Information System). Incorporating most of SIAVE's subsystems, SIGER benefitted strongly from the development of SIAVE.

3.5.3 Planning, monitoring, and evaluation and the SIGER architecture

The Embrapa Planning System (SEP) used the conceptual framework for planning, monitoring, and evaluation adopted as the basis of SIGER. Its general conceptual model is described in Embrapa (1995).

The elements of the SIGER structure are the programming instruments and articulation mechanisms that comprise the Embrapa Planning System. The structure also takes into account the information needs of its customers—the managers in charge of these instruments and mechanisms (Goedert et al. 1994).

The following planning instruments are considered in SIGER:

1. **The strategic plans.** These instruments define Embrapa's or a center's course towards the accomplishment of its mission. They are managed by Embrapa´s Board of Directors and by the heads of centers, respectively.
2. **The annual working plan.** This synthesizes the annual programming work, which is defined in projects and subprojects of each unit. The plan is managed by the heads of the research centers, with the help of internal technical committees.
3. **The program.** This defines the institutional polices in certain themes or national priority issues and is managed by technical program commissions.
4. **The project.** A project describes actions to solve customers' and users' priority problems. These problems have a broad and complex nature. Projects are generally multidisciplinary and involve more than one institution or research center. A project is managed by a project leader.
5. **The subproject.** This is an auxiliary instrument for organizing the necessary activities to solve problems of a more specific nature, embedded within the problem aimed at by the project. It generally requires only one discipline and has local reach. It is managed by a subproject coordinator.

There is a hierarchy among the different levels of decision making (and between the several planning instruments). This hierarchy determines that there should be consistency among the strategic plans, the tactical plans, and the operational plans. The lower levels of planning constitute the requirements for the upper levels. This is made operational, in MIS, through the increasing aggregation of information. Thus, most of the necessary information for the different subgroups of customers can be obtained through the aggregation of information originating in the lower levels of decision.

Figure 3.4 presents the architecture described above. It also presents the flow of information that characterizes SIGER. The entry of information into the system occurs, mostly, at the base (subprojects) and proceeds to the planning instrument immediately above.

The system also includes a flow in the opposite direction, with each upper level informing the immediate lower level on its performance, either in the planning, execution, or accomplishment of established objectives.

3.5.4 System attributes

Attributes of the system must be selected in order to help design the concept, create instruments, and informatization[19] and implementation. These attributes play an important role: they help minimize, from the first moment, resistance to and dissatisfaction with the system on the part of customers and users.

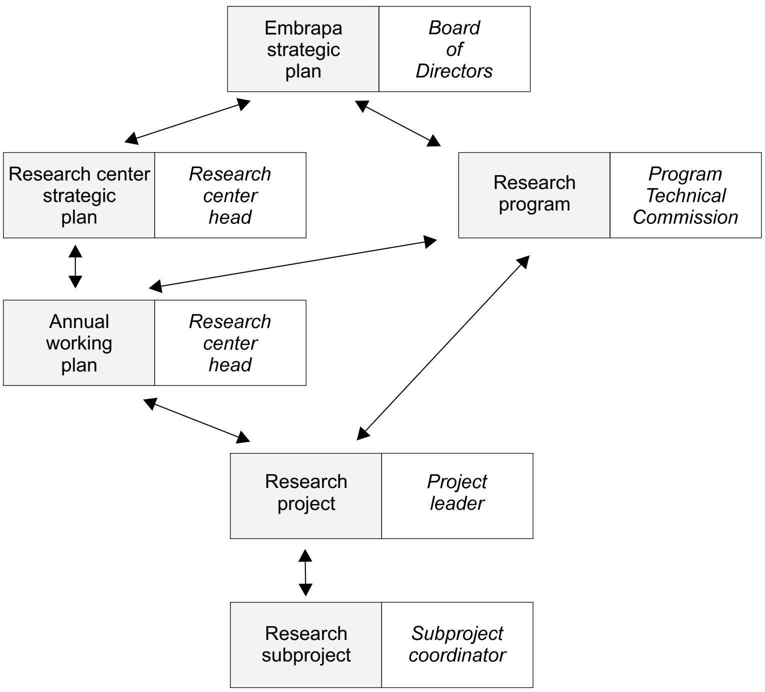


Figure 3.4: Architecture and flow of Embrapa´s MIS

19. See the glossary in Annex 1 for an explanation of the term informatization.

Users will especially resist the system if they perceive it as being of little value in the decision making process. This may be caused by (1) badly formulated program, planning, and evaluation (PM&E) concepts, (2) excessive amounts of information being collected, without clear application for the administration of projects and programs, (3) information collected in an unstructured, open-text format, hindering data collection, processing, and retrieval by users, (4) disproportionate amount of time required for the entry of information, e) complex and user-unfriendly software, (5) technically poor software without facilities for word processing and graphics, and (6) poor system documentation. Table 3.7 presents some MIS attributes, based on Bolivar et al. (1997). Some attributes are more important for design and others for system implementation. SIGER incorporates all these attributes as guiding principles for its design and implementation.

3.5.5 SIGER conceptual design

SIGER was developed in five stages:

I. Assessment of information needs of customers and users

This stage started with identifying subgroups of clients who would require information generated by SIGER. These clients are the same as those defined for SIAVE (see above): Embrapa managers at all levels, subproject coordinators, project leaders, research centers' technical internal committees, program technical commissions, heads of research centers, and Embrapa's Board of Directors. These subgroups constitute the internal clients of SIGER.

The system was also designed to take into account the information needs of external clients, that is, institutions (or their representatives) to which Embrapa is accountable: those with which it has partnerships, national and international financial institutions, nongovernmental organizations, the national congress, the media, and society at large. They all demand corporate information, be it about plans (strategic or tactical), projects in execution, results obtained, or resources invested. This kind of information is generated, typically, by the PM&E processes that SIGER supports.

Table 3.7: Attributes for MIS Development and Their Importance for Design and Implementation

Attributes	Design	Implementation
Linkage to PM&E	High	Medium
Customer driven approach	High	High
Evolutionary approach	Medium	High
Decentralization	Medium	High
Integration with other systems	High	Medium
Flexibility and simplicity	High	High
Informatization	High	High

The information needs of the two groups of clients of SIAVE and SIGER were assessed in 1994 (Paez et al. 1994). Table 3.8 gives examples of the different kinds of information needs at the various management levels (and of external clients).

II. Definition of limits and objectives of the system

The objectives and limits of SIGER are part of the general conceptual model of the system (Embrapa 1995). The model establishes that SIGER will focus on the following: (1) PM&E of Embrapa's R&D, (2) a database of clients and their demands, (3) a database of technologies, services, and generated products, (4) a subsystem for impact evaluation of research, and (5) a subsystem for individual performance appraisal. Only the first four are included in the MIS developed; subsystem 5 is being developed independently of SIGER (although it also will be capable of communicating with SIGER).

Table 3.8: Examples of Information Needs of SIGER Customers

Client category	Information needs (examples)
Society at large (politicians, media, agribusiness, etc.)	a) research results aimed at specific solutions b) costs and benefits of investment in research c) environmental impact of technologies developed
Coordination, control, and accounting organizations	a) R&D impact on agribusiness b) performance of R&D in development programs c) relationship between demand for R&D and supply of solutions
National and international R&D donors	a) social and economic impacts of technologies developed b) physical and economic performance indicators c) technologies, products, and services commercialized
Embrapa Board of Directors	a) contribution of programs and projects to accomplishing Embrapa's mission b) research centers annual working plan evaluation c) budgetary and financial flow by program and research centers working plan
Program managers	a) relationship between demand on program and supply of projects b) relationship between planned and executed projects c) social, economic, and environmental impacts of projects within the program
Annual working plan managers	a) relationship between planned and executed projects b) research center annual investment in R&D c) research center budget and financial flow
Project leaders	a) relationship between planned and executed projects b) project management quality c) financial execution of projects

Source: Paez et al., 1994.

III. Definition of variables, performance indicators, and instruments

This stage corresponds with the detailed conceptual design of the MIS, which serves as a guide for its informatization. The stage includes identifying component variables of SIGER, performance indicators, and instruments for data entry and output.

The variables were selected taking into account the R&D model and Embrapa's Planning System. "Free" entry of information was avoided; there are many menus to classify the information, so that it can be easily recovered, quantified, and cross-tabulated with other variables.

Performance indicators correspond with variables that refer to PM&E properties and can be used to measure the performance of each SIGER instrument as follows:

- **Planning performance indicators.** These refer to the end results of a project or subproject, when a finished product is ready to be delivered to clients. Established during planning, they will be used again during evaluation to compare, for instance, actual performance to planned performance. There are three types of these indicators in SIGER, each addressing a different subgroup of managers:
  - R&D performance indicators
  - institutional development and management performance indicators
  - production and commercialization performance indicators.
- **Planning effort indicators.** These are indicators of the efforts that are made to accomplish the end results defined through performance indicators. These indicators describe the various methods that can be put to work toward the end results, such as research, validation, diffusion, marketing, and building skills in teams. Each method has its own set of effort indicators. During monitoring, these indicators will be focused on seeing how the execution is progressing.
- **Initial ("ex ante") evaluation indicators.** These can be categorized as follows:
  - *Linkages to planning* is the degree to which the project or subproject complies with what was established in the instrument to which it is subordinated.
  - *Linkages to demands* is the degree to which the project or subproject addresses the demands of the instrument immediately superior to it.
  - *Technical and managerial feasibility* consists of the fit between planning and several technical and/or managerial factors (team technical capacity, managerial capacity of the responsible person, proposed calendar, budget, support offered by the unit) that may facilitate or hinder its execution.
  - *Expected impacts* consist of estimates of economic, social, environmental, and institutional impacts expected to happen as a consequence of projects and subprojects results.
- **Monitoring indicators.** SIGER's instruments for monitoring adopt some of the indicators widely used in the literature about evaluation (Sbragia 1984), as well as some specific ones for Embrapa. These indicators may be categorized as follows:

- *Management quality* is defined as the degree to which management was efficient in providing conditions and making decisions in regard to necessary actions for the execution of projects, subprojects, programs, and plans.
- *Observance of deadlines and costs* is the degree to which, in the execution of a subproject or project, its completion date and real costs match the planned deadlines and costs.
- *Potential of obtained partial results* consists of the degree to which partial results identified during monitoring may generate sales, patents, or adoption by customers.

- **Final ("ex post") evaluation indicators.** SIGER adopts the following indicators in evaluating projects and subprojects:
  - *Management quality* - (see above)
  - *Observance of deadlines and costs* - (see above)
  - *Technical quality of the results* is the degree to which the results obtained by a subproject or project match planned outcomes, given the means that were available to reach them.
  - *Contribution of results to knowledge advancement* consists of the degree to which the results obtained from a subproject or project contribute to the progress of knowledge in a specific area of study.
  - *Impact on the institutional sustainability* is any benefits, in terms of capacities (human, physical, or financial) or commercial or technical transactions that are obtained as by-products of a subproject or project and that contribute to institutional invigoration.
  - *Economic impact* is a potential benefit, in terms of increases in efficiency or reduction of costs in processes or agricultural products, facilitated by the results of a subproject or project.
  - *Social impact* is any benefit, in terms of generation of employment, income increase or improvement of life quality, for the target population of a subproject or project.
  - *Environmental impact* is the potential effect on the natural resources of an area, generated by the results of a subproject or project.

IV. System informatization

The system asks for a description of the obtained final results (characterized in terms of commodity chains, productive systems, natural systems, geographical reach, and the disciplines in which the result may be classified), the situations in which they can be applied, and the conditions in which they were tested and validated.

SIGER has available four functions that support appraisal, planning, monitoring and evaluation respectively. These functions are applied to subprojects, projects, programs, working plan, research center strategic plan, and Embrapa strategic plan.

Hence, the system offers 24 instruments (four functions times six programming instruments) for the various groups of customers and users.

In SIGER, a detailed and thorough design of the information flow and the instrument managing the flow allow the automated control of the whole process of PM&E.

There are also established rules for reading and writing access for different clients and users. Since the system will operate on networks (Internet or satellite communication), security of information is a major concern.

SIGER is being developed to use a hardware architecture consisting of an integrated network of workstations (or PCs) connected to each other and to common peripherals; in this case, all levels of clients may have up-to-date information available at any time, and updating is done by a single data entry. The interfaces of the system are developed in Delphi, using a Windows platform and with the editing facilities of the better-known word processors.

The basic system architecture was first developed by Pedroso et al. (1995) and its detailed configuration is fully described in Leite et al. (1998).

V. Verification and validation of SIGER

This is the final stage of system development before it is implemented and consists basically of the following:

1. A verification test to check the consistency of the electronic version with the conceptual design of the system. Based on this test adjustments are made in the electronic version.
2. A validation test of the corrected, electronic version with a sample of clients, in real conditions. This test will suggest improvements to the electronic version.

These two tests have been done on SIGER and demonstrated without a doubt that the system could enter into operation with some small adjustments, since it was complying with client's information needs to a degree well beyond expectations. The tests also showed that SIGER behaved as expected when the system was evaluated in the light of the attributes with which it was drawn (see above). Table 3.9 presents validation results for some of the attributes of the system. The conceptual framework and methodology used in SIGER's validation are described in Castro et al. (1998).

The validation of SIGER showed that all the characteristics described above were well received by the clientele; the conceptual design, particularly, emerged from the validation as the strength of the system, although it is also the point that demands more care during SIGER implementation.

A more detailed account of SIGER conceptual design is given in Lima et al. (1998).

3.5.6 SIGER's main features and outputs

The following are SIGER's main features:

- It responds to the information needs of its clients, guaranteed through needs assessment, validation, and future system monitoring.

Table 3.9: Average Scores of Evaluations Concerning Critical SIGER Attributes

Critical attribute	**Client**	
	Project leader	**Subproject coordinator**
Ease of use (navigation) – Flexibility and simplicity	8.2*	8.4
Ease of use (word processing) – Flexibility and simplicity	7.7	8.7
Ease of use (data input) – Flexibility and simplicity	8.2	8.2
Conceptual clarity – Linkage to PM&E	6.4	7.1
Information needs attendance – Client-driven approach	7.3	7.7

* 10 = very good; 1 = very bad

- It covers the total of PM&E functions, such that the process of PM&E is considered as an integrated whole.
- It integrates PM&E functions, such that information that is used in more than one function can be recovered and/or processed as required, facilitating user interaction with the system and allowing for the adequate management of activities.
- It integrates planning instruments at the strategic, tactical, and operational levels in a way that will allow for increasing aggregation of information in an upward direction and that will provide consistency between the three levels in a downward direction.
- It integrates with other Embrapa systems, allowing for customized use of information entered in different systems and requiring only one input of a given information item from users.
- It is consistent with SIGER's conceptual framework and Embrapa's planning system conceptual framework. Some of the principles and concepts that guide research at Embrapa can be checked against reality from time to time.
- It automates the PM&E process as a whole, facilitating management of the process.
- It emphasizes feedback flows in the system, allowing all users to benefit from information entered.
- Written and on-line help focuses heavily on PM&E concepts rather than on computation details.
- Required input information is menu driven. Text answers are required only occasionally, enhancing the system's capability to offer customized reports that require a combination of the several variables entered into the system.
- Several useful facilities are available in SIGER: word processing, tables, graphics editing, etc.

- It operates through networks (Internet or via satellite). This guarantees on-line updating and storage of information for all clients, strengthening the system's quality and reliability.
- A less sophisticated version is available for clients who cannot interact with the system through these networks.

Table 3.10 presents instances of effectiveness and efficiency information, which will be available from SIGER's evaluation instruments. These examples demonstrate the usefulness of these types of information for management, at different levels.

SIGER facilities for output of aggregated information allow for improved links among operational, tactical and strategic planning levels in Embrapa. Managers at these levels have a strong source of information. An improved level of R&D management

Table 3.10: Instances of Information Related to Efficiency and Effectiveness, Available in SIGER, Generated by Ex Post Evaluation

Available information	**Instrument**			
	Project/ subproject	**Program**	**Annual work plan**	**Strategic plan**
1. General statistics	X	X	X	X
2. Listing and quantity of projects, subprojects, by product, discipline, geographic region, ecosystem, Program and Research Center		X	X	X
3. Obtained versus planned end results	X	X	X	X
4. Time for reaching end results (actual versus planned)	X	X	X	X
5. Data base of technologies, products and services generated by projects and subprojects, by Program, by Research Center	X	X	X	X
6. Publications and other forms of diffusion	X	X	X	X
7. Technical quality of execution and results	X	X	X	X
8. Contribution of obtained results to institutional sustainability	X	X	X	X
9. Contribution of obtained results to knowledge advancement	X	X	X	X
10. Social, economic and environmental impacts from obtained results	X	X	X	X
11. Global performance evaluation	X	X	X	X

Source: Castro et al. (1998)

can be expected and, therefore, greater internal efficiency and institutional effectiveness.

3.5.7 SIGER implementation

Conceptualizing SIGER took a long time because the environment in which the system is to perform is complex and because it is an innovative system. The environment comprises 500 R&D projects and 2000 subprojects, 17 national R&D programs, 56 R&D centers and institutes, and the professional activities of about 4000 agricultural researchers spread throughout Brazil (2000 from Embrapa and 2000 from the Brazilian Agriculture Research System).

Obviously, conceptualizing a system to cover the information needs of a number of clients of this magnitude is a tremendous task. Actually implementing the system is an even greater undertaking. A good strategy is therefore needed to successfully replace the MIS currently in use (SINSEP) by SIGER.

The first problem was replacing SINSEP. At least five years of R&D activities are recorded in the SINSEP database. There are on-going PM&E activities that have to continue without interruption. Furthermore, the new system has several conceptual innovations in its design, making an automatic transfer impossible. Therefore, moving from SINSEP to SIGER requires some efforts from researchers at some point of the process.

A strategy was set for implementing SIGER. The main strategic points were:

- The use of a massive internal campaign for pointing out the advantages of using SIGER, at the individual, managerial and institutional levels.
- A large training process, covering researchers, R&D managers and computer support staff and comprising the PM&E concepts and the computer aspects that each one should be able to cope with, in order to fully operate SIGER.
- Production of support material, in written and visual format, aiming at information needs of users. There will be a thorough and detailed user manual, bulletins, and videos, with the necessary information for using the system.
- Revision of the R&D management structure aiming at full use of SIGER potentialities in R&D management.
- Diagnostics of a computing support system in the R&D centers, bringing up-to-date the available hardware, software, and personal skills demanded by the implementation of SIGER.

In the formulation of the implementation strategy, several assumptions and requirements were considered. The operation of the system should take into account:

1. the current rules in the planning system (SEP)
2. the established Embrapa planning calendar
3. the existence of internal and external R&D partners

4. the direct use of the system by its users (since previous MIS had been managed by third parties)
5. the smallest degree of disruption of the ongoing R&D planning
6. the diversity of computers skills in Embrapa
7. the different aptitudes and existing capacities in planning and R&D management

Several methodologies were used in the implementation, aiming at motivation, conceptual and instrumental training in system use and user assistance. The main methodologies used were: (a) training of instructors, users, and computer technicians in system use and operation, (b) project leaders' training in projects management, (c) SIGER´s homepage with instructions on its use, (d) discussion lists, (e) videoconferences addressed to users and computer technicians, and (f) bugs and suggestions recording. Additionally, the user was given written support material, a list of clues on the system, and general instructions through individual e-mails. During the whole implementation period an assistance service by phone was also available, answering to managerial, conceptual and computer use questions.

This whole methodological apparatus was designed to reduce technical, managerial, and cultural difficulties generated through the implementation process, minimizing potential resistance to the changes thus introduced.

SIGER implementation was initiated in November 1998, with the training of 25 so called "multipliers." It continued in the first semester of 1999, when the multipliers trained some 1300 project leaders (from Embrapa itself and from the Brazilian NARS). The training focused on SIGER's basic concepts and use. Part of the leaders group also received a course in project management.

The participants evaluated all training events. The evaluation was made with the help of Likert scales (five-point scales), with 1 the most negative end of the scale and 5 the most positive. The variables considered in this evaluation referred to (a) training planning, (b) training development, (c) imparted information, (d) instructors' knowledge, and (e) general results of the training. These groups of variables measured the reaction of the participants both to the training and to SIGER.

The general averages obtained in this training were 3.9 for the training planning, 4.1 for its development, 4.3 for the imparted information, 4.7 for the instructors' knowledge, and 3.8 for the performance improvement that might be expected with the system's use. These averages presented small variations among the research centers that participated in the training, due to, probably, differences in the skills of the teams of instructors. The results indicated that SIGER implementation began very favorably, but very cautiously, because the emergence of problems and the resistance during actual system's use were predicted and expected.

3.5.8 Evaluation of limitations and difficulties

In his book *How to implement information systems and live to tell about it*, Fallon (1995) points to the difficulties in implementing information systems in a network architecture, especially in situations characterized by scarce resources and heterogeneous capacities, hardware, and software. According to the author, systems that operate in a stand-alone fashion are far away less complex, when compared to those designed to operate through a network.

SIGER is a system projected to operate in these two ways: in local networks and stand-alone. The use of these two configurations confirmed Fallon's warnings. In situations with a network configuration, SIGER implementation presented a significantly larger number of problems than in those that used a stand-alone configuration.

The implementation team evaluated the impact of the identified problems on SIGER credibility. In this evaluation, some 60 Embrapa R&D managers answered questionnaires. The results of this evaluation indicated the following categories of problems, occurring in the first year of SIGER implementation:

- **Conceptual** (13% of total number of problems indicated). Difficulties related to the conceptual model of planning, monitoring and evaluation (PM&E) adopted by SIGER.
- **SIPF / SIPJ** (6%) . Difficulties in registering external personnel and organizational partners, in the systems SIPF and SIPJ, SIGER's information suppliers.
- **SIGER´s use** (13%). Difficulties in applying system concepts correctly (this category differs from the "conceptual" category because it refers to a false conceptual assumption adopted by the user).
- **Computer use** (41%). Difficulties related to the use of the several SIGER's software components.
- **Patches** (3%). Impact of SIGER updating.
- **Training** (3%). Lack of skills in SIGER's use.
- **Manuals** (4%). Difficulties of access to SIGER manuals.
- **Planning rules** (3%). Difficulties in using the system, due to some outdated rules of the planning system;
- **Partners** (4%). Difficulties felt by partners of Embrapa, in the use of SIGER.
- **Printing** (8%). Problems with SIGER printing module.
- **Machines** (1%). Lack of appropriate hardware to execute SIGER.
- **Suggestions** (1%). Suggestions for improvement in any of the previous items.

Contrary to what Fallon predicts, problems of a technological nature seemed to be more intense in this first implementation stage. Computer-use problems, patches, printed reports and use of SIGER, all somehow related to technological aspects, were indicated more frequently by R&D managers than coordination and communication problems. A final dimension explored was the degree in which these problems affected R&D programming. Fortunately, the R&D managers' perception points to a small or medium repercussion of problems over R&D programming.

3.5.9 Lessons in the implementation of management information systems in R&D organizations

There is no question that the implementation of an MIS has a great impact on an organization's culture. Few doubt the importance of information for management. Also, there is much consensus on the need for strategic R&D management. Problems begin when people need to change their behavior and routine in order to create a true information culture. Everybody seems to agree that they need to receive information, but few are motivated to give information that is demanded from them. To create a flow of information, input and output are both necessary, and the creation of a true information culture in any institution needs to begin with changes in the attitudes of individuals. Users must therefore be motivated and mobilized, to create an awareness that providing relevant and appropriate information is important, and that it needs to be provided in a way that makes the information useful for decision making. Resistance emerges not because of the inherent aspects of the MIS, but because of the cultural change it implies.

On that point, when implementing an MIS, it is essential to determine which problems are genuine deficiencies in the system and which simply reflect resistance to innovation. The aim here is to first solve the former ones and to develop involvement and communication strategies that address the latter. This dilemma has been a reality during the implementation of SIGER. Some factors were restrictive, while others propelled the implementation of the system. In table 3.11, the principal restrictive and driving factors identified by SIGER´s implementation team— including the ones found in the evaluation described above—are presented.

For each restrictive factor, a countermeasure was determined: elucidation of new concepts which were not well understood; new formats for printed outputs; new training, at an advanced level, for the chosen "information manager" at the R&D centers; new training for researchers from organizations engaged in relationships

Table 3.11: Main Driving and Restrictive Factors in SIGER Implementation

Driving factors	**Restrictive factors**
Innovative conceptual model.	Innovations noticed as threats.
Training process and motivation.	Hardware platform and software.
User involvement in the design, validation, and implementation phases.	Competition with other managerial innovations being implemented simultaneously with SIGER.
Documentation of the system and support offered to the users (manuals, help desks, technical and managerial support).	Size of the coordination team and development.
Inappropriate performance of previous MIS.	Negative attitude in relation to MIS performance.
Communication between coordinating team and users.	Complexity of the concepts introduced by the system.

Source: Castro et al. (2000)

with Embrapa. Most of the limitations found have already been solved, as is shown by the fact that all projects planned to begin in 2001 were programmed within SIGER. Some of the limitations, however, are more difficult to overcome, since they are at the core of SIGER: the instability caused by the Data Bank Manager System (DBMS) used for SIGER can only be eliminated through the adoption of another DBMS. This decision is already made, but not yet operational.

The difficulty imposed by the integration with other corporate systems (SIPF and SIPJ, for short) is also a problem not totally overcome. There are several barriers to implementing the needed solution here, without compromising information reliability. One of the causes for the problems is the architecture of both systems and the availability of these systems for Embrapa's partners. A possible solution is to have these systems communicate with the users through networks.

Another important limitation—the size and skill mix of computer team—at the moment of implementation is being addressed through the incorporation of external members in the original team. In such a complex system, however, the socialization of new members takes time.

It is important to note here that a big push toward complete and successful adoption of SIGER was obtained when the first outputs allowed by the system were made available to clients and users. In the words of one of these clients, the system "made it possible to recover, in a 15-minute period, statistics which used to take me one month to put together." The realization of the gains to be obtained with SIGER was a decisive step towards its successful implementation, completed in the past year.

3.5.10 Conclusions

After almost five years of work in the conceptual design, computerized development, validation, implementation, and continuous improvement of an MIS in a national R&D system such as Embrapa and associated partners, it is possible to draw some useful conclusions:

- **Team formation and conceptual model.** It is essential to constitute multidisciplinary teams, with professionals in R&D management, computer science, and information, and with users' representatives. The project coordinator must manage potential conflicts among these disciplines, creating a common conceptual framework for conducting system design and development.
- **MIS conceptual model should be as detailed as possible.** It should also be based on careful studies of managers' information needs. Institutional strategy is a good reference against which to select and rank these information needs.
- **During computerized development,** it is crucial to maintain as a guideline the conceptual model, as opposed to focusing on hardware and software solutions, even when the conceptual specifications are difficult to implement in the hardware and software platform adopted (or available).
- **Opportunity and feasibility of MIS development.** The decision to build and implement an MIS should be evaluated carefully, taking into consideration the costs,

risks and the long-term outlook of this enterprise. Short-term expectations for MIS results will only end in frustrations and failures.

- **Need for top-management support.** An MIS has a profound impact on the organizational culture, and it may affect formal and informal power structures. Therefore, it can generate active and passive resistance, which needs to be identified and addressed. In this process, support from top management is a key factor.
- **MIS implementation and competition with other organizational innovations.** An MIS potentially affects all employees. In the initial phase of implementation, the occurrence of problems with the use of the new system is stressful for all involved. Therefore, it is not wise to implement simultaneously other innovations of such a broad nature, in order to keep the stress under control.
- **Need of different teams for implementation and user support.** In the initial phase of implementation, support to users is to solve actual problems in the system, and problems generated by the users' limited knowledge of the new software. This is the work of the support team and the implementation team. Consequently, it is very important to maintain two teams working in an integrated manner, but with different responsibilities, so that none of the functions is neglected.
- **Communication and documentation management.** As said, communication with the user is vital. It is important to communicate concepts, logic, reasons, strategies adopted in the implementation, operation procedures, uses, and purposes of the system. All this information should be generated, organized, documented, updated and quickly delivered to users. Information management is crucial, and it should be systematic and user oriented. In this area there is much scope for managerial innovation, in the choice of communication media, or else in methods and alternative processes. The experience with SIGER showed that many problems might have been avoided, if communication had been more agile. It also showed that printed communication is expensive, slow, difficult to update, and often inefficient, as it doesn't always reach the user.

The whole process of overcoming the difficulties emerging during implementation is a slow one, due to the diverse nature of the causes. The process requires lots of negotiation, communication, and firmness on the part of those who are in charge of making decisions about the system and those who are to implement it. Time, empathy to users' complaints, efforts in the direction of adjustments, and patience are all good virtues for all those involved. To introduce such a system is to introduce a big organizational change, one that will affect the work of everybody within—and, in many cases, even outside—the organization, putting into work the known circle of insecurity, fear, and resistance. With all the challenges embedded in such an enterprise, it is our conviction that it is worthwhile, from the viewpoint of gains to be obtained by the organization.

References

Bolivar, H., A.M.G. de Castro, A.H. Sánchez, M.G.O Aguilera, and M.A. Sarmiento. 1997. Sistemas de información gerencial (SIG). Quito: International Service for National Agricultural Research.

Castro, A.M.G. de, S.M.V. Lima, J.R.P. Carvalho, and A. Freitas Filho. 1998. Desenho conceitual e validação do sistema de informação gerencial (SIGER) da Embrapa. Anais do XX Simpósio de Gestão de Inovação Tecnológica. São Paulo: FEA/USP

Castro, A.M.G., S.M.V. Lima, J.R.P. de Carvalho, and E. Bacarin. 2000. "Implantação do SIGER na Embrapa - Estratégia, Resultados e Lições." Anais do XXI Simpósio de Gestão de Inovação Tecnológica, São Paulo: FEA/USP.

Embrapa. 1995. Modelo conceitual do sistema de informação gerencial da Embrapa - SIGER. Brasília, Mimeo.

Fallon, H. 1995. How to Implement Information Systems and Live to Tell About It. New York: John Wiley & Sons.

Flores, M.X., M.L.D. Paez, J. de S. Silva, and A. Freitas Filho. 1994. Planejamento estratégico em C&T: Teoria e aplicação. In Gestão em Ciência e Tecnologia: Pesquisa Agropecuária. Edited by W.J. Goedert, M.L.D. Paez, & A.M.G. de Castro. Brasília: Embrapa-SPI.

Goedert, W.J., P.J.C. Genú, A.P.M. Galvão, J.R.R. Peres, and J.R.F. Cabral 1994. Sistema de planejamento. In Gestão em Ciência e Tecnologia: Pesquisa Agropecuária. Edited by W.J. Goedert, M.L.D. Paez, & A.M.G. de Castro. Brasília: Embrapa-SPI.

Leite, A.A.A., M.C. Visoli, S.R.M. Evangelista, and A.A. Souza. 1998. Informatização do sistema de informação gerencial da Embrapa. Campinas, SP: Embrapa-CNPTIA

Lima, S.M.V., A.M.G. Castro, C.H.S. Ayres, and A.F.D Ávila. 1998. O sistema de informação gerencial da Embrapa (SIGER): Desenho conceitual. Brasília, DF: Embrapa-DPD

Paez, M.L.D., A.M.G. de Castro, J.R.P Carvalho, and M.A.G. Martins. 1994. Sistema integrado de informação gerencial em P&D: Acompanhamento e avaliação. In Gestão em Ciência e Tecnologia: Pesquisa Agropecuária. Edited by W.J. Goedert, M.L.D. Paez, e A.M.G. de Castro. Brasília: Embrapa-SPI.

Pedroso Jr., M., M.A.A. Leite, P.C. Oliveira, and M.C. Visoli. 1995. SIGER - Implementação. Encontro Anual de Gerentes da Embrapa. Brasília, DF.

Rossi, P.H. and H.E. Freeman. 1989. Evaluation: A systematic approach. Newbury Park, CA: Sage Publications.

Sbragia, R. 1984. Avaliação de desempenho de projetos em instituições de pesquisa: um estudo empírico dentro do setor de tecnologia industrial. Revista de administração da USP, 19 (1): 83-93.

3.6 Lessons learned from the MIS case studies

This section reviews the five MIS case studies to identify lessons that can be learned and applied elsewhere. These derive particularly from reported difficulties, but also from successes and other comments made by the writers. We have attributed these issues to the three main components in implementing an MIS, as identified by Hobbs (1996) the MIS itself, the implementation process, and institutionalization (see figure 3.5).

One consequence of institutionalization may be to change the way the NARO is managed. Such a dramatic outcome is increasingly an expectation of information systems (see also section 7.2.1 "Characteristics of information systems").

The case studies also offer lessons for the issue of impact. This is probably the climax of the whole endeavor and will be dealt with after examining the components of implementation.

3.6.1 Information tool: The MIS

PROINPA originally developed its MIS on the flat-file, Reflex-based INFORM system. The flexibility of INFORM and its logical concept allowed PROINPA to adapt it to the situation of the program without depending too much on ISNAR.

Later PROINPA obtained an English copy of the relational INFORM-R prototype. It wished to upgrade to this relational model but had two problems with converting INFORM-R directly to its needs. First, INFORM-R uses quite a lot of database "programming" (Visual Basic) to deliver its more complex outputs and processes. As plain database management skills are more readily available than are programming skills, this prevented local adaptation of the system. Second, it needed a Spanish version.

ISNAR has long recognized that using programming in the product is a potential danger to local adaptation, but it took the view that some programming was necessary initially to deliver all the requested functionality. It long intended to break the relatively sophisticated model into simpler component parts or modules, some of which at least would have no programming. An INFORM Light for project management would be an early contender for this strategy. Unfortunately, funds have proved very difficult to obtain for this work.

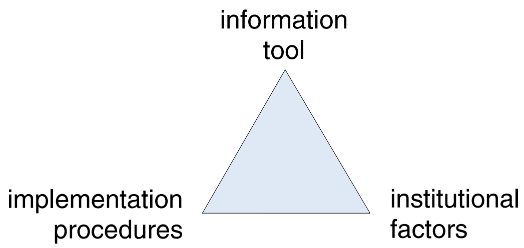


Figure 3.5. The implementation triangle

NARO (Uganda) has also made very good use of the original flat-file INFORM. However, after their national research information officer tested INFORM-R, NARO regards the flat-file nature of the INFORM as "a great limitation" and has described several of its drawbacks compared with the new version.

The Ghana paper describes their experience with the old INFORM, but having seen INFORM-R, the author of the paper provides the following disadvantages of the former:

- the need to enter certain data twice due to the lack of linkage in a flat-file system between the two main files, research activities and scientists
- difficulty in learning all the commands of a system without a menu front end
- system too simple to handle the increasingly multidisciplinary and inter-institutional nature of the research program

Zambia is the only case where the new model, INFORM-R, has been used. Ironically, this was the one case where there were practically no funds to implement an MIS once the product had been developed. On the positive side, the prototype was tested with real data for four years, and there are extensive reports of how it met the needs of management. However, Zambia has yet to benefit from it pending the finding of funding sources. One interesting comment was that INFORM-R is targeted at managing the research program, whereas the management was reported to have a greater immediate interest in information on finance, personnel, and inventory. ISNAR has developed no tools for these functions. These three areas are more typical in generic information systems from the business world, though this would need to be examined further. Many NAROs clearly would need help in adopting and adapting such systems.

The Embrapa paper focuses on its own large and locally developed systems. It has been through a number of development projects, accumulating experience along the way, so its remarks are potentially very valuable.

The paper draws attention to the need to conduct a systematic information needs assessment among users before designing an information system. The information requirements of different management processes, such as PM&E, should be considered as an integrated whole and not in isolation. In many instances different processes will use the same information items, so duplication in data capture and storage must be avoided.

Organizations that have several different information systems should follow common standards such as naming conventions, as well as software and hardware connectivity to facilitate communication between them later.

In Embrapa's earlier systems, too many fields allowed free text entry. This made aggregation of such data almost impossible. One way of avoiding this is to offer a standard set of likely entries to select from.[20]

20. See the descriptions of fields for commodities, disciplines, and specialisms within disciplines in Annex 2 "Database Tables and Fields in INFORM-R."

The user interface and the time required for data entry and retrieval of some of the early systems proved to be unfriendly for users and led to resistance and dissatisfaction.

In only one country was the INFORM budget coding system (Gijsbers 1991) used, but then this is the only site where project budgeting was strongly developed. No doubt the influence of the parent organization (CIP) was important here as they have a strong budget and expenditure ethos, lacking in most NAROs. In the face of the general lack of the budgeting component, ISNAR will have to reassess its approach to managing financial data. There is a call from Ghana for better handling of management accounting, including time allocation across successive years.

The INFORM and INFORM-R facility of keywords has only occasionally been used beyond its introduction in initial training courses. PROINPA is one such case, which had a skilled documentalist in INFORM to enhance the usefulness of the key words. It seems that this feature will do best under more sophisticated managements. Alternatively this feature may become redundant with the development of sophisticated search engines as a part of database management systems.

3.6.2 Implementation

PROINPA reported two lessons related to implementation. First, training, at different levels, proved very important. Training for the national MIS coordinator could have been more thorough and was restricted to only one person, so that when after three years that person left, her replacement had to learn by doing, which weakened the implementation. Also, users should have had follow-up training to give them a greater understanding of the system.

Second, there was a danger of trying to capture too much detail. Research details were initially collected at activity level, which proved to be very difficult to manage. There was a tendency to be scrupulously detailed, which may make the implementation process more difficult. Later, cost detail was collected at the subproject level, which was a definite improvement. PROINPA is unique amongst these cases in their success in marrying the system with the much larger system of their parent organization, CIP. There may be valuable lessons here for other research organizations taking up system integration.

Implementation in NARO (Uganda) has on the whole been quite successful. The writers attribute this to the methodical sequence employed in its implementation, paying attention to the involvement of managers from the start, training for MIS practitioners and directors at every station, and the supply of new computers for each station.

The writers point also to a problem of insufficient funds for consumables needed to maintain the system. They link this to the MIS practitioners' role and work, which is being insufficiently recognized by management at all levels. The study recommended that the operational management of the MIS be recognized as a line item in the station budget.

The Ghana paper draws a number of lessons for implementation. CSIR felt a need to distribute the MIS function across many people—at HQ and the stations. It had learned a very clear lesson of the danger of excessive centralization (at HQ) and of dependency on a single information system specialist—the system came to an abrupt stop when this person was lost and only restarted two years later when new staff were appointed to this work. Much ground had in the meantime been lost.

The system had to be integrated into the annual research management cycle. Data collection was hampered because it coincided with the start of the rains, when scientists are at their busiest with field activities. Time allocation data proved more difficult to obtain than other personal data. Finally, it pointed to the need for more involvement of users, including directors, program coordinators, and scientists, in the design and implementation of the system, and more on-the-job training.

Implementation in Zambia has been quite problematic. The writers attribute this to the following:

1. **Training:** lack of funds for a comprehensive training program, particularly to support decentralization of the system to the stations.
2. **Computers:** problems and delays over procurement.
3. **Finance:** (a) lack of funds for traveling and (b) ISNAR running out of MIS development funds before completion of the new version of the MIS.

In Brazil, the writers commented that there needs to be a strategy for implementing the system, giving users a clear view of the system and of its benefits for research management.

An encouraging point is that in two of the four INFORM studies (Bolivia and Ghana) only limited help was requested from ISNAR in implementing the system. All four have subsequently maintained it with little or no formal assistance.

3.6.3 Institutionalization

Institutionalization we take to include things that change the way the NARO is managed, as this is increasingly what information systems are expected to do (see section 7.2.1 “Characteristics of information systems”).

Two features of institutionalization stand out in Bolivia’s PROINPA case. First, using researcher time allocations as a proxy for resource allocation caused something of a culture shock with the researchers—they had never been held accountable for time spent. Perhaps more effort has to be given to explaining this aspect to them. Moving to monthly reports on time allocation data and including technicians proved excessive, and the latter was abandoned.

Second, the use of cost centers for indirect costs proved a better way of handling indirect costs than the INFORM guidelines’ recommended method of attributing these to each research project. PROINPA developed four cost centers specifically for indirect costs:

- operational costs linked to the PROINPA management and administration such as institutional cooperation, postgraduate training, and services
- operational research costs linked to subprojects

- administrative costs
- communication costs including publication and the documentation center

The advantage of this is that it gives project managers a more realistic and intuitively right feeling for the costs of their projects. However, it can be argued that distributing indirect costs across research projects provides management with a forceful, if somewhat unwelcome, reminder of the real costs of the "administrative overhead" of a NARO. Experience in at least one other country suggests that managements are not yet ready to accept this message, or at least not when delivered in this manner.

Institutionalization in Zambia faces two serious problems: lack of a strong patron to provide support and failure by management to adopt the project budgeting system.

The clear impact that the MIS has had in Uganda is attributed in part to the early recognition of involving management in institutionalizing an MIS. Some of the success of the implementation was attributed to the initial demonstration of the system to senior management and their inclusion in the training program. The writers also identified two areas of difficulty in institutionalization. First, recording scientists' time to research, management, extension, training, studies, and other time sinks proved problematic initially. The concept of recording how scientists spent their time was quite new and evidently disconcerting. (See also the PROINPA experience.) Second, insufficient funds have been available for MIS activities both operationally and to upgrade what is now perceived as a somewhat obsolete system. The writers clearly wanted the system to be upgraded to the relational model, but funds were not yet available for this transition.

The Ghana paper shows that we still have much to learn on institutionalization. It offers an excellent graphic (figure 3.1) of priorities versus resource allocation that tracks the link over three years. It shows very clearly the gap between priorities and resource allocation, and in many cases the gap is getting wider instead of smaller, as one would expect had this data been made available to and used in the research planning process. The power of the system is clear, but so is the need to involve research managers so that they may act on its implications. This points, as a reviewer of these case studies remarked, to another lesson for ISNAR—it needs to research why these changes are happening in apparent contradiction to agreed priorities. What other factors are driving these changes and how do they operate? Determining the mechanism for change is far more important than preaching for change that experience has shown does not happen.

The Embrapa case stands out in two respects. First, it comes from an exceptionally large NARO (56 R&D centers and 4000 scientists), and the system, system design, and development teams also had to be large. Second, it has attempted to implement an MIS a number of times, and with success now evident it is a rich source of lessons. Some of Embrapa's earlier systems were developed without a clear linkage to the organization's planning system. It concluded that it is important to ensure that any information system relates closely to the relevant R&D management processes. Where necessary these processes need to be fully developed before embarking on a related information system. Much the same applies to the need to ensure that any information system is in line with the organization's strategy, mission, and objectives. Such recommendations from experience of actual information system building have contributed to the methodologies described in chapter 4.

3.6.4 Impact

In Uganda, the INFORM MIS had been integrated into the day-to-day running of the research business of the NARO. The writers report that research management has improved mainly due to access to accurate and timely information for decision making. It seems this has been recognized at station level too: "A series of training sessions for program leaders and other scientists at the request of institutes is a clear manifestation of the importance that the research managers attach to the MIS." In support of this the writers give examples of a number of ad hoc questions that managers had applied to the MIS to get answers for decision making.

Impact in Bolivia has also been positive. The INFORM project database was "especially useful for the directors of the program. The graphical illustration of combined factors such as cost of projects, number of persons by project, personnel costs, etc., allowed the directors to obtain a global view of the situation of the project and could follow-up more efficiently." The time allocation data, though difficult to introduce, improved accuracy and were helpful to estimate the personnel time and cost requirement in planning activities. "The personnel database was less used and should have been updated every year, for example to monitor the short-term training activities by scientist."

PROINPA is unusual in that it has linked INFORM to their accounting system. This has enabled them to "monitor more accurately the direct as well as indirect costs of the subprojects. It was also useful to address the reporting requirement of the different donor agencies PROINPA had to deal with and estimate its research investment by geographical and political region." It clearly uses the MIS to manage the research program.

The MIS has had both a positive and negative impact on Zambia's Research Branch. On the positive side, the Research Branch has four years of records of detailed research activities and scientists. It has a remarkable 50 years of summary records of research activities (see table 3.1), which can easily be searched to determine what has been done in the past in any commodity or discipline. On the negative side, funds have been quite inadequate to implement the system outside HQ and institutionalize it. Closely linked to this is the lack of interest from management. It has probably been insufficiently sensitized to what the MIS has to offer. A lesson for ISNAR is that its MIS work has been seriously compromised by not having other partners to test the new MIS system.

In Ghana the system was used extensively in 1994 to generate information for priority setting during the preparation of the national agricultural research strategic plan and later for the preparation of a human resource development plan. With the untimely death of the sole officer managing the system, who managed it in a highly centralized manner, the system inevitably fell into disuse until new staff were appointed in late 1996. From December 1996 the system has been used to monitor progress (achievements) of research activities. In 1997 the system's data collected over a number of years was used to compare resource allocation between commodities with a previous set of priorities. A remarkable output illustrating this information (figure 3.1) is provided as a very significant output, but it also points to a significant gap in impact: over the years that are covered by this information, management has seemingly been unable to act on it and to reduce the obvious gap between perceived priority and resource allocation.

Embrapa's new MIS is being implemented at this time and it is too early to record impact. In due course it will be well worth revisiting as this new system builds on many lessons learnt from the earlier attempts, and much care and planning has been invested in the new one.

References

Gijsbers, G. 1991. Revenue and Cost Codes for Use with INFORM. INFORM Guidelines Part 3. The Hague: International Service for National Agricultural Research.

Hobbs, H. 1996. Strengthening the use of information tools in NAROs: The implementation triangle. A conceptual framework. Unpublished paper. The Hague: International Service for National Agricultural Research.

Part 2. For the Information Manager

Chapter

4 Building a Management Information System

This chapter introduces the design, development, and maintenance of an MIS, including an introduction to the relational database model.

National information system coordinators are therefore the primary target readers. Those research managers with responsibility for supervising information management in the organization will find sections 4.1 "Issues in the implementation of an MIS" and 4.2 "Institutionalization and the management of change" of interest. MIS practitioners at station and national level may also find these two sections useful.

4.1 Issues in the implementation of an MIS

Although many organizations greatly benefit from their MIS, developing and implementing such systems successfully is far from certain. The literature on this subject contains many and sometimes very expensive examples of failures. Examples are those of the London Stock Exchange (see box 4.2, page 140), the London Ambulance Service, and the Denver (USA) airport baggage handling system. Such cases have led to much study of the reasons for failure and the factors that contribute to success. Here we will look at the issues that have emerged as important to information system development.

4.1.1 The people

This section discusses the involvement of those groups that the system is to serve, particularly research managers and researchers, as well as ministry officials and donors.

Top management

In establishing an organization-wide information system, the managers of three organizational areas usually need to be involved:

- the core or business of the organization
- finance department
- information technology unit

In a NARO, a chief research officer may be the appropriate representative of the first, and the head of finance of the second. Information technology often suffers from its newness in NAROs. There is often no tradition of a senior post in this discipline, greatly hampering the adoption of modern information management concepts and the harnessing of the benefits of information technology. The position is therefore often represented by a senior scientist with a particular interest in and some knowledge of computing. If no such person is available it may be that a librarian, involving the discipline of information science, is involved. Another choice is a senior biometrician. In any case the position needs to be accorded senior management importance.

Middle management

Middle management will provide the human resources most involved in the routine operation of an MIS. NAROs face a number of problems with respect to these staff.

Scientists who manage the MIS at station level may do so at the expense of their scientific research work, which may lead to reduced prospects of promotion. The solution is to formally recognize that researchers managing the MIS are gaining valuable research management skills, which should benefit both the NARO and the individual. Managing the MIS should be recognized by both management and researchers as a career opportunity. Wherever possible, two scientists (ideally one senior scientist and one junior) should be appointed to manage the MIS at each station.

The MIS scientist may also face reduced opportunities for higher-degree studies. Only recently have some more enlightened NAROs begun to accept that information management is an acceptable subject for higher degree studies.

Another dilemma for scientists managing information is the dichotomy between their original discipline, such as plant breeding or animal health, and the new information science. Time spent on managing information is time taken from pursuing their conventional scientific discipline, and they may therefore suffer reduced prospects of advancement in their scientific field. They have two choices: retaining a position in their original science or switching to the new science. The station director can help here by providing them with sufficient support staff, ensuring that they allocate only a small proportion of their time to MIS tasks. If the prospects for information management in the organization look good, and if they prove to have an aptitude for it, then they may seek to specialize in the new science. As this may become apparent only over time, management should give them time (at least a year) to make their decision.

Finally, NAROs are likely to face the threat of competition from the private sector for staff that have received specialist information management training. The private sector is generally ahead of the public sector in recognizing their value. Those NAROs that neglect the issues raised above and fail to offer attractive terms of engagement and prospects for promotion will be the first to suffer such losses. In addition, with their computing and information management skills, MIS specialists enjoy an attraction from a wider range of industries, such as banking, not given to plant breeders or veterinarians. If a NARO wishes to retain its MIS people, then it will be especially incumbent on top management to ensure good working conditions and scope for professional advancement for its information specialists.

Teams

Mankin et al (1996) identify three areas critical to implementing an information system: (a) the organization and the changes it needs to face, (b) teams, and (c) information technology. Successful organizations are abandoning the old organization structures where a client tracking down some information is shunted from department to department. Instead, individuals have access to pooled knowledge that can be harnessed to solve problems. Information technology allows individuals to link to each other and to information repositories. This is likely to involve change in the way the organization operates. The contribution of teams is seen particularly in internal projects, such as the development of information systems.

User committee

A user committee enables the views of all the intended groups of users to be represented (see table 4.1). During the establishment of an MIS, the user committee meets a number of times to monitor and advise on the system and the establishment process, ensuring that users' needs are well addressed.

When the system is operational, the committee needs to meet less frequently. Its role then shifts to monitoring how the system serves its users and agreeing on behalf of all on any changes to the system. Any suggestions are passed to the system administrators for implementation. External users, such as representatives from other ministries, donors, and farmers' organizations, may best be served by occasionally sending them a questionnaire followed up by consultation by one or two interviewers in their own office.

Table 4.1: Potential Users of an MIS for Agricultural Research

User	Interests
Research managers	Resource allocation, program planning, monitoring, at the station level.
Research scientists	Checking what research has been done before and what is being done now, before proposing new work.
NARO leaders	National resource allocation in relation to national research priorities and those at the level of the agroecological zone; research program planning and monitoring.
Ministry of agriculture	National resource allocation in relation to national agricultural priorities and those at the level of the agroecological zone.
University agricultural faculties	Matching university research program with the NARO program and national priorities; opportunities for collaboration, and avoidance of unintended duplication between university and NARO scientists.
Ministry of finance	National resource allocation in relation to national priorities.
Farmers' organizations	National resource allocation across commodities, and constraints within commodities.
Donors	Targeting individual donor support with respect to national agricultural research priorities.

Technical responsibility for the system

An MIS for agricultural research normally has to operate at more than one level, e.g., at research stations and at the national or institutional HQ. Most data is entered at station level in an annual or seasonal cycle. Annually or seasonally the station systems are passed to the HQ, where they are aggregated to form the national system, which is then redistributed back to stations. This requires that stations do not introduce significant changes to their system independently, but rather pass suggestions for change to the national user committee for consideration. Approved changes are then passed to the HQ technical unit for implementation.

This HQ unit has technical responsibility for the system. It should not be a single person, as the consequences of one person's absence, due for example to travel, sickness, or resignation, are too serious. The unit should be closer to research management than to a computer or other technical service facility. A PM&E unit is a good home for this responsibility. Two or more members of such a unit should be trained in the MIS and its software in order to be able to modify the system and implement new outputs on request from the user committee.

Examples of titles of the person in charge of this work are national research information officer and national MIS coordinator, equating to the database administrator of commercial companies. They would take the lead in developing, introducing, and maintaining the system (see section 4.4 "Managing the system").

Role of the biometrician

The biometrician has critical roles in the quality of research activities in being involved in their design and the analysis of their results. In fulfilling these roles the biometrician inevitably collects research activity proposals (protocols) and complete data sets of results. This represents a second repository of such information, the first being the MIS.

In the absence of a PM&E unit or similar, the relationship of the biometrician to researchers and experiments may make the biometrics office an alternative home for the MIS. An additional advantage is that a biometrician is usually computer literate. Additional clerical staff will need to be provided for data entry and routine data processing.

4.1.2 Choice of system and software

Staff charged with the task of implementing an MIS have a number of alternative routes available:

- procuring an existing MIS
- using a generalized application package, such as database, spreadsheet, or project management software, to develop such a system
- writing a new system for their institute using a programming language

Ready-made system

Ready-made systems may be available for some information-management needs. There are many accounting packages available on the market that are targeted at small and medium-sized businesses, and some of these could very probably be adapted to the accounting needs of a NARO. But there are very few generalized MIS for agricultural research. The rather specialized requirements and small market size seems to make the commercial development of such a product uneconomic.

Advantages of a ready-made system are that many customers share the development costs; it has been extensively tested; a user community may provide a source of advice; and it is likely that in the face of such a user community, the developer will continue to support and enhance the system (Roberts 1996).

Generalized application package

Another option is to select a generalized application package such as a database or spreadsheet software and use it to build the MIS. This obviously will take longer and call for more skills than procuring a ready-made system, but it offers the prospect of tailoring the system to the organization. The initial cost of the software is low, and even when multiplied by the number of stations that need their own copy, the greater cost will be in developing the system and training staff throughout the NARO in using it. If a widely used brand of software is selected, then there is a good chance that the institution will have or acquire the skills to use it. This is very important for the future maintenance of the system. According to Brooks (1995), a rule-of-thumb for the life-time cost of maintaining a widely used program is 40% or more of the cost of developing it.

"Off-the-shelf" project management software falls between the two cases above. It is less flexible than a database or spreadsheet package and therefore quicker to bring into use, but it needs to be selected with care to ensure that there is sufficient flexibility to capture all the desired project attributes and obtain the desired reports. For example, an off-the-shelf project management software comes with a number of built-in reports but few, if any, of these will apply directly to agricultural research management.

Custom-made system

The third option is to build a new system from scratch using a programming language. However, it is likely to take longer than the previous method and will cost considerably more, as the development cost is borne by a single client. Also, an organization is less likely to be competent in maintaining the system, and if it is, such competence is probably costly and somewhat distant from the business users. External programmers can develop the system in a programming language, but this creates a permanent dependency on the supplier. And should the supplier go out of business, it is probably difficult or expensive to maintain the system further.

An argument for developing an own system is that every organization is different, and a single product cannot accommodate all such differences. Using a standard generalized application package will help considerably because there is a reasonable prospect that the organization has the competence to maintain and adapt the system in the future. This is important as every organization changes along with its information needs.

An organization may have to adapt its practices to the software system. This is not necessarily a bad thing; acquiring a major information system may offer an organization an opportunity to streamline its management and operational processes.

Whatever software is chosen, its ease of use is very important. Since the mid-1990s, off-the-shelf software has become very userfriendly. Users can now focus on managing data rather than managing the software, and do so in days or weeks rather than months. This is clearly important for the information systems developer, who can greatly help users towards this position by (1) making clear what data is available, and (2) making it easy to find the data of interest to them.

The system's "front-end," usually a set of carefully crafted menus, is the primary device for delivering these facilities and is therefore a crucial component, that needs to be carefully assessed when selecting a system.

Very large private-sector organizations have in the past tended to choose to invest in custom-made systems, but this trend is changing (see box 4.1).

Box 4.1: Telecommunications Industry Shifts Away from Tailor-Made Systems

Until recently, large protected telecommunications monopolies have been able and willing to spend large sums to develop their own proprietary systems in-house. Recently this has changed with a shift to the procurement of cheaper off-the-shelf software.

One reason is that in the former case the research and development costs are amortised over just one client—the company itself—instead of sharing the cost over several customers. Another is that in-house IT departments do not always charge a competitive price. As one telecommunications consultant, Frank Owen, put it, "If I wanted to charge £10 million for a billing system, I wouldn't be able to sell it. But if I was an IT manager providing the system internally to an industry giant, such as Britain's BT, France Telecom, or Deutsche Telekom, I'd get the money."

This is now changing. Ten years ago, US computer chip manufacturers like Texas Instruments and Motorola had their own computer-aided design systems. Today, they all use external products like Cadence and Mentor Graphics, which are sold world-wide. A third advantage is that an external supplier should have a good understanding of features in use in other parts of the world.

This is a summary of Shillingford, J. "Open Systems Lure Operators." Financial Times*. 17 March 1998.*

To build in-house or to outsource

A NARO that decides to embark on developing its own information system has three options:

1. using one or more members of staff with information management skills
2. using the organization's IT department. Currently only large NAROs have such a department, e.g., Brazil's Embrapa (see chapter 3.5).
3. outsourcing the design and development of the system to consultants outside the organization

Mankin et al. (1996) provide some guidelines on the potential advantages and disadvantages of each (see table 4.2).

Installing hardware may need some technical assistance initially, but PCs have become rather dependable in recent years. With some precautions, such as protection against variations in the electricity supply and software viruses, the hardware may perform reasonably with little further attention. Communications technology, including the installation and maintenance of local networks, is still complex, and specialist help is probably necessary. Moreover, the whole ICT area is continuing to develop so that some specialist technical advice is needed to ensure that significant opportunities are not missed.

The choice of system usually emerges from discussions, but if there is disagreement between a number of choices, a scoring system can help the organization reach a decision. A simple scoring system can convert qualitative assessments into a common denominator and facilitate comparison. It does not provide an absolute measure of value, such as financial cost or number of months needed, but a relative measure between similar options. Scoring systems cannot address all relevant factors, nor is this necessary. They focuses on just some of the most important issues to better inform those who are to make the decision.

This approach is probably more effective than a single person's direct decision because the final choice incorporates several people's perspectives of several relevant factors. In practice the process of agreeing on the important factors may be as important as the resulting score (Laudon and Laudon 1994).

Table 4.3 gives an example. After some discussion, managers have agreed on a set of factors or criteria that they find important in the choice. They have attached a weight to each factor, reflecting its relative importance. Each system is then scored against each factor. Each score is then adjusted by multiplying it by the respective weight. A total for each of the systems suggests the most valued one.

4.1.3 Defining a system boundary

In developing an information system there is a natural tendency to continuously extend the boundary and include additional data items or even topics on the grounds that they might be useful. This should be avoided because it

- risks delaying the release of an operational system to end users

Table 4.2: Comparison of Different Resources for Constructing an Information System

Option	Advantages	Disadvantages
Internal, team-based experts	• Often highly motivated. • Expertise in both IT and the business. • More prone to use off-the-shelf, easier-to-use software than IT professionals and hence likely to get systems up and running more quickly.	Excludes the internal IT department, which (a) know best the IT infrastructure; (b) is responsible for maintaining corporate IT standards and (c) may one day be asked to maintain the system. Conclusion: it is usually important to involve the IT department.
Internal IT department	• Product is likely to be compatible with other IT systems in the organization.	• Many systems professionals do not understand the specific needs of users and have little or no knowledge of users' work. • Some users have sufficient IT competence to recognize when they are being inadequately served, which results in dissatisfaction with the IT department and a preference to develop their own systems. • Tends to subordinate users desires to their own views.
Outsourcing to external consultants	• For small organizations: more cost-effective than own IT department; purchase what you need when you need it. • Larger organizations: do this when there is a lack of time or particular skills in IT department. • Can offer a combination of range of expertise and experience not available in the IT department. Many focus on particular industries and have worked on many sites in that industry (though not often in agricultural research). • An outsider can introduce more easily unpopular changes deriving from system development. • IT department believed to be too technocratic for user-empowered systems.	• Unless carefully managed, runs the risk of large expenditure without a functional system at the end.

Based on Mankin et al. (1996)

Table 4.3: Example of a Scoring Model to Choose among Alternative Information Systems

Factor *	**Weight %**	**System A**		**System B**		**System C**	
Ease of use by users	25	3**	75	2	50	3	75
Initial cost of ICT	10	3	30	2	20	2	20
Comprehensiveness of information	15	1	15	3	45	2	30
Chance of success	40	3	120	3	120	4	160
Compatibility with other systems	10	1	10	1	10	4	40
Total			250		245		325

* Different factors will be selected as the most important in different situations: these are just examples.

** 0 = quite unacceptable; 4 = excellent

- adds progressively to the cost of initial development and of subsequent maintenance
- increases the system's complexity which in turn (1) increases the risk of failure and (2) adds to the user's learning curve

The case study from Bolivia (section 3.4) is a good example of the recognition of how too much detail was being included in developing the system and how moving to a simpler model benefitted the development. While the original specifications cannot always be adhered to (requirements cannot be completely predicted), they should be kept in view. A useful rule is to restrict the addition of items to those for which a user claims a need and demonstrates a use within the general subject area of the system. Developing and implementing an information system successfully is a difficult undertaking for any organization, and size and complexity add substantially to the risk (see box 4.2).

Crowe and Avison (1980) give several reasons why databases in particular are best limited in size. One reason comes from the difference between traditional transaction data-processing, where economies of scale are apparent, and database systems, which typically have a much wider range of data structures and processing requirements. This calls for a large number of tables and their relationships, which increase exponentially as the boundary of the system is enlarged. The capacity of hardware and software may be enhanced to cope with such increasing complexity, but the ability of those who have to design and maintain the system is finite.

From practical experience in Africa, Peterson (1994) urges systems to be developed for selected critical tasks in selected units of the organization, which can operate independently yet fit organizational boundaries. He goes further and advises that integration of systems be limited. One advantage of several component systems is the prospect that they offer of experimentation and the acceptability of failures. Such

Box 4.2: Example of Failure of a Major Information System Development Project

In the early 1990s, an information system titled Taurus was sold as a multi-million pound project that would revolutionize the London Stock Exchange. Since the unification of European Community, London—a traditional financial center—found itself increasingly competing with other cities for the high finance business. The financial district found itself competing with more modern exchanges: organizations that use technology to help them make swift transactions.

The London Stock Exchange decided to modernize its operations by introducing new technologies both within the exchange itself and among the organizations associated with it. The project started with all the optimism of any new venture sold as revolutionizing an industry. However, in 1993, after years spent in development, the head of the London Stock Exchange announced that the Taurus system had failed. It never reached the operational stage; at its demise parts of it had not been implemented.

In the initial days after the announcement people searched for answers: why had Taurus failed? A highly regarded firm of computer consultants, Ovum Consultancy, suggested that the failure of Taurus was a direct result of poor configuration management practices. Configuration management involves identifying the components of a software system and tracking the changes made to them. It also involves maintaining information about how to assemble the components into systems. In practice developers and organizations find configuration management activities very difficult, because the software components have technical relationships—called dependencies—that must be coordinated by the people working on that code.

Taurus was a highly distributed system; teams of developers worked together on individual parts of the projects. At the same time Taurus required that all the organizations linked to the London Stock Exchange build Taurus-compliant systems. In this highly distributed environment the different developers and organizations struggled to coordinate their efforts with each other. Technically, it was difficult to align the distributed development efforts so that all the systems worked together. Socially, it was hard to maintain communications among the different developers and organizations working on the project so that everyone understood what changes were taking place and why.

Software engineering researchers know that dependencies exist between pieces of code. However, little is known about how technical dependencies among modules of code create and reflect social dependencies among the developers, teams, and organizations working on them. The story of Taurus clearly illustrates that the problem of trying to coordinate these technical dependencies is a managerial problem.

Source: Grinter (1996)

failures are then limited so that they are rich sources of learning rather than serious set backs. New and more successful versions can be developed in a comparatively short time. This is very much in line with the concept of prototyping, in which a simple system is developed rapidly and put into the hands of users to get early feedback on its usefulness and problems.

On the other hand, developments in information technology in recent years have allowed large, well-resourced organizations to develop very large data repositories called data warehouses. The main advantages of a single large system covering all areas of an organization are better access for all to organization-wide information and no duplication of stored data. Until recently the complexity of such large systems, for the reasons listed above, has been a serious limitation. And the progress that is now being reported is still feasible only for large organizations with exceptional managerial, information technology, and financial resources. For now and for some years to come most NAROs are likely to progress more quickly with a discrete system for each main function, but agreeing on a common set of names, using common software so that data interchange is easy.

4.1.4 Critical success factors

Critical success factors are management aspects believed to be crucial to the success of an operation (Cashmore and Lyall 1991). They are sometimes described as necessary albeit not sufficient conditions for success. The "escape clause" indicates that we do not yet have a complete understanding of the reasons of success in each situation, but we may draw some general conclusions. Critical success factors in information systems in general have been reviewed by Li (1997) and by DeLone and McLean (1992). Several authors have identified critical success factors that are specific to an MIS or executive information system (EIS). Rainer and Watson (1995) list the following conventional keys to EIS success:

- The EIS should be developed in response to and / or support a specific business need.
- An executive sponsor who champions the project is crucial.
- EIS support staff must have the necessary technical, business, and interpersonal skills.
- Effective data management is essential for the data to be relevant, accurate, and attractively presented.
- A prototype must be released soon to capture executive interest while it is still high.
- The system must be easy to use and navigate, requiring virtually no training of users.
- Response time must be extremely fast.

They point out that "although many organizations receive significant benefits from their EIS these still are high-risk / high-return systems, as shown by the large number of EIS failures. Not only are EIS prone to fail, they can fail at any point in the system's lifetime." This should prompt us to take the issue of critical success factors seriously in developing our own MIS.

Critical success factors for the agricultural research MIS

The following suggested critical success factors for the agricultural research MIS take the above analyses into account. They also derive from the experiences in a number of implementations in Latin America, Asia, and Africa of ISNAR's MIS INFORM (Nestel 1991; Vernon 1994). We cannot yet say whether all these factors are critical, in the sense that they are necessary conditions for success, but we believe this is the case for most of them.

- **There needs to be a perceived need for such a system within the organization.** See section 4.2 "Institutionalization and management of change."
- **Implementation requires a sequence of several different interventions that go beyond a one- to two-week training workshop.** ISNAR held three regional MIS training workshops in Asia and Africa, each for two weeks and each with representatives from many countries, who were taught in some detail how to operate the system and apply it to research management. There was very little impact with only one country that subsequently implemented the system. In this country, the system was implemented by a team of three who were not present at the workshop.

  One Asian institute has run many MIS training courses, each with single delegates from several agricultural research institutions. The course and the training materials are impressive, but the course leaders knew of no institution that had taken up and implemented an MIS as a result of a staff member attending such a course.
- **There needs to be a patron or sponsor in top management.** Convinced of the need for the system, the sponsor is able and willing to demand that the required inputs for the system are made available and that outputs from the system are delivered on time. He or she does not necessarily need to be a direct user but will use, and demand, its outputs. A sponsor needs to have a good overview of the system, what it entails in staff time, and what it can do for the organization, its institutes, directors, and scientists.
- **There should be an operational sponsor usually from senior management, for supervision of the day-to-day development and management of the system.** This operational sponsor should have the authority to demand outputs from the system, ensure their use in the regular cycle of research program planning and monitoring, and intervene in the resolution of operational problems. This person should be an active user of the system, and may be a member of the user committee (see section 4.1.1.)
- **Researchers need to be integrated into the development and implementation and use of the system.** Researchers are the main source of data for the system, and the MIS needs data from every researcher each year. These requirements will not be met unless the researchers are, and feel, part of the system and are allowed to benefit from it. The system needs to be demonstrated to them before it is implemented to show the benefits relevant to them alongside the cost, i.e., their time to

provide information. They must be included in the distribution of outputs from the system. It is desirable that they be offered easy access to the system on-line. It is of course also essential that a small number of researchers be part of the system development process, representing the user community, and be involved in testing the system.

- **At least two "MIS practitioners," who are also scientists, should be appointed at each station or institute.** Having two practitioners means joint assistance and support in implementation and continuity in case one is unavailable due, for example, to sickness, a training course, or maternity leave. They should normally be scientists, as professional knowledge of agricultural research is needed in managing such a system. It is helpful, especially on larger stations, if clerical assistance is available for data entry. Such persons need to be included in systems training.
- **The director's cooperation is needed in selecting trainee MIS practitioners, particularly in choosing researchers likely to remain in post for at least the following year**. There is merit in selecting one senior researcher and one junior researcher. While able to stand in for each other, the senior researcher will bring experience in research management and the latter may be willing to contribute more to day-to-day management of the system. The senior partner is less likely to leave for long-term training, leaving responsibility for the MIS to a single MIS practitioner. A director with the ambition to be promoted will find that his or her opportunity will be enhanced if he or she can nominate a suitable successor; a senior MIS practitioner is likely to be a good candidate because of the excellent experience in research management he or she has gained in managing the MIS. Section 4.4 "Managing the system" gives sample terms of reference for this position.
- **Scientists selected for MIS practitioner training must have good access to a computer at the time of the training.** Promises of computer procurement in the future are of little use because the value of computer training declines rapidly if the skills are not applied very soon after the training. The introduction of an MIS to a NARO is a significant undertaking. One aspect is the need for a sufficient stock of PCs, and often this means new procurement. It is helpful to use such newly procured stock for the MIS practitioners' training, at the end of which the practitioners return to their stations with a complete working system.
- **"Keep It Small and Simple."** Many organizations tend to succumb to the temptation to include more and more data items, or whole new modules, into the system. This temptation should generally be avoided. Systems are more likely to fail due to too great a complexity than to simplicity. One useful rule is not to admit new entities unless both a use and a user are identified.
- **The MIS implementation should itself have goals, defined qualitatively, quantitatively, and by time.** These will vary in different implementations but a suggested list is offered in the section on system implementation (4.3.3).
- **Training must be provided for all research managers: program leaders, station managers, heads of programs, and top management.** They all need to have a good knowledge of the system's contents and functions in order to exploit the system well in the research management processes of planning, budgeting, monitor-

ing, report writing, and evaluation. Senior managers are best trained in using the system on their own, as some are likely to be new to PCs, and it is not prudent to expect them to take their first steps in the presence of junior staff.

- **The MIS should soon produce useful outputs that can be used by station managers, institute managers, national managers, and scientists.** This is the whole purpose of the system. If neglected, the support of these constituencies will wane as most MIS information declines in value with time. Practitioners should therefore establish strict deadlines for submission of information and not wait for late information, or the system will take on the time-scale of the slowest.

- **Once established, the system needs to be institutionalized, i.e., integrated into the management processes of research planning, program formulation, and monitoring.** This is crucial because it fulfills the primary reason for having an MIS. It takes time—perhaps several annual cycles of the research program.

4.1.5 Agreement on management information requirements

The developers of the information system need to consult research managers widely and need to agree on managers' broad information needs for managing the research program. This provides an outline of what the system needs to include. Box 4.3 gives an example of research managers' information needs. It summarizes some results from a survey that was part of a wider information needs study by ISNAR.

At its simplest, a system may provide very basic inventory information: a list of research activities at each station and a list of researchers. This would not usually qualify the system as an MIS, but the state of information management in some places is so inadequate that even this information is a useful step forward. ISNAR's DOCRA system, introduced in the early 1990s to a number of West Asian and North African countries, operated at this level.

More typically, an MIS for agricultural research will aggregate data on individual research activities to provide information on the distribution of resources between, for example:

- commodities (crop and livestock species) and noncommodity (production) factors such as soil, water, post-harvest processing, economics, and sociology
- scientific disciplines
- geographical areas such as agroecological zones
- basic, applied, and adaptive research

This information helps management compare resource allocation with agreed research strategies and priorities. Even if these strategies are not in place, experience in the business world has demonstrated that a new information system can improve the way a business is conducted.

Resource allocation can most simply be quantified by the number of experiments. Accuracy is progressively improved by (1) using researchers' time as a percentage of a person year, (2) the same but multiplied by annual salary, (3) by tracking experiment-related expenditure, and (4) by attributing overhead costs to experiments in a

Box 4.3 Survey of Research Managers' Identified Information Needs

The survey used a technique borrowed from information system development in industry. Managers were asked to imagine that they returned to their desks after a considerable absence and to state what important information they would want on their desk to bring them rapidly up to date on the performance of the institute. This would be prior to the arrival of an important donor delegation. In practice most of the responses indicate material that is well known to the experienced information systems practitioner, but from time to time there are valuable new perspectives or suggestions. Additionally, when the responses are classified by topic and subtopic and analyzed quantitatively, as shown here, they indicate in which areas the greatest demand for information lies. This survey is one of a number of information needs assessments carried out by ISNAR.

Topic	Subtopic	No. of information suggestions
Research activities	Research reports; archive	52
	Current research activities	44
	Program planning	40
	Research management	24
	Priority setting	22
	Collaboration	20
	Monitoring	17
	Evaluation	16
	Beneficiaries	5
	Project appraisal	4
	Total	**244**
Personnel	General	44
	Researchers	31
	Training	11
	Support staff	8
	Technicians	3
	Remuneration	1
	Total	**98**
Finance	General	36
	Budgeting	26
	Donors	18
	Budgeting, capital	10
	Government	6
	Total	**96**
Physical resources	General	20
	Inventory	19
	Vehicles and machinery	11
	Infrastructure, facilities	8
	Land and soils	7
	Buildings	3
	Total	**68**
Research administration	General	45
	Total	**45**
Institutional structure	General	22
	Total	**22**
Information management	General	12
	Hardware	1
	Software	2
	Costs	1
	Training	1
	Total	**17**

suitable proportional manner. These stages incur costs in terms of management overhead, which have to be balanced against perceived benefits.

A decision is needed on the extent to which the research MIS covers personnel and financial areas. Ideally, it covers research-related data on the researchers and budget allocations to research activities. Information beyond that may be more properly the domain of separate personnel and accounting systems.

Other useful areas to include in the extended MIS are data on research activity proposals, completed activities (archive), and the actual results of experiments. Including the last greatly facilitates the production of annual research reports. It may be sensible to leave these enhancements until a simpler system is successfully established and integrated into the research program management cycle.

There is a tendency, derived from civil service bureaucracy, to use the MIS for routine monthly and quarterly reports. Such reports, however, contribute little to the management of agricultural research programs. If the MIS is intended to be a research information system, then it will deal largely with aggregated data on key inputs and outputs of research. This data does not usually change significantly during a cropping season. Major data collection and reporting is therefore best done only annually or, where there are two crops per year, seasonally.

Changes that normally will need to be entered more frequently are those regarding new or departing personnel and the time allocation of researchers. When estimates are made in advance (ex ante) an annual judgment is probably the only option. Where an attempt is made to capture direct values of this parameter, monthly or even weekly returns may improve accuracy. The extra cost of such frequent data collection will need to be justified in terms of greater value of the outputs.

4.1.6 Allocating resources to managing the MIS

Ideally the cost of management of an organization should be kept within about 20% of the total cost of the organization. However, this is often exceeded by burgeoning bureaucracy, always at the expense of the core research. Management of the MIS has its own cost, which adds to the above cost of management. This is a further reason to constrain the system within reasonable limits.

Having said that, it is crucial to the value of the information system that sufficient skilled resources be allocated to its development, implementation, and operation. This is a particularly difficult task in developing economies, where skilled information practitioners are hard to find and if trained within the organization are hard to retain.

4.1.7 Quality assurance

Quality assurance is the pursuit of standards that seek to guarantee, as far as practicable, that the final product satisfies the user. There are internationally accepted quality assurance standards and procedures such as the ISO 9000 suite for manufacturing industries, which includes ISO 9000-3 for software development. These are more appropriate to larger systems than we are considering here.

Careful attention to the main issues discussed in this text will go a long way to assuring quality. Some (e.g., Tober 1997) identify four determinants of software systems quality:

- usability
- stability
- documentation
- maintainability

The first (usability) should derive from early and continued user testing and comments from users. The second (stability) and fourth (maintainability) will be encouraged if a relational database software is used by addressing relational principles (see section 7.5 “Elements of a database system”). The third (documentation) is discussed as part of the structured method of systems analysis and design in section 4.3.1.

4.1.8 Managing the MIS development project

Traditionally, a project is required to be (1) completed within time, (2) completed within budget, and (3) delivering the right product. To an extent these are conflicting or competing demands, and part of the task of a project manager is maintaining a reasonable balance between them.

There is a range of project management software on the market. The simplest programs enable the manager to list the project’s several components and track them through estimates of their start dates and duration as well as monitor their individual progress to completion. Typically this is done through Gantt charts and the Program Evaluation and Review Technique (PERT) (see section 4.3 “Systems analysis and design”). Project components can usually be entered in a hierarchy, such as project, tasks, and activities. For very large and complex projects, such software support tools are a valuable management aid. They are less useful for small projects as they incur a considerable overhead in time initially in learning how to use them and subsequently in entering data and generating and using the reports.

A person familiar with database software could construct a basic project management system that records all tasks and activities, their start date and estimated duration, and their completion. Such a system may lack the Gantt and PERT charts of the purchased product but can still serve project management well. An additional module can record the “bugs” and user comments that start emerging on the release of the first prototype. Such a project tracking database system would record for each problem or bug an estimate of the time needed to resolve it and the percentage resolved to date. This enables the system to report at any time the outstanding resources still required to completion. The outstanding resource needed and the number of items provide a useful indicator of progress: both should show a steady decline if the project is moving forward satisfactorily to conclusion. The flow of user comments and reports of system faults will initially increase if this process is going well, but then in time also decline as the system moves closer to users needs and a fault-free state.

4.1.9 Introducing an existing MIS to a new client

As agricultural research organizations across the world have much in common, ISNAR developed a generalized management information system—INFORM (Information for research management). First designed in the late 1980s, INFORM can be adapted to local conditions. Taking advantage of developments in relational database technology in the mid-1990s, ISNAR turned it into a more powerful and userfriendly system, renaming it INFORM-R.[1] This section provides an overview of the procedures to implement such an existing system in a new site. Most of the critical success factors in section 4.1.4 still apply, which is essential to institutionalizing the system. What is saved in using an existing system is much of the analysis, design, and development work of the system described in section 4.3.

The following is a typical sequence of events in introducing an existing system to a new site:

- Demonstrate the MIS to senior management and representative program managers and researchers. Provide outline estimated costs (in both financial and staff-time terms) of implementing such a system throughout the institution. Open the floor to discussion. The object is to provide an overview of what such a system has to offer and what it will cost both to implement and subsequently use. It is also an opportunity for the demonstrator to learn of any special features of and concerns in this location. The organization has then to make an intuitive cost-benefit analysis to decide whether or not to proceed. If affirmative the next steps follow.
- Assess (1) existing information flows and processes, e.g., the existing annual research management cycle, key research planning meetings, research reports, (2) users' information needs that are not satisfied under current systems, and (3) information needs not yet covered by the MIS to be introduced.
- Develop a plan of action: Agree on the (modified) system to be implemented, the procedure for implementing it, and on its integration into the research management cycle.
- Organize initial training of national coordinators and data-entry personnel if available.
- Visit all stations that are to be included at this stage and collect data on researchers and their research activities. Enter this data into the system.
- Provide training to MIS practitioners from each station, typically for one or two weeks, using the data just collected.
- Sensitize senior and station managers: typically this is for one or two days and overlaps with the last day(s) of the practitioners' course; their joint presence will be very useful.
- Follow-up: usually two or three visits at intervals of a few months, to all stations, to resolve unforeseen problems, monitor progress, and learn from the experience.

1. The 'R' refers to the *relational* database model used: section 7.5.3 describes this model.

4.2 Institutionalization and the management of change

This section has been adapted from H. Hobbs (1996, 1999)

The history of computer-based information systems includes a disturbing proportion of failed implementations. Any new MIS implementation project needs to strive to avoid a similar fate. In a review of the literature on the management of change, Hobbs (1996) identified three requirements for the successful implementation of an information system:

- The information tool (in our case, the MIS) has to work. It has to be well designed, useable, and understandable.
- The implementation procedures have to be organized and scheduled in such a way that the NARO staff at all levels become active, willing, and system-competent participants in the process of using the information tool for their particular research management role.
- The local situation has to be explored and well understood before attempting an implementation. It will influence the appropriateness of the tool, the need for its local adaptation, and the best approaches with which the tool can be institutionalized (fully integrated into the research management cycle). This is essentially the process of managing change.

These components are represented in figure 3.5 (page 122). The implementation process and local institutional factors depend on how the organization manages change. Management of change is crucial to the implementation of an institution-wide information system.

Hobbs (1999) identifies five phases in change management, which can occur in different sequences: "Different situations require different solutions, so the phases (...) should not be thought of as strictly linear. Rather, they serve as markers or reminders of the various processes that need to be performed to manage change successfully."

- recognizing the need for change
- creating pressure for change
- deciding what to change
- creating the conditions for implementation
- achieving change

4.2.1 Recognizing the need for change

There are many warning signs that tell managers that the NARO may need to change. In the case of the implementation of an MIS in a NARO, these signs might include the following:

- annual reports not being produced
- difficulty in getting information of previous research

- difficulty in getting information of current research
- little or no linkage between the research program and national goals
- no budgets, or only very unreliable budgets
- new directors unaware of what research is being done
- little evaluation of researchers' work
- research files being stored in piles on desks and window ledges

It is essential that *managers*, not just MIS practitioners, see these signs and recognize the need for taking action to correct them. The MIS must not be offered as a solution, unless and until management recognizes the problem. It may be that management has little idea of what researchers are doing, or that management is unable to say how much of the budget remains unspent, or so much time has to be spent collecting the same information and repackaging it for different reports. When management has a clear view on the problem, a dialogue on MIS can be started.

4.2.2 Creating pressure for change

Change requires pressure; without pressure there is no change. However, pressure alone does not make change happen. The NARO manager has to translate this pressure into an internal sense of urgency. This is changing from stakeholders saying "You must change" to NARO managers saying "We must change." Internal urgency is accepting that the new and untried is less dangerous than doing nothing.

One way of building an internal sense of urgency is building a strong "change team" to guide the change. This is a group of people with substantial influence in terms of position in the organization, information, access to resources, expertise, reputation, and relationships.

But an important lesson is that the change team should focus on the process and not on the content. The team does not decide who needs an MIS or what the MIS should contain. It builds a coalition for change by asking staff and other groups what they need and want from an MIS and bringing these different groups into the process of establishing an MIS.

4.2.3 Deciding what to change

Information is power. An MIS can both create this power and shift the location of power, possibly leading to conflict. So an MIS must openly deal with the following issues:

- **Structure:** how is the MIS organized, and who controls it?
- **Staff:** who manages it?
- **Procedures**: how it is implemented, how it is used?
- **Linkages**: who works with whom?
- **Outputs:** what is to be produced and who will get it?

- **Culture:** how are things really going to work?

4.2.4 Creating the conditions for implementation

In creating the conditions for implementing the MIS, negotiation skills play an important role as the issues listed above will need to be negotiated with all the affected or interested groups.

The implementers are likely to encounter resistance in implementing the MIS. This needs to be recognized not as something to be overcome but rather as a clue to what is going wrong and a signal to implement in different way. Resistance is a relative term: it means that one group is behaving differently from how another group wishes them to behave, or that one group does not like what others are demanding of it. This may actually be healthy and may point towards the need for new procedures and attitudes. Whether these changes are required in the "resistors" or in the implementers can emerge from a process of analysis, discussion, and negotiation. A forum and rules for negotiating permit different groups to give their views in a structured and constructive manner. It gives them the opportunity to assess the strengths and arguments of the differing views and creates an environment where compromises or agreed solutions can be worked out.

In negotiating it is useful to present the new initiative as a proposal and not as a project. A proposal sends the signal that negotiations are possible; a project says that the important decisions have already been made.

To have a successful negotiation process, it will also be important to bring to the surface the assumptions of different groups of the need for an MIS, the reasons for promoting it, who will benefit, and who will lose. Any change process therefore needs to have a vision that answers the following questions:

- Why do we need an MIS?
- What will our MIS look like?
- How are we going to get there?

4.2.5 Achieving change

In this phase, perhaps the most difficult of the change process, we begin to produce results. For an MIS it is important to recognize that change will probably not be one big, sudden event, but rather a sequence of small events. It is best to plan for this, to make a plan of transition of small steps building to a greater change. At each step, achievements should be celebrated to let all stakeholders know that the change effort is succeeding.

These issues surrounding the introduction of an MIS into an organization, and the myriad of institutional factors that interplay between people, need to be borne constantly in mind through the design, building, introduction, and management of the system described in the remaining sections of this chapter.

4.3 Systems analysis and design

by Dickson Baguma and Richard Vernon

The phrase "systems analysis and design" refers to the processes through which the existing practice of information management is studied. The results are used to redefine the needs for information, which form the basis for the design of a new system to meet these needs. This section describes two main approaches: the structured method and prototyping.

The structured method is a formal approach with well-defined stages. It has been widely used in the past and is still used for large and complex systems. It requires a combination of two rather different skills: a knowledge of the business enterprise, in our case the research organization, and a knowledge of information systems and their software and hardware technologies.

Prototyping is the result of a trend towards providing users with a working system early in the process, rather than detailed descriptions of what is planned. There is a merging of some of the stages of the structured method. It is more marked in medium- and smaller-sized systems. It is supported by increasingly powerful and easy-to-use software.

In the context of this book, we offer an overview of the stages and some of the key tasks that each stage involves. This will help research managers appraise proposals for information system development and monitor a system development project. Very large NAROs will still need specialists to assist them in developing their systems. In such cases this section will provide the research manager with a useful overview and explanation of what the specialists are doing.

4.3.1 The structured method

The structured method consists of the following stages:

- terms of reference
- feasibility study
- systems analysis
- system design
- system development

Terms of reference

The terms of reference (ToR) is a formal document to guide the analyst and the users with a broad view of the system development project. It specifies the objectives, scope, any special constraints, material to be used, and resources available (cash limits, time deadlines, persons committed). When an outside consultant is used, the ToR provides the basis of a contract. The ToR may be provisional pending the findings of the feasibility study.

Feasibility study

A feasibility study is a preliminary investigation to determine the prospects of the proposed information system and whether it should be developed. An analyst makes a rapid study of current practices and systems and of any problems and gaps. Alternative solutions are examined and may include merely solving specific problems, broad improvement across the existing system, or moving to a new system. A cross section of users is consulted. It results in a feasibility report that advises whether a project for a new system is worthwhile or whether work should be stopped at this stage. Whatever the advice, it must be substantiated with facts and explanations. These should include an indication of resources needed and the costs and benefits, even though it will be difficult to determine these with accuracy (see section 7.2.2 "Some cost-benefit considerations of information systems").

If the advice is to proceed with the project and develop a system, then it should specify the outline of a solution. It may offer a choice between two or more options, and it may specify ways in which the original ToR need to be adjusted.

Systems analysis

Systems analysis involves an in-depth and documented study of current practices in information management. Based on this study, a proposal for a revised or new system is developed. The process has three parts: gathering data, analyzing it, and developing from this analysis the system proposal, which will be the starting point of the next major stage: system design.

Gathering data

In consultations with all levels of users, the analysts will seek to identify all important data entities and information flows, i.e., their origins and destinations, and record and analyze the information obtained. In a NARO the users typically include research managers, program and project leaders, station directors, research information managers such as the librarian and biometrician, and representative research scientists.

This stage uses several sources of information:

- **Documents.** These include annual reports of the national institution and individual stations and/or programs, research protocols, and any routine reports used for research planning, budgeting, monitoring, and reporting.
- **Questionnaires.** These need to be designed carefully to avoid misinterpretation. An accompanying letter can explain the purpose and give a return date. Note that there are limits to the accuracy that respondents can provide, particularly with numeric data such as hours per week spent on a particular task.
- **Interviews.** These are often the most productive source of facts and the best way to tap people's views and experiences of the current systems and practices.
- **Observation.** This comprises visits to key offices and operations to view their processes, check previously recorded facts, and look for new (unrecorded) facts and

for bottlenecks. Good powers of observation are a valuable trait of the good analyst.

The order in which the methods are listed may also be useful in applying them. Initial study of appropriate documents can provide the analyst with a good initial view of the current systems and assist in developing an effective questionnaire. A well-thought-out questionnaire is likely to prompt users to think about their information use and needs, which can enhance a subsequent interview.

Holohan (1992) and Cashmore and Lyall (1991) describe formal methods for the identification of actual data items They are probably best used within an institution in interaction with the management being served.

Analyzing data

At this point, the systems analyst is not looking for a solution so much as a clear understanding of the current systems and any problems with them, and for opportunities for greater effectiveness and efficiency in information management. There are well-established tools to assist in a methodical systems analysis, such as data flow charting, data entity modeling, and the function-organization matrix. They become important particularly in large and complex systems. The interested reader is referred to specialist systems analysis and design text books for further information on these tools.

This phase may also encounter resistance from users, who may have legitimate fears that are best addressed by open and frank discussion between managers, analyst, and users on plans and prospects.

Developing a system proposal

A system proposal specifies at a high level (i.e., without detailed properties, which will be developed later in the system design) the outputs (e.g., printed reports, on-screen charts), the inputs (e.g., data collection forms), and the main database tables and their relationships. It also lists the major tasks of the project to implement the proposed system, with their duration, and the resources needed.

Outputs reflect the perceived information needs of users. An MIS output only has merit if a management decision might be made on the basis of an output. Difficulty in identifying any such decision casts doubt on the value of the output.

A formal report with the analyst's recommendations is presented to management. This report should be accompanied by an oral presentation attended by all involved parties. Based on a review of the report and the presentation, management makes a decision whether to continue to the system design phase (i.e., the proposed system is feasible), revisit the systems analysis phase to meet areas of concern, or terminate the project.

Management may look at operational, technical, and economic aspects of a system proposal and at its development schedule:

- Operational capability ensures that the system can perform its intended functions within the existing organizational environment with its current (or planned) personnel and procedures, i.e., that the system will be used once it has been developed and implemented. If the end users resist a new system, the system might not be exploited to its full potential. The operational aspect is largely a human relations problem.

- The technical area comprises hardware, software, and personnel to develop (or purchase), install, and operate the system.
- From the economic standpoint, the analyst has to determine and demonstrate that the benefits to be derived from the system are worth the time, money, and other resources required to implement it.

Finally, management has to be convinced that the proposed implementation schedule is realistic and acceptable. The analyst may find Gantt charts and the Program Evaluation and Review Technique (PERT)[2] helpful to demonstrate the tasks of a project and their sequence and relationships.

If management is convinced that the project is feasible and should continue, then the project is taken to the next phase.

System design

The system design stage specifies the details of outputs and inputs and of the storage and processing, which are at the center of the system. As in the previous stage, the designers look for opportunities to maximize efficiency and effectiveness in information management. The system should be able to deliver what users need by collecting and entering into the system a minimum amount of data and providing outputs that are coherent, easy to obtain and read, and recognized by users as valuable.

Outputs and inputs

Outputs are mainly printed reports and on-screen charts and tables. These are all specified down to every data item, along with the source of each—typically one or more fields in a stored table (see also section 4.3.5 “Use of graphical outputs”). The same is done for inputs that are mostly data collection forms and data entry screens, along with the destination (which field in which table) of each item. Data collection forms should be printed with unchanging base data already in place. For example, users should not be required to fill in their name and other personnel information again every year unless it has changed. Nor should they be required to enter data that is available elsewhere or that can be calculated by the system. Calculated data

2. A Gantt chart (named after its developer, Henry Gantt) plots the several tasks of a project against time, showing the estimated start and end of each task and its overlap with other tasks. It is a simple and yet powerful device for planning and monitoring the progress of a multiple-task project such as the development of an information system. The Program Evaluation and Review Technique was developed from the Gantt chart by disaggregating tasks into smaller units, each ending in a milestone and linking related milestones in different tasks into a network. This network can in turn be used for critical path analysis, to determine for example the best order of tackling many tasks. A good review of these techniques is given by Koontz et al (1984) pp. 592-5.

should not be stored in the system anyway; it should be calculated as and when needed in a report, at that time. The extensive use of codes was formerly justified because it saved computer storage space. As this is no longer a constraint, users are better served by not having to learn codes.

Instead, the system should make extensive use of look-up tables. These enable a data entry clerk to put in items from a standard list, such as commodities or research stations. This list is presented on the screen, and the data entry clerk merely has to select an item with the mouse (or type the first few letters, after which the system will immediately enter the remaining text). Figure 4.1 shows an example of a look-up list of disciplines. Last, alternative input and output formats may be specified for testing for user preference. The front end or user interface is specified in terms of, typically, menus and their items and the function of each.

Storage and processing

Data in such a system is best stored in a relational database, the main building block of which is the table.[3] All tables that are needed to meet the requirements identified in the systems analysis report have to be defined in terms of their constituent parts,

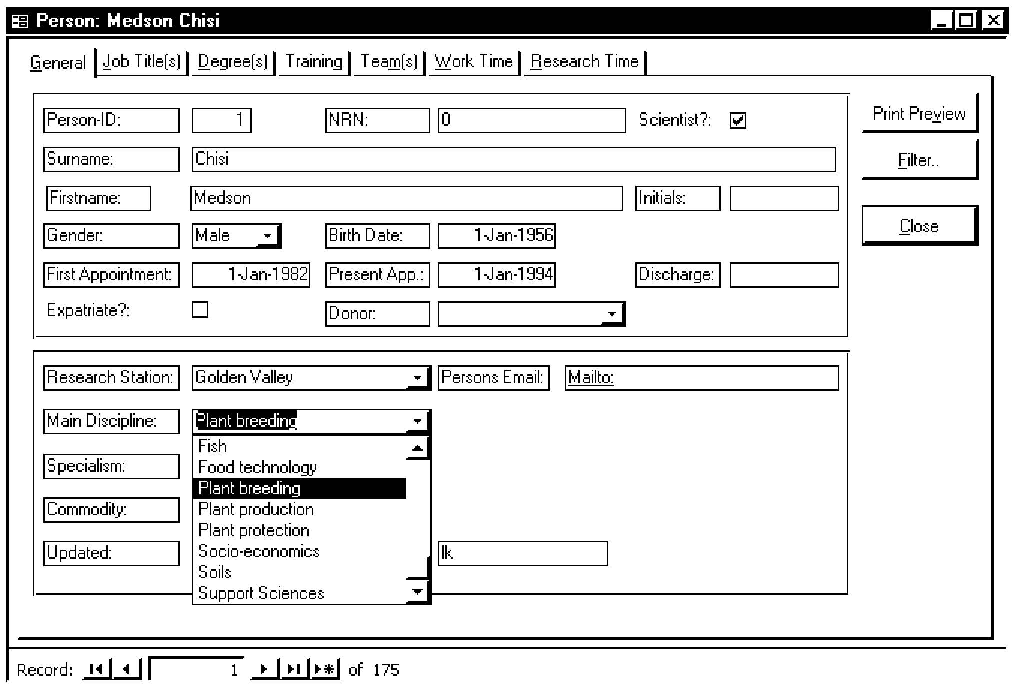


Figure 4.1: Example of a data entry window with a look-up list of disciplines

3. For more detail see section 7.5 "Elements of a database system."

the fields, field descriptions and properties, and table primary key.[4] Field properties include validation rules that provide some control against the input of invalid data, such as a person's year of birth earlier than, say, 70 years ago. Other rules declare that some fields are obligatory: a person record must have a person name, and a project record must have a project ID and a title. The relationships between tables have also to be defined.

Manual procedures such as the entry of data have to be defined, as do automatic features such as aggregation of data by the system to generate high-level reports, the conversion of approved research proposals to operational projects, and the creation of new data sets on moving to a new year. If a separate archive system is planned to store records of completed research activities, then the mechanism for transferring records needs to be specified. Procedures to test the system, complete with appropriate test data, need to be developed, and back-up procedures to guard against loss at both station and HQ levels need to be described.

The system may include built-in (on-line) help facilities, which need to be written out in full. The hardware (PCs, printers, modems), software (database management system), and any network and telecommunications systems have to be specified in detail. The time schedules and cost estimates provided in the system specification report should be reviewed and where necessary revised. Finally, all these activities have to be fully documented in a report for users, management, and the system developers.

System development

This phase sees the translation of the designed system into an operational system. User testing will inevitably call for changes. If these changes are central to the operation and requirements of the system as originally defined, they need to be addressed. If they significantly extend the originally planned boundary of the system, there needs to be a formal review with management to determine whether to proceed first and implement the system as planned.

Testing the system

A poorly tested system is a frequent cause of problems during and after systems implementation. The system should not become fully operational until it has been tested for errors. Although testing is tedious, it is essential that it is carried out thoroughly. It is useful to develop a formal test plan, which should include for each test: its purpose, definition of test inputs, detailed specification of test procedure, and definition of outputs expected.

A system's components are first tested individually, using previously established test procedures and test data. Then the system is tested in its entirety to ensure the components are compatible and function as intended.

4. An example of tables and fields and their definitions is given in annex 2 "Database Tables and Fields in INFORM-R."

Testing must include users: this will not only identify further errors but also provide indispensable feedback on the users' reaction to the system and the degree to which they find it meets their requirements. Testers should include users who are not expert in the system or its software—experts will avoid user's errors that more normal users will be prone to make. Testers will need clear guidance on what they are expected to do. They also need a good medium in which to record test results, an unhurried environment, and motivation to test thoroughly.

Examples of categories of items that need to be tested and how they should be tested are shown in table 4.4. A checklist of actual items needs to be compiled, along with an appropriate form to record test results and note any necessary corrective actions.

All new manuals, new hardware, and all system interfaces must be tested to ensure that they meet the design specifications.

File conversion

File conversion is the activity of changing any existing paper or electronic files into a form in which they can be used by the new system. Both paper and electronic forms will often be structured very differently from the new system. Paper files will need to be keyboarded into electronic files. It may be worth using an intermediate electronic file that matches the existing paper format to ease the chore of this keyboarding effort, and which the systems staff can later process into the new system. Electronic files from a prior system may be processed into the new system although the batch process files that are needed for this are difficult and time consuming to develop. Considerable manual intervention is usually required, especially if the data is inconsistent and incomplete. Users may have to be asked to update and correct their data, but as it is a time consuming task, they may need to be persuaded of its necessity.

Documentation

Documentation is the material that explains what the information system does and how people interact with it. All documentation must be clearly identified and dated. Documents must be readily understandable and should not refer to other documents as far as possible. While a number of documents were prepared as part of the systems analysis and design processes described earlier, here we look at those to be released as soon as the system is developed. The documentation should comprise the following:

- a user guide, which enables an inexperienced user to learn how to use the system for the major functions for which it is designed
- a user reference or operations manual, which provides more detailed information for advanced users, including the routine management of the system
- a systems reference manual, which provides a record of the system structure and design in sufficient detail in order to enable new staff to continue the maintenance of the system and make future enhancements in case the original system developer(s) depart

The first two may be combined in a single document. Increasingly such assistance is provided by built-in, on-line help facilities within the system software. The systems

Table 4.4: Examples of System Tests

Category of item	**Test**
Menus	Every menu item should be tested to check that: • the title makes sense • its position with respect to items above and below is logical • on selection, it takes the user to the correct place Testers may be invited to suggest alternative menus or parts of menus that make more sense to them. Alternative suggestions should be compared to find common views.
Reports	Data needs to be in the system before reports can be tested. It should include some records with all nonobligatory fields empty, as this can upset inadequately designed reports. Every report needs to be generated and checked that: • on-screen report generation works • all its on-screen functions work as expected • calculated fields, totals, and subtotals calculate correctly • on-screen version of reports close properly • reports to paper do actually print out • paper print out layout is as intended
Data entry	Data collection forms need to be printed, completed, and used as a source for data entry. All data collection forms should be used this way, and data should be entered into all tables. • Enter deliberately invalid data: the wrong type of data, e.g., numeric into a date field, alphabetic into a numeric field, impossible dates such as 30 February and planned end dates in the past, oversize, undersize, out of valid range, zero, and negative amounts. • Enter real data from actual records: choosing some at the extremes, e.g., longest records, large number of keywords, longest experiment title and longest objectives. • Delete some records if this is permitted; try to delete if it is not permitted • Change the data in some records where changes could occur in real life, e.g., add a new location to an experiment, change the surname of a female researcher as in marriage. • Test data should include records that have empty fields. • Try to devise and test invalid combinations of data. As a source for data entry, use the list of outputs for checking all data entry forms that the system can print out and all data entry cases described in the system's user manual. Check the systems specifications or data dictionary for a list of all tables and check that all tables have had data entered into them.
Data processing	Test all processes, ensuring that: • all procedures represented in the menus are included in the tests • all those described in the manual are included in the tests • all those you would expect in the research project life cycle (1) exist in the system, (2) are represented in the menus, and (3) are described in the user manual • all processes yield correct results after running them; this will require careful examination of the results • accidental, erroneous actions cannot damage the system Additionally: • Try month-beginning, month-end, year-beginning, and year-end procedures. • Live (current) data can be checked against previous results. The system's user manual may cover many of the processes but the system itself should be studied on-line to check all of them.

reference manual includes a data dictionary, which lists and defines all tables and their fields.

4.3.2 Prototyping

The previous section explained how the structured method of system development consists of a sequence of distinct stages. This is in contrast to the prototyping method. Prototyping consists of the rapid development of a preliminary model of a system (or part of a system) for an iterative cycle of testing by users and refinement. Users' comments are fed back to modify the system to better deliver what users want. This method recognizes the difficulty of prescribing in detail a system's requirements on paper prior to its construction, and the value of enabling users to develop their statement of needs dynamically through repeatedly commenting on an evolving prototype. Prototyping thus integrates systems analysis, design, and implementation.

Prototyping is particularly useful for systems where requirements are difficult to specify (e.g., decision support systems) or for solutions for designing the human interface, such as the screens for data entry, selection menus, and reports (Laudon and Laudon 1994). A valuable aspect of prototyping is that it encourages users to become interested in the development of a system. This greatly reduces the risk of implementing a system that is based on incorrect analysis of users' real needs, and it avoids the associated cost of major redevelopment. Prototyping is more useful for small systems; large ones therefore need to be broken down into small components. However, it remains important to undertake at least a basic systems analysis prior to starting. Care should be taken to fully finish the system and write the documentation.

4.3.3 System implementation

After developing the system it must be integrated into the routine operation of the organization. This will include at least the following steps:

- consultation with management on the details of an implementation plan and schedule
- notification of staff throughout the organization of this plan
- appointment of MIS practitioners at each station
- installation of any new hardware and software and train staff in using it
- progressive integration of the system into the program management cycle

The MIS implementation should itself have targets, which will vary in different cases. The following targets are a suggested set using logical framework terminology of qualitative (Q1), quantitative (Q2) and time (T) indicators:

Target 1: A few days or weeks (T) after MIS practitioners have returned from an MIS training workshop to their research stations, they are in a position to issue the first reports.[5] These could start with a directory (Q1) of all scientists (Q2) and a list of all current research activities (Q2). Selected management reports using data aggregation could follow, e.g., resource distribution across commodities, scientific disciplines, and AEZs.

Target 2: Such MIS outputs are used by research managers in the next cycle of research program planning and made available to research scientists either as printed outputs to each researcher or in electronic form.

Target 3: The MIS is integrated nationally and its outputs used in the next cycle of national or AEZ research program planning.

Most MIS information declines in value with time. Output from the personnel and project files should continue without waiting for project budgeting output, which tends to be more complicated and may be included in the next round of output. Absence of some data such as that from a particular program, station, or individual should not hold up outputs beyond a declared date. It is best to proceed without the missing data this time round, or the system will take on the time scale of the slowest. Adherence to a prompt but realistic timetable from the start is well worthwhile. It sets a standard that those missing it the first time round will be anxious to meet the next time, to ensure their work is not left out of reports known to be distributed widely and on time.

Training

Training staff in using the system is essential. There are different training needs for different categories of users: research managers, the national MIS coordinators, station MIS practitioners, and scientists (see also Annex 5 "A Syllabus for Training in the Implementation of a Management Information System"). It may be necessary to enlist external expertise for some of this training. If the NARO has procured a ready-made system or used an external organization to develop a local system, then the source of the system or expertise is likely to be a useful provider of initial training. In such cases it is important to pass such expertise on to local staff at an early stage through a "train the trainer" program to ensure the system's sustainability. Any appointed training officer(s) within the NARO should be closely involved.

The nature of an MIS implementation in a NARO lends itself well to the transfer of expertise. The implementation involves a number of workshops, the first of which is typically led by external trainers. These trainers train local staff not only in using the system but also in training future users. During the first workshop, the learners can even take on some of the training, certainly acting as "demonstrators" in practical sessions. Local and external trainers can work together in delivering the second workshop. By the third, local staff can take the lead with external trainers, if present at all, merely backstopping.

5. This of course assumes that real data has been captured prior to and for use in the training. We strongly recommend this, because the training is then much more meaningful. See the suggested sequence of events in section 4.1.9.

It is very important that all those undergoing such training are in a position to apply it immediately. It is quite different from learning to ride a bicycle—skills learned in computer training are lost rapidly if not employed. So the training has to be fitted into the whole MIS implementation plan in such a way that hardware and software are already in place. If new equipment is being procured, then it should be used for the training workshops; trainees can take the new equipment with them back to their station, with the MIS installed and operational, for immediate use.

Systems changeover

At this stage, all the prior system implementation activities have been completed, and the new system should be ready to take on all its planned functions. Sometimes a newly procured or developed MIS replaces an existing system. This is typically the case in sectors such as banking and commerce, which often have established procedures for such changeover. There are a number of recognized ways of making the conversion, the choice of which depending on the communication required during the changeover, the nature of the system, and the skill of the staff responsible. A pronounced winter or dry season offers a relatively calm period in the annual program management cycle to have the conversion take place, particularly if it is well planned. Four ways of changing over from an existing system to a new one are the following:

Parallel conversion. The old and new systems are run side by side for a fixed period of time. Outputs from each system can be compared to check for errors in the new system. The causes of errors can be investigated and revisions made in the final design of the new system. The duration of parallel processing may have to be extended if an excessive number of errors need to be corrected.

- *Advantage:* The old system is available to fall back on in case there are serious difficulties with the new one.
- *Disadvantage:* The high cost, mainly in staff time, of operating two systems simultaneously, and in the time and skills needed in running the cross-system checking.

The case study from Brazil in section 3.5 provides an example of parallel conversion.

Direct conversion. The old system is dismantled immediately after the new system has been put into place. It might be the preferred choice if the old system has so many weaknesses that a parallel conversion would serve no useful purpose or if the new system is very small or simple and does not involve major risks.

- *Advantage:* It avoids the cost of running two systems simultaneously.
- *Disadvantage:* Should the new system fail, there is no system currently operational.

Pilot conversion. The new system is installed initially in only part of the organization, such as one or a few research stations, where it can be thoroughly tested in all its aspects before it is extended to other stations.

- *Advantage:* The old system is available to fall back on in case there are serious difficulties with the new one. It is less disruptive and there is more time for training and for detecting and correcting errors, and it imposes lower peak demands on users and the system operators.
- *Disadvantage:* The organization is left working with different systems in different parts of the organization.

Phased conversion. Throughout the organization only part of the new system is implemented initially. When that part is proven to work efficiently and users are comfortable with it, another part may be introduced, again for organization-wide adoption. In a NARO an obvious application would be to first introduce a module to manage research projects, then one for project budgeting, and finally a research personnel module.

- *Advantage:* The old system is available to fall back on in case there are serious difficulties with the new one. It is less disruptive and there is more time for training and for detecting and correcting errors, and it imposes lower peak demands on users and the system operators.
- *Disadvantage:* The modules have to be able to be implemented and operated independently and, later, collectively.

4.3.4 System links and limits

The system specification report (see system design phase above) will have identified any other systems in the organization with which the new system will have to exchange data. The design needs to address this through, for example, using similar field names and properties. For example a research program-based MIS will need some information on researchers and research budgets. It may be feasible to build links with any existing personnel and finance systems, but these will need to be defined. There is a useful rule in information systems that duplicate copies of data should be avoided as far as practicable. It is more efficient to build links between the various sources so that all systems requiring a particular data item obtain it from only one stored source (see section 7.5 "Elements of a database system" for a discussion of this issue). In the absence of computerized finance and personnel systems, it may be necessary to include a limited data set of these areas in the research program MIS.

It will also be necessary to consider the distribution of outputs of the system to external parties. If significant personal information is included in any such outputs it will be important to check that this does not infringe the rights or concerns of the persons concerned, nor any rules set by the government on the reporting and distribution of personal data.

Limits to MIS functionality

There is an old adage that says "the map is not the country;" it is just a representation of a very small set of data about the country. It can be very useful in working out where you are and which is the best way forward. But it may be unable to tell you whether you will find a place to eat in the next village or whether it will rain on the way there.

The same applies to an MIS: it may report precisely the numbers of scientists with an MSc or higher degree currently working in plant protection, and it may list past experiments that have looked at fertilizer effects in bananas. It is silent, however, on the relative leadership capabilities of station directors, and it is weak on causal relationships, such as why there have been three times as many experiments on fertilizers than on plant diseases. It may flag an apparent issue for management, such as a major difference between agreed priorities and actual resource allocation, but it may not provide the explanation for the difference. The key issue is that an MIS provides managers with facts to make better management decisions, but managers still need to weigh up the evidence, from the MIS and elsewhere, and make the decision.

4.3.5 Use of graphical outputs[6]

Designing useful outputs from an information system requires not only statistical skills but also a good knowledge of the business. Graphical outputs also need some technical skill in graphics. The discussion below is directed primarily at the printed output but in many cases also applies to on-screen displays.

Graphical outputs offer an efficient way of passing to the user a rapid view of an aspect of the research program. Especially with large sets of data, managers can much more quickly derive conclusions or grasp the meaning of the underlying data when they are represented graphically. See, for example, figure 2.2 (page 25) , which captures resource allocation data from hundreds or even thousands of experiment records and compares it with a country's research priorities. It is possible that these conclusions are more easily remembered when seen in graphical representation.

Tables usually outperform graphics on data sets of 20 numbers or less. A two-way table or crosstab can be used to show the variations of two factors in relation to a third. See, for example, figures 2.11 (page 52) and 2.12. If the data set is large, a multiple (as in figure 2.13) or stacked histogram may be preferred. Additionally, simple tables and pie charts can show the breakdown of one factor into its constituent lower-order factors. For example, see figure 3.3 (page 101). The physical representation of numeric data on a graphic should be directly proportional to the numerical values.

The efficiency of a graphic is influenced by its design. The information in a graphic lies mainly in the data, so the viewer's attention should be attracted to this part and not distracted to excessive structural components. The number of dimensions of the graphic should not exceed the number of dimensions in the data as they contribute no extra information and distract. Superfluous dimensions also limit the size of the data set and can lead to error or ambiguity in the reader's perception. Only some of the total amount of ink used in a graphic represents data, so the use of ink for nondata purposes (such as borders) should be minimized. These aspects of efficiency affect people's absorption of information, the rate at which printer consumables are used up, and the rate at which graphical data is transmitted through electronic media.

Some graphics distinguish between categories of data by different shades or patterns of cross-hatching. The disadvantage of representing data this way is that it requires

6. Tufte (1983) describes this subject in detail.

the reader to consult a separate key to see what each shade or pattern represents. A graphic that offers direct labeling of each category is easier to interpret, but if a category is dispersed over the graphic, then a separate legend may be necessary. Labels are more easily read in lower case or mixed upper/lower case than in upper case only. Sans serif should be restricted to titles and labels.

If direct labeling of data is difficult and keys have to be used, then shades of gray are easier to comprehend than pronounced cross hatching. Bold black and white stripes can have an unpleasant effect on the eye; they are a distraction if the purpose of the graphic is to inform. Similarly, grid lines and crosses are usually best omitted. If they are needed to see the precise position of data, they should be muted by being thin and gray rather than black.

In a report, small informative graphics can be embedded in the text or printed alongside it.

The following guidelines may help produce more informative and efficient outputs:

- As with illustrations in a scientific paper, an output should be largely self-explanatory without reference to the text.
- All outputs should include certain base information about their origin:
  - the information system from which the output is produced
  - date of report
  - the year(s) or season to which the report refers
  - the geographical base, e.g., name of research station or area to which it applies
  - whether it is based on (a) ex ante plans (estimates) or on (b) ex post data
  - the units used; in long time series monetary values should be deflated.
- There are usually better sequences than alphabetical, for example ordering by content or by data values.

4.4 Managing the system

At the start of this chapter we discussed the roles of key persons and groups of persons in implementing and managing the system. Now that we have seen how a system is built, we are in a better position to examine the responsibilities of those charged with maintaining it. It is much preferred to locate such people in the domain of the research program and its management rather than in a separate computer or information unit.

From time to time there will be suggestions at both station level and HQ for changes to the system. A distinction must be made here between components of the system:

1. **The underlying database, comprising the tables and their fields and relationships.** Changes should not be made to this component unilaterally at the station level, because this will result in incompatibility between models of the system at different stations. Rather, suggestions for change should be put to the user com-

mittee (see section 4.1.1), which will consider them and, if found worthwhile, will authorize the database administrator (see below) to make the necessary changes. These will then be provided to all stations, perhaps at the following (annual) system distribution.

2. **The data contained in** 1**.** The data in the system should normally be entered and updated at station level. It should conform to any standards agreed to as part of the system, for example a standard set of scientific disciplines applicable to both scientists and research activities. If the allocation of research activity ID numbers is made nationally, these must not be changed at station level as they are heavily used by the whole system.

3. **The "front-end" or user interface.** This comprises the built-in data entry devices and the built-in data outputs (reports and their underlying queries, etc.). The user interface should usually be treated as for component 1.

4. **Supplementary outputs (queries and reports) to meet ad hoc needs for information.** These outputs are usually created (temporarily or permanently) at station level and wherever users are interrogating the system for information and need to go beyond the built-in reports. This is a valuable component of any flexible information system. With modern userfriendly database system software, it should be open to any user willing to devote some time to learning the basics of queries and reports. The MIS practitioners at each station will normally be trained in covering this facility and should be able to provide for any ad hoc requests for information, particularly from the station director. A mechanism needs to be put into place to allow such ad hoc outputs to be saved and not overwritten by the annual distribution of the updated national data set.

4.4.1 The system managers

MIS practitioners at the station level and the MIS national coordinators at HQ have special responsibilities regarding the management and maintenance. They work closely with the user committee.

The MIS national coordinator (database administrator)

Elbra (1982) points out that an organization cannot afford *not* to have a responsible person in charge of the system, and that it cannot afford to assign responsibility solely to one person (see also the earlier section 4.1.4 on critical success factors). There should be at least two people trained to this position. Normally they would be part of a central planning, monitoring, and evaluation unit. The tasks attached to such a position include the following:

1. Collect, check, and aggregate all the annual (or seasonal) data sets of the station into the national system, and redistribute this to all stations.

2. Generate all required reports from the national system and distribute them as appropriate to all stations, senior staff at HQ, relevant ministries and donors (see figure 5.1, page 174), and all appropriate events such as national program planning. The system will normally provide the core material of the organization's an-

nual report, for which the MIS national coordinator may have a major contributing responsibility.

3. Maintain the data definitions. Ultimate responsibility for the substance of each definition rests with the owner, supplier, or main users of the data. In practice this may be realized by maintaining an active national user committee that establishes a reputation of responding to new ideas and suggestions of change. The database administrator is responsible for ensuring all user-agreed definitions are incorporated into the system and added to the data dictionary.

4. Provide technical and advisory support to research stations and their practitioners, convening debate (a) between directors and program leaders on how the MIS can best meet their information needs and (b) among practitioners on maintaining consistency of terms and completeness of the national database.

MIS practitioners at the research station

MIS practitioners should be agricultural scientists in order to manage this kind of information. Especially in a large research station they may be supported by trained clerical or data-entry staff for entering data and other routine tasks. Suggested terms of reference for MIS practitioners are the following:

1. Follow appropriate training in implementing an MIS and attend subsequent follow-up seminars as may be arranged.

2. Collect data. Once a year or once every season collect and enter data on researchers, research activities, and, if appropriate, budget data for the institute or station. Collect monitoring information at appropriate intermediate times.

3. Process data. Within an agreed time period (typically three weeks from when data should be supplied by researchers) prepare and pass to the director an agreed standard set of reports, even if some data are incomplete (see also section 4.3.3 “System implementation”). Discuss with the director and program leaders this provisional output and agree on any additions and revisions. This can be done in the presence of a PC running the MIS; different reports can then be generated and explored.

4. Produce final outputs (1) on paper for distribution to the director general, all institute directors, and all researchers and (2) on diskette (the original database files for researchers, research activities and budget, and derived reports) for distribution to NARO HQ for compilation of national output.

Annex 5 gives a suggested curriculum to meet the specialist training needs of these positions.

4.4.2 Collecting data

Data collection is often one of the more problematic tasks in any information system because the data usually has to be provided by people rather than automatic mechanisms. Experience suggests the following two practices to assist in the process.

Collecting initial data

In the initial implementation, scientists at each research station are gathered in a meeting room and given a presentation of the MIS to show them what it is, what it has to offer to them and to management, and what it requires from them to make it work:

- Benefits to scientists include access to details of all research projects nationwide and if the MIS includes an archive module, to past research also. This is of considerable assistance in the planning of new projects and the preparation of research papers for publication. Having their own projects in electronic form also helps them prepare their own annual reports; the MIS may well include a module specifically for this purpose.
- The input needed from the scientist is to complete or update, usually annually, personnel and research activity data forms.

This is followed by a discussion that can help resolve any issues. Finally, personnel data collection forms (sheets) are given out for immediate completion. Any questions are dealt with on the spot. At the end, the forms are collected and one of the main components of data collection has been completed. Research activity forms can be distributed and explained at the same event, but it usually takes longer to complete them. Researchers may best do this in their own time later.[7]

The completion of both forms can be simplified by obtaining beforehand from the station administration a list of scientists and their experiments, keying them into the system, and have the system print out customized forms for each person and each experiment for filling in the rest of the data.

Collecting subsequent data

The second device to ease the data collection process comes into play in the second and subsequent round of data collection, usually the following year or season. At this point the MIS should print out data collection sheets complete with standing data, leaving blank only those few fields that are likely to change. These sheets, for each person and each experiment, can then be completed very quickly by each individual.

4.4.3 Achieving accuracy and timeliness of data

Achieving accuracy and timeliness of data are two closely linked issues. Data needs to be sufficiently accurate for the purpose, and no more. For example, the time distribution of researchers may be requested to the nearest 2% (= approximately one week). This is quite sufficient for the purpose of informing management of the broad resource allocation across commodities, AEZs, disciplines, etc.

At least some outputs of interest to those providing the bulk of the data—the researchers—need to be released quickly to both researchers and managers. If this is done, peers can be expected to exert a strong pressure on those who missed some in-

7. I am indebted to Ghazi Hariri, a former ISNAR staff member, for this approach to collecting initial data.

formation, exaggerated, or otherwise gave misleading data; their peers will pick up many such errors ("What do you mean by saying you are working on my bean variety trial?"). It does no harm also for those supplying data to know that soon afterwards the resulting information will be in the hands of their peers, research managers, senior ministry officials, and donors. These effects are realized in the second round, when data suppliers are guided by the experience of the results of the first.

It is therefore important to get some outputs out soon, even if the data is incomplete and not fully verified. Delaying outputs to the timetable of the slowest is to set standards to those with the lowest (see also 4.3.3 "System implementation").

The importance of timeliness is that data loses value with time, and delays in delivering outputs results in a drop in interest of both the data providers and data users.

4.4.4 Addressing very large sets of data

The enormous gains in data-storage and -processing capacities of recent years have led to some interesting prospects for deriving new information. Methods are being developed to harness the power of modern computer systems to search very large data sets. The assumption is that it is possible to derive small but valuable items of information from them. Some have called the process "data mining," analogous to the process of extracting small quantities of valuable minerals from large bodies of rock or ore (Fayyad et al. 1996).

The process starts with cleaning up the large data set in order to bring it into a uniform state to facilitate processing (Mathews 1996). This processing may include a search for trends, common features from data of disparate sources, and patterns that may contribute to a greater understanding of the business and of behavior of people.

The technique appears to have potential but there are some difficulties. There have been some reported successes from the commercial world, but there are also occurrences of false results, which need to be guarded against. It calls for considerable investment in computing hardware, software, and skills, and of course large data sets. In many cases these exist in unsuitable formats or even as paper and microfiche, and they are not immediately accessible to computer processing. ISNAR has made an initial study of its holdings of the data in MIS of NAROs, and the work shows some promise in identifying information of interest to agricultural research management[8] (see section 2.3.6 "Regional information sharing").

References

Brooks, F.P. 1995. The mythical man-month: Essays on software engineering. Anniversary Edition. Reading, USA: Addison-Wesley

Cashmore, C. and R. Lyall. 1991. Business information, systems and strategies. London: Prentice Hall.

8. ISNAR regards data from NAROs as the property of NAROs and will seek their permission before making public any such material for which the origin is apparent.

Chege, M. 1997. Paradigms of doom and the development management crisis in Kenya. In *The Journal of Development Studies*, Vol. 33. No. 4, April 1997.

Crowe, T. and D.E Avison.1980. Management information from data bases. Basingstoke: MacMillan Press.

Date, C.J. 1990. An introduction to database systems. Volume 1. Fifth Edition. *The Systems Programming Series*. Reading: Addison-Wesley.

DeLone, W.H. and E.R McLean. 1992. Information systems success: The quest for the dependent variable. *Information Systems Research*, 3(1) August 1992. pp. 60-95.

Elbra, T. 1992. Database for the small computer user. Manchester: The National Computer Centre

Grindle, M. 1997. Divergent cultures? When public organizations perform well in developing countries. In *World Development*, Vol 25, No. 4.

Hobbs, H. 1996. Strengthening the use of information tools in NAROs: The implementation triangle. A conceptual framework. Unpublished paper. The Hague: ISNAR

Hobbs, H. 1999. Ten tools for managing change in national agricultural research organizations. ISNAR Research Management Guidelines No. 5. The Hague: ISNAR.

Holohan, J. 1992. Use of executive information systems in measuring business performance. *Journal of Information Technology*. 7:177-186

Kean, S. A. 1993. The institutional politics of agricultural research policy in Zambia: A model of contingent innovation. Doctoral thesis. Norwich: School of Development Studies, University of East Anglia.

Laudon, K.C. and J.P Laudon. 1994. Management information systems. New York: Macmillan Publishing.

Li, E.Y. 1997. Perceived importance of information system success factors: A meta analysis of group differences. *Information & Management*. 32:15-28.

Mankin, D, S.G. Cohen, T. K Bikson. 1996. Teams & technology: Fulfilling the promise of the new organization. Boston, MA: Harvard Business School Press.

Markus, L.M. 1983. Power, politics, and MIS implementation. *Communications of the ACM* (Association for Computing Machinery). Vol 26, No. 6.

Najam, A. 1995. Learning from the literature on policy implementation: A synthesis perspective. Working Paper. Austria: International Institute for Applied Systems Analysis.

Peterson, S.B. 1994. Saints, demons, wizards and systems: Why information technology reforms fail or under-perform in public bureaucracies in Africa. Development Discussion Paper No. 486. Cambridge, MA: Harvard Institute for International Development.

Rainer, R.K. Jr. and H.J. Watson. 1995. The keys to executive information system success. *Journal of Management Information Systems*. Vol 12, 2:83-98.

Roberts, C.B. 1996. The impact of information technology on the management of system design. *Technology in Society*. 18, 3:333-355.

Tober, B. 1997. The quality factor: How do you ensure better software quality? *Byte Magazine*, April. 40IS:15 - 22.

Vernon, R. 1986. Do you ever wish you were better informed? The role information technology can play in development programmes. *International Agricultural Development*. January-February: 12-16.

Vernon, R. 1994. Evaluation of ISNAR's implementation of INFORM (Agricultural Research Management Information Systems) to date and proposals for the future. ISNAR Discussion Paper No. 94-12. The Hague: ISNAR.

Vernon, R. 1995. INFORM: An information system for agricultural research management. ISNAR Briefing Paper No. 21. The Hague: ISNAR.

Vernon, R. 1999. The information needs of agricultural research management. *Quarterly Bulletin of the International Association of Agricultural Information Specialists*. Vol. XLIV (1,2): 45-55.

Chapter

5 Information Systems for Administration

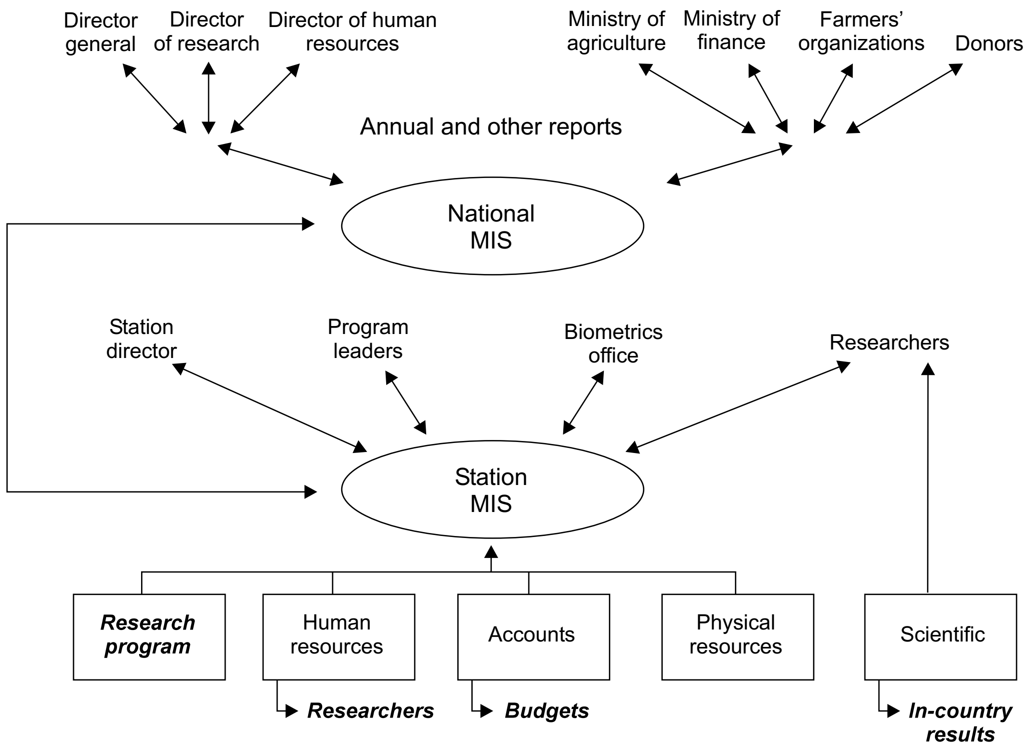
This book focuses mostly on the information needed in the management of research programs, as distinct from the administration of research organizations. Research programs, after all, are the raison d'être *of the agricultural research organization. But as figure 5.1 shows, the research program is part of a much larger whole. This chapter focuses on the information needs of managing the three supporting resources: finance, personnel, and physical resources. The chapter is therefore firstly addressed to the heads of the services respectively responsible for these resources, viz. (chief) personnel officer, (chief) accountant and perhaps, for the last, the farm manager, head of stores, and head of workshops. In larger organizations those responsible for the day-to-day administration of these services are also included.*

The chapter offers ideas and guidance on the management of information for each of these key support services.

An early question that arises when planning the management of information for administration is whether to use a single large system or to use separate (and possibly linked) systems for the management of personnel, finance, and physical resources. Arguments for the two options are presented in section 4.1.3. In practice, individual systems are commonly developed in an ad hoc fashion and in isolation from each other. Only a few organizations are able to manage the whole information set in an integrated way, as it takes considerable resources, commitment, and skill.

In the absence of such integration it is helpful to include, within the management system of the research program, some information from these other areas. Examples are financial information such as budget allocation (and possibly expenditure) against individual projects and basic information on the scientists. ISNAR's INFORM and INFORM-R systems follow this "middle road" approach.

In this chapter, specialists in the fields of scientific information, research program budgeting and accounting, and personnel management offer additional guidance in these areas.



Note: ***Bold italicized*** components are found in ISNAR's MIS INFORM-R.

Figure 5.1: Typical information flows within a NARO

5.1 Financial information in an MIS

by Govert Gijsbers and Bruce Fraser

Introduction

Finance is an important component of the MIS of every research organization. It requires excellent management and reporting, especially as the funding for research in many developing countries is now much reduced. Additionally, both national and international donors demand more detailed information on the use of "their" funding. Thus research organizations must not only be able to use funds effectively to achieve institutional objectives, but also to demonstrate that fact to outsiders. An MIS can assist in providing financiers and other stakeholders with accurate and timely information.

A comprehensive MIS will cover the activities of a research organization (programs and projects) and take into account the resources needed, in particular finance and staff. The scope of the MIS will vary greatly depending on the specific needs of the organization, the legal requirements of financial reporting in the country's public sector, and the availability of other information systems (e.g., accounting, human resources). From the perspective of information technology, all data should be registered only once in a shared database, with all information systems using that same data and the same revisions, avoiding problems of multiple versions. In practice it will be difficult to achieve this ideal situation, and there is a question how desirable it would be; if a system is tightly integrated any problem in the system may affect many functions. If a system is "loosely" integrated, then individual components can proceed to function in spite of problems elsewhere.

Research is done through programs, projects, and experiments, and the MIS should be capable of producing reports that combine data from all levels. It is essential to identify the basic building block, and for any project-based organization this will ordinarily be a project. This may be further subdivided into activities or geographical areas. There is also often a need to consolidate data from the various projects, which can often be captured in "programs" (see section 2.1.2 "Key components of an agricultural research MIS").

The financial information contained in a NARO's MIS will therefore need to be tailored to many specific needs. A single specification covering all of these cannot be established, but some general guidelines do apply. In the following sections we will first discuss the integration of financial information in an MIS and then some of the ways in which it may be used to improve research management and accountability.

5.1.1 Terminology

In financial analysis there is no generally accepted set of terms, but the following concepts are important in developing the financial component of the MIS and especially to ensure consistency with the terminology used in other financial information systems.

Cost center is an accounting term for a location where costs are gathered for a common goal. There are two categories of cost center: (1) for research projects and (2) for administrative departments (e.g., the computer department). The cost center is managed by the budgetholder, who is also responsible for financial performance.

Operational costs (sometimes called direct costs) are the out-of-pocket expenses incurred in carrying out a project, e.g., costs of seed and travel. These are usually easily identifiable as belonging to a particular project, e.g., through a supplier's invoice and staff expense claims.

Indirect costs are costs incurred in carrying out a project that arise from institutional expenses, e.g., utilities and premises. Staff costs can fall into this category although some organizations consider this an operational cost. Mechanisms need to be put in place to allocate personnel and "overhead" costs to individual projects. For example, a time registration system can ensure that salary costs are reallocated to the various projects that an institute is managing.

Cost categories provide a description of the nature of a cost and include both operational costs and indirect costs. The totals of these are usually reported in the annual financial statements.

It is very important that the different financial systems within an organization use consistent terminology and common codes for expenditures, organizational units, etc. The system(s) should be consistent with the institution's chart of accounts, a reference system used by accounting departments incorporating both cost centers and cost categories. In some countries the chart of accounts is subject to governmental control.

5.1.2 Cost types

A project-based research MIS needs to include information on operational costs, research staff costs, and indirect costs. If a particular cost can be directly related to a research project, then it is referred to as a direct cost. If it cannot be directly charged to a project, it will be allocated initially to a general cost center (e.g., administration), and then reallocated as appropriate.

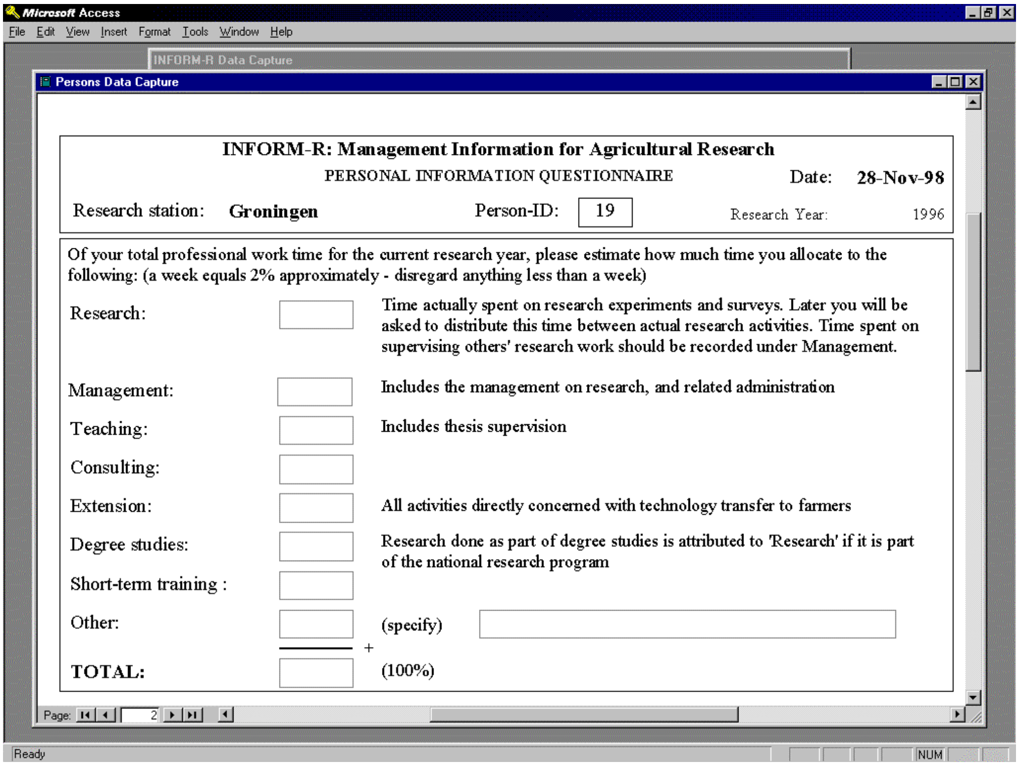
Cost type 1: Project operational costs

Operational costs include those of materials, supplies, travel, consultants, publications, etc. Whenever possible, operational costs of individual research projects should be charged directly to the project. The cost of inputs such as fertilizers and chemicals can usually be fairly easily related to projects. Other costs, e.g., fuel, are often more difficult to relate to individual projects and would be treated as indirect costs. The reasons for this are mostly practical; it would be too costly or time consuming to generate the necessary information.

Cost type 2: Personnel costs

The efficient use of human resources is crucial to the success of agricultural research. Staff time is very often the largest "cost" component of an organization, so it is important that mechanisms are in place to properly allocate this cost to the various departments or projects of the organization. Thus, information on the activities and the cost of research staff is a key feature of an MIS. Traditionally, public-sector organizations do not account for the use of personnel as stringently as does the private sector, although this is changing rapidly as research becomes more commercially oriented and increasingly held accountable for how resources are used. For billing purposes the staff costs of research projects provide essential information.

Staff time can be allocated to projects either through an estimation of the time allocated to the different research and nonresearch activities by each researcher, or by a time registration system using time sheets that record the number of hours or days spent on the various activities. Staff may be reluctant to having to account for their time, so the solution depends greatly on the culture of the organization. Figures 5.2a and 5.2b present an example from INFORM-R of how annual time allocation information may be obtained in a two-step process, first by allocating time to research and nonresearch activities, then by allocating the research time to individual projects.


Microsoft Access
File Edit View Insert Format Tools Window Help
INFORM-R Data Capture
Persons Data Capture

INFORM-R: Management Information for Agricultural Research
PERSONAL INFORMATION QUESTIONNAIRE
Date: 28-Nov-98
Research station: Groningen
Person-ID: 19
Research Year: 1996

Of your total professional work time for the current research year, please estimate how much time you allocate to the following: (a week equals 2% approximately - disregard anything less than a week)

Research: Time actually spent on research experiments and surveys. Later you will be asked to distribute this time between actual research activities. Time spent on supervising others' research work should be recorded under Management.
Management: Includes the management on research, and related administration
Teaching: Includes thesis supervision
Consulting:
Extension: All activities directly concerned with technology transfer to farmers
Degree studies: Research done as part of degree studies is attributed to 'Research' if it is part of the national research program
Short-term training :
Other: (specify)
+
TOTAL: (100%)

Page: 2
Ready
NUM

Figure 5.2a: MIS questionnaire for researcher's time allocation

To calculate the cost of a researcher's participation in specific research activities, corresponding staff cost information is required. The total annual cost of a researcher to the institute is made up of salary plus benefits. A share of this total can be allocated to represent the researcher's participation in specific projects or activities (figure 5.2a). The total annual costs can be divided by the number of expected working days in a year to arrive at the daily rate. As it is impractical to calculate this for every individual researcher, in practice a standard cost is usually calculated for a group or category of researchers, e.g., salary grade. Research time is charged to projects, and time spent on nonresearch activities (e.g., training and administration) can be charged to general cost centers (see below).

Cost type 3: Indirect costs

Some expenditures benefit the institution as a whole and not just an individual project or group of projects. Often referred to as **overheads**, some examples are the cost of administration, facility, utilities (water, electricity), services (library), and infrastructure.

These institutional expenses are often allocated to projects by way of a single estimated percentage, e.g., "25% overhead on all other costs." But more precision is often possible. The cost of a greenhouse, for example, may be reallocated to those

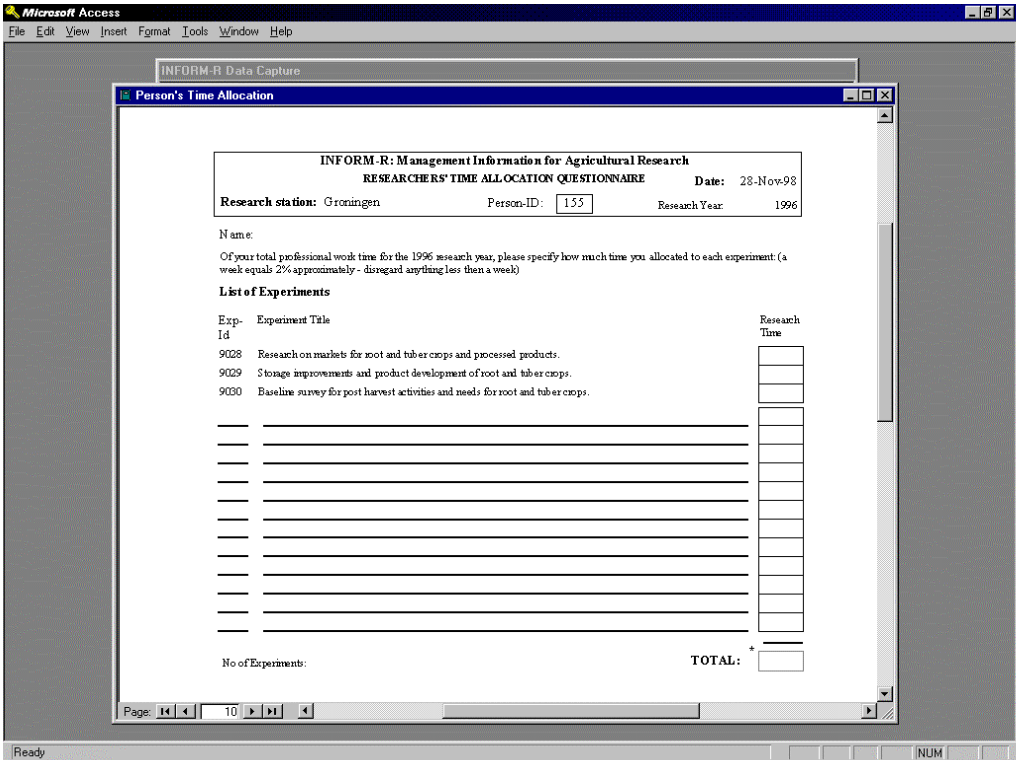


Figure 5.2b: MIS questionnaire for researcher's research time: breakdown across projects

projects that use the greenhouse in proportion to the number of square meters that each project uses. The decision of which costs to reallocate back to projects is a matter of purpose and presentation. On the one hand, each project should carry a share of institutional overheads. For example, the institution's electricity bill can be distributed across all projects. In this way, donor funding for projects contributes also to the cost of running the institute. On the other hand, if too many overhead costs are charged to projects, then the responsible researchers and directors will have difficulty to relate the numbers to actual project expenditures. Adding unrealistic overheads to projects will also make it difficult to obtain external funding.

5.1.3 Cost centers

Cost centers, whether for projects or institutional purposes, are the primary units of detailed financial information. So as not to overwhelm research managers with financial data, MIS developers should limit the financial information in the research MIS to summary information. Detailed financial information should lie only in the accounting system. Agreeing on and establishing the boundaries of the research MIS and the accounting system is essential in order to avoid duplication of effort and thus inefficiency. If it is agreed that the research MIS will only contain summary information, it still has to be decided what data fields are needed. The most important fields are actual expenditures (current and previous year), budget (current and next year), and sources of funding (donors current and next year).

Research projects as cost centers

An MIS will attribute as many costs as possible to the cost center "research project." In defining a project a balance has to be struck between the desire for detailed information and what is practical. The cost of collecting detailed data at the project level is often prohibitive, and resistance of researchers to collecting and providing the data may well make the effort counterproductive. Thus, budgeting and financial analysis for a large number of projects and experiments is usually not practical. For example, the Nepal agricultural research budget management system recommends to define a project as wherever separate budgets and budget-versus-expenditure reports for a collection of activities are required (NARC 1996).

General cost centers

The choice of general cost centers should be based on the information requirements of the institute and can be any unit for which cost information is desired. General cost centers may be specified for institute overheads such as administration, they may be related to a nonresearch function such as training, or cover the cost of a specific unit or service (e.g., the greenhouse or the library). For budgeting, analytical, and reporting purposes it may be decided to present the information on the different cost centers; alternatively, costs may be reallocated back to some or all of the research projects.

The process of allocating and reallocating costs to cost centers consists of the following four steps:

1. **Allocate all operational and researcher costs to projects.** This normally includes all costs of inputs for which information is available (through vouchers or claims), the researcher cost of participation in projects (based on time allocation information), and direct labor.
2. **Allocate operational and staff costs for which no data is available to general cost centers.** This will include researcher staff time not used on projects, support staff for which no time allocation data is available, indirect labor used for general purposes, etc. Operational costs in this category are those inputs used for general institute purposes.
3. **Allocate indirect costs to the appropriate general cost centers.** The amount of effort that should go into dealing with indirect costs depends on the size of indirect costs compared to the direct costs. For example, when low, all electricity costs may be allocated to "utilities;" when high, it may be worthwhile to allocate the electricity to cost centers such as the greenhouse (from where a fair share may be allocated to those projects that use the greenhouse).
4. **Reallocate indirect costs to projects.** As mentioned under the previous point, the need for this varies and depends to a large extent on the purpose. If a project pro-

posal is developed for a greenhouse-based project it is important that all relevant costs are presented in the budget.

5.1.4 Including financial information in an MIS

An MIS can cover cost issues in four different ways:

1. **Operational costs only.** This enables a basic project information system to track operational expenditures against budgets. It covers only those inputs that are directly used in the project (e.g., fertilizer and casual labor) but misses many other expenditures. Often an MIS is limited in its coverage to these operational costs because they require cash expenditures by the organization, unlike salaries, which are usually paid by headquarters.
2. **Staff costs only.** The allocation of staff time to specific research projects provides important information on how an institute's most valuable resource is used. The information can be obtained in different ways depending on the detail required. ISNAR's INFORM system uses estimates of the number of weeks per year spent on research, but weekly or daily timesheets can be used if more information is required.
3. **Staff and operational costs.** Combining options (1) and (2) provides a complete view of the costs of research and other program activities undertaken at a research institute.
4. **Total-cost approaches.** These include options (1) and (2) as well as the indirect costs to provide a complete picture of an institute's finances. This is the most comprehensive approach and fully covers the research organization's finances. It is also the most complex and requires considerable financial expertise.

5.1.5 Using an MIS

Financial management consists of a variety of tasks, including planning, financial strategies (resource mobilization), budgeting, financial execution, financial monitoring and control, evaluation, and accountability (Bruneau 1994). A variety of tools and software are available to accomplish most of these tasks. An MIS is not normally the best tool, as it is primarily a research management instrument. However, an MIS is very useful in the following three important activities: preparing budgets, monitoring project implementation and expenditures, and analyzing resource allocation.

Preparing budgets

Implementing a project budgeting system is a complex task that requires considerable management skills to deal with the different institutional levels and their different time horizons. A NARO also needs to coordinate with parent institutes such as the ministry of agriculture, the planning department, or the ministry of finance. Special project budgets may need to be developed for international donor organizations. The MIS should therefore be able to extract data on all projects supported by a donor. At the national level, a NARO normally receives an overall budget and needs to submit drafts of the budget at given times of the year.

Ideally, a NARO has a neatly structured hierarchy of activities at the levels of station, program, project, and experiment, but in reality the portfolio of activities is often rather disorganized. This complicates a meaningful allocation of research resources to specific activities and makes the budgeting process a highly time-consuming matter. An MIS stimulates the NARO to resolve the logic of this structure, but such resolution has to occur prior to implementing the MIS.

Most countries have an annual planning/budgeting cycle to prepare and submit new activities, review ongoing activities, and reconcile activities with expected resources. Sometimes funding is available for a longer-term period, particularly in the case of foreign-donor assisted projects. Integrating the different time cycles may be difficult. In most countries the agricultural year does not coincide with the financial year, which adds to the complexity. In addition, financial requirements for most projects are highly time-bound; if funding is not available at the right time to purchase the necessary inputs, the project will need to be postponed or abandoned. Finally, budgeting for agricultural research is complicated because information has to be obtained on activities in different administrative and agroecological regions.

These complexities explain to a large extent why project budgeting and accounting are complicated processes. Many NAROs find it difficult to institutionalize a structured process, and many planning units spend a disproportional amount of time on this activity. An MIS that includes project-based budgets can be an important instrument to prepare drafts of the budget, which can be shared with the relevant organizations.

Monitoring

There are two types of monitoring: (1) comparing budgets with expenditures and (2) monitoring project implementation progress versus expenditures. The first is essentially a financial management function and the second is a research management task requiring milestones or indicators against which to measure project progress. In practice, however, it may be difficult to neatly separate the two. An example of financial monitoring is provided in figure 5.3.

Agreed-upon indicators or milestones can be used to monitor project implementation. These may present percentage completion or value judgments, such as on-schedule, delayed, or even abandoned. This information on project progress can be related to the actual and estimated use of funds.

Resource allocation reports

In preparing budgets and monitoring expenditures the MIS is used as an operational tool. The financial information contained in an MIS can also be used very profitably for more analytical purposes. Relational databases help present resource allocation information in many different ways. Each project or experiment represents one record in the database, which provides descriptive information on discipline, commodity, station, region, and type of research, as well as financial information. If this information is available for all projects, it can be aggregated to produce summary budget and expenditure information by program, department, commodity, region, or by any other descriptor in the system. This information can be used to generate

PBS Report 2: Experiment Expenditure against Budget & Allocations to Date							
Budget year: ____________ Station: ____							
Program ID	Project ID	Exp ID	Budget (for the year)	Total allocation to date	Total expenditure to date	Balance of budget unspent	Balance of allocations unspent to date
Total this project*							
Total this program							
Total this station							

* Totals for project, program, and station included only where these links and/or the base data are available.

Figure 5.3: Example of a report for project budget and expenditure

standard reports, for example on the allocation of resources by commodity or by program, which may be used in annual reports, program reviews, and evaluation meetings. In many cases this will generate information previously unavailable, for example the cost of maize research nationally across all institutes. The MIS can also answer ad hoc questions such as the cost of research in specific sectors, e.g., dryland areas or biotechnology.

References

Bruneau, H. 1994. Managing financial resources. Principles and recommendations for agricultural research managers. ISNAR Discussion Paper No. 94-13. The Hague: International Service for National Agricultural Research.

Gijsbers, G. 1991. Methods and procedures for the development of INFORM. INFORM Guidelines, Part 2. The Hague: International Service for National Agricultural Research.

Gijsbers, G. and R. Vernon. 1996. Report on the first ISNAR mission to assist in the development of a project monitoring & evaluation system for the Agroenterprise & Technology Systems Project of the Nepal Agricultural Research Council. Unpublished. The Hague: International Service for National Agricultural Research.

NARC. 1996. NARC budget management system: Procedures manual. Kathmandu: National Agricultural Research Council.

5.2 Personnel information for research management

by Ted Brush

Introduction

Human resource management is a key factor in the performance of research programs. Planning for recruitment and training, assigning tasks to staff, and monitoring and evaluating the performance of staff are especially important responsibilities of research managers. These are crucial processes in order to develop and implement management strategies for achieving the goals of a NARS. The objective of this section is to help NARO managers understand how personnel information may be used to support their management processes. We focus on needs of research managers in information about researchers since they are key staff in the execution of research activities.

To examine information about researchers, we describe what, why, and how information is useful to research managers and when and where managers might use the information. First, we consider uses of personnel information in general in NAROs in order to distinguish information that is particularly relevant for research managers. Next, we discuss the collection, processing, validation, storage, retrieval, and reporting of this information. Then, examples of two NARS that used personnel information to develop and implement management strategies are introduced. Finally, we indicate further reading on personnel information for research management.

5.2.1 What personnel information is relevant for research managers and why

Most NAROs are public-sector organizations in which personnel administration is governed by civil service regulations; commonly, these regulations are not particular to scientific organizations. Personnel management is usually relegated to administrative offices. However, many of the requirements for personnel information in these organizations have more to do with general administrative practices in the public sector than with managing research staff in particular. Administration focuses on the implementation of staff regulations, for example, to make sure staff are paid the authorized amount on time, to correctly allocate staff benefits (insurance, leave, retirement, etc.) and allowances (housing, transportation, etc.), to report staff information required by civil authorities, etc. While there is overlap, some information about personnel that is required for administration is not useful to research managers and vice versa.

Managers are interested in the performance of their research programs, so they want to know if the programs' inputs and outputs are progressing as planned. With respect to inputs, managers need information about the qualifications of researchers to be able to recruit researchers with appropriate skills and to plan necessary training. They also need access to information about the amount of time that researchers commit to the programs to better schedule research activities in their units, to monitor implementation, and to better justify requirements for additional staff time if necessary.

Regarding program outputs, managers need information on the status of outputs that researchers are expected to produce (e.g., publications and other products of research activities) and routine reports such as interim, annual, or experiment completion reports. This helps them better direct staff to achieve those outputs and better evaluate researchers' achievements and to reward performance where appropriate (Brush and Kramer 1997). Performance evaluation itself also generates information in the form of written statements and/or performance scores produced by supervisors. Information emanating from personnel evaluations may be important for various personnel actions such as promotion, recognition, and reward for performance. Information from staff evaluations can be sensitive, particularly if it effects compensation, and must be handled appropriately.

What specific information about personnel might enhance managers' capacities for planning, monitoring, and evaluation? Data on the academic degrees held by researchers, on recent participation in training courses, and on the employment experience of researchers provides important information about qualifications. Information on the disciplines, specialties, and dates of degrees, courses, and experiences is crucial for ascertaining the skills and capacities of potential and existing researchers. (See table 2.4, page 37, for examples of personnel information in one MIS.)

Disciplines are comprised of specialties; for example, the discipline socio-economics includes specialties such as sociology, economics, anthropology, and so forth. Setting up a taxonomy of disciplines and specialties can help research managers understand more about the domains of knowledge in their organizations. Since NAROs are knowledge organizations, this understanding can contribute to strategies for managing these organizations. Data on disciplines and specialties enable managers to analyze and aggregate information in order to make recruitment decisions and assess needs for training of researchers.

Dates of employment reflect the scope of experience and the potential mastery of skills. Information on most recent dates of training can indicate when upgrading of training may be necessary.

Information on researchers' time allocation is crucial for managers to develop and implement strategies to manage research activities in their units. Figure 5.4 is an example of time allocated by researchers in one laboratory. Data can be for the time researchers plan to spend and/or the time they have in fact spent in various activities. The former, called ex ante data, is useful for planning inputs to programs and it allows managers to seek additional time for activities that might be understaffed. The latter, known as ex post data, is useful for monitoring and evaluating program inputs; it allows managers, for example, to concentrate their efforts on evaluating activities that have consumed most of researchers' time. Whether time data should specify hours, days, weeks, or other increments may be an issue. For the purposes of planning, monitoring, and evaluating research activities, data on weeks of researchers' time are sufficient for most managers.

It is helpful to request researchers to account for all of their time; that is, to indicate their time for research as well as for other job-related activities. This enables managers to understand the full use of time by researchers and, for example, attempt to real-

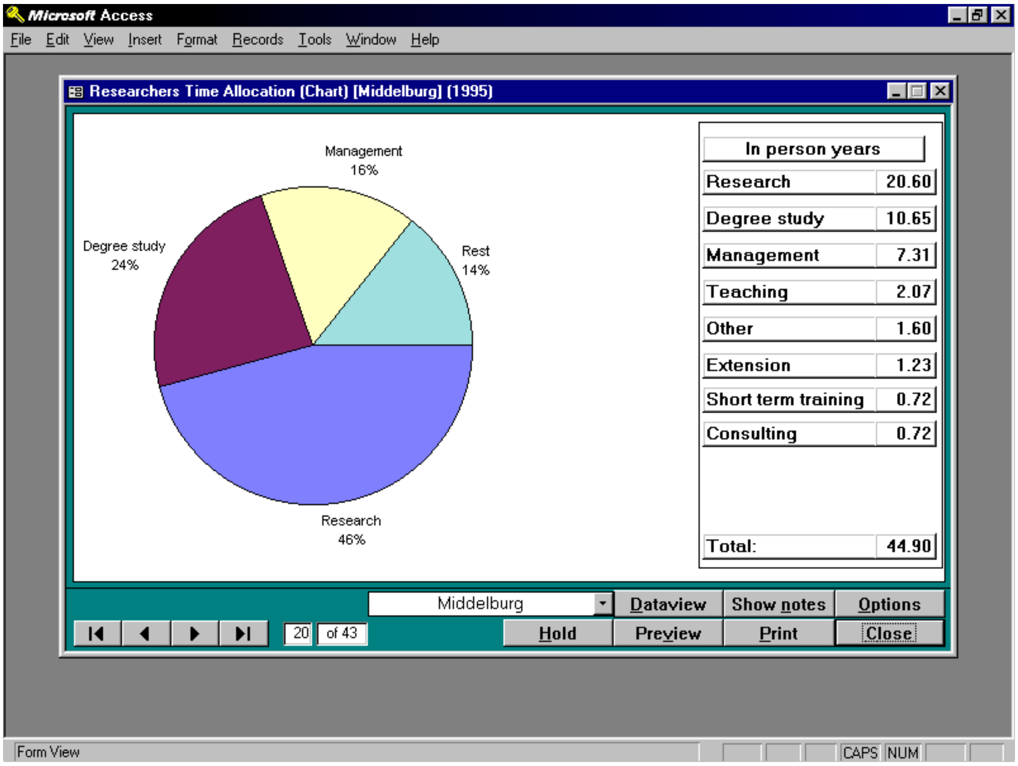


Figure 5.4: Example of an MIS output on researchers' time allocation

locate time where necessary. Data on time allocated to specific experiments or projects within separate research programs is useful for planning and evaluation. Creating a classification of programs, projects, and experiments can be a valuable learning experience for managers. Time in other activities should also be divided among various categories, such as time for administrative duties, training, and extension. Again, creating a standard list of such activities can help managers understand more about their organizations. We referred earlier to researchers' outputs; here too, building a taxonomy of outputs helps build understanding of the organization.

5.2.2 Developing and using personnel information for research management

Where in a research organization should the responsibility reside for collecting, processing, storing, and reporting personnel information for research management? There are several possibilities: one is on the administrative side and another is on the program (technical) side of the organization. As mentioned above, personnel matters are customarily an administrative function, and personnel information is typically managed by the office of administration. Since it already manages personnel information, it is logical that this office would also handle the information described here.

However, research managers, who are the primary users of the information discussed here, may not be normal clients of the administrative office responsible for personnel information. For example, the administration officer may not issue regular reports about personnel to managers of research programs. If the information is not received by such users, the information system will be ineffective. Faced with a gap between the administrative and program sides of research organizations, developers of some MIS systems have tried to locate these on the program side. Although this may bring the information closer to primary users in research management, programs may lack the staff and the ability to manage personnel information. Moreover, such duplication of information across two systems—one for administration and one for research management staff—is not only inefficient but, as described elsewhere in this book, leads to data inconsistencies, which in turn provoke further problems.

One solution might be found in a single system where information users (e.g., program managers) participate in the oversight of the information function of the NARO. This might be achieved by establishing an "information committee" to help guide the information function and bridge gaps that may exist between users of personnel information and those responsible for collecting and reporting on this information. In any case the information requirements of users must be analyzed and they need to be consulted before the system is implemented. It may be that two systems (e.g., administrative and program databases) are the best solution. However, steps must be taken to ensure data can be exchanged between the systems, and mutual updating procedures need to be established in order to make the systems always consistent. And source persons such as researchers need to be reassured that they will be asked only once for the same information needed for the systems.

The usefulness of some personnel information is affected by age. While some data are permanent (e.g., dates of birth), other data loses relevance over time (e.g., researchers' allocations to various activities and projects). Time stability of data affects requirements for information management; data with transient values need to be collected and reported punctually. Many research organizations have found that time-allocation data must be collected and reported on an annual basis at a minimum. Others collect time-allocation data more frequently, for example on a weekly basis, but care is needed not to impose data-capture tasks on users any more than is really needed. Not only is this wasteful, but it also risks alienating users whom the system will depend on for much of its essential information.

It is not uncommon for researchers to disagree with the need for data on their use of time and to resist the time-keeping attempts of their organization. Some maintain that due to the nature of their jobs, time cannot be allocated reliably in advance through ex ante data collection; others claim that collection of time-allocation data, whether ex ante or ex post, conflicts with their need to be creative in their jobs. Setting up a time-keeping system, even on an annual basis, must be done with sensitivity to researchers' attitudes. Resistance may be eased by positively conveying the benefits of this data to researchers. Enabling researchers to participate in developing and implementing management strategies might be the best way to convey these benefits. It can also help if responsible supervisors counsel their researchers about time management, perhaps, in conjunction with verifying the time allocations of researchers.

5.2.3 Examples of personnel information for research management

In 1993, a research institute in West Africa with 150 researchers (110 national and 40 expatriate scientists) in 12 centers undertook to improve the management of its human resources. As part of this effort, the director general asked the chief of human resources to set up an information system that could reveal how the researchers were using their time. The chief developed a system for the institute based on ISNAR's INFORM model. In collaboration with one research center, the chief first completed a prototype system that specified nine types of activities for time allocation by scientists: research in the institute, consulting outside the institute, studying, teaching, extension, production (e.g., seed multiplication), conference attendance, administration, and leave. After developing a successful prototype, the chief extended the system to include all centers. Several important results ensued for the development and implementation of management strategies in the institute. For example, the amount of time allocated to extension activities, 6%, was far below the institute's target of 20%; the result was to stimulate researchers to allocate more time to extension. The information system also showed that some researchers allocated time to more than 10 projects annually. This information enabled the institute to develop a management strategy to help researchers focus their project commitments thereby enhancing opportunities for individual and program achievement.

In another case, a research institute in Eastern Africa developed a five-year plan for staff training in order to improve its capacity to achieve its goals. As there was a freeze on new positions, a key objective of the plan was to upgrade the qualifications of researchers; the target was 30% with a Doctorate degree, 60% with a Master's, and 10% at the Bachelor's level. The institute used a variety of personnel information to develop the plan. Dates of birth were used to identify researchers retiring during the planning period and to develop a profile of the retirees, including information on their qualifications, disciplines, specialties, and research programs. This allowed the institute to also develop a profile of the work force after the departure of the retirees. Combining these profiles with information about the priorities of its research programs, the institute was able to plan degree training, in particular disciplines to upgrade qualifications of staff in priority programs and to help fill gaps resulting from staff retirements. Information from previous staff evaluations was available to help select staff for training. This example illustrates the advantage of combining personnel and program data in order to implement management strategies.

References / Recommended reading

Brush, E.G. and C.A. Kramer. 1997. Staff performance assessment and reward in international agricultural research centers. ISNAR Research Report No. 12. The Hague: International Service for National Agricultural Research.

Describes the processes (goals, responsibilities, and criteria) used by the international centers of the Consultative Group on International Agricultural Research to assess and reward the performance of their staff, analyzes and presents examples of protocols for assessment and for reward, and draws lessons for the design and development of these processes for human resource management.

Peterson, S., C. Kinyeki, J. Mutai, and C. Ndungu. 1997. Computerizing personnel information systems: Lessons from Kenya. *International Journal of Public Administration*, 20(10): 1865-1889.

Case study of implementing a personnel information system in the Ministry of Agriculture, Livestock Development and Marketing; presents a framework for implementation that uses support from top management (so-called "saints") to meet challenges from opponents ("devils") and encourage the input of technical experts ("wizards").

Cascio, W.F. 1992. Managing human resources: Productivity, quality of work life, profits (International Edition). Singapore: McGraw-Hill.

A text on personnel management describing information requirements for all aspects of the personnel function whether in private or public organizations. Follows the earlier but still useful text by the same author: W.F. Cascio and E.M. Awad. 1981. Human resources management: An information systems approach. Reston, Virginia: Reston Publishing.

The Website of the International Association for Human Resource Information Management (www.ihrim.com) has links to professionals for discussion and advice on personnel information systems and to a variety of commercial software products for these systems.

5.3 Physical resources

Earlier we included physical resources as an input into agricultural research (see section 2.1.2 "Key components in an agricultural research MIS"). But the management of physical resources is usually delegated to a number of separate cost-center managers such as the station manager (buildings, field management), transport officer (vehicles), and laboratory scientists and technicians (laboratory equipment). Each of these traditionally has operated their own paper-based record system. In recent years we have seen the computerization of some of these systems, still usually as separate systems, although some integrated systems are also being developed. Table 5.1 shows physical resource categories from one system along with indicative annual depreciation rates. Table 5.2 illustrates typical fields for a simple physical resource item.

Table 5.1: Examples of Categories of Physical Resources in an Integrated System and Their Depreciation Rates

Resource category	**Annual depreciation rate***
Field equipment	0.25
Laboratory equipment	0.10
Office equipment	0.10
Computer hardware and software	0.33
Vehicles	0.25
Workshop equipment	0.20
Buildings	0.02

* Rates taken from developed countries where economic factors may prompt earlier replacement than in developing countries

Table 5.2: Examples of Data Fields for Physical Resources Objects

Field	Field Description
Institute ID	A unique number by which an institute is identified.
Project ID	A unique number by which a project is identified.
Class ID	A unique number by which a category of asset is identified. *
Asset ID	A unique number by which an asset is identified.
Asset name	The name of the asset.
Person/office responsible	
Model/model number	
Serial number	Usually a unique number provided by the manufacturer.
Date allocated	The date from which the asset has been allocated to that project.
Start value	The monetary value of the asset at the time when it was purchased or provided to the institute.
Procurement source	The origin of the funds that purchased the asset or the donor of the asset.
Source	Name of supplier
Condition	1= good; 2= fair; 3 = inoperative; 4 = scrap
Comments	These are remarks about the asset as provided by the in charge of the asset or the storekeeper.

* see table 5.1 above

From a perspective of an MIS for research program management we will look at two categories that have a more direct relevance: information technology equipment and the research station.

5.3.1 Information technology

As the amount of hardware (computers, printers, modems, networks, etc.) and software becomes larger in a NARO, it will become increasingly worthwhile to establish a system to record their details. There is merit in starting with a simple system, restricting it to one that delivers only what is needed. If good standards of system development are followed, then expansion to meet new needs should not be difficult. Such a system will be useful in ordering spares for computers and printers, implementing agreed upgrades of software versions, ensuring compliance with software licenses, checking hard disk capacity for new software, and reducing the risk of theft and improving the recovery of stolen items.

In larger organizations the term "configuration management" is now applied for managing information technology resources, and may include staff with IT skills. Because hardware and software resources will play a crucial role in developing and operating an MIS, a configuration management system (CMS) may be useful in an MIS in larger NAROs. Configuration management applies also to the project of information system development and maintenance and may include version control of systems, document management, and change management.

5.3.2 Research station

The "research station" is one category of physical object that is of more direct interest to research program management. An MIS could usefully include research-relevant data on research stations such as local mean rainfall, which AEZ they fall in, etc. Annex 2 "Database Tables and Fields in INFORM-R" includes a database table for research stations that shows suggested fields.

Chapter

6 Scientific Information

This chapter consist of four papers that consider the management of scientific information. The first three will be of interest to all scientists and those responsible for managing information: librarians and other information specialists. The chapter ends on the specialist topic of geographical information systems and an assessment of what they have to offer us in agricultural research.

6.1 Library and documentation systems

by Mónica Allmand, Peter Ballantyne, and Margaret Ngwira[1]

6.1.1 Managing information for effective research

In general, information in research organizations is intended for two purposes: first, to improve the management of research and the organizations that do the research, and second, to add value to the actual research and experimentation that research organizations are mandated to deliver. The first purpose is the main focus of this book. This section focuses on the second: the information and the knowledge that is the product of research worldwide.

The complexity of the role of research managers cannot be underestimated. They manage people, funds, equipment, vehicles, etc. The most pervasive resource they manage, however, is information: management information as well as scientific information, the lifeblood of the research process. Quality decision making and quality research cannot exist in an information vacuum. Stewart (1997) trumpets this in the opening of his challenging book Intellectual capital: the new wealth of organizations:

"Information and knowledge are the thermonuclear competitive weapons of our time. Knowledge is more valuable and more powerful than natural resources, big factories or fat bankrolls. Success comes to the companies that have the best information or wield it most effectively—not necessarily the companies with the most muscle."

1. Mónica Allmand is Head, Information and Documentation Service, ISNAR. Peter Ballantyne is Manager, Knowledge Sharing, International Institute for Communication and Development (IICD). Margaret Ngwira is Senior Librarian, Agriculture and Natural Resources, University of Namibia.

This explains why today's managers generally spend a great deal of time on information issues. They are eager to find ways of managing this organizational asset in the most effective and efficient way. They do not only look for ways of diminishing the information costs, but they also try to maximize the total information value of the output of an institutions' performance. Research managers and scientific researchers have different information needs, which can be summarized as follows:

Information on management issues is often not easily accessible to organizations in developing countries. This information is generally generated by profit-oriented companies, and the cost of accessing it is high, even if produced in an academic environment. Journals (paper or electronic), books, and abstracting services are necessary tools to keep managers up to date on management procedures. Some information is available free of charge via the Internet, but in many cases it is in the form of citations, and the full text has to be obtained elsewhere. Recent information about new motivational theories, human resource management, new technologies for management applications, information on policy and economic issues, and the impact of globalization are some types of information that managers may need.

Researchers are interested in scientific information to plan and carry out their research. Past results are essential for future experiments; the work of other researchers gives new insights; and information on the problems and challenges facing clients is necessary to orient research to their needs. Currently the options open to scientists to access scientific information are expanding, with the information services (including libraries and documentation services) playing an important role in the process. Researchers and those who add value to incoming and outgoing information together form the modern knowledge organization.

In this section we focus on the content of the research. We look at information as an input to the scientific process itself—its use in research, in problem solving, and in policy advice. We look at systems and mechanisms to manage this information and ways in which researchers and research managers can maximize the benefits to be gained from investments in "scientific information." We also look at ways of maximizing the impact of a research organization by using new media to increase the exposure of the institutions' research results. Thus, we provide managers with some basic approaches that may help them develop management information policies that fit their own organizations.

6.1.2 Information management: A key concern of researchers and managers

Gaining access to information has long been a priority for scientists. Modern notions of science are based on a long-evolving process by which research builds on what is already known, and in which each research project adds incrementally to the global sum of all knowledge. A system of peer review assesses proposals for new research and ensures that they satisfy methodological standards, including reference to prior work by others. Locally, research managers normally add a "relevance" criterion, asking for evidence that new research addresses national, sectoral, or other priorities and satisfies the needs of target groups.

Before any research can be planned, the first step is the process of literature review. Contrary to the thinking of some scientists, Baldensperger (1993), drawing on his experience with the International Foundation for Science (Sweden), suggests that reviewers of research proposals for potential funding may consider the literature review as even more vital than the research concept or design. Without an adequate literature review, the scientist is sailing in uncharted waters.

The result: "good" researchers and "good" research proposals need to show how the current knowledge base contributes to new research activities. They also need to demonstrate how local needs will be addressed. One of the key challenges for researchers is to find and gain access to the global knowledge base on the problem they are addressing. They also require effective information systems about local problems and priorities, as well as any information on what is already known, in the local context. Thus, Levey (1996) writes: "a good research proposal demands a thorough review of current journal and scholarly literature [and a] plan to include funding to support information services in your proposal."

The key task for the research manager is to build up and support a mix of information mechanisms suited to the needs of the research staff and tailored to their particular research interests. One of the priorities for such a research manager is to ensure that an organization's information effort is coherent, that the various elements act in a complementary manner, and that an organization's information activities are linked to cooperative information-sharing initiatives at the local and international levels.

In this context, it is valuable to refer to the work of Röling (1990), who refers to agricultural knowledge information systems (AKIS), describing them as the process of generating, transforming, integrating, storing, and retrieving knowledge. Organizations are involved in this process and they have to understand its role in each one of the processes.

6.1.3 Information management trends

In this new millennium, the information world faces an era of great changes, which influence directly the way scientific information is produced, processed by intermediaries, distributed, and accessed. Information communication and technology and especially the Internet have made a huge impact. Information services, traditionally responsible for managing this information, are passing through a process of change and have to redefine their role. NARO managers have to understand this process of change in order to facilitate the adaptation to it and make the best use of the available resources.

Until the mid-1990s, the Internet and the World Wide Web (WWW) (see also 6.1.5) were hardly known outside a few scientific communities in the USA. Researchers found the information they needed in libraries and documentation centers and through various research and information networks. They made much use of their own personal contacts—friends and colleagues—and attended meetings and seminars to update their knowledge. Electronic mail was just emerging. Information management was considered to be a specialized, dull, but necessary task best left to specialist computer managers, librarians, archivists, records managers, database designers, and engineers.

The information environment has changed profoundly in the last decade. Several sweeping changes are transforming the ways in which knowledge is generated, communicated, and absorbed.

First, we are seeing dramatic changes in the way in which organizations operate, especially those in the private sector. Partly driven by technologies, but also by issues such as globalization, organizations are adopting all kinds of knowledge-based approaches centered on strategic alliances, outsourcing, and a focus on core competencies. The effects are also being felt in the public sector. "Knowledge management," for example, is one of the driving forces behind recent reorientations in the way the World Bank is building up networks of expertise, in- and outside of the Bank. This signals the emergence of "virtual organizations" and projects in which people are not tied to any physical location, but still draw on and contribute to the organization's activities.

Second, we can see profound changes in the attitudes of managers towards information. Today, information, and information and communication technologies (ICTs), have captured the attention of influential institutions and individuals. Suddenly, information initiatives are in every donor portfolio and national and multilateral development leaders are lining up to affirm their commitment to the benefits from investments in ICTs. This is an enormous change from the situation even five years ago. Just as striking, we have begun to see researchers, managers, and scientists "appropriating" information or knowledge management into their own work, giving it a much higher profile. In this, they are transforming our previous definitions of information management and broadening the notion of an information specialist.

The third change is perhaps the most significant as it underpins the other two factors: the technologies used for information exchange and management have been transformed. The most visible change is the Internet. In 1992, some 700,000 computers world-wide were linked to the Internet, by early 1996 this number had reached 10 million, and over 300 million in 2000. Of course, these technologies are more widespread in industrialized countries than in Africa, the Caribbean, or the Pacific. Nevertheless, we are seeing a massive explosion of investment in information technologies, the Internet, and related applications in developing countries. There is a feeling that this is a technology that can, if used intelligently, help developing-country organizations become more efficient, effective, competitive, and ultimately more successful (Ballantyne 1998).

The process of change in the use and applications of information and communication technologies by information services started in the early 1980s. The first significant change was the appearance and widespread use of the personal computer, which helped a large number of information services access to the new technology. The steadily declining cost of hardware and software popularized the use of information technology. Improved communication technology (e.g., the development of modems) also made distant access to information possible. Large collections of databases were stored in services like DIALOG in the USA, which could be accessed online. CD-ROMs, with their huge storage capacity, became increasingly widespread in the late 1980s, also in the information services of developing-country NAROs. With CD-ROM drives attached to personal computers, the cost of communications could be avoided, and ease of access to information held in databases was greatly enhanced. This combination of in-house national databases and access to CD-ROM databases of wider information sources made the task of carrying out a comprehensive literature review potentially less demanding for scientists. In the last years, a new revolution has changed the work of the intermediaries of scientific information: the Internet has come to integrate tens of thousands of interlinked, independent electronic networks. Access to this information is not available in all parts of the world, but the situation is improving rapidly.

6.1.4 Managing scientific information

For scientific knowledge to be transformed into information, knowledge needs to be made explicit and communicated through hard-copy or electronic publication material. Scientists may store, retrieve, and share this information among each other, and specialized information services may add value by facilitating efficient retrieval. As titles of documents alone do not normally represent the contents sufficiently, documents are more accurately retrieved through this value-added summary. Information specialists (e.g., documentalists and librarians) therefore classify and index documents (e.g., books, articles, research reports) and extract relevant information. They give descriptors of the document, using a recognized thesaurus, and an abstract and enter this value-added record of the document into a database. They can also pass this information on to international agricultural information services such as the FAO-coordinated AGRIS or to CABI for inclusion in its CD-ROM databases. This immediately makes remotely generated research information accessible to an international audience and raises the profile of the institution.

Ballantyne (1991) conducted research on the information needs of small countries. He concluded that one of the great challenges faced by researchers in developing countries is access to relevant information and knowledge that may facilitate and validate their own research process. If little or no information is available, the researcher has to resort to experimenting, which may be a waste of time if similar research has already been done elsewhere. NARO managers can help make their organization's research more relevant, efficient, and effective by providing researchers with easy access to available information in the scientific community. This can be done by facilitating ICT and strengthening the NARO's library and documentation services. Ballantyne mentions three aspects of information flow and management:

- the demand for information
- the potential sources of information
- the mechanisms used by the research system to identify and acquire information

Information management is the process of matching the demand for information with the available or accessible supply of information. On the demand side, researchers have a variety of information needs to carry out research programs or activities. The kind of information needed depends on the researcher, the research programs, and the way the research is conducted. On the supply side, access to knowledge sources or systems depends on the type of organization, the resources available, and the potential to network with other specialists.

There are several mechanisms to assess and interpret the demand, identify and evaluate suitable knowledge sources, and obtain and disseminate the necessary information. These mechanisms may vary according to specific circumstances; the ultimate objective of establishing these mechanisms is to ensure that researchers receive the information that they need—and, equally important, that they receive it on time.

In small counties, the local sources of information may be limited. Most information may have to be obtained from external sources. Researchers may have to rely mostly on their own networking with colleagues and on the support received from other NAROs and agricultural research centers with a similar field of interest.

6.1.5 The Internet and the World Wide Web

The Internet is now widely accepted as an important source of information. It is expected to expand to cover all possible topics and all disciplines in an increasing number of languages. While it shares many features with the traditional sources of information (library catalogs, bibliographic databases, full text databases) the Internet is notoriously disorganized, which makes it is difficult to conceptualize, browse, search, filter, or reference the system. This raises a number of questions. How can the Internet be better organized for more effective retrieval? Can it be controlled and if so to what extent? How can the traditional principles of storage and retrieval be applied to the best advantage? How can the existing structures, in particular standard classification systems and other classificatory structures, be utilized? It is difficulty to find satisfactory answers, but there is evidence in the research literature and on the Internet itself that these questions are beginning to be addressed. Williams (1996) has proposed a classification to provide greatly improved access to resources on the Internet. The task of classifying would fall on the information professionals, who were traditionally charged with this work on paper-based information (Williams 1996).

Practical aspects of Internet information: Access and quality

In NAROs with Internet access, researchers need to know how to find information efficiently, as on-line time can be a precious commodity. Investment in some initial training can be cost-effective. Including relevant "links" on the institution's Website, for example to the CGIAR centers and the Food and Agriculture Organization of the United Nations, can save time. Specific information on the Internet can be found by using so-called search engines (e.g., Yahoo, Alta Vista) that scan millions of pages on the Internet in seconds, or through indexed Web pages and subjects.

Perhaps as important as being able to access the Internet for information is being able to determine how reliable the information is? It is very easy for anybody to set up a Website and post information on it. But the normal tools that information specialists use to control the quality of the information, e.g., peer review and reputation of publishing house, are difficult to apply. While Websites maintained by organizations of international repute can be expected to present mostly good-quality information, it is very difficult for a user to assess the quality of information on the vast number of Websites now available. Some guidance is becoming available however. For example, there is a group of information experts "scouting" the Internet, and for the last three years they have been evaluating information on the Internet and gave it some form of seal of approval (Scout report 1998).

Beyond the annual report: Websites and home pages

A NARO's marketing flagship is its annual report. It reports on completed research, research in progress, collaborative activities, funding, facilities, staff, and outreach, and it lists the organization's publications for the year reported. But more often than not, the production of an annual report is delayed due to financial and/or staffing constraints.

A NARO's Website can be an excellent tool in annual reporting by offering an electronic, text-only version of the annual report, which saves printing and distribution costs. It can also post the finished parts of the annual report when the hard-copy report proves difficult to complete and produce. The Website may contain similar material to the annual report but may be updated (for example monthly). The Website provides a highly visible and much-needed exposure for an organization. It should contain links to partner organizations and have e-mail links to management and other key staff that interested people may contact. Maintaining an attractive presence on the Web need not be expensive.

The information manager may be well positioned as "Web master" to coordinate the input into the NARO's Website. If the site is managed by another section, the information manager should arrange for an information center page on the Website and contribute to the institution's presence on the Web. Training in developing Web pages is becoming more easily available. Information managers should following such training as well. If an institution does not yet have the capacity to have a Website, it should try to make arrangements with another institution to do this on its behalf. ISNAR does this for a number of NAROs; some of these may have the capacity to develop their own pages, but for various reasons, e.g., inadequate local telecommunication services, they are unable to maintain a Website on the Internet.

6.1.6 A new role for information services

As noted above, communication networks have changed the way organizations work internally and also their relations to the external scientific community. Borders are rapidly expanding. The role of library and information services has changed accordingly. Fouché (1997) presents something of this change in tabular form (see table 6.1).

Organizations are increasingly evolving from traditional, paper-based institutions to ones where information is located in "cyberspace:" the virtual space of telecommunications and electronic networks and the electronic data stores that they provide access to. A growing number of organizations are developing their internal cyberspace, an in-house electronic network generally referred to as the Intranet, storing electronic data, information, and knowledge, and an external one, the global network known as the Internet. An organization's information services are part of the internal cyberspace. Increasingly they serve as a gateway to the external cyberspace, the world-wide scientific community.

Table 6.1: Comparison of the Service and Resource-Sharing Characteristics of Industrial Age and Information-Age Libraries

Traditional, industrial age libraries	Evolving information-age libraries
Tied to a physical location, the library.	Not tied to a physical location. The library is the "virtual" information network with its home page as the front door.
Focus remains on print collections housed on site.	Focus increasingly on digitalized information sources in multiple media format which exist on computerized databases and Web sites around the world, linked to the home page of the information network.
Batch processes.	Real time, concurrent and interactive processes.
Users visit the library for service.	Users access information services by PC from work, at home, and even while commuting.
Libraries share information resources through interlibrary loans.	Sharing of information resources and expertise (staff) by participating members of the network to respond to user needs world wide.
Service model: based on the notion of the librarian as an intermediary who uses information retrieval and dissemination tools to provide users with access to published information sources.	Expanded collaborative model: goes beyond linking "users" to information sources. Participating members in the network collaborate to create an interactive managed knowledge and communication platform for knowledge workers with mutual interest to accumulate, share, and apply knowledge for mutual benefit and goal attainment

In this context librarians are increasingly becoming "cybrarians." In their new role, they not only carry out the traditional tasks of searching, collecting, processing, and distributing paper-based information, but also collect electronic information and contribute in organizing and distributing it. Thus, they have to develop expertise in "navigating" the Internet, finding relevant information on it, and delivering it to the end user. For example, they will have to search Internet-based discussion lists for relevant information and play an important role in assisting researchers in finding electronic documents.

However, many information users will continue to prefer to use printed information. It is likely that for some time to come, information specialists will have to help them by printing out, processing, circulating, and archiving electronic documents. In future, documents will increasingly be filed electronically, and the researchers and managers will get used to dealing with electronic materials.[2]

Revaluing library and information services

Cost effectiveness is vital in any modern organization, and cost effectiveness increases as the number of transactions on an existing resource increases. Research managers need to get maximum value from their information service and therefore need to consider whether their information service is giving an adequate rate of re-

2. The steady improvement in hardware and software, particularly in the screens of personal computers, will make this easier.

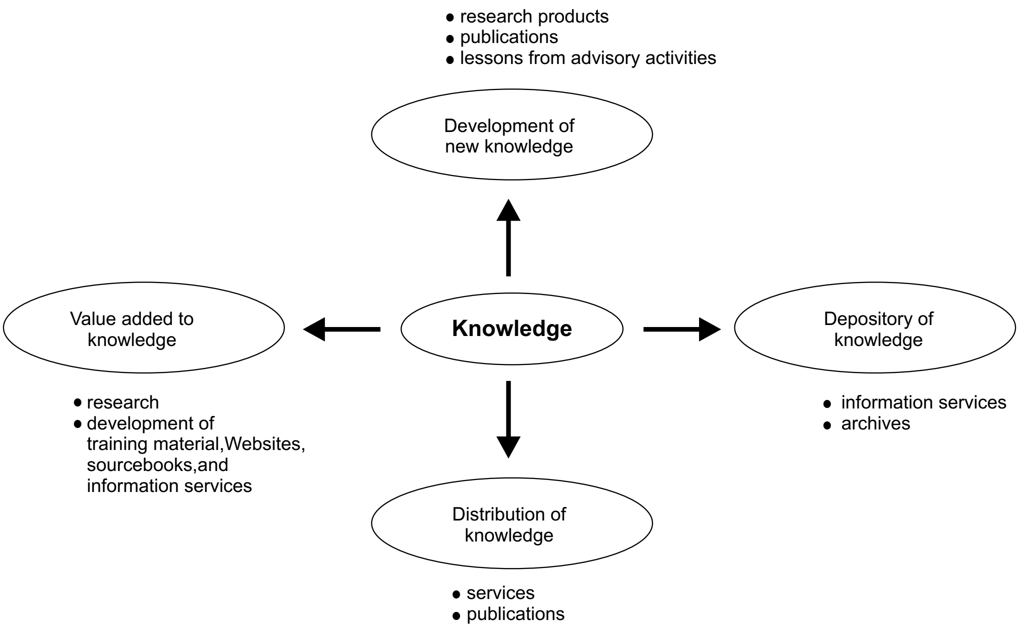
turn, and if not, why not? Is it well positioned within the organizational structure, feeding logically into the research process, or is it linked with "general services" along with plumbing and paper-clip supplies? Do key documents, which would attract scientists to the library, find their way immediately from the manager to the library or do they go to other line managers and end in the library almost by chance, if at all?

A study on the flow of information in social science and economics in eastern and Southern Africa (Hicks 1997) showed that in a survey of almost 200 scientists, the majority ranked lack of adequate access to scientific literature (particularly journals) as second only to lack of funds for conducting research. The same survey showed that in spite of the scarcity of libraries in sub-Saharan Africa they are still the most important source of information for the majority of researchers (Nsubuga in Hicks 1997). The study further showed that proximity of the library is significantly correlated with domestic and international publishing; those far away from an information service will publish less. Even on-site, managers should where possible plan to locate the library where it can be maximally utilized and not, as is sometimes seen, upstairs and out of the way of the normal traffic flows. If research quality is to be maintained and improved in these dynamic days, managers will support with enthusiasm the forces that are discussed in this chapter that are enabling their information services to deliver more for less, and to play an important and visible role in marketing the research center, its personnel, and its outputs.

The Internet and Intranet technologies are changing the concept of information management within organizations. Libraries and information services feel the impact due to the increase of information accessible via the Internet. Library managers have to dedicate part of their resources in order to take full advantage of the access to this world-wide system of distribution of information. At the same time, the speed of the generation of new knowledge within the organization requires the support of these professionals to facilitate its access through concepts, or clusters of concepts. Within the knowledge organizations, the roles of libraries and information services have to be reinvented, and one possible way is to understand their roles in the processes of adding value, distributing, and facilitating the depository of knowledge and information.

There is a fast-growing body of management literature dealing with the concept of the knowledge organization. Classic examples are *The knowledge creating organization*, by Ikujiro Nonaka and Hirotaka Takeuchi (1995), and *Intellectual capital*, by Thomas A. Stewart (1997). Speck and Spijervert (1997) present a conceptual framework that can easily be adapted and used by research institutes such as national agricultural research organizations (see figure 6.1).

The essence of this figure is that it shows how the activities of different parts of the research organization can be viewed as an integrated whole, as a complete knowledge process. All departments/divisions are essential chains in this process, where knowledge input is transformed into knowledge output by creating new knowledge or by adding value to it. It shows that knowledge enrichment is, essentially, not the work of enlightened individuals, but the outcome of an integrated organizational performance.



Note: Based on Speck and Speijervert (1997)

Figure 6.1: Four basic processes of knowledge management

In order to perform effectively, information managers must be perceived as collegial leaders by other team leaders. This means that their qualifications should be at the post graduate level. It may be necessary to consider further study opportunities to bring them to par. They should be consulted during budgetary preparation. They should undergo appraisal for their contributions to the research process as well as for their management skills. They should be mandated to develop or have developed information products relevant to national problems, e.g., local databases, web pages, annotated bibliographies on topics such as key crops, land degradation, and marketing opportunities. They may be called upon to participate in research teams in terms of assisting with initial literature analysis and review for projects, and in managing the documents and outputs generated by the research process.

Conclusions

The introduction of ICT has helped information specialists/managers to become a gateway for researchers to the worldwide body of knowledge. But ICT also directly affects the work and the role of researchers in many different ways. Researchers have become knowledge workers, and knowledge creation has increasingly become a group or team process, to be described in organizational terms. This in turn has far-reaching consequences for the role of managers of research institutions and for the perception of managers of the work of the information services. Managers will no longer merely look at the aggregated performance of individual researchers but will have to focus on the knowledge output of the organization as a whole. Information services are not just a facility from which individual scientists obtain information. In the knowledge organization, they form an integrated part of the total value-adding process of knowledge creation and enrichment-the core business of a research organization.

References

Allmand, M. 1998. El papel de la información en la estrategia de ISNAR. Paper presented at the V Reunión de Bibliotecas de Ciencias Agropecuarias y Veterinarias y II Reunión Latinoamericana de Bibliotecas de Ciencias Agropecuarias y Veterianrias. Buenos Aires, 14-15 Abril 1998.

Ballantyne, P. 1991. Managing the flow of scientific information for agricultural research in small countries. The Hague: International Service for National Agricultural Research.

Ballantyne, P. 1998. Communities and Internetworking in a development context. Background note for the ECDPM/CTA/IAALD Seminar Cooperative Internet Strategies for a Developing World: A State of the Art. Freising, Germany, 1 June 1998.

Baldensperger, J. 1993. How to draft a grant application for a research programme. Stockholm: International Foundation for Science.

Fouché, B. 1997. Knowledge networks: A new paradigm for promoting access to information access and exchange. Paper presented at IUCN Workshop: Environmental Documentation and Resource Center Network, November 1997,Windhoek, Namibia.

Hicks, E.K, N. Nsubuga, B. Tidjani, D. Rahamato. 1997. The flow of information: Social and economic science in sub-Saharan Africa. Research report 1997-02. Groningen, Netherlands: University of Groningen.

Lancaster, F.W. 1978. Towards paperless information systems. New York: Academic Press.

Levey, L. 1996. Know your needs: Including e-mail and information support in grants and projects. Washingon DC: AAS.

Montviloff, V. 1990. National information policies: A book on the formulation, approval, implementation and operation of a national policy in information. Paris: United Nations Educational, Scientific, and Cultural Organization.

Nonaka, I. and H. Takeuchi. 1995. The knowledge-creating company. How Japanese companies create the dynamics of innovation. Oxford: Oxford University Press.

Roling, N. 1990. The agricultural research-technology transfer interface: A knowledge systems perspective. In Making the link: Agricultural research and technology transfer in developing countries. By D. Kaimowitz. Boulder: Westview Press.

Scout Report. 1998. http:/wwwscout.cs.wisc.edu/scout/report (visited 17 Nov 1997).

Speck, R. van der and A. Spijervert. 1997. Kennismanagement: Intelligent omgaan met kennis. Kenniscentrum CIBIT.

Stewart, T.A. 1997. Intellectual capital: The new wealth of organizations. New York: Doubleday.

Williams, N.J. 1996. Knowledge structures in the Internet. Faculty of information studies. Toronto: University of Toronto.

6.2 Information networks

Networking in the context of agricultural research is an agreed set of activities by which a group of research scientists with a common interest regularly exchange information on their research programs. The benefit is mutual and lies in the shared information and the low costs of such an exchange mechanism. Networking scientists expect that the shared information will contribute to more effective research through, for example, avoiding unintended duplication. Cost lie mainly in the time and other resources that researchers commit to the exchange activities. For the network to be successful the value gained from the shared information must be judged to justify the cost of contributing to it. Networking can overcome some of the problems of isolation faced by scientists on small stations in remote locations.

A network usually has a focus on a particular theme: a commodity, group of related commodities, or a scientific discipline, such as beans or agroforestry. The exchanges may include both ex ante information on research plans and proposals and ex post information on research results. Formerly, only written or printed material was exchanged, but members of a network now increasingly use electronic communication, notably e-mail. Table 6.2 gives some examples of networks supporting agricultural research.

Table 6.2: Some Examples of Agricultural Research Networks

Network	Core interest	Main media	Contact
ILEIA: Information Center for Low External Input Agriculture	Low external input agriculture, indigenous farmer technologies.	Newsletter, theme workshops, encouragement to local documentation centers.	Ann Waters-Bayer, P.O. Box 54, 3830 AB Leusden, Netherlands
AgREN: Agricultural Administration (Research and Extension) Network	Organization and management issues in agricultural technology development and research-extension links.	Biannual newsletter, discussion and network papers, register of members.	Overseas Development Institute. Portland House, Stag Place, London, SW1E 5DP, UK
ATNESA: Animal Traction Network for Eastern and Southern Africa	Development and promotion of animal traction in the region.		Dept. of Agricultural Engineering, Sokoine University of Agriculture, P.O. Box 2003, Morogoro, Tanzania
ASFRE: Association for Farming Systems Research-Extension	Promotion of the development and dissemination of results and methods of participatory, on-farm systems that merge research and extension.	Biannual newsletter.	Dr Namil Ranaweera, Min of Agriculture, Lands & Forestry, Sampathpaya, Battaramulla, Sri Lanka E-mail: minagr@slt.lk

Nelson and Farrington (1994) give a good overview of the topic. They list a number of preconditions for successful networking. The most commonly cited requirement is the existence of a common issue or topic amongst collaborators. There needs to be a "realistic strategy for working towards solutions." Members must have the "capacity to contribute resources, time or information" and use their individual or institution's competitive advantage to complement those of other members. Skill development should be an outcome, either as enhancement of members' knowledge or skills just from being a member, or through organized training such as workshops. Balances between structure and flexible management styles as well as between potentially unequal partners, need to be cultivated. Incentives that motivate members' regular communication and sharing in management will help solve the inevitable problems that arise in the development of a network.

Coche et al (1997), proposing a new network to serve librarians, specify guiding principles:

- participation means partnership in technical decisions, in adapting the network to changing technology and evolving needs, and in sharing the wealth of data gathered
- duplication of the work done for existing national or regional networks should be avoided, but such existing networks should be complemented in the specialism of the new network

Coche et al (1997) also describes the following objectives of networks:

- stimulate the collection and organization of information relevant to the network subject
- improve access to and dissemination of existing information resources
- harmonize the tools and methodologies for sharing this information throughout the region
- train and share the expertise of professional staff by means of courses, workshops, meetings, bulletins
- ensure that the results and findings of research and development activities are incorporated into international as well as regional information systems
- ensure that repositories of relevant publications are available in the subregions
- produce bibliographic outputs which will benefit users in the whole region

Starkey (1997) provides a detailed description of the development and operation of a global network based on the experience of the International Forum for Rural Transport and Development. Further aspects of the management of networks are given by Fouché (1997).

Networks differ from the traditional provision of information in their interactive or two-way communication. In the last 20 years or so networks have successively adopted new electronic information and communication technologies as these have emerged (see also the previous section).

References

Coche, A. and J. Collins, in collaboration with A. Ziéhi. 1997. Support aquaculture development in Africa: Aquatic farming systems information network. CIFA (Committee for Inland Fisheries for Africa). Occasional Paper No. 22. Accra: FAO.

Fouché, B. 1997. Knowledge networks: A new paradigm for promoting access to information access and exchange. Paper presented at IUCN Workshop: Environmental Documentation and Resource Center Network, November 1997,Windhoek, Namibia.

Nelson, J. and J. Farrington. 1994. Information Exchange Networking for Agricultural Development. Wageningen, Netherlands: CTA.

Starkey, P. 1997. Networking for development. London: International Forum for Rural Transport and Development.

6.3 Information services to agricultural research management at the Technical Centre for Agricultural and Rural Cooperation

by Thiendou Niang[3]

The Technical Centre for Agricultural and Rural Cooperation (CTA) was established in 1983 under the Lomé Convention between the African, Caribbean and Pacific (ACP) States, and the European Union Member States. The center's goals are to contribute to poverty alleviation, food security, and sustainable resources management. It does this through two principal objectives: (1) to improve access to information on agricultural development and rural development, through promoting contact and exchange of information among partners and providing information on demand, and (2) to strengthen the capacity of ACP countries to manage information and communication. This includes strengthening facilities to produce, acquire, exchange, and utilize information and supporting strategies development.

These objectives are translated into activities such as books and monographs, seminars, and technical meetings as well as study visits. The activities also include information policy analysis and partnership development, and documentation services such as the Selective Dissemination of Information (SDI) service.

6.3.1 Information use environment

The agricultural research institutions in ACP countries are facing a number of constraints, the biggest of which is the budget issue. National public funding to research activities has declined and international donors are facing reduced budgets. Yet, ACP country institutions depend for 85% on external support. The NARS of the ACP countries are composed of small NAROs. The typical NARO is about 15 years old, largely publicly funded (90%), and is doing applied research. In West and Central Africa, a typical research institution has 500 members of staff, of which 85 are qualified researchers. In the Indian Ocean countries (Comores, Djibouti, Madagascar,

3. Head, Information and Capacity Development Department, Technical Centre for Agricultural and Rural Cooperation (CTA), The Netherlands

Mauritius, and Seychelles) a typical research institution has 200 members of staff, of which 50 are qualified researchers. In 1997, in the member countries of the Association for Strengthening Agricultural Research in Eastern and Central Africa, the 10 NAROs employed a total of 1994 scientists, with only 1296 with qualifications of MSc and higher, (cf. 1200 in Malaysia, 2000 in Indonesia, over 40,000 in each of China and India) (Mrema 1998).

In spite of limited human capital in the ACP countries, there is competition among scientists, which discourages sharing of technological information. At the same time, the financial allocation for information and documentation services ranges between 0% and 5% of research institutions' budgets. The ACP research institutions therefore have difficulty in meeting the costs of access to information. They have not yet established appropriate information and communication policies and strategies. As a result, there is low scientific productivity; the average researcher publishes only one or two articles every four years.

The main sources of agricultural information for researchers are their institutional libraries, but many libraries are impoverished. The other sources of information are the agricultural research centers of the CGIAR operating in the ACP countries, informal networks of colleagues, and FAO and CTA. Researchers attend international meetings and conferences, but external travel has become prohibitively expensive. The lack of adequate knowledge of information on sources and procedures for acquisition and the inadequate information services are the major constraints to accessing information. The inadequacy of internal information infrastructure reflects the low level of funds allocated to information activities and to training of staff.

6.3.2 The Selective Dissemination of Information program

CTA's Selective Dissemination of Information (SDI) program is a tailor-made service that provides specific information to a researcher or a group of researchers. The program's objective is to improve research capability by allowing researchers to stay abreast of the latest scientific and technical information. The purposes of the SDI are as follows:

- improve access, use, and adoption of appropriate agricultural technologies
- improve priority setting and program formulation
- enhance knowledge of funding sources and mechanisms
- strengthen national and regional research networking
- improve integration of scientists into the worldwide agricultural research community
- enhance capacity of the NARS to manage research programs and to promote institutional reform process for higher performance level of research organizations,
- reduce duplication of research activities and programs through sharing of experiences and lessons.

Origin

A seminar on scientific and technical information for agricultural and rural development held in France in 1984 recommended that CTA offer selective distribution of information to research organizations. A pilot project was launched in 1988 to provide research institutes with easier access to scientific and technical information and results of works carried out by national, regional, and international bodies. The pilot served 71 research programs covering Benin, Burundi, Cameroon, Ethiopia, and Madagascar.

The evaluation of the pilot project (1988–90) revealed that the SDI services provided essential information for linkages and created an atmosphere of liveliness and scientific stimulation in research. On that basis, CTA decided in 1990 to extend the service to all ACP countries. Recognizing that there were at that time more than 6000 agricultural scientists in the ACP countries who were potential beneficiaries of the SDI service, CTA developed a strategy that targets the service to senior scientists with the expectation that a multiplier effect at the institutional level would enable others to benefit through sharing the outputs. The distribution of the SDI during the first phase was based on a quota system to ensure that all countries got the opportunity to benefit from the service. Each country received at least four "personal" profiles, which are for the specific subject interests of individual users, in addition to the "standard" profiles, which are of wider subject coverage and designed to serve multiple users.

Operation

Every two months, beneficiaries of the service are sent bibliographic records that include abstracts (up to a maximum of 150 records per personal profile per year and 200 records per standard profile per year). Profiles are run on multiple databases (AGRIS, AGRITROP, and CAB-Abstracts) to ensure optimum recall in terms of timeliness and subject, geographic, and linguistic coverage. In the case of personal profiles, for subjects insufficiently covered by the above databases, searches include other appropriate databases, e.g., AGRICOLA, Chemical Abstracts, BIOSIS/, PASCAL, and FSTA. Priority is given to the beneficiary's language preference, with the following range: Dutch, English, French, Portuguese, Russian, and Spanish.

Additionally each user has a quota of five "document units" per year: a unit may be an article copied from a journal, report, or book, or from an item of grey literature. Records retrieved by each SDI run are de-duplicated to ensure that, as far as possible, only one record is supplied for any given bibliographic item within that batch of output. All outgoing communications are despatched by airmail or e-mail.

Starting in 1995, CTA took steps to maximize the effectiveness of the service. The supply of profiles is now more focused on research priority programs than on individuals, and the activity is being more closely monitored to ensure that information provided is relevant and that it reaches the intended target groups. The total number of profiles at the end of 1997 was 1142 profiles (141 personal and 1001 standard). The notable change from previous years is an increase in the number of standard profiles and a reduction in the number of personal profiles.

Based on previous requests, the core subjects are on agricultural extension, agricultural research management, agroindustrial by-products such as feeds, draught animals, farming systems, groundnuts, integrated pest management, maize breeding, oil palm, on-farm trials, post harvest, poultry, small ruminant, soil fertility, soil, and water conservation.

Governance and structure

The SDI is managed at four levels:

1. The national and regional research institutions identify the priority information themes and promote the use of the service within their institutions to reach as may beneficiaries as possible.
2. The service providers (CABI and CIRAD) have been selected on the basis of their experience with tropical agriculture. CTA assesses their information resources (database, subject coverage, volume, quality, annual update, publications, library services, including networks) the information technology infrastructure, staff qualifications and experience, especially in service delivery in developing counties.
3. The users and the partner organizations.
4. CTA's role is to define policies and strategies, and to coordinate service delivery. It is also CTA's role to coordinate monitoring, evaluation, and impact assessment activities.

Monitoring and evaluating the service

To assess the information priority themes for the SDI services, CTA implemented a strategy of continued interaction and dialogue with national research institutions, subregional organizations, the CGIAR, and its European partners. This is further informed by the specific recommendations of the SDI external program and management review as well as the conclusions of the evaluation forum of the SDI. CTA implemented a continuous users' needs survey to identify through diagnostic questionnaires the specific information requirements of scientists. The value of a careful users' needs analysis is to reduce the risk of providing inappropriate service and to reduce cost. The specific areas of concerns are priority information themes, technology and media, and partnership and networks.

As part of the monitoring process, all recipients of the SDI services were sent a questionnaire. The purpose of the exercise was to update the existing list of beneficiaries and to analyze the relevance of the profiles they received and the use made of the information. Recipients were encouraged to cite any research area in which they found the information useful. This exercise also helped eliminate irrelevant and out-of-date profiles and update and adjust the service according to user needs.

In 1994, CTA commissioned an evaluation of the SDI to determine the extent to which it was being used, its relevance to user needs, and its impact on target groups. The evaluation was also intended to identify ways of reducing costs, without affecting the scale and quality of SDI services and to make recommendations on the future operation of the services. The evaluation team concluded that the products and services provided were of high quality (multidatabase searches, provision of abstracts, and full-text documents on request). They recommended that the service be targeted to research programs and that research managers at national level should play key roles in determining the priority subject coverage, target group, and duration of the service.

Communication, promotion, and training

Since its inception, CTA has publicized the SDI service by producing and distributing a special brochure and by publishing articles in *Spore*, CTA's two-monthly bulletin. In addition, CTA staff visit a number of research institutions and attend major strategic meetings on agricultural research in ACP countries.

In the context of its initiative to progressively decentralize the SDI program through closer linkages with regional and national institutions, CTA has organized courses to develop SDI management capacities in ACP countries. The courses have focused on the means of organizing, establishing (including needs analysis and defining profiles) database and CD-ROM retrieval techniques, managing an SDI service, and promoting techniques and networking. The first course, organized in collaboration with the West Africa Rice Development Association in 1995, brought together 16 participants working in agricultural research institutions. The second one, organized in collaboration with the International Center for Research in Agroforestry and the International Livestock Research Institute in 1996, was attended by 15 information professionals. Organized in 1997 in cooperation with the Pacific Regional Agricultural Program, the third course was attended by 12 professionals.

Initial contacts were made with regional institutions (the Caribbean Agricultural Research and Development Institute and the South Pacific Commission) and national programs in Chad and Ethiopia to look into the possibility of decentralizing some of the SDI activities. The main aim of the decentralization is to assist national centers to develop their own capacities in managing SDI services and to ensure continuity of services.

Experiences and lessons learned

The experience of working with ACP agricultural research and natural resource management institutions over the past 10 years has taught CTA some interesting lessons.

The first is that information alone will not solve the agricultural research constraints in ACP countries. At the same time, these problems cannot be solved without appropriate access and use of information in research programs and institutions' management. This requires institutional sensitivity to information and communication activities, which should be translated into increased funding: 5–12% of the institution's annual budget is suggested. There is no alternative approach to long-term sustainability.

The second lesson is that the successful management of an SDI program depends primarily on core management and governance capacity. This includes definition of strategic objectives, priority setting, program planning and implementation, as well as program monitoring and evaluation. This also implies effective communication with key stakeholders (researchers and research managers, service providers, and partner organizations) with diverse demands and expectations.

The third lesson is that the management of an SDI program involves the development of partnerships and alliances worldwide. Central to this issue is networking and resource sharing at national and regional levels. There is for example a need for strong linkages, using information and communications technologies, in order to be able to download upstream research and to adapt research methodologies and techniques. These linkages should involve the CGIAR centers, especially those established in ACP countries, universities, and advanced research institutes, as well as the nontraditional research partners, including farmers' associations, NGOs, and the private sector. There is also a need to link SDI with research activities of policy networks such as Réseau d'Etude et d'Analyse des Politiques Agricoles and the Eastern and Central Africa Program for Agricultural Policy to address broader and long-term policy issues that will affect agricultural and rural economies. These issues include input policies and technology adoption, economic viability of technology packages, subregional and international trade and natural resource management in a changing economy.

Finally, during the development of the SDI program, CTA analyzed copyright, licensing, and other legal and contractual issues relating to SDI, with particular regard to duplication of electronic distribution. It was concluded that unfair use of data represents a threat to the databases producers' sustainability. To guard against this there are restrictions on database users imposed by the standard license agreement, to the effect that "the product itself may not be given away or sold; abstracts from the products may not be sold, published or distributed to users on any basis beyond the one site in the institution to which the product was originally supplied." Another issue was that output from an on-line search is intellectual property owned by the information provider and protected by copyright law. This has been addressed by licensing schemes which have been developed to enable copying and redistribution to multiple customers; information brokers are permitted to act as intermediaries in the supply of original records and as agents in the declaration of redistribution requested by a single customer.

References

Mrema, G.C. 1998. Current and planned activities of ASARECA. Presentation at the 18th SPAAR plenary session. 26th February 1998. Entebbe: Association for Strengthening Agricultural Research in Eastern and Central Africa.

6.4 Geographic information systems

by Jeffrey W. White, Adriana Rodriguez-Aguilar, David P. Hodson, A. Dewi Hartkamp[4]

Introduction

A geographic information system (GIS) is a computer-based system for storing, manipulating, and presenting data in a spatial (geographic) context. A GIS is much more than a computerized drafting system for producing paper maps. Spatial data can be analyzed with diverse methods, and outputs may be presented in formats ranging from summary statistics or tables to animation of maps varying with time. Although use of GIS shows explosive growth in such differing fields as marketing of consumer goods, tax collection, disease control, and forest management, the analytic capabilities of GIS hold special promise for agricultural research. Agriculture is one of the most spatially intensive activities of humankind, and spatial variation in production conditions, whether driven by climatic, edaphic, biotic, or socioeconomic factors, is of paramount concern in agricultural research. Furthermore, increasing concerns over the impact of agriculture on natural resources, both on-farm and in adjacent regions, dictate that agricultural research focuses even more on spatially varying processes.

This chapter first presents an overview of basic concepts of GIS, then examines applications of GIS to agricultural research and, finally, considers why GIS has not yet realized expectations in agricultural applications.

6.4.1 An overview of GIS

For a full introduction to GIS, readers should consult excellent texts such as Aronoff (1995) and Burrough and McDonnell (1998). This brief review will only attempt to highlight some concepts and issues that have particular relevance to agricultural scientists faced with a decision on whether to apply GIS in their research.

Data sources

Data may be entered into a GIS from sources including digitizing tablets, scanners, and satellite imagery. Historically, digitizing has been a major activity of many GIS laboratories, but this situation is changing rapidly as data become available through the Internet or in other digital forms. Examples of spatial data sets readily available for some regions include those derived from climate, soil, land cover, and elevation (table 6.3).

Use of the Geo-Positioning System (GPS) merits special attention since this system offers the prospect of adding a spatial component to research at a very low cost. Hand-held GPS units process signals from a system of navigation satellites maintained by the US government and calculate the user's position with an accuracy of approximately 100 m (Herring 1996). Larger, more costly units can provide accuracy at submeter levels. GPS units are now widely used by germplasm collectors, pa-

4. Natural Resources Group, International Maize and Wheat Improvement Center

Table 6.3: Examples of Data Sets of Spatial Factors Relevant to Agriculture

Variables	Region	Grid cell size or map scale	Source	Comments
Max/min temperature, precipitation, potential evapo-transpiration	Central America & Mexico	2.5 km	J. Corbett, personal communic. 1997	Available as monthly data and for various types of seasons (e.g., coolest three months, driest three months).
	South America	9 km	J. Corbett, personal communic. 1997	Available as monthly data and for various types of seasons (e.g., coolest three months, driest three months).
	Africa	5 km	Corbett (1995)	Available as monthly data and for various types of seasons (e.g., coolest three months, driest three months).
	Asia	4 km	IIMI (1997)	Global 16km IWMI (2000) available as monthly data.
Elevation	Global	1 km	USGS (1997a)	
Soil	Global	1:5 million	FAO (1995)	Includes pH, depth, water holding capacity, but small scale is a major limitation.
Land cover	Global	1 km	USGS (1997b)	Derived from remote sensing (very approximate).

thologists, soil scientists, and others, and one can anticipate that use of GPS will be as routine as taking photographs.

Data from satellite-based sensors or other forms of remote sensing are frequently cited as promising sources for data on land use, soil degradation, and climatic or weather conditions. The potential of these methods is illustrated by the effort of the United States of America Geological Service (USGS) to create a global land cover database, working primarily from weather satellite data (USGS 1997b), which provides data on over 200 types of vegetation formations or other land covers using an approximately 1 km-square grid cell size. Nonetheless, many projects have found that handling primary data from remote sensing is extremely difficult.

Spatial data are stored either in a vector or raster format. In the vector format, data are recorded as coordinates of points (e.g., locations of a town or site of a germplasm collection), lines (roads, rivers, or drainage paths) or polygons (boundaries of a field or a body of water) along with associated attributes. The spatial entities represented thus correspond roughly to objects that would be found on a map. In the raster, or gridded, format, a region is subdivided into regularly spaced cells in a rectangular array. Data values are assigned to specific row and column positions within the array.

Consideration of which formats to use is an important early step in developing a GIS project. The vector format has advantages in terms of storage efficiency and the ease with which map scales and projections may be adjusted (table 6.4). It is particularly suited for data that is inherently discontinuous, such as locations of political boundaries, roads, rivers, and other linear features. The raster format is better suited for handling quantitative data that will vary continuously over a region. In agricultural work, where researchers frequently wish to examine variation in quantitative data such as precipitation, soil pH, and slope, this format is especially useful. Fortunately, most GIS systems now allow both vector and raster data to be manipulated together in a single project.

Selecting an appropriate map scale often requires balancing research objectives that require a high resolution (large scale) against constraints of data availability, storage and processing. Table 6.5 provides some guidelines on how different map scales relate to research objectives.

Since the Earth is a spheroid, any attempt to present data on a flat surface requires that the spatial data be transformed. This transformation process basically involves trade-offs between conserving relative areas and preserving the shapes of regions. Depending on the expected final product or use, different projections such as Uni-

Table 6.4: Comparison of Vector and Raster Data Formats

Characteristic	Vector	Raster
Efficiency of data storage	High	Moderate
Handling of point and line data	Excellent	Poor
Handling of quantitative data showing continuous variation	Fair	Excellent
Handling of quantitative data with discontinuous variation	Good	Excellent
Ease of overlay operations	Low	High
Ease of representations of topological relations	High	Medium
Ease of re-projection	High	Low
Ease of re-scaling	High	Low
Ease of use with digital imagery	Low	High

Table 6.5: Comparison of Different Map Scales

Map scale	Arcs	Equivalent grid cell size	Useful for studies within
1:5,000,000	1.5°	25 km	continent
1:1,000,000	20′	5 km	country
1:250,000	4′	1 km	region within country
1:50,000	1′	200 m	province/state level
1:10,000	10′′	40 m	watershed/municipality
1:500	0.5′′	2 m	single field

° = degrees; ′ = minutes; ′′ = seconds

versal Transverse Mercator (UTM) and Robinson will be required. For initial data processing, many groups prefer to manage data using latitudes and longitudes, rather than projecting them. Unfortunately, many data are supplied with incomplete or incorrect information on map projections.

Analytic tools of GIS

Aronoff (1995) classifies analytic functions of GIS as including data retrieval, overlaying, neighborhood analyses, and connectivity functions. The full classification is listed in table 6.6 with examples relevant to agriculture. Simple retrieval of data, classification and measurements are examples of retrieval functions. Layers of spatial data may be overlaid using either arithmetic or logical relations. Neighborhood analyses include such functions as searches based on distance criteria, determining whether points or lines are within polygons, and interpolation procedures. Connectivity functions emphasize the spatial relations among features and include tools for evaluating proximity and paths along networks of roads, streams or other linear features. While this framework has clear heuristic value, few GIS projects will select these functions explicitly. Rather, they will be invoked in highly dynamic and interactive fashion as part of the research process.

Presentation of results

While conventional paper maps are still an important product of GIS, results are increasingly provided in electronic formats. This may be as simple as a series of map images stored in electronic form or, as is increasingly the case, an intermediate software and data product that allows end-users to view and manipulate the maps to suit their particular needs. The Internet now has many examples of applications that allow users to view data on maps interactively. Advantages of this approach include reduced publication costs, ease of updating, and maximizing the utility of the data products.

6.4.2 Applications to agricultural research

It is difficult to imagine activities in field-oriented agricultural research where GIS could not be applied with benefit. In research planning, GIS can assist in identifying target regions, prioritizing sites either as "hot spots" or as representative of key regions, and defining efficient sampling frameworks. During actual research activities, GIS may provide the foundation for a data collection and management system, as well as allowing careful monitoring of research progress. And during analyses, the diverse analytic tools of GIS can be exploited as well as the ability of GIS to link spatially referenced experimental results to data on climates, soils, and other factors. In the sections that follow, three examples of applications of GIS to agricultural research are provided. For further examples, the report compiled by UNEP GRID-Arendal (UNEP 1997) contains 43 case studies on a wide range of topics and covering diverse regions and spatial scales.

The Kenya Maize Database Project

In 1992, the Kenya Agricultural Research Institute (KARI) and CIMMYT established the Kenya Maize Database Project to harness GIS in assessing the successes and fu-

Table 6.6: Summary of GIS Analytic Functions and Agricultural Applications

Function	Description	Examples of agricultural application
Retrieval		
Retrieval	selecting data without modifying spatial relations	determining soil pH of a polygon or grid cell
Classification	grouping a set of features based on retrieved attributes	defining land-use suitability classes in a watershed
Measurement	measuring lengths or areas or counting features in a given class	determining the are of a region identified as suitable for zero tillage
Overlay		
Arithmetic	combining layers through arithmetic operations	creating a layer representing total annual precipitation by combining 12 monthly precipitation surfaces
Logical	combining layers through logical operations	creating a layer showing regions with soil pH > 6 and annual precipitation ≤ 1200 mm
Neighborhood		
Search	assigning a value to target features based on criteria related to a search area	identifying farms within a 10 km radius of a fertilizer distribution center
Point-in-polygon	determining whether a point is contained in a given polygon	determining which germplasm accessions were collected from sites with soils of volcanic origin
Line-in-polygon	determining whether a line passes through a given polygon	determining whether an irrigation network crosses all farms in a district
Topographic functions	defining traits such as slope and aspect	creating a surface showing variation in slope within a watershed
Thiessen polygons	defining "areas of influence" around points	in the absence of other information, interpolating data from various weather stations over a region
Interpolation	predicting values at specific locations based on known values at neighboring locations	interpolating data from various weather stations over a region but assuming certain statistical relations and including effects of elevation
Contour generation	creating contour lines from point data	creating a contour map of soil depth based on point samples
Connectivity		
Contiguity	evaluate traits of features that are connected	determine the total area of uninterrupted buffer zone (e.g., not crossed by roads) along a river
Proximity	measure the distance between features according to specific rules	locating buffer zones for nonagricultural use along a river margin
Network	analyzing traits of paths defined by interconnected lines	finding the shortest route for distributing seed samples to multiple farms
Spread	a more complex form of proximity analysis where that evaluates a function that accumulates with distance	finding a cost-effective route for distributing seed samples by accounting for road quality, terrain, and distance
Seek	performs a directed search according to specific criteria	trace the probable flow of water based on a digital elevation model
Intervisibility	identifying areas that can be seen from a specific view position	locating a food processing facility where it will not be visible from an important archaeological site
Illumination	portraying the effect of illuminating a surface from a nonvertical position	representing a surface of expected soil lost in a visually dramatic manner for policymakers
Perspective view	portraying a surface from a non-vertical viewing position	representing a surface of expected soil lost in a visually dramatic manner for policymakers

Note: Based on Aronoff (1995)

ture needs of KARI's maize research (Lynam and Hassan 1998). GIS was applied both to define an efficient sampling network to assess KARI's impact and to characterize maize systems in Kenya. Climate surfaces were created, and a cluster was conducted to define eight maize production zones (Corbett 1998). The project was successful not only in documenting the impact of KARI's maize germplasm, but also served in highlighting many opportunities to increase impact both through germplasm and crop management. In particular, it identified a dry mid-altitude to highland region where farmers lacked adequate germplasm (Hassan et al. 1998). This led to changes in maize breeding priorities and a revision of the initial climatically based zone classification.

African Country Almanac—Putting GIS technology within reach of NARS

Information systems such as GIS are often unavailable to NARS due to lack of the appropriate software/hardware, data in digital format, or technical expertise. This section describes a new research resource for several sub-Saharan Africa countries: the African Country Almanac, which overcomes these problems but still delivers powerful GIS functionality. The African Country Almanac is a joint development project between John Corbett's group at the Integrated Information Management Laboratory, Texas A&M University, and CIMMYT's Natural Resources Group. The Almanac arose from the UNIX-based Spatial Characterization Tool (SCT), developed by the Texas group (Corbett and O'Brien 1997). Unlike the SCT, however, the Almanac is a completely stand-alone, easy-to-use system that requires no specialized GIS software or expertise. The Almanac was created using MapObjects (ESRI, Inc.) and Visual Basic (Microsoft, Inc.) and is distributed on CD-ROM. Users of the Almanac only need a PC running Windows 95 or NT, and a CD drive. Almanacs are currently available for Angola, Ethiopia, Kenya, Liberia, Nepal, Sierra Leone, Tanzania, Uganda, Zambia, and Zimbabwe, and several other countries are in preparation.

The Almanac provides three main categories of analytic tools: visual display and querying, site similarity analyses, and zonal analysis. Over 50 layers of spatial data are available for visualization and query with the Almanac. These include climate, soils, land use, population, elevation, infra-structure, political boundaries, and environmental data. Datasets in the public domain are used in nearly all cases. In addition, users may add their own existing spatial data. The interpolated climate surfaces, created by Corbett and coworkers (Corbett 1995) using tri-variate thin plate splining techniques (Hutchinson 1995) are an integral part of the Almanac. These provide long-term (30-year) monthly averages for a range of climate variables (precipitation, evapotranspiration, maximum and minimum temperature) and incorporate elevation as a covariable at a resolution of 3 arc-minutes (approximately 5 x 5 km square). Various climate models have been derived from these climate surfaces including: "Optimal season": the five consecutive months with the highest precipitation to potential evapotranspiration ratio (P/PE); 'Trigger season": the longest run of consecutive months where the P/PE ratio is greater than 0.5; "Dry season": the longest run of consecutive months where the P/PE ratio is less than 0.5; and "Quarters": the three wettest, driest, warmest, and coolest months.

The Almanac allows rapid display of any of these layers of data and also the creation of customized maps (e.g., figure 6.2), which can then be exported and incorporated

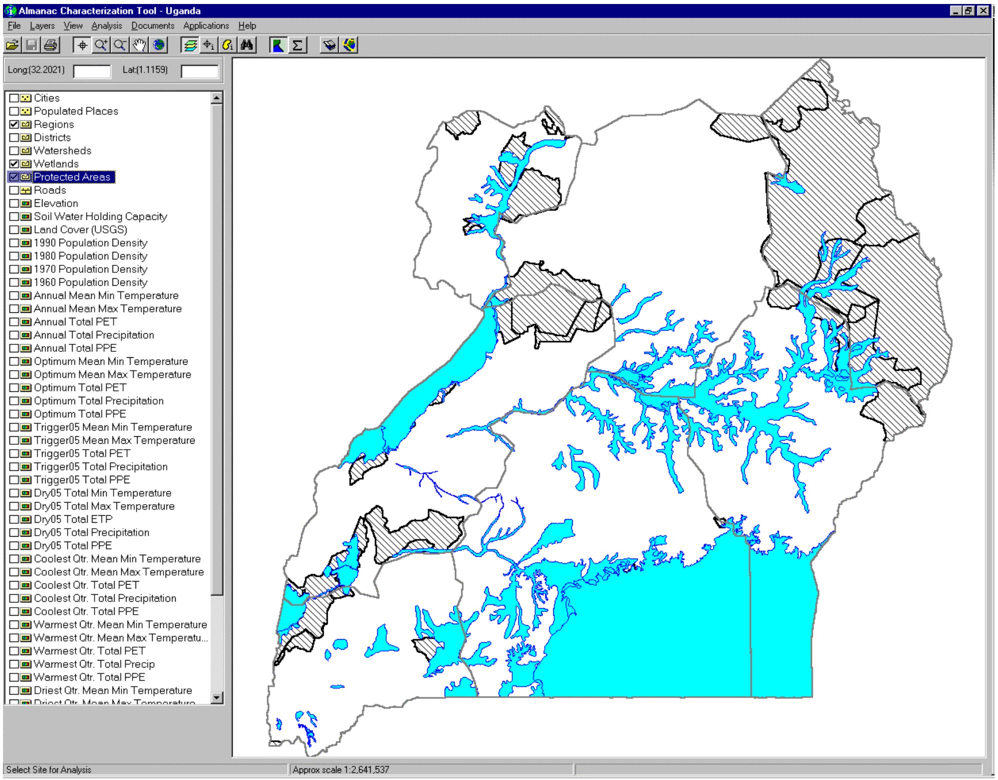


Figure 6.2: Almanac screen display illustrating protected areas, wetlands, and regional boundaries for Uganda

into other software packages, including word processing and presentation packages. Data from a specific location, either obtained by entering a latitude and longitude or simply by pointing and clicking with a mouse, can also be viewed and exported.

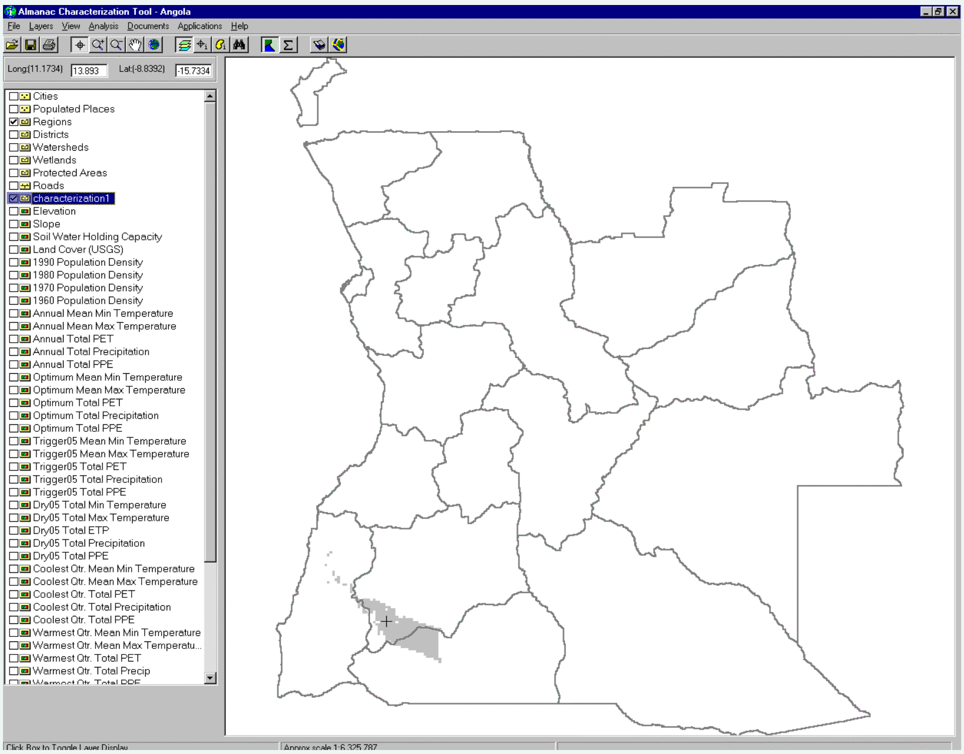
The site comparison function allows users to determine regions of similarity (e.g., in terms of climate during a particular season) to a user-specified site (see box 6.1.)

The zonal analysis function allows users to define agroecological zones by specifying upper and lower limits for climatic, edaphic, demographic, ecological, and topographic variables. Zone characterization may be as complex or simple as the user requires. Using the Almanac, it is just as easy to determine either of the following agroecological zones: (1) where in Uganda has favorable season precipitation less than 1500 mm, an evapotranspiration range of 500–1000 mm, medium or high soil water-holding capacity, landuse that is not forest, and an elevation range of 500–1500m; (2) where in Angola has an annual precipitation less than 2000mm and an elevation less than 1000m.

Box 6.1: The African Country Almanac: A Hypothetical Case Study Illustrating the Application of the Site Comparison Function

Examination of data from a CIMMYT International Maize Trial at Chiange, Angola, (searchable data contained within the Almanac documentation section) reveals that the three highest yielding varieties in this trial were Poza Rica 8022, Guaymas (1) 8022, and El Paraiso 7929. As a result of this trial, Angolan maize researchers might want to determine where else in Angola these varieties may perform well. The site comparison tool of the Almanac enables them to select the location Chiange and then generate a zone of similarity, in this case +/- 10% of values for favorable season climatic conditions [but users have total flexibility in terms of how much variance and which climate model to use]. The result of this analysis is illustrated in figure 6.3. In addition to the zone being mapped, the Almanac generates a range of statistics relating to the zone (e.g., total zone area; mean, min., max., and standard deviation for all variables) in exportable text file format. The inclusion of ArcExplorer software (ESRI, Inc.), enables users to display and query such a similarity zone in many different ways. For example, the climate similarity zone for Chiange could be displayed in terms of differing population density, which may assist researchers in determining not only where climatic conditions may favor the new varieties, but also where they may have maximum impact.

Figure 6.3: Almanac screen display illustrating the favorable season climate similarity zone (+/- 10% of climate variables) for Chiange, Angola (+), and district and regional boundaries



Additional examples of applications of the Almanac tools include the following:

- assisting in the selection of research sites that are most representative of particular regions within a country
- determining where promising new agronomic practices or germplasm may be deployed with the highest probability of adoption
- locating agroecological zones which may be suitable for germplasm adapted to certain environmental conditions
- setting breeding program priorities in cases where it is discovered that large agroecological zones exist for which there is no suitable germplasm currently available

The Almanac is being further improved to expand the number of countries covered and increase the functionality and types of data. Subsets of major databases being developed at CIMMYT—International Wheat Information System (IWIS), Sustainable Farming Systems Database (SFSD), and International Maize Information System (IMIS)—will be incorporated into the Almanac. Additional analytical functions may include enhanced querying facilities, within-polygon statistical capability, transect analysis, and a batch tool to extract data from multiple sites simultaneously.

Interfacing agronomic models to GIS

Agronomic models, particularly crop simulation models, are logical partners of GIS since their point-based predictions can be usefully extended over regions. Typically, a GIS is used to provide input data for the models, the models are then run over the appropriate regions, and then GIS is again used to display the simulation outputs. Since a large amount of data are exchanged in this process (often many megabytes), data input and output structures have to be carefully defined. The International Consortium for Agricultural Systems Applications (ICASA) has proposed standards to facilitate such work (Ritchie 1995). Two examples of software systems that use ICASA standards to link GIS and models are Spatial Analysis Tool (Thornton et al. 1997) and AEGISWIN (Engel et al. 1997).

We have used a similar tool developed at Texas A&M (J. Corbett and S. Collis, personal communication) to simulate the potential yields of rainfed maize for the state of Jalisco in Mexico (figure 6.4). Climate data from approximately 200 locations were first interpolated using the spline method of Hutchinson (1995). The interpolated data were then subject to cluster analysis to identify sets of environments with similar climates. These data were then overlaid with soil profile descriptions linked to soil data from the FAO Digital Soil Map of the World to use as inputs to the CERES Maize model. Although only yield limited by possible water deficits was simulated, future work will include effects of tillage practices and residue management both on crop yields and on runoff.

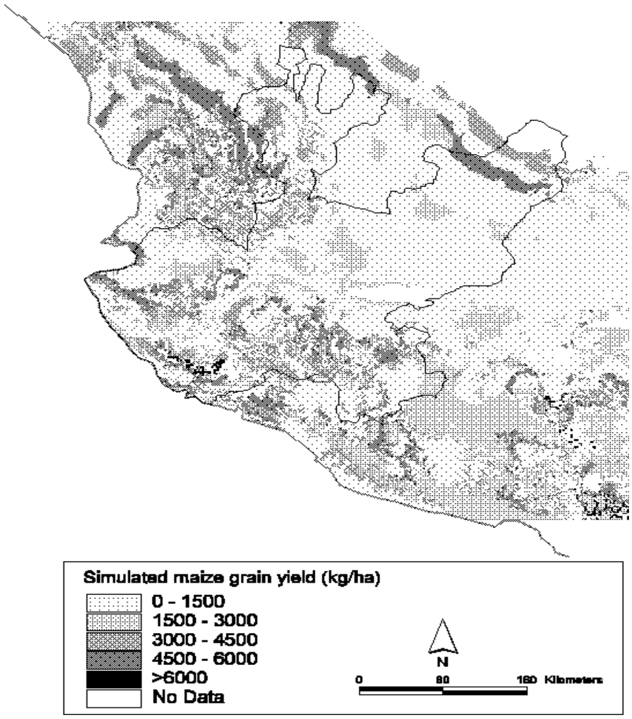


Figure 6.4: Variation in yield potential of summer rain-fed maize in Jalisco, Mexico

6.4.3 Limitations to effective use of GIS

Notwithstanding some notable successes, there is widespread sentiment that the impact of GIS in agricultural research has fallen well below expectations. We cite four major factors that influence this:

1. GIS awareness. Because GIS is a recent development, many agricultural scientists simply lack an awareness of how GIS can benefit their research activities. Experience at CIMMYT is that “awareness building workshops,” of one of two days duration, are invaluable in orienting scientists.

2. Data availability. Although the availability of agronomically useful data is increasing rapidly, there are many surprising deficits. Most notable are the lack of more detailed soil data than provided in the FAO Digital Soil Map and of data on crop distributions and production levels.

3. Hardware and software costs. Hardware costs are decreasing rapidly, especially storage devices and computers per se. However, specialized equipment such as plotters and digitizers are still too expensive for many groups. Software vendors have generally preferred to increase the functionality of GIS packages while maintaining prices at relatively high but constant levels. Nonetheless, products such as ArcExplorer and the Country Almanacs, which are distributed free of

charge, offer the possibility of minimizing the number of full GIS packages that a research project must maintain.

4. The institutional setting of GIS capabilities. Some degree of centralization makes sense for efficient management of base data and the more expensive software and hardware. However, rapidly declining costs and increased availability of data argues for increasingly placing GIS capabilities directly into the hands of researchers. The role of GIS specialists thus should evolve to one of developing new tools and key data sets and of providing support for particularly difficult GIS applications. In this case, access to specialists might include problem specific collaborations with other research institutions or outsourcing to commercial services.

References

Aronoff, S. 1995. Geographic information systems: A management perspective. Canada: WDL Publications.

Burrough, P.A. and R.A. McDonnell. 1998. Principles of geographic information systems. Oxford: Oxford University Press.

Corbett, J.D. 1995. Monthly evapotranspiration surfaces for Africa, v1.1. Climate coefficients created using ANUSPLIN (CRES, ANU, Canberra). (CD-ROM). Nairobi: International Centre for Research in Agroforestry.

Corbett, J.D. 1998. Classifying maize production zones in Kenya through multivariate cluster analysis. In Maize Technology Development and Transfer - A GIS Application for Research Planning in Kenya. Edited by R.M. Hassan. Wallingford: CAB International.

Corbett, J.D., and R.F. O'Brien. 1997. The spatial characterization tool. Texas A&M University System, Blacklands Research Center Report No. 97-03. (CD-ROM).

Engel, T., G. Hoogenboom, J.W. Jones, and P.W. Wilkens. 1997. AEGIS/WIN - A program for the application of crop simulation models across geographic areas. Agronomy Journal 89(6).

FAO. 1996. Digital soil map of the world and derived soil properties. (Version 3.5, November 1995). Rome: Food and Agriculture Organization of the United Nations.

Hassan, R.M., J.D. Corbett, and K. Njorge. 1998. Combining geo-referenced survey data with agroclimatic attributes to characterize maize production systems in Kenya. In Maize Technology Development and Transfer - A GIS Application for Research Planning in Kenya. Edited by R.M. Hassan. Wallingford: CAB International.

Hutchinson, M.F. 1995. Interpolating mean rainfall using thin plate smoothing splines. International Journal of GIS 9: 305-403.

Herring, T., 1996. The global positioning system. Scientific American. Feb.: 32-38.

Koontz, H., C. O'Donnel, H. Weihirch. 1994. Management. McGraw-Hill.

IIMI. 1997. World Water and Climate Atlas: Asia Version 1.0 CD-ROM. Colombo: International Water Management Institute.

Lynam, J. and R. Hassan. 1998. A new approach to securing sustained growth in Kenya's maize sector. In Maize Technology Development and Transfer - A GIS Application for Research Planning in Kenya. Edited by R.M. Hassan. Wallingford: CAB International.

Ritchie, J.T., 1995. International Consortium for Agricultural Systems Applications (ICASA): Establishment and purpose. Agricultural Systems 49: 329-335.

Thornton, P.K., H.W.G. Booltink, and J.J. Stoorvogel. 1997. DSSAT Version 3.1 Spatial Analysis. In DSSAT version 3. Draft volume 4. Edited by G.Y. Tsuji and G. Uehara. Honolulu, Hawaii: IBSNAT, University of Hawaii.

UNEP. 1997. GIS in agricultural research: Awareness package. Edited by N. Denisov, C. Heberlein, L. Czaran and O. Simonett. UNEP/DEIA/TR.97-9. (Also available at http://www.grida.no/prog/global/cgiar/htmls/awpack.htm)

USGS.1997a. GTOPO30 Documentation.
http://edc.www.cr.usgs.gov/landdaac/gtopo30/README.html.

USGS. 1997b. Global Land Cover Characterization.
http://edcwww.cr.usgs.gov/landdaac/glcc/glcc.html

Chapter

7 Aspects of Information Science and Technology

This chapter serves three groups of readers:

- *First, it provides the MIS practitioners at station and national level who come from an agricultural research background with an introduction to their new specialism of information management.*
- *Second, it provides those research managers who will have to approve procurement of new ICT equipment with a source of reference in layman's language.*
- *Third, it helps researchers and research managers who find the use of ICT interesting and useful to their work to find out how to get the most out of such equipment.*

7.1 Information science

7.1.1 Information

Humans are said to have gone through a hunter-gatherer age (from 30,000 BC), an agricultural age (from 3000 BC), and an industrial age (from 1800 AD), and are now said to be in an information age. People have always used information and began to process it formally when the alphabet was developed around 1500 BC and mechanical printing in the first century AD. Since the first scientific journal was written in the early 1600s, the amount of published scientific information has approximately doubled every 15 years. But the development of the modern electronic computer in 1946 is a milestone in information management, ushering in an age of large-scale automated data processing that might be regarded as the start of the information age.

Knowledge is our remembered perception of reality. We can never know the whole of reality. One reason is that there is too much of it for one person to observe. Another is that the perception process differs from one person to another: two people describe the same object or incident differently (see figure 7.1). Our view therefore is a product of that part of reality that is seen and that is our perception process. In designing a data management system we need to recognize the limitation of the first and try to avoid the variation of the second by rendering the data capture process of a system as objective as possible.

Philosophers have argued down the centuries about the nature of knowledge, particularly on the respective importance of deductive reasoning based on principles, and perception using the senses, as the source of knowledge. As researchers today we rely heavily on perception in observations, but in experiments we also deploy deduction.

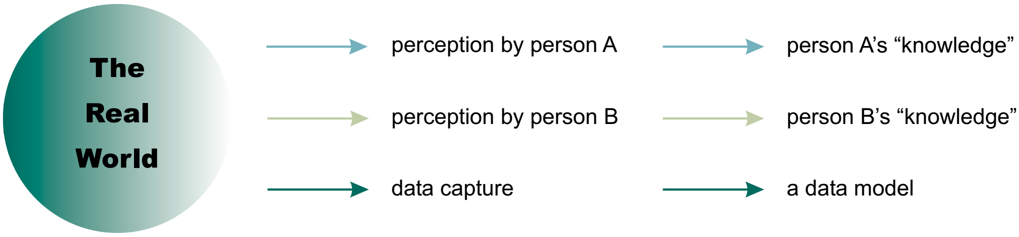


Figure 7.1: The real world and our views of it

This book is concerned with information systems and hence with that part of information that is formally processed. This has to be recognized as only a portion of information: societies carry a large pool of information that is not captured within a formal information system but that nonetheless plays a crucial role in the affairs of people. The same is true of an organization such as a NARO, which holds much knowledge in the minds, experiences, and skills of individuals. Section 7.1.3 explores this further.

7.1.2 Data, information, knowledge, and wisdom

Data is a representation of small components of the real world. It is often collected and stored in large quantities and is in itself usually of little value. However, when passed through processing that attributes meaning, it is transformed into information. For example, data that indicates the positions of the two hands of a clock can be combined with train timetable data to provide the information that we have missed the last train of the day.[1] Information is needed to make decisions—in the cited example the information can be used to decide to start walking or take a taxi.

While the boundaries between information, knowledge, and wisdom are indistinct, each represents further processing and an increase in value from the previous state. A significant proportion of data is often numerical and this proportion declines as one moves to information, knowledge, and wisdom. There are similar declines along this sequence in ease of computer processing and in quantifying. Menou (1993) defines knowledge as "information that has been meaningfully aggregated into a reservoir of facts and concepts that can be applied." Wisdom has been described as "distilled knowledge derived from experience of life, as well as from the natural and social sciences and from ethics and philosophy" (European Commission 1997).

The value in these three higher levels remains just potential until the information, knowledge, or wisdom is applied, typically in influencing decision making. This is a crucial question in deploying information systems: will managers make better decisions as a result? We will see shortly how we can apply criteria of quality to judge value, but this is in itself of little worth unless and until the outputs of an information system are actively used.

1. Modified from Checkland (1990)

Information systems have long had much to offer in the management of data and information. More recently they have progressed further in attempts to offer knowledge through the use of artificial intelligence, expert, and decision support systems.

7.1.3 Informal information

In this book we focus on information that can be stored and processed in an organized manner. Such information is stored either on paper or, more usually nowadays, in electronic form. In either case it can be measured in terms of storage capacities: pages or bytes[2]. But we need to recognize that there is a large body of information that may not be captured in any such structured system but that may have an influence on management decisions. Much of this may be legitimate in the sense that although it is not captured and processed formally, because for example it is too transient or too varied, it may still be to the benefit of the NARO and its goals. Indeed, after exploring the relative importance of computer-mediated information versus conventional noncomputer-mediated information, Lan and Scott (1996) conclude that managers today still rely on their personal knowledge for organizational decision making more than they do on any other information media, including computer-mediated information. They provide an interesting assessment of the relative importance of information types (see table 7.1), which gives personal knowledge nearly twice the importance of computer-mediated knowledge. But the latter does seem to be increasing: they quote a 1983 study in which only 3.4% of information flow in an organization was found to be computer-mediated, compared with 22% in 1996 (see table).

This personal or informal information is the "tacit" knowledge of Nonaka et al. (1996), who distinguish it from "explicit" knowledge, which has been externalized and is hence able to be committed to a storage system. They identify four business processes that facilitate movement between these two types of knowledge in ways that create new knowledge for the organization:

Table 7.1: Relative Importance of Different Sources of Information

Information source	**Use (in %)**
Personal knowledge	40
Computer mediated	22
Interpersonal contacts	17
Upper-management formal documentation	19
Other	2
Total	**100**

Modified from Lan and Scott (1996)

2. A byte is equivalent to one alphabetic character.

- **Socialization** encompasses the informal exchange of experiences and ideas between people. The organization can facilitate this by providing appropriate opportunities for such informal contacts. It sees the movement from individual to group tacit knowledge. Information technology plays little role at present in this process, but developments in video conferencing may change that.
- **Externalization** is the process of distilling explicit knowledge from tacit to yield firm concepts. "Group-ware" such as Group Decision Support Systems[3] can assist here.
- Through the **combination** process, new and existing explicit knowledge are used to develop new prototypes or products, and information technology has much to offer in managing such explicit knowledge with databases providing a repository and "outline processor" systems supporting the development of new combinations.
- The cycle is completed through **internalization**, through which individuals take explicit knowledge into their own internal tacit knowledge base. Organizational management is greatly influenced by information flows and we have much to learn in this area. Nonaka et al. (1996) provide a framework for further understanding.

Informal information clearly has an important role to play in an institution. Management can facilitate this role by actively encouraging collaboration and providing facilities and mechanisms for its interchange. Mankin et al. (1996) emphasize the importance of such collaboration by advocating teams—the overarching premise of their book is that "achieving high levels of organizational performance requires the integrated development of information technology, team structure, and the overall organizational context. The [institutional] development processes should be interrelated so that designs in all three areas will be mutually reinforcing."

Some informal information that may play a major role in managers' decisions may be regarded as mischievous rather than legitimate, as the information may be to the personal benefit of the individual manager, who may not wish such information to be available to others. One of the difficult roles of top management is to create an environment where the need for such information, and the opportunity for using or misusing it, is minimized, along with the more encouraging actions described above.

7.1.4 Information as a resource

Information has a remarkable characteristic that marks it out from other resources (money, people, oil, timber) in that its deployment to a particular use does not, or need not, reduce its value or content or availability for other uses: it can in principle be replicated indefinitely. When word of mouth or, later, the printed word, were the only means of replication, these media were limiting factors of such reproduction. Modern ICT provides remarkably powerful and cheap alternative means of dissemination. Insofar as information is useful, the low cost of replication and the ease of

3. Discussed further in section 1.1.5 "Information systems."

wide dissemination provide good reason for public institutions to recognize information as an important resource, a public good, and a valuable agent for development.

Unfortunately modern information technology is still restricted to a minority of the world's population and, therefore, so is access to public information. Tribe (1994) observes the following:

> "80% of the world's new knowledge is currently the preserve of less than 10% of its population. More than 900 million adults throughout the world [out of a population of 5.5 billion] are virtually excluded from the benefits of the growth of knowledge because their illiteracy and poverty combine to lock them into their present state of ignorance. And the situation is becoming steadily worse. For example, present levels of annual national investment in science and technology range from $700 per person in Japan to 22 cents in a country like Nigeria. Inequalities in the distribution of knowledge are already wider even than those in the distribution of wealth, and present policies are ensuring that they will increase further in future."

On the other hand, the impressive advances made in recent years in ICT offer some grounds for hope. It may prove possible for less developed nations to leapfrog over some of the stages through which the technologically advanced have worked their way. Low-orbit satellite communications may in future enable remote regions to have access to major information resources that are currently largely the preserve of those with good telephone line connections (see section 7.3 "The main technologies"). But today large gains in a more equitable distribution of information remain elusive.

The value of information

The value of information is difficult to measure. This is at least in part because information depends on content, accessibility, and context.

Information content is difficult to measure partly because of the wide diversity in information storage media: libraries, private files, people's minds, and other people's skills, company reports, and electronic storage devices. The choice of a unit of information is not simple. That used in storing and transmitting electronic data—the binary digit or "bit"—can be measured and is useful in this context but, like the printed word, bears little relationship to the information content of data or its value.[4] Even within electronic media there may be huge variation: a terse electronic message may approve a multimillion dollar commercial transaction in a few hundred bytes (each equivalent to one alphabetic character), whereas a picture of little value can take up many thousands of bytes.[5] Such difficulties in the measurement of value in electronic data have worked against proposals considered by the European Union to apply a tax to electronically transmitted data, the so-called "bit-tax."

Information accessibility greatly influences its value. Scientists have been known to invest considerable resource in research that, unknown to them, had largely been

4. A computer usually uses eight *bits* to represent an alphabetic character.
5. A *byte* is a group of usually eight bits, representing one alphabetic character.

done before. The problem has been lack of access to good records of earlier work, both at home and abroad. If good reports of such earlier work had been available and accessible, then, in many cases, researchers could have realized outcomes of value to farmers in a much shorter period of time. New information and communication technologies help improve accessibility but, so far, only to part of known information and only for that part of the population with good connections to the new electronic information repositories, such as those that may be reached via the Internet. Much agricultural research information previously recorded on paper, especially the "gray" literature of annual reports and local extension advice, is still not available even to those well connected via the Internet. And as described above there are many with no such connections. So information may simultaneously be of great value to those who are able to access it and worthless to those who cannot.

The context of information also influences its value. A farmer may receive information about new technologies that are not locally available or, because they were not adequately tested in on-farm research in that locality, are inappropriate. The same information in the hands of a farmer in a neighboring country with a more developed infrastructure may prove highly profitable. Directing information to locations, situations, and people where it is most useful is an important function of information systems and information practitioners.

The medium in which information is transmitted can also affect its value: a book may be of less value to the illiterate than the spoken word of an extension agent or a radio broadcast; instructions on the side of a pesticide container would be more useful in a local language than in English.

The realized value of information depends on many factors, making it difficult to measure it. Nevertheless there is wide acceptance that much information is valuable, provided it can be made available in the right place, time, and form. The acceptance of inherent value in information presents a challenge to information practitioners to demonstrate benefit from increased use of information, namely impact. Menou (1993) describes an international consultation amongst some three dozen information specialists to identify the impact of information on development. He reports that "no positive, explicit mention of a possible contribution of information to development, beyond the established cliché, could be traced in the statements justifying most information-related programs and policies in developing countries. The expected benefits are expressed in such broad and general terms that it is difficult to translate them into potential, observable results." We would make the claim here that an MIS for agricultural research helps make better decisions on resource allocation. This leads to greater efficiency in terms of more outputs of new technologies for a given investment. This claim still awaits confirmation, however. Several studies have shown high rates of return from investment in agricultural research (e.g., Cruz et al. 1982, USDA 1996). Similar studies are lacking on the return of investment in supporting MIS.

7.1.5 Other attributes of information

Four attributes of information are important particularly when designing an information system and its associated data capture processes, if the intrinsic benefits of information are to be realized:

- **Relevance.** It is common to find information that is collected but not used. This wastes effort in collection but more seriously distracts users (e.g., the managers) from the more important information they need. A useful rule is to exclude any data for which a clear use and user cannot be identified.
- **Accuracy.** There is an enormous range in the precision expected of information. If engineers from different continents are each asked for the coefficient of expansion of copper, they are likely all to give the same answer to several decimal places. Expectation of precision declines as one moves from engineer to biologist, sociologist, and management consultant. The required precision depends on the nature of the entity that is the object of the information and the purpose to which it is to be put. The level of accuracy therefore needs to be assessed for each situation; a minimal level is needed but too high a level is wasteful. The allocation of national priorities across commodities is a good example. Figures for research allocation to each commodity probably need not be more accurate, at this time, than within, say, 10% of actual in order to guide research planning. Indeed the ranking order is more important than absolute figures. In contrast, yield data of an experiment testing whether a new variety yields more than existing varieties is typically reported to three or four figures. These differences in accuracy need to be addressed in the design of information systems.
- **Currency.** The value of much information declines with time. This is particularly true of management information, so that an MIS needs relatively swift data capture, processing, and output subsystems to enable such outputs to be of maximum value. In an agricultural MIS, a maximum turnaround time of a few weeks is often a realistic target. (Scientific information, such as the output from agricultural research, retains its value for much longer; the life cycle of an insect pest may still be valid and applicable decades after its elucidation.)
- **Accessibility.** An information system must present its contents so that it is quick and easy to use by managers who may in many cases be new to computers.

Wang and Strong (1996) in analyzing these attributes further, urge that information system professionals view them more from the users' view. For example, they found that it was insufficient for the information system staff to ensure that data was accurate; users needed some additional information, for example on the source of the data, to be able to judge the accuracy for themselves. The lesson is that users need to be consulted on the application of these quality criteria.

7.1.6 Artificial intelligence and expert systems

As the processing capacity and speed of computers have dramatically increased in the past decades, attempts have been made to have computers simulate some of the mental abilities of humans. So far, success is largely restricted to narrow, clearly bounded information areas such as mathematical calculations, diagnosis of certain medical conditions such as human skin diseases, and analysis of geological data. General human intelligence and "common sense" have proved vastly more complex than at first realized, and computer replication has proved elusive. Even full-text language translation has not been attained, due in part to the extensive ambiguity of languages. People resolve such ambiguity by intuitive reference to the context and to large stores of linked facts in their own memory.

Some expert systems have successfully mapped the knowledge of experts in narrow specialized fields and by applying a few hundred or thousand of derived rules can provide useful conclusions, such as crop disease diagnosis from a set of observations fed into the system. For example, farmers in a number of countries have access to a system that advises when to spray a potato crop with fungicide against potato late blight (*Phytophthora infestans*), given a number of input variables. These systems depend on the fact that for late blight to develop, the spores require the potato leaves to be wet for a critical minimum period (effective blight hours [EBH]) and above a minimum temperature and relative humidity. In one system in the Netherlands, EBH start to accumulate after 12 hours with rain and conducive temperatures and relative humidity, or after 16 hours without rain but with conducive temperatures and relative humidity. Different parameters may apply to different potato cultivars as they exhibit different levels of resistance to late blight.

Keane (1995) describes a system that has served farmers in Ireland for some years. It is based on the input variables of hourly temperature, relative humidity, and rainfall figures, from which it calculates the number of hours with temperature not below 10°C and relative humidity of at least 90%. This data is matched against base data on the conditions known to be favorable for late blight development, and advice is given to farmers when it is appropriate to spray. The alternative for farmers is to follow a fixed interval spraying regime, but this usually results in more and unnecessary sprays.

7.2 Information systems

This section describes general features of information systems and focuses on management and related information systems applicable to the agricultural research environment.

7.2.1 Characteristics of information systems

Information systems may be entirely paper based or to a varying degree supported by computers. They all share the characteristics of any system: something goes in, is processed, and a product comes out. In the case of an information system, data is fed in and is processed in order to provide the output of information useful to decision making by the users. The processing may include validation against specified standards, verification (see section 7.5.4 "Some characteristics of data" for descriptions of both of these) and, where necessary, correction, and then various transformations and aggregations to provide the desired information in the desired form for each category of user.

In early days, computers and their software were rather inflexible and users were often forced to change their way of doing things to suit the requirements of the information technology. In the last decade, this situation has changed. First, systems became flexible enough to do what the user was accustomed to do. Then systems became powerful enough to facilitate new and more efficient and productive ways of doing business. Now, a successful information system implementation is expected to change the way a business is conducted, and clearly the change should be for the better.

In some organizations an introduction of a new information system has had even greater consequences and has actually transformed the nature of the whole business. When the American Airlines company introduced SABRE—an interactive transaction processing system for airline reservations and ticket sales and extended to hotel reservations and car hire—in the early 1960s, the company soon earned more profits from the system than from its original core business of air transport.

We do not expect an MIS to change the nature of a NARO, but we should expect it to *facilitate*, in conjunction with appropriate management initiatives, a transformation in the way it operates, with tangible gains in productivity and efficiency.

7.2.2 Some cost-benefit considerations of information systems

We are here concerned primarily with computer-based information systems. A cost-benefit analysis may support a decision on whether or not to proceed with the establishment of such a system and to judge between different systems. In practice the costs and benefits of information systems and information technology generally have proved notoriously difficult to quantify.[6] A minimum step that should be taken when considering whether to take on a computerized information system is to ask how well would we get on without it.

Costs

All information systems incur costs additional to doing things manually. Costs may be broken down into the following:

- **System acquisition:** purchasing and adapting or in-house development and implementation. The purchase cost of software may range from modest, if existing off-the-shelf software and systems are used, to high if a tailor-made system is developed specifically for one organization. Even a ready-made system has to be adapted to the organization's needs and processes, incurring staff and sometimes consultant's time.
- **The supporting ICT:** computers, printers, telecommunications systems. The procurement costs of new hardware are obvious, but applying them to a particular information system can be complicated by the sharing of existing ICT by more than one system.
- **User training:** formal and self-learning. This cost is often overlooked but can be large (Baker 1997; Hochman et al. 1994). Therefore, an MIS must be easy and intuitive to use, and allowance must be made for user training and learning in the implementation part of the project. The self-learning component is more difficult to cost than formal training.

6. For a wide-ranging discussion of the role of information technologies on the economies of people, companies and countries see The Economist, 28 September 1996. Supplement: "A survey of the world economy: The hitchhiker's guide to cybernomics."

Another complicating factor is that in addition to their main, initial cost, all three categories have an ongoing maintenance cost.

Benefits

In a major survey, Brynjolfsson and Hitt (1993) found that the returns from investment in information systems made a substantial and statistically significant contribution to company output. They found that gross return on investment (ROI) for computer capital averaged 81% for the 367 large firms in their sample. But they were puzzled by the inconsistency of other ROI studies reported in the literature. They examined separately the effects of information technology (IT) on three factors: productivity, consumer value, and business performance.

They found that the effects of IT depended heavily on which of these is measured (Hitt and Brynjolfsson 1994). They found that computers led to higher productivity and created substantial value for consumers, but that these benefits did not result in measurable improvements in business performance. The same writers identified another complicating factor in such studies that restructuring and cost-cutting were often necessary to realize the potential benefits of IS: some companies that had not yet embarked on these processes and had not shown benefits of IT investments did show such benefits later when these processes were addressed.

A further problem is that many benefits have no direct link with finance or other easily measured variables. This does not, as some claim, make them "intangible": most are clearly recognizable and valuable but not readily quantified in absolute terms.

Measuring benefits

There are a number of quantitative techniques for estimating ROI that can be applied to information systems. Details are beyond the scope of this book but some are listed here to enable the interested reader to explore them further. Figures in parentheses indicate the percentage of responding companies reporting having used them in a survey reported by Marsh (1998): the payback method (65%); net present value (53%); internal rate of return (42%); and economic value added (16%).

As with assessments of research projects (see section 2.2.4 "Appraising the research proposal"), these quantitative methods address only part of the issue and usually depend on simplifying assumptions that necessarily separate the model to some extent from reality. While they may still provide valuable insight into the potential value of a proposed information system, they should not be assumed to provide a complete picture and should not be solely responsible for a decision to proceed with system implementation.

7.3 The main technologies

Information technology comprises hardware and software; each is discussed in turn here, followed by a separate section for communications technology (telecommunications).

The field of information technology is developing so quickly that we address here only some selected issues that are likely to be relevant to persons responsible for managing information in the NARO. The rapid pace of change means also that any quantitative references are probably soon out of date; capacities and speeds are increasing, in many areas doubling every 18 months, and yet prices are falling.

Countries of the South are witnessing the same phenomenon but with a significant time lag, which applies differentially to different technologies. Transfer of communication technologies lags behind that of personal computers and software due to its dependence on large national infrastructures, usually still in the hands of government monopolies (see table 7.2). But this situation is changing: in 1995 there were few Internet host computers in African countries other than South Africa, but by January 1999, there were more than 10,000 (Computers in Africa 1999). This spread of information technology represents a rapid relative increase from a very low starting position, so that today it still affects only a small part of the population. For example in Ghana, one of the more IT-progressive African states, less than 1% of households were recorded to have a PC in the home (Ahiabenu II 1998).

Table 7.2: Some Differences between North and South in Access to Information Technology

Technology	**South**	**North**
Personal computers:		
• distribution	• some staff: mostly professionals, some support staff, few managers, often held in secure central room	• nearly all staff, on own desk
• models	• wide range of old to latest with consequent service problems	• fairly modern: old are discarded
Direct access to:		
• internal telephone lines	• most staff	• all staff
• external telephone lines	• restricted to senior staff	• all staff
Local area networks or intranets	• exceptional in smaller organizations	• common, connecting almost all staff
Access to the Internet via reliable telecommunications	• few, and only in larger towns	• most staff
Qualified information specialists	• few • hard to retain in government service	• many • not so hard to retain
Mains electricity: 24 hours / day, constant voltage	• supply not assured, technology needs to be protected from voltage variations, some stations need (standby) generators	• assured

7.3.1 Hardware

The hardware consists of the various components of the computer (keyboard, monitor, etc.), the printer, and other physical devices. The whole, together with its software (described in section 7.3.2), comprises an information system and, as is normal for a system, is equipped for inputs, processing, and outputs. There are matching hardware components for each of these functions and for data storage, as shown in table 7.3. Each is described here along with an indication (where appropriate) of typical capacities. These are summarized in terms of the features typical of a complete personal computer that one might order today, in table 7.8.

The personal computer

The personal computer (PC) is a system for receiving, storing, and processing data into more useful outputs, on a scale that is appropriate to a single person. It has been central to a revolution in information technology in the last two decades. The PC was originally designed in 1981 by the computer company IBM and manufactured by them alone until 1987, when many competitors began to make it, following in large measure the original design. Originally it used the DOS operating system[7], which was later replaced by the Windows operating system. This standard dominates the market today with many manufacturers of the hardware, nearly all using the same Windows operating system from Microsoft. Some regret the lack of competition inherent in such a market dominance for the operating system, but it has had the advantage that people with a wide range of hardware can use the same software.

The second major type of personal computer is the Apple, first made by Apple Computers in 1977. It developed the Apple Macintosh, the iMac, and more recently the Macintosh G3 and the Cube G4. Apple has maintained a monopoly over this standard including its own proprietary operating system. It is used particularly by those working on graphical applications and desktop publishing. Recent Apple models are able to run some Windows-based application software.

Table 7.3: Components of a Personal Computer Sorted by Function

Input	Storage	Processing	Output
Keyboard Mouse (input of instructions) Scanner Modem	Main memory (RAM) Secondary memory: • hard disk (fixed disk) • diskette • CD-ROM • tape backup	The central processing unit (CPU) e.g., Pentium, 486, 386	Monitor or screen Printer Modem

7. Operating systems are described in section 7.3.2 "Software."

The keyboard

The keyboard is the main input device: a person types in the data, which is "echoed" on the screen, and, on instruction from the user, "saved" to the hard disk inside the computer system box. Keyboards are also used to pass commands to the PC. Those who are to enter data into an MIS as a major part of their work are well advised to learn to touch-type, a skill that greatly increases the speed and reduces the labor of entering data. There are self-teaching books and computer programs to help with acquiring this skill.

Keyboards do not often cause problems. The most common problem is a single key failing to function, which may be due to dirt. Keyboards should be periodically cleaned. Turn it upside down and tap gently so that loose dirt is dislodged and falls out. This should be done regularly anyway in a dusty environment. A vacuum cleaner (with key board upside down) will give a better result. If a particular key is not working, and the above methods have not helped, the key cap can be pulled off to provide access for better cleaning of the key mechanism; computer stores sell a tool for removing chips, which can also be used to grasp a key cap and pull it off.

The mouse

The mouse is a device to pass instructions to the computer and its software, usually by making selections on a visual display unit (VDU), or monitor, by moving the screen cursor and clicking with the mouse button on the chosen item. The mouse is held under the hand and moved around on a flat surface; the movements are encoded into X and Y coordinates and this data is sent to the PC via the mouse cable. Any flat surface will do, but a special mouse pad with a textured surface facilitates smooth rotation of the mouse ball.

Most mice have two buttons[8], the primary and the secondary. For a right-handed person the left button is configured as the primary button. The primary button is used for selecting items on the screen, for dragging an item to a new position, and, by holding down the button while moving the cursor, to select a group of items or area of text. The secondary button, in some operating systems[9] such as OS/2 and in Windows from Windows 95 onwards, provides access to context menus; additional options and information can often be found by clicking with the secondary button on screen toolbar and menu items.

It is possible to do quite a lot in Windows and Windows-supported software without a mouse by using the shortcut keys: these are often quicker for experienced users for making selections of choices than using the mouse. Typically a shortcut key consists of the Ctrl key, or the Alt key, along with one letter key—often shown as an underlined character on the tool bar or menu item to be selected. Excessive use of the mouse can lead to wrist pain, and in this circumstance the shortcut keys may provide an alternative.

There are two common connections of a mouse cable to the PC, and they are not interchangeable except in a dual-mode mouse. The model that came out earlier uses a

8. The mouse of the Apple personal computer has only one; some other mice have three buttons.
9. Operating systems are described in section 7.3.2 "Software."

9- or 25-pin male plug to connect to one of the (usually two) PC's serial ports (Com 1 or Com 2); it usually does not matter which as the system discovers this for itself on booting up. More recent PCs have a separate mouse port. Note that this may be physically identical to the keyboard port, but the two are not interchangeable, so ensure each is in its correct designated socket. If necessary consult the manual for a diagram showing the position of each, or check for symbols marked adjacent to each.

If the cursor movement becomes erratic it may indicate dirt in the mouse mechanism. The mouse can be opened from the underside and the ball removed. The ball and its contact rollers should be inspected for any signs of dirt, which can be cleaned with soapy water or alcohol, and dried thoroughly before reassembling.

The central processing unit

As the name implies, the central processing unit (CPU) is the "brain" of the personal computer. It is the device that implements instructions from a software program or from the user via the keyboard and processes data accordingly. It accepts each line or instruction of a program, executes the instruction, and looks for the next instruction.

Central to this processing is the transistor, which is a solid state (no moving parts) electronic switch. Many transistors mounted on the same semiconductor base comprise an integrated circuit. The CPU is such an integrated circuit, with, in modern examples such as the Pentium, some five million transistors. Most PCs today have a 386, 486, or one of the several Pentium processors, all made by Intel. Two other companies, AMD and Cyrix, make CPUs for PCs, and these are fully compatible with those of Intel. MMX (multi-media extensions) Pentiums have additional components that enhance performance for all software and have specific advantages for video, audio, and graphical applications. Apple computers have used processors from Motorola. The Apple iMac and Cube models have Motorola G3 and G4 processors respectively.

The effectiveness of a CPU is determined largely by its speed, measured in megahertz (MHz, or million cycles per second). However different CPUs have other internal characteristics that influence performance, and a CPU may be faster than another at some tasks, e.g., word processing, and slower at other tasks, for example the mathematical functions of a spreadsheet, or the manipulation of graphics. Popular computing magazines periodically survey the market and compare current widely available processors. The speed at which a computer performs tasks depends largely on the CPU, although the size of main memory (see below) has an influence too. Today's CPUs operate at speeds of 33–100 MHz for a 486 processor to between 100 and 600 MHz for the various Pentium models. The Apple iMac has a 233 MHz Motorola processor and the Apple Cube G4 processor runs at 500 MHz. A recent innovation is to record a unique identity number on to each individual CPU chip in order to discourage theft and enable the origin of any unit to be identified. The Pentium III has this built in as a user-selectable option. Some groups concerned with the protection of an individual's rights for privacy have objected to this device on the grounds that when such a PC is connected to the Internet, the number of a coded processor can be detected by others and could be misused. The manufacturer's response in the face of such opposition is to ship identifiable chips with the coding facility switched off. A user can then choose to switch it on.

Storage devices

There are a number of devices used to store data and a description of the more common ones follows. Table 7.4 provides a summary and a comparison of capacities and requirements.

Main memory (RAM)

A computer's main memory stores the code of the program currently running and the data that is being processed by the CPU. To do this it must be able to operate at a speed matching that of the CPU, and it has to have all its storage sites immediately accessible by the CPU at random. For this reason it is often called random access memory (RAM). RAM is fast and therefore good for data processing, but it is also expensive and it requires electrical power to operate, which means it loses its contents when the computer is switched off. So there is secondary or backup memory (see next section) that provides long-term storage that remains in place on switching off the computer, and provides a much larger store and at a much lower unit cost.

When finishing a piece of work it needs to be "saved": this copies the current data from RAM to secondary memory, usually the hard disk. Much modern software saves automatically before closing, and some do this at intervals also: database software typically saves a record when the user moves to another record. When such automatic saving is not available, and particularly when the electricity supply is unreliable, it is wise to manually save one's work at regular intervals such as every 10 minutes, and again before switching off the PC.

The amount of main memory is important because if there is too little, the CPU is not fed with data fast enough and consequently slows down the performance of the computer. The amount of RAM typically found in a PC is steadily increasing as hardware costs decline and software becomes more complex and demanding of processing capacity. Today 64 MB (units are described in the footnote to table 7.4) of RAM is sufficient for many applications, but already 128 MB is needed by some software.

Secondary memory

Secondary memory is typically a magnetic hard disk (sometimes called a fixed disk), a removable diskette and, in recent years, a removable optical disk, the CD-ROM. One can usually see when the CPU is accessing any of these devices because a light on the front of the system box is switched on during access. It is advisable not to remove any device while it is being read or written to, or when a PC is using data or using a program stored on a removable medium. In the case of diskettes it is safer to transfer data or programs from the diskette to the hard disk and access them there.

Hard disk

The hard disk is the main storage for both data and programs. It is therefore a key component. Hard disks do occasionally fail, and sometimes a particular file may be damaged or accidentally erased. It is even possible to accidentally reformat the

Table 7.4: Electronic Storage of Data

(a) Examples of capacities

Item	Typical capacity today (in bytes*)	Notes
3.5 inch diskette	1.4 MB	
CD-ROM	600 MB	
Hard disk	1+ GB	Also called the "fixed disk" as it is fixed inside the systems box. The others in this table are "removable media."
DVD-ROM	4.7 GB	Digital Versatile Disk – read only
DVD-RAM	2.6 GB	Digital Versatile Disk – read & write

(b) Examples of requirements

Item	Storage requirement (in bytes)	Notes
A4 sheet of typescript	4 KB	complete sheet of plain text, stored as ASCII code. A short letter may be less than half of this as plain text. A word-processed document may be twice this per page plus an initial overhead for the document..
A4 sheet sent by fax	48 KB	stored as 300 dpi bit map image.
A one page graphic	1,200 KB	saved as a 300 dpi bit map, but much less than this for some graphics formats such as GIF and JPEG (see Graphical Formats section below)
A one-second audio clip	8 KB	voice quality
	176 KB	CD quality
A single video frame	90 KB	in compressed JPEG format (640X480)
An hour of full-motion video	1 GB	

*A byte is equivalent to one alphabetic character or two numerical characters.
1 KB (kilobyte) = 1024 bytes
1MB (megabyte) = 1,048,576 or approx. a million bytes
1GB (gigabyte) = approx. a thousand million bytes

whole disk and lose everything on it. For these reasons it is important to make a copy of the data stored on the hard disk at regular intervals (this is discussed further in the piece on magnetic tape below).

Typical sizes for the hard disk today are between 2 and 10 GB. This is much larger than was typical five years earlier but today's software is also larger, and if graphical data (images) are stored this occupies much space. Eventually most people come up against the limit of space. There are some ways of easing this problem:

- Be selective in what you store. E-mails, and/or any attachments, whether stored locally on your PC or on a file server (which will also risk filling up), have a tendency to accumulate. It is useful to set up one or two new directories ("folders") in your e-mail system, for example one called Archive, which is reserved strictly for things you are convinced have long-term value, and another called Pending, for those that need action but can be erased when acted on. The current "inbox" can then always be maintained with just the day's new mail. Sometimes an e-mail may need keeping but not its attachment(s), which can occupy much space.
- In the Windows operating system there are three other accumulations that should from time to time be cleared. One is the Recycle Bin, which is the subdirectory C:\Recycled\. The second comprises files with a .TMP file extension, which can accumulate in more than one place but typically in C:\Windows\Temp\. Close any programs that are running before deleting the .TMP files. The third set applies only to a PC connected to the Internet and comprises the Internet files found, for example (with Windows Internet Explorer) in C:\Windows\Temporary Internet Files\.
- It is sometimes helpful to reorganize the whole file collection on a hard disk to re-assemble files that have become fragmented into pieces stored in different parts of the disk. To do this in Windows 95/98, select Start on the task bar, and then Programs, Accessories, System Tools, Disk Defragmenter. This facility first tests the state of defragmentation of the hard disk and offers a recommendation whether or not to proceed.

Both hard disk and diskette are covered with magnetic material and rotate at speed. Data is stored by magnetizing very small points in linear concentric circular tracks through read-write heads, similar to that used in a domestic cassette recorder. The read-write head is held very close to the magnetic surface of the spinning hard disk. The hard disk normally rotates continuously while the PC is switched on, although on some PCs it can enter into a low-power standby mode in which the disk may stop rotating.

A network computer, or diskless workstation, is a PC without a hard disk but with a connection via a network cable to a central file server from which data and software programs are "borrowed" via the network just for the time the user wishes to use them. Such a system costs less than an ordinary PC, particularly if applied to a large number of users on a network. The prospect is being actively marketed by some large companies such as Oracle, but it is too soon to say how successful it will be.

Diskettes

Diskettes are removable (unlike the hard disk) and are convenient for passing information to others; they can readily be sent in the post. They are also useful for making backup copies of small amounts of data.

The common 3.5 inch high density (HD) diskette has a storage capacity of 1.44 MB; there are also 760-KB double-density (DD) and 2.88-MB extra-high-density (ED) versions. Older PCs may still have a 5.25 inch disk drive; the commonest 5.25 inch ("floppy") diskette stores 1.2 MB. The read-write head of a diskette drive actually touches the magnetic surface so there is a danger of dust and fingerprints causing interference. This was more of a problem with the 5.25 inch diskettes: the 3.5 inch diskettes have a retractable cover to protect the read-write area from dust when the diskette is removed from the PC.

Some organizations have eliminated diskette drives from all their networked PCs to lessen the misuse of the machines and in particular to reduce the risk of computer viruses (see later), which are mostly introduced by diskettes being passed between users.

Optical discs[10]: The CD-ROM

A CD-ROM (Compact Disc - Read Only Memory) is another form of removable storage but has a much greater capacity than diskettes. It is widely used to distribute software programs, which tend to be rather large. It is also used to distribute data on a wide range of topics, both for public sale and for the internal information needs of companies.

A typical 120 mm CD-ROM holds 550 or 680 MB of data. Data-transfer rates from optical discs are usually slower than from a magnetic hard disk but faster than from a diskette. The speed of the original version, known as "single speed" or "1x", was 150 Kilobytes per second (KBps). Nowadays 24x (3,600 KBps) or 32x (4,800 KBps) are common.

CD-ROM is a form of optical disc and stores data by altering the reflective property of very small points on a spiral track in a layer within the disc. Such changes are detected by the read head, which directs a laser light beam on to the disc and measures the reflected beam. It is the same device as the audio CD: only the stored information is different: one represents sound and the other, data. Most new PCs today have a CD-ROM drive fitted as standard.

Optical discs do not usually give problems provided the lower surface is kept free of dirt and scratches. Minor scratches and fingerprints may not cause difficulty but deeper scratches and dirt may interfere with the reading of data. Dirt can be removed with a soft damp cloth, if necessary with the assistance of a window cleaning fluid as long as it does not damage plastic. It is possible to buff out scratches with a plastic cleaning fluid or polish. If you are working in a very dusty environment it may be worth using a drive that accepts a CD caddy—a case to keep the CD free of dirt and dust and that opens automatically for reading when mounted in the optical disc drive, rather like a 3.5 inch diskette.

There are other types of optical disc. CD-R (CD-Recordable) discs can be used for backing up large data sets or distributing large files. A CD-R drive can "burn" the data into a CD-R disc. This is a type of WORM disc (Write Once Read Many), which can be written to only once and gradually fills up like an exercise book and cannot be erased or over-written. CD-R drives operate more slowly than most CD-ROM drives but when writing to them it is necessary, particularly on earlier CD-R drives, to ensure that the PC is able to supply the data at sufficient speed: if there is a gap in the data flow the recording session may fail. A fast magnetic drive and a large RAM

10. The spelling "disk" is often used for magnetic media; "disc" for optical.

reduce the chance of this problem. Such a CD-R disc can be read by a conventional CD-ROM drive and so this technology is an economical way of passing a large file to others. CD-R discs are now quite cheap at around $2 each.

CD-RW (CD-Rewritable) discs can be written to and then over-written later with new data. This is proving an alternative to magnetic tape (see below) as a cost-effective way of storing archive material. Some earlier CD-ROM and CD-R drives are unable to read a CD-RW disc but newer ones, with a MultiRead head, have this function included.

DVD (Digital Versatile Disc) is a new standard that comes in a data version (DVD-ROM) and a video version (DVD-Video). Its first release has a capacity of 4.7 GB; larger versions are planned. There are also writable (DVD-R) and rewritable (DVD-RAM) forms.

Magnetic tapes; backing up data

Tapes provide a quick and efficient way to store a spare or backup copy of large amounts of data though this role can also be met with a recordable or rewritable CD (CD-R, CD-RW) as described above. It is very important to make such backups regularly as the original set of data on a computer can be lost easily through any of several causes: e.g., a spike in the electrical supply, accidental erasure of a file, accidentally formatting the hard disk instead of a diskette, failure of the hard disk, and theft. The consequences of the loss of the only copy of data can be very serious, making it well worth while to regularly make a backup copy. A tape is a useful device for this purpose. The hard disk can be regularly backed up and the tape removed and stored in a safe place elsewhere.

A new PC may come with an internal tape drive already built in. If not, the following factors should be considered in choosing one: internal or external, storage capacity, compatibility, cost, and the capability of the drive and associated software. An internal drive occupies less space and is cheaper. An external drive is more easily shared between a number of PCs. It is connected to one of the PC external ports. A capacity should be selected that will allow an entire drive (or partition of a partitioned drive) to be backed up to a single tape. This allows a simpler backup procedure without having to change tapes part way through.

Compatibility is a factor when you have previous tapes, in which case backward compatibility with those earlier ones might be needed, and when there are several computers to be backed up: it is cheaper and simpler if all use the same tapes. There is a wide range of tape drives and tape cassettes and although there is a standards committee for such tapes (the so-called Quarter-Inch Committee, QIC) they have approved a range of QIC quarter-inch tape formats based on two sizes of cartridge: a standard cartridge and a mini-cartridge. Additionally there are two other standards: the 4mm digital audio tape (DAT) and 8mm videotape. While externally the cartridges for these two are similar to the common music cassettes and video recorder cassettes respectively, the method of recording is in each case different.

When users are connected to a network, a service for backing up users' data can be provided centrally. Users' data may be normally stored both locally on each PC's hard disk and centrally on a file server's hard disk. Arrangements can be made for both such stores to be backed to a central, large tape or hard-disk facility.

Monitor (and the Video Adapter Card)

The monitor displays data as it is typed into a PC via the keyboard and displays the results of any editing or processing. An MIS typically has many built-in (i.e., predesigned) reports, which can be seen on the monitor and/or printed to paper. The monitor is therefore one of the two main tools to get information out from a computer (the other being the printer). Choice of a good model of monitor will make it easier and more comfortable to use and reduce the risk of eye strain.

The great majority of monitors for desktop PCs are based on a cathode ray tube (CRT), which is also the device used to display the picture of a television. A "video adapter" card in the computer generates the signal that is sent from the computer to the monitor. So the choice of this video adapter also affects the display seen on the monitor. Nowadays a personal computer bought from one of the better manufacturers is likely to have a good balance between the monitor, the video adapter card, and the inherent properties of the computer, including its installed operating system. But any personal computer catalog or price list is likely to show a range of models with different monitor specifications. Box 7.1 describes these features and may assist in coming to an appropriate choice. Nowadays many monitors allow the user to select the resolution from a set of in-built standards. Many users buying a complete PC from a reputable manufacturer may have little need to consider more than the screen size. The norm today is 17 inch. Older monitors that offer only the VGA standard may be unable to display many Internet sites properly.

A cathode ray tube would be much too bulky for a portable computer. Instead these use a quite different technology for their screens—liquid crystal display (LCD). LCD screens are more or less flat and use much less power than do CRT screens (around five watts instead of 100 watts or more). Up till now LCD screens have not been used much for desktop computers because they are considerably more expensive, particularly for a large screen, than a CRT. But costs are falling, and it is likely that LCDs will become more common as the display unit of the desktop computer too. LCDs use the whole of their front surface for display without the "lost edge" of the CRT monitor. This means that a 15-inch LCD is roughly equivalent in picture size to a 17-inch CRT display. Typically it comes with a comparable screen resolution, e.g., 1024 x 768 pixels. The resolution of a LCD is fixed, unlike many CRT units, where it can be varied. There are different types of LCD: active matrix displays (using a thin film transistor array and therefore referred to as TFT) provide a clearer and brighter picture than passive matrix displays (also called Dual Scan displays) and can be read more easily at an angle, which is useful when two or three people are looking at the same screen.

Printers

Several technologies have been developed for printers. The main ones are dot matrix, laser, inkjet, and daisy wheel, although the last is seldom seen now.

Box 7.1: Characteristics of Monitors

Size of screen is given as the diagonal size of the cathode ray tube on which the screen is projected; perversely the actual visible screen (the "active display area"), which is what really matters to the user, can be up to 1.5 inches* less. Thus the common 15-inch monitor has an active display area of about 13.5 inches (343 mm) diagonally across. 15 inch has been a common size but nowadays the benefits of a larger screen are appreciated and 17 inch is becoming the standard: it offers a significantly improved image and is not much more expensive. Moreover some developments such as Websites on the Internet have been designed for a 17-inch screen, so this is becoming the minimum standard for use without eye strain when accessing the Internet. Those working particularly with graphical images such as a geographical information system, desktop publishing, and computer aided design (CAD) programs, may wish to consider a 21-inch screen, but one does have to pay rather more for this larger screen.

Screen resolution controls the sharpness of the image and is measured by two parameters. The screen is composed of a large number of picture elements (shortened to "pixels"), rather like the dots in a newspaper photograph. The first measure of screen resolution is the number of pixels horizontally across the screen and vertically from top to bottom. The larger the number of pixels, the sharper the image. Older standards are VGA with 640 x 480 pixels and SVGA with 800 x 600. Now 1024 x 768 is becoming a standard and is useful for viewing Internet pages. Table 7.5 lists common resolutions and the recommended minimum screen size for each.

The second measure of resolution is the "dot pitch" and applies only to color monitors. Color on the screen is generated from three components in each pixel, one each of blue, green, and red. The dot pitch is the distance in millimeters between these "triads:" the smaller the distance, the better the resolution. Nowadays a dot pitch of 0.28mm is the maximum recommended, with 0.25mm or less preferred.

Color monitors are almost invariably included with a PC nowadays. Color is useful in graphical interfaces such as the Windows operating system but is not needed for text processing. The alternative is monochrome, which may be amber, white, or green. A monochrome monitor is cheaper and might be considered for a file server.

Emission of electromagnetic radiation from the monitor has been raised as a potential health hazard. Several studies and reviews have suggested that the risk, if it exists, is small. Some add that while this is so, more information needs to be gathered. A good review of the topic can be found in IEEE (1997). Some manufacturers have responded to this concern by providing monitors with reduced levels of radiation, although these cost a little more. Others have developed anti-glare screens. Radiation emission declines rapidly with distance so it may be wise not to sit very close to the screen. There is less emission from the front than from the sides and rear of a monitor so when there are several PCs in a room there is merit in arranging them in such a way that users are not sitting close to the rear or side of another's PC.

Power saving is enabled on some monitors through software control that can be set so that the monitor automatically goes into a standby mode after a set period of time that

(continued on next page)

(Box 7.1 cont'd)

the PC has been inactive. This standby mode reduces the monitor power consumption to about a quarter, e.g., from around 100 watts for a typical monitor to about 25 watts. Not all PCs have this facility; it is more common in laptop computers, which need to conserve battery power.

Two **frequencies** are quoted in the specifications for a monitor. The frame rate or vertical frequency is the number of times per second that the whole screen is refreshed. 60 Hz (cycles per second) is often regarded as the minimum for an acceptable image without a noticeable flicker, and 70 is better. (People vary in their sensitivity to flicker, so while some may be comfortable with a refresh rate of 60 Hz, others may find it disturbing.) Some monitors allow selection from a number of frequencies: in Windows, select Start, Settings, Control Panel, Display, Settings, Refresh Frequency. The second frequency is the horizontal frequency, which is the number of times per second the electron beam inside the monitor's cathode ray tube sweeps from one side to the other; this is approximately the vertical frequency x the number of horizontal lines. Provided one has selected an appropriate resolution (see table 7.5) and there is a satisfactory vertical frequency, one has little need to address the horizontal frequency. An exception to this might be when procuring a video card separately from the monitor; it will then be necessary to ensure these various factors match up.

* The traditional unit of measurement for PC screens is inch.

Table 7.5: Monitor Resolutions

Resolution (pixels) (horizontal x vertical)	Title	Minimum recommended monitor screen size* (inches, diagonal)
640 x 480	VGA (video graphics array)	13
800 x 600	SVGA (super video graphics array)	15
1024 x 768	XGA (extended graphics array)	17
1280 x 1024		21
1600 x 1200		

* Adapted from Mueller (1998).

Dot matrix was for many years the commonest. It is relatively simple and cheap, both to buy and to run, but rather noisy. Graphics capabilities are limited. It is still used where multiple or carbon copies are required, as for invoices and salary slips. It can do this because unlike laser and inkjet, dot matrix is an impact device: small pins strike an inked print ribbon (as in a typewriter) to press the image of a character on to the page. This impact causes the noise, which can be irritating. It is possible to reduce the noise by housing the printer inside a container. Dot matrix typically print to continuous sheet (fan fold) paper running over a roller and driven through sprockets fitting holes on each side of the paper. They may also be fed individual sheets. They come in 9-pin and 24-pin print head models with the latter giving the better print quality.

The laser printer has become the market leader for normal document and letter printing. It is quicker, quieter, and gives a sharper print than dot matrix, and has much better graphics capabilities. It is also quicker than inkjets and, in terms of consumables, slightly cheaper per page. The laser printer "ink" comes in a toner cartridge, which has to be replaced after between 1000 and 5000 pages. Laser printers, if left on continuously, use more electricity than do dot matrix or inkjet printers; some have an automatic standby mode into which they move when no printing has been done for a set period of time, and use less electricity in this state. When a laser printer issues a "toner low" message you may extend the toner's use a little by removing the toner cylinder and shaking it gently from side to side to redistribute the toner powder across the cartridge.

Inkjet printers are used particularly when color is needed and when space is limited; they are convenient as a portable printer. Modern inkjets provide a print resolution close to or as good as that of laser printers, but cost of consumables is a little higher. Models with a separate cartridge for the black ink and another for the colored inks allow replacement of one when it is finished without having to discard the other while it still contains ink. There are two types of inkjet technology. Thermal inkjets use a high temperature to generate a vapor to force the ink through the jets. Piezo inkjets, developed more recently, use piezo-electric crystals; removing the need for high temperature allows alternative inks to be used and extends the life of the jets.

Both laser and inkjet printers are page printers; they accept and process from the computer the full content of a printed page as a whole, making use of a page description language (PDL) to put together the page layout before commencing to print it. There are two major PDLs on the market: Hewlett Packard's PCL (Printer Control Language) and Adobe's Postscript. If electronic files are regularly passed to an external printing company it is worth checking the preference of the company for which PDL they prefer and ensuring new printers meet this requirement.

Printers are brought into use for only a small proportion of the time that a PC is in use, so it is economic, when a number of PCs are in use in the same or nearby rooms, to link them to a single printer. When the PCs are connected by a network, the network can be used to connect them to a shared printer. A number of stand-alone PCs can also be linked to a single printer through a simple printer switching box. For a single user a laser printer with an output of 4-6 pages per minute (ppm) is appropriate; for a small office 10 ppm, and for a networked printer serving between 10 and 25 people, 20 ppm is recommended. Most inkjet printers print at less than 8 ppm for monochrome text and 2–4 ppm for color.

Cost of printing comprises the initial capital cost of the printer and the running costs, which comprise the paper and the printer consumables. The capital cost of printers has steadily declined in recent years. The cost of consumables, even apart from paper, may equal the cost of the printer within two years in a busy office and is therefore worth considering. It may equal 2–3 US cents per page for laser printers and 2-4 cents for inkjets (for a monochrome text page, more for color and/or extensive graphics). Prices may be higher in developing countries but the ratio is likely to be similar.

Modems

A modem is a device that enables data to be exchanged between two computers via the public telephone system. Computers use data in a digital form whereas the telephone line was designed for voice transmission using an analog form. A modem converts the computer's digital data to analog for this transmission. At the destination, where the telephone wire delivers the data to the remote computer, another modem converts it back into digital form. The two main uses for this connectivity provided by modems are access to e-mail and to the Internet.

Modems can be fitted internally into a PC or attached externally and have a twin wire connection to the telephone system. A significant feature of a modem is the speed at which it can transmit data. At the time of writing most modems operate at 28.8 or 33.6 Kilobits (thousands of bits) per second (Kbps). Modems of 56 Kbps are available but require particular circumstances to deliver such a speed, circumstance unlikely to be available to most NAROs soon.

The great majority of modems on the market today can also be used to transmit faxes. This means that, with the appropriate fax software on the PC, a fax can be prepared on a PC and transmitted via the modem just as with a dedicated fax machine.

Scanners

A scanner is a device that converts text and pictures on paper (or other materials such as overhead transparencies and 35mm slides) into an electronic form. This function is of interest to agricultural research because many countries have a half century or more of potentially useful agricultural research history largely recorded on paper that could be preserved, distributed, and accessed more effectively in electronic form. See box 7.2 for an explanation of these advantages.

Scanners focus a beam of light on to the paper and record in its reflection (or transmission in the case of a transparency, 35mm slide, or photographic negative) the difference between image or text components and the blank page. A sensor carriage carries a row of many sensors down the page, pausing briefly at each step to capture another row of small images of the page. The result is an electronic file made up of a grid of large numbers of tiny dots or pixels. Each pixel is represented by an electronic record of whether that spot on the paper was black or white or a shade of gray or, in the case of color scanners, the color. This file is, like a fax message, an electronic image of the paper original, but it is not in the form of ASCII text (see glossary in annex 1) that would enable it to be read and edited by a word processor.

The part of the image representing text can be processed into a form that is suitable for word processing by using optical character recognition (OCR) software. This examines the image of each character, identifies which character it is (i.e., A, B, 1, 2, ;, ?, etc.), and stores it as the ASCII code for that character. The resulting file can be read and edited by a word processor. It occupies less storage space than its original image format, and for this reason it is faster to transmit electronically. The disadvantages of OCR are that it takes longer than scanning just as an image and the output is likely to have some errors, which have to be checked for and corrected manually particularly when scanning poor-quality documents.

Box 7.2: Advantages of Electronic Documents over Paper Documents

Preservation. Paper documents, particularly items of gray literature (which has not been officially published), are often held in libraries as a single copy, which is rather vulnerable. Someone may borrow the only copy in the library of the 1975 annual report and fail to return it. A worse scenario, but one that has happened several times, is that a major research library is burnt down. The collection of records of the past half to one century of research that has been conducted in many countries is steadily eroding (Vernon 1986).

In contrast, electronic files can be copied in seconds and copies sent to every research station in the country, to ministry HQ, and to collaborating institutes in other countries. This quickly provides immunity against the loss of a single paper copy. Electronic files occupy less storage space than their paper equivalent.

Wider access. Electronic documents are more easily made widely accessible, to all in a country or even to many around the world. A copy can be made and e-mailed in seconds to anyone requesting it. A copy can be regularly sent to all collaborating institutes for local access. And a copy can be mounted on a Website, where anyone can read it remotely or download a copy for local study and if desired printing to paper.

Easier searching. Consider the question: what research has been done in the past on non-pesticide methods of disease control in beans? With paper documents one would have to manually look through the section on beans, and/or plant pathology, in every annual report, or local agricultural journal since records were kept, if they were indeed all still available. With electronic storage the same search could be done in minutes.

The following factors may prove useful when selecting and using a scanner:

Connectivity. Connecting a scanner is easier if it has a USB interface, rather than SCSI or parallel port (described in the next section). Your PC must also have USB and Windows 98 (or higher) to use this feature. For SCSI, a SCSI card and cable are required. If your computer has none of these, the alternative is to run the scanner from a parallel port.

Ease of use. The control buttons to operate the system may have various degrees of complexity. The problem is that scanners, particularly the more expensive ones, use many techniques that can improve the quality of the final image, and these several techniques can be adjusted to get the best image. There is therefore scope for developing some skill in using a sophisticated scanner. Some models offer a "quick" mode of usage, where the user selects the purpose of the scan, and the machine makes the more complex decisions based on this information and on a preliminary scanning of the target to assess its nature and quality.

A **preview facility** enables a quick scan to present a small preview of the image on-screen. This enables the scanning area, the resolution, and the bit-depth to be adjusted before the main scan. Preview scans may take from a few seconds to a minute or more in some models.

Resolution, bit-depth, and color. Resolution is an indication of how detailed an image is and is measured in terms of the numbers of pixels horizontally (the x axis) and vertically (the y axis) used to represent the image. Optical or hardware resolution is often reported in the form of two figures, e.g., 300 x 600. The first, the more significant, usually represents the number of scanning elements per inch in the sensor; the second the number of movement steps per inch that the sensor carriage makes for each exposure. The optical or physical resolution is often enhanced by the software interpolating additional information derived from the average of adjacent measured data. Scanning resolutions higher than 600 dpi are usually achieved through such software interpolation.

Scanners mostly allow you to set the resolution at which they scan. The best choice is influenced by the quality of printer that the image may later be printed with. It is rarely necessary to use a resolution greater than 240 dpi for color on paper, and 72 dpi for display on a computer monitor including Web applications.

The amount of data used to record each pixel is known as the bit-depth. The very simplest image has a pixel bit-depth of 1: it records merely whether the pixel is black (represented by a 1) or white (a 0). Shades of gray need 4 bits for 8 shades: 8 bits allow 256 shades. Recording color typically uses 24 bits: 8 bits to represent the degree of each of red, blue and green. At the time of writing, most scanners use 36 bits per pixel for recording color. Bit depth directly affects the size of the resulting file.

Resolution and the quality of the image are also affected by the quality of the optical components such as lenses and the mechanisms used in the scanner, with better results offered by the more expensive machines.

Sensor mechanism. There are two main types of sensor mechanism in today's desktop scanners: (1) CCD (charge-coupled device), found in most desktop scanners and at the time of writing gives a better output, and (2) CIS (contact image sensor), a newer technology with fewer parts and therefore more compact, more energy efficient, and cheaper.

Optical character recognition. The conversion from a bit-map pixel image of a page of text to recognized alphabetic characters is a complex process involving a number of stages, and it has proved difficult to achieve a 100% error-free result. This is in part because of the diverse range of text fonts. Several decades of work in this area have led to great improvements, particularly in the last few years. Late in the process the file is run through a spell checker to at least ensure all the text comprises real words, and then through context sensitive analyses to select, where there is some ambiguity over a letter or word, the most likely option. Well-proven optical character recognition software includes Caere Omnipage and Xerox TextBridge.

Scanning for OCR processing, which is the case for the bulk of stored agricultural research information, is best done by selecting "grayscale" on the scanner. This will be quicker than color scanning and results in a file typically one third of the size. Line-art mode gives an even smaller file sizes but often loses detail, reducing the accuracy of subsequent OCR processing.

35mm slide scanner adapter. When planning to capture 35-mm color slides, which can be valuable for example in assisting with the recognition of plant diseases and mineral nutrition systems, an adapter for this purpose is available in some of the better models.

TWAIN compliance. TWAIN provides a standard in image acquisition for connecting TWAIN-compliant applications with TWAIN-aware devices. It specifies how image acquisition devices such as scanners, digital cameras, and other devices transfer data to software applications. Selecting a scanner that is TWAIN compliant is likely to make connection to PCs, printers, and other devices easier. TWAIN compliance is common these days.

Speed. Speed becomes important if one has to scan a large number of documents. The figures below come from a manufacturer's sales brochure[11] in late 1999 and give some idea of the time needed to scan a single sheet:

- up to 9 seconds, 4 x 6 inch photo
- up to 45 seconds, 4 x 6 inch photo to Microsoft Word
- up to 60 seconds, OCR letter size black and white text to Microsoft Word
- up to 45 seconds, black-and-white line art to Microsoft Word

When speed is important and when the end product is a text document rather than an image, it makes sense to avoid unnecessarily high resolutions, which greatly add to the time taken for scanning.

Automatic document feeder. An automatic document feeder is helpful for scanning a large quantity of material. It is sometimes provided as an optional extra that can be ordered on the more expensive models. A typical feeder may hold 25 sheets. Formerly there were on the market scanners that could only be fed individual sheets but nearly all models now are "flat bed," in which the document or book is laid on a glass surface.

Input/output interfaces

Input/output interfaces facilitate connection and communication between a PC and peripheral (external) devices such as printers, mice, modems, and scanners. Common interfaces include the serial port, parallel port, SCSI, and USB.

Serial and parallel ports have long been provided on PCs. The former sends data over a single wire. The eight bits of a byte needed for each alphabetic character follow each other in succession down the wire, whereas in the latter there are eight wires, one for each bit all moving in parallel. A serial plug typically has nine pins but only two carry the data, one in each direction. Older versions had 25 pins, based on a standard (RS 232) that first enabled otherwise incompatible devices to communicate in two directions. Parallel ports send data over two sets of eight wires: parallel plugs may have 25 pins or 36 pins. Parallel port connection originally enabled a faster transmission rate and was originally developed and is still widely used to send data from a PC to a printer. Serial connection has been used particularly for modems, which connect to the two-wire twisted pair of the telephone system. Normally paral-

11. Hewlett Packard's Scanjet 6300

lel port cables should not be longer than three meters, whereas a serial cable can be up to 12 meters.

SCSI (small computer system interface) enables several (up to 8 or 16) external devices to be "daisy-chained" together, i.e., to be strung together like beads on a necklace. In this way it offers great expandability. It requires a SCSI adapter card to be installed in the PC and configured. It is a parallel system and consequently its connecting cable is best limited to three meters or so.

One would think that parallel systems inevitably are faster than serial, but in practice they encounter many difficulties, and there is a move to serial systems for greater speed and functionality.

USB (universal serial bus) is representative of this tendency and offers significant advantages over those so far discussed. It allows new peripherals to be connected to a PC without the need to install a card and reconfigure the PC. Indeed, one can add a peripheral to a PC while it is running with no need to reboot the PC. This is achieved through the USB system coupled with the plug-and-play facility of Intel processes (Pentium III onwards) and a broad range of drivers in the Windows operating system (from Windows 98 onwards but not Windows NT). The result is a much easier task for users when connecting new devices to their PCs. Most new PCs come with a USB, and if not present it can be purchased and installed.

7.3.2 Software

Software comprises the programs that contain instructions to the personal computer on how to process, move, and store data. Software may be divided into the *operating system*, which controls the operation of the computer, and *application software*, which is designed to meet particular needs of users.

Application software

Application packages allow users with no programming skills to accomplish a range of tasks with their personal computer. Common types of generalized application software include the following:

Word processors are used for letters and longer documents. Common examples are WordPerfect and Microsoft Word.

Spreadsheet and database systems manipulate data. Spreadsheets are designed particularly to deal with numeric data and usually offer good graphical displays for their outputs. They have a structure of rows and columns and a large set of mathematical functions for a wide range of numerical processing. Modern spreadsheet software enables links to be maintained between several individual spreadsheets. Common examples of spreadsheets are Microsoft Excel, Quattro Pro, and Lotus 1-2-3.

Database software is suitable for storing and processing both numeric and textual data and can handle very large data sets that would be beyond the functionality of a spreadsheet. A database can handle relationships between various data more effectively than can a spreadsheet and can largely avoid duplication of stored data, even in large data sets with multiple uses and users of the same data. This greater power of the database means that it is more complex, so for smaller tasks where the data is largely numeric a spreadsheet is more appropriate. Database software is particularly useful for an MIS (section 7.5 “Elements of a database system” looks at it in more detail). Examples of database software for personal computers are Microsoft Access, FileMaker Pro, Imprise Visual dBase, and Microsoft Visual FoxPro. There are related products for handling text such as AskSam. Larger databases management systems for mini-computers, which may be more appropriate for very large NAROs, include Progress, Oracle, and Sybase.

E-mail enables users to easily generate and transmit electronic messages to other local PCs via an institution’s local area network (LAN) or to remote PCs via the public telephone system.

Statistical application software is now available to support a range of statistical analyses of research data. Geographic information system (GIS) software extends data management to spatial characteristics and may have application in the agriculture field, which is a spatially intense activity (see section 6.4 for a detailed discussion of its application for agricultural research).

Operating systems

An operating system provides an interface between application software and the hardware on the one hand and between the user and the hardware on the other. It also performs the “housekeeping” tasks of managing files and managing the computer’s memory. Table 7.6 lists a number of operating systems found on PCs, along with some of their features.

MS DOS and PC DOS (collectively “DOS”) were the standard operating system for most users of PCs for many years. They are command driven and users need to remember the commands for the more common tasks that they wish the computer to do. Apple computers pioneered a different approach using a graphical user interface for its operating system. Microsoft’s Windows operating system has brought the graphical user interface to PCs and has largely replaced DOS. Many application software products are no longer being developed in DOS versions.

Compared with DOS, Windows has the following features:

- It has a “point and click” system, in which options from on-screen menus are selected by pointing to the chosen item with a screen cursor, which is moved by moving the mouse and clicking with the mouse button.
- It provides a common environment for running a wide range of application software, any of which can easily be identified and started from the Windows screen.
- It has enhanced computer memory management facilities and has become increasingly multitasking in its later versions, enabling more than one task to be addressed simultaneously. However, this uses more memory.
- It enables a number of software applications to be running simultaneously and allows information from one to be transferred to another.

Table 7.6: Common Operating Systems for PCs

Operating system	Hardware	Notes
PC DOS	IBM PCs	DOS stands for Disk Operating System. An early operating system for microcomputers from IBM, designed originally for only 640 KB of memory for program use.
MS DOS	IBM PC clones	Microsoft's equivalent of PC DOS.
OS-2 Warp 3 OS-2 Warp 4	IBM PCs	From IBM. 32-bit, multi-tasking; used more across large organizations, government departments, and utilities than small business.
Apple System 7, 8.5, and MacOS X.	Apple computers	Operating system from Apple Computers; multitasking, powerful graphics and multimedia functionality. See also www.apple.com
Windows: a) Windows 3.x b) Windows 95 c) Windows 98 d) Windows NT e) Win 2000	IBM PC clones	From Microsoft Corporation, the dominant operating system for PCs today. a) Still used on PCs with a 386 or 486 processor. b) Part 16-bit, part 32-bit; multi-tasking c) 32-bit OS. Will take advantage of Intel's new CPU, Katmai. (a), (b), and (c) all retain a legacy function to run DOS software. This is finally abandoned with (d) and (e). d) 32-bit operating system used mostly on workstations and network file servers; multitasking, multiprocessing. e) The next version of both Windows NT and Windows 98.
UNIX	Workstations and minicomputers, network file servers	64-bit OS, multitasking, multi-user processing, networking; can operate on many different computer CPUs.
Linux	As for UNIX, and beginning to be used on PCs too.	A recent alternative to UNIX that has started to become more widely used from 1998, for example by Internet Service Providers, but also becoming used on desktop PCs. Like UNIX but unlike Windows, Linux is an "open-source" operating system—available from several vendors and even available free on the Internet. Some vendors, such as Red Hat Software, have added a user friendly graphical front end to it. Some PC manufacturers are now offering PCs with Linux OS installed instead of Windows. Website: www.redhat.com and www.caldera.com
Java OS	Network computers	From Sun Microsystems, which makes workstations and file servers; especially for networks and heavy-duty data-entry applications. Website: www.sun.com
BE OS	PCs	Developed originally for Apple Macintosh but a PC version was released in 1998. Still early in development but may prove an alternative to Windows in due course. Used by multimedia developers who need powerful graphical functionality. Website: www.be.com

- It facilitates the addition of new software and hardware devices. Its plug-and-play feature recognizes a wide range of new devices and automatically enables their operation.
- It requires machines with more memory, a faster processor, and a good video graphics card to operate at its best. New versions of Windows continue to appear: current versions in use are Windows 3.1, 95, 98, NT, and 2000.

Although table 7.6 presents a number of operating systems currently available for personal computers, more than 90% of PCs use the various Windows systems, and this dominance is unlikely to change greatly in the next few years. This means that for most NAROs expanding their computing facilities the choice is largely between Windows 98 and 2000 for users' PCs and Windows NT or Linux as an alternative for a file server for a network. Many people recognize the existence and gradual spread of alternatives as providing, in due course, some much-needed competition to Windows. Table 7.6 gives their Websites, so readers can examine the current position. Independent views should be consulted in the personal computer press.

When upgrading from an older to a new version of an operating system, and especially when changing the operating system, it is wise to backup to tape or CD-ROM the whole hard-disk contents. Some advocate following this with removal of an old version and making a fresh installation of the new, or even reformatting the hard disk first, rather than just upgrading the existing system in situ.

Data

Data is the computer's coded representation of letters, numbers, and other symbols that the computer can process to provide information.

The unit of data is the binary digit, or bit. There are only two types of bit: zero (0) and one (1). Eight bits are used to represent an alphabetic character: for example the letter A is represented in the ASCII code (the most common code used in microcomputers) as 10100001, and the letter Z by 10111010. Such a string of eight bits is known as a byte. All data stored in a computer, whether representing text, numbers or even pictures and sounds, is stored in these basic units of zeros and ones. The main, perhaps the only, reason for us needing to know about bits and bytes in the context of information management is that they are the units of measurement in storage capacities. In deciding what hardware to buy we can be assisted by estimates of our storage needs. Similarly in choosing communication equipment we will need to be aware of the speed at which it can transmit data and this is expressed for example in bits per second (see tables 7.4 and 7.7 for examples of data capacities and transmission rates respectively).

Data formats

Plain ASCII text may be regarded as the simplest representation of data. Different applications then apply various formatting to this basic text. Word processors format text to indicate **bold**, *italic*, and <u>underline</u>, and a wide range of fonts or character styles and sizes. Spreadsheets apply a structure based on columns and rows, and database software provide a structure that usually includes fields and records; both use formats that distinguish between text, numbers, dates, etc. The significance to us here is that each software may use a different formatting system, which can cause a problem when we wish to transfer data from one system to another. Usually plain ASCII text can be transferred without difficulty between softwares even from different manufacturers, but this is not always the case for formatted data. Often a software manufacturer will provide conversion tools that facilitate ready exchange of data between, for example their own spreadsheet and database packages. The more common application packages also provide for data transfer between different manufacturers' packages.

Graphical formats

Apart from text and numerical information, computers are often used for the manipulation of graphical information, such as illustrations and graphs in publications and on screen, for geographical information systems (GIS) and computer-aided design (CAD). The use of illustrations and graphs is discussed in section 4.3.5 "Use of graphical outputs." Section 6.4 describes the use of GIS. Here we look at the graphical formats that computers use to display data.

Graphical information can be stored in many different formats, often the proprietary formats of certain types of software or hardware, such as image scanners. There are two main groups: bitmap or raster images and vector graphics. Metafiles are formats that may contain either raster or vector graphics data.

Bitmap or raster images store information for each pixel of the picture on the screen. These formats are often used for scanned images and photographs, discussed earlier.

The size of a bitmap graphics file is determined by the screen resolution (the number of pixels vertically and horizontally) and by the number of colors for each pixel. A screen resolution of 1024 x 768, which is common nowadays, has 786,432 pixels. If 8 bits (=1 byte) are used for each pixel to provide 256 colors, the storage needed for a single refresh of the screen is 786,432 bytes or 768 KB (a kilobyte is 1024 bits). This is the minimum amount of RAM memory one would need on a video card adapter. It also decides the size of a bitmapped file.

On the Internet two graphical file formats are widely used. JPEG (Joint Photographers Expert Group), a true-color format, gives the best results for photographs. GIF (Graphical Interchange Format) has a selected pallet of colors and is good for any other kind of picture. For a given image it gives a smaller file size than JPEG. The figures in this book's chapter 2 "The Agricultural Research Management Information System," are mostly GIF images, captured from screen images with the Windows Alt-Print Screen command or, for better selection of part of the screen, using the SnagIt image capture software program. CorelCapture and many other screen-capture utilities do a similar job.

Vector files contain data described as mathematical equations. They are typically used for line art, GIS, and CAD. A drawing program for vector graphics uses a library of graphical objects ("primitives") such as circles, lines, rectangles, ellipses, arcs, or curves, and the numerical coordinates of the different primitives that are used in the picture. Computationally, it is much easier to scale or otherwise manipulate such a picture than a bitmap image. This is useful in CAD and GIS applications, where pictures frequently need to be resized or shown from a different angle. They are also used in computer games and other animation for three dimensional displays.

Multimedia

Throughout most of this book we have considered the management of information in a text format, augmented where appropriate with graphical formats such as histograms and pie charts. Such a graphical format enables the viewer—the research manager or scientist—to gain a better or more rapid view from data than would be offered by just text or even tables. Figures 1 and 2 in the Executive Summary of this book are examples.

Multimedia uses not only text and graphics, but also still images, sound, animation, and motion video media. It is used for presentations, training (especially self- and distance-learning tutorials) and simulations. Its use in training is of potential interest to us. Some claim that in order for multimedia to be an effective system of communication it must be interactive, i.e., it must provoke or demand responses from the user. It is claimed that it reduces learning time, it maintains a consistent quality with no "off days" like a human teacher, it can be replicated and re-used endlessly (or at least until it is out of date), and it provides more thorough learning—students can learn at their own pace.

The costs of multimedia are twofold. First, it is much more demanding of data storage and processing capacity, and, if transmitted, of communication bandwidth (see table 7.4). CD-ROMs can meet the high data storage need most effectively. For transmission, full-motion TV-quality video can now be achieved with data rates of approximately 1.5Mbps and sometimes even less, but this is beyond the facilities available to most NAROs. Most PCs these days are capable of displaying motion video, although the resolution available depends on the power of the computer's video adapter and CPU. For sound, a separate sound card and loud speakers are needed.

A more serious cost is its very high demand of skills. As multimedia is a relatively young phenomenon (having developed largely through the 1990s), these skills are still scarce and expensive. Most NAROs find it difficult to find, recruit, and retain information specialists in the more basic media of text and graphics.

We may also refer here to the use of the audio function. Although not multimedia, it can be used for example to enable two users to have a voice conversation using the Internet as a carrier of the digitized voice. When both are linked to a local Internet service provider, this can be much cheaper than using the telephone. Video conferencing systems take this a stage further and enable a single person or a small group of people at each of two locations to hold a meeting where all are both seen and heard by all. It can dispense with the cost and time of traveling to one location.

7.3.3 Telecommunications for NAROs

Telecommunication is the transmission by electrical means of information, which may originate in alphabetical, numerical, or pictorial form, from one place to another. The process usually starts with the conversion of the information from its original form into signals suitable for electrical transmission, and ends with the conversion back into a form intelligible to people. Transmission is usually via electrical cable or optical fiber (e.g., for traditional fixed-line telephone and fax) or space (radio, as used for cellular telephones and satellite-supported communications).

An early milestone in telecommunications was the laying of a trans-Atlantic submarine telegraph cable in 1865. It could transmit 12 words per minute (128 bits per second), which may be compared with recent transmission rates shown in table 7.7. Voice transmission came in 1876 with Graham Bell's discovery of the telephone. Submarine telephone cables appeared only in the 1950s. Nowadays they comprise optical fiber; recent trans-Atlantic cables provide 5 Gb per second on each of four fibers per cable and carry 300,000 simultaneous speech conversations.

Today, some 120 years after its discovery, the telephone is still not readily available on many agricultural research stations in developing countries. The problem is illustrated by the variation in telephone density—the number of telephones per 100 persons. The average for the World is 13, that for Europe and the USA around 50, but that for Africa is less than 2. Many journeys are made at a considerable cost of time, fuel, and vehicles that could be eliminated if a reasonable telephone system was in place. Inadequate telephone connectivity is a block not only to spoken telephone conversations but also to access to fax, e-mail, and the Internet, which usually use the same wires. Some major problems have been the maintenance of long-distance telephone wires and the reluctance of governments to liberalize the telecommunications market. New means of transmission and recognition by governments of the national benefits accruing from a competitive private telecommunications sector may help resolve this problem.

Radio communication dispenses with the need for a wire connection. Short-wave radio has long been used for point-to-point communication in special projects, but this provides connection only to one or a few other sites. In recent decades short-wave radio has been used in various new ways to provide access to the public telecommunication system through mobile ("cellular") phones, satellite communications, and packet radio.

Cellular systems allow an individual to connect to the public telephone system by means of a hand-held mobile phone, which transmits to and receives from a local short-wave radio base station. Base stations are distributed regularly so that a person is always within range of one. The range is usually a few kilometers, which makes the system suited to well populated areas. Base stations are all connected to a central controller, which in turn is linked to the conventional wired (fixed-line) telephone system. A mobile-phone user can in principle speak with any fixed-line telephone user or any other mobile phone user. As a user moves for example with a car, bus, or train, the system automatically switches the user's mobile phone to successive base stations. The world traveler working in cities in different continents may encounter different systems in different countries, but cartridges are available to enable adaptation to the local system. GSM (Global System for Mobile Communication) is a European digital standard, which has now spread to many parts of the world. But cellular systems tend to be confined to regions of higher population density; they and the fixed lines are estimated to cover only 20% of the Earth's land area at the end of the century, so those living in remote areas, which includes many research stations, may need to access satellite communications systems.

Table 7.7: Some Examples of Data Transfer Rates

Device	Typical rates* (bits per second)	Notes
Human speech:		
• reading	• 40 (360 wrds/min)	• A trained fast reader can read at 800–1000 words a minute.
• speaking	• 15	
PC reading:		
• diskette	• 380 K	
• hard disk	• 500 K	
• CD-ROM single speed	• 150 K	• The original standard for CD-ROM.
• CD-ROM 32 speed	• 4,800 K	• One of several faster modes now available.
• DVD-ROM disk	• 2,760 K	
• DVD-RAM disk	• 1,380 K	
PC to PC via local area network:		
• Ethernet 10Base-T	• 10,000 K	• Twisted pair wires.
• Ethernet 100Base-T	• 100,000 K	• Twisted pair wires or fiber optic cable.
PC to PC via 33.6K, 56.1K modem and public telephone system (copper wire)	33.6K; 40 K	
PC to PC via Integrated Services Digital Network (ISDN)	128 K	Uses two analog copper wires.
PC to PC via ADSL	640–9,000 K	Asymmetric Digital Subscriber Line.
PC to peripherals:		
• Serial port	• 115.2 K	• (Original RS232 standard: 19.2 K).
• Parallel port	• various	• Originally 960 K.
• SCSI	• 5–10,000 K	• Small Computer System Interface.
• USB 1.0	• 12,000 K	• Universal Serial Bus, 4-wire cable.
• USB 2.0	• 480,000 K	
• IEEE-1394 standard (FireWire)	400,000 K	• 6-wire cable.
Mobile phone:		
• GSM cellular phones, when used for data transfer	• 9.6 K - 64 K	• In Africa restricted mostly to urban areas.
• 3rd generation mobile phones in urban areas	• 255 K+ by 2002?	
Coaxial cable connection	8–10,000 K	
Satellite:		
• INMARSAT - geo-stationary	• 64K	• 10 satellites cover the Earth.
• LEO satellite systems	• 2.4K upwards	• Low Earth Orbit.

* Note that, except for human speech, these are steadily rising year by year.

Satellite communications enable a person to make telephone calls, send e-mails (see below) and faxes, and connect to the Internet, from anywhere. Some units are as small as a brief case, weigh only 3.9kg, and connect to a laptop or desktop computer to provide all these services, at a data transmission rate of 64kbps. Potentially they offer an alternative to poor local telephone and Internet services, but at the time of writing, cost, particularly usage charges, limit their use for NARS.

Satellite communications are of various types. The first to become available made use of "geostationary" satellites. These rotate around and with the Earth every 24 hours and therefore remain above the same point on the Earth's surface, at an altitude of some 36,000 kilometers. Each such satellite therefore has its own "footprint" of the Earth's surface, in which two-way traffic flows between it and ground stations. They lie in the equatorial plane and three satellites can cover the world, excluding the polar regions. The user connects to a geostationary satellite by means of a portable short-wave radio transmitter with a parabolic reflector aimed at the satellite. Such a device is the size of a briefcase and costs around $3,000; calls may cost $3 per minute, plus a monthly subscription, but these costs are likely to reduce with competitors coming on to the scene. The distance between Earth and satellite causes a delay in message delivery quite noticeable in telephone conversation.

Low Earth Orbit (LEO) satellites orbit the Earth several times per day, at a much lower altitude, typically 780 km. At this closer distance smaller transmitter-receivers can be used, in the form of a hand-held phone with a small antenna, though these may weigh 300–500 grams compared with the 150 grams or so for a cellular phone. Subscribers to a single LEO satellite system use a store-and-forward process: a message is passed to the satellite as it passes over the sender, and is stored on the satellite until it passes over the recipient, when it is transmitted down to a nearby ground station into the local telephone system and on to the recipient. Multiple satellite systems allow instantaneous ("real-time") two-way communication by automatic switching of a message between successive satellites. Globalstar, for example, uses 48 operational LEO satellites. Dual purpose handsets are available that can be used with the terrestrial cellular system and that switch automatically to satellite connection when the user roams outside of the cellular coverage area. LEO satellites have a shorter life expectancy (5–7 years) than geostationary units.

Medium Earth orbit (MEO) satellites orbit the Earth at an altitude of around 10,000km. In one configuration 10 satellites are distributed equally between two planes, each inclined at 45 degrees to the equator and at right angles to the other, to provide complete, continuous overlapping coverage of the Earth's surface. Each satellite will circle the Earth approximately once every six hours.

Both geostationary and orbiting types of satellite retransmit received messages back to a ground station on Earth, which passes them into the conventional telecommunications network and, if appropriate, into another local cellular network. The geostationary satellites are generally sited over parts of the world with high telephone and data traffic but the LEO orbits inevitably pass over both high-volume traffic regions mostly in the North and over low-volume areas in the South. Their operators may offer lower rates in the latter to make at least some use of their systems while they pass over low-traffic regions. Many NARS may find benefit here. Contact with satellites is best with direct line-of-sight as buildings impede the signal. Satellite systems are still expensive, but this may change as more systems become operational. Transmission times and therefore costs can be reduced by compression techniques built into modern systems (box 7.3 gives an example). Original commercial plans for satellite systems have proved to be optimistic with Iridium filing for bankruptcy protection in August 1999 and the competing companies adopting more conservative plans as a result.

Satellites are also used to determine the position in terms of longitude and latitude, or a map reference, of a ship or vehicle or person carrying a GPS (global or "geo"-positioning system). The NAVSTAR system uses 24 satellites, each of which broadcasts time and position data continuously. GPS use in agriculture is described in section 6.4 "Geographic information systems."

Packet radio combines radio and computing technologies in a low-cost system that lets personal computers communicate with each other using radio frequencies (Zipp 1994). It is appropriate for bridging the "last-mile" gap, where no telephone line exists to remote sites. It works best up to a range of 150 kilometers but this can be extended to 1,000 or 3,000 km in some circumstances. The CGIAR-CGNET system uses packet radio for its links to Burkina Faso, Cote d'Ivoire, Ethiopia, Kenya, Mali, Niger, and Zimbabwe, and several NGOs use them when working in remote areas. The example in box 7.3 describes how such technologies have provided scientists with access to a range of scientific journals in a library some 3000 kilometers away.

Radio links are generally more suited to voice than data communications such as e-mail, fax, and Internet access, where the slow speed, usually not greater than 6kbps, compares poorly with typical speeds of 33.6kbps and 56kbps for modems on fixed-wire links. Where the latter is not available, radio can of course make the difference between no connection and a limited connection.

E-mail

Electronic mail is a method for sending electronic messages (i.e., messages that have been typed into a computer) over electrical cable such as telephone wires and computer communication networks. It is usually much quicker than postal mail, in both preparation (no need to address, stamp, and post an envelope) and transmission (seconds or minutes usually instead of days or weeks). It may not always be as quick as a direct telephone call, because the transmitted message is merely deposited in a remote computer (a file server), where it waits until the intended recipient connects their own computer to the file server to check for waiting mail. On the other hand, if someone is out of their office, a telephone call may miss them altogether (unless a message is left on a recording/answering machine).

An e-mail can have a file attached to it: a word-processed document, a spreadsheet, a data base file, or even a picture. This provides a very efficient way of passing information. This facility has its limits; some sites are connected by lines with rather low data-transfer rates, and a large file can cause serious delays to other mail. For this reason some transmission services impose a file size limit, for example of 50KB or 1MB; files above this size are rejected.

Box 7.3: How Communications Technology Provided Scientists with Access to a Remote Library

On 14 May 1998, Ester Lwanga-Semakula, Coordinator/ARIS, notified Prof. J. Mukiibi, Director General of Uganda's National Agricultural Research Organization, through the AfricaLink network that Namulonge Agricultural and Animal Research Institute (NAARI) was now on e-mail: "I have the pleasure of informing you that NAARI is now accessible on E-mail: via radio communication equipment, which was installed at NAARI on Monday, 11 May, by Bushnet." The high-frequency (HF) radio link was the crucial connection between the research station, NAARI, and Bushnet, an Internet service provider, in Kampala, Uganda.

On the other side of the continent, at the International Institute of Tropical Agriculture (IITA), Nigeria, the head of IITA's Information Services Program seized the opportunity to test the potential of AfricaLink. The e-mail network was established by the US Agency for International Development to facilitate exchange of research information, especially between international agricultural research centers and national research systems such as NAARI. He sent an e-mail to NAARI inviting a request for information from IITA's library, one of the strongest in sub-Saharan Africa. He soon received a response from a NAARI scientist for specific pages from eight scientific journals.

The request was immediately sent to the IITA library, where a librarian conducted a search of their database. Six of the eight journals were in the collection. The specified pages were photocopied and delivered to the Multimedia Unit for preparation for sending though the Internet. In the Multimedia Unit, text was converted to small files using an optical character reader. Tables and figures, however, posed problems.

E-mail by HF radio can be costly, mainly because an expensive radio system is used to send and receive messages, and the data transfer rate of HF radio is rather slow. It might require two to five minutes to receive a single small image such as a table from a journal article. In comparison, a computer communicating over a telephone line can ordinarily receive the same image in only four to 10 seconds.

The Multimedia technical team studied the problem and devised means to reduce the file sizes of both images and texts so as to produce the smallest possible file sizes for transmission. After assessing the various combinations of file formats and compression systems available, as an example, a 200-kilobyte file was reduced to a mere 18. During this experimental period Bushnet donated its services, so NAARI paid nothing for the file transfer. Compression allowed transmittal of the article for approximately US$5. Had compression not been used it would have cost 12 times as much.

IITA has subsequently responded to additional requests from NAARI and from other research institutes in sub-Saharan Africa. Requests have been passed on to other international agricultural research institutions, including the University of Georgia, USA.

Note: based on Philpot (1998)

A single e-mail can be sent simultaneously to several addressees, either by typing in the e-mail address (see below) of each intended recipient, or using a previously established group of addressees.

Connection

At its simplest, e-mail may be used within the confines of the organization, i.e., within the organization's own local area network (see the next section below). But increasingly organizations provide access through a "gateway" to the Internet (see below), which enables communication on a worldwide basis. In either of these cases connection of a person's personal computer is through the local network. A stand-alone computer, for example at one's home, is connected to the public telephone system through a modem (explained earlier in section 7.3.1) to send e-mail. In all cases it is necessary to have e-mail software, called a mail client, such as ccMail, Microsoft Exchange or Outlook, or Qualcomm's Eudora, or an Internet browser that performs the mail client function. Additionally one needs an account, analogous to having a telephone number. All e-mail accounts have a unique address, which follows a prescribed format: <name>@<domain name>. The <name> is the account or account holder's name. The <domain name> identifies the file server where the account holder is registered and where his messages are stored, rather like a post office box, pending his logging in over the network to collect them. The domain name structure is similar to that used on the Internet (see below).

Local area networks

A local area network (LAN) is a high speed communications system between a number of computers and other data processing devices within the confines of a building or a group of adjacent buildings. Typical speeds are between 1 and 100 Mbps, and most LANs have a range of up to 800 meters. The medium is coaxial cable (as often used for connecting a television to an aerial), a pair of twisted copper wires as used for telephones, or, in larger LANs, an optical fiber.

A LAN can offer the following:

- exchange of data, e-mail, files, etc. between computers within an organization
- sharing of expensive items such as a laser printer. Typically a group of users in a single room or nearby rooms can share a single printer. If there is a problem with that printer then it is possible to send material for printing via the LAN to the next nearest printer.
- access to large repositories of data such as the library catalog to find out if the library has a particular publication, and even to find out its current loan position: in or out
- access to special application programs residing on a file server
- access via a "gateway" to external systems such as other LANs (as part of a wider area network or WAN), or public networks such as the Internet, or the telephone system for sending e-mail and faxes from the desktop computer.

These facilities all contribute to greater productivity.

Installing and maintaining a LAN is a specialist technical function and is beyond the scope of this book. Lauder (1994) gives a good overview.

Intranets are the results of the integration of Internet paradigms (see next topic) and standards with an organization's existing network, desktop, and server infrastructure to create more effective business management systems within the organization.

The Internet

The Internet is a worldwide collection of networks, communication protocols, and software applications. It is emerging as a low-cost means of information sharing between almost any two computers that are connected to the public telephone or other telecommunication system. Some believe that ready access to information will become the catalyst that transforms economic and social structures around the world and supports fast-paced sustainable development. The Internet is one way in which such access could be provided. Others point out that similar aspirations were prompted by earlier technological innovations such as the development of the telegraph and the railways. Great advantages did accrue but were far from universally or evenly distributed.

There are two main ways of communicating through the Internet. First, any connected computer may send and receive e-mail to any other. Second, significant numbers of large computers act as hosts (called servers) for large repositories of information on a wide range of topics. The owners allow these servers to be accessed by the public using the telephone system to transmit requests for data into the server and transmission of the requested data back to the inquirer. Very many public, educational, and commercial organizations provide such a service.

Some servers have structured their contents to provide for discussion forums, which nowadays cover a wide range of subjects. It can be very useful for example if one has a question on database design. There are forums covering most of the more common database softwares, and a question posted there one day is very likely to prompt some useful answers from more experienced users over the following days. One can also learn much from connecting to such a discussion forum and reading the questions and answers of others on topics of interest.

World Wide Web. The "Web" is that part of the Internet served by more sophisticated protocols that enable hypermedia to be used. A hypermedia facility allows various formats of documents on a remote computer to be transferred to the user's computer and presented on the screen in a readable display. It makes use of the hypertext transfer protocol (HTTP, shown in Website addresses). The Web is a more recent development but it is rapidly becoming more widespread across the Internet, replacing the older methods of access and interaction. It is sometimes used as a synonym for the Internet.

Connection. Local commercial Internet service providers (ISP) usually provide connection to the Internet through a monthly or annual subscription. In some countries, ISPs provide such a service free of charge to the public; revenue is derived from advertising. When an organization with its own internal LAN or intranet (see earlier) connects to an ISP through a gateway, all users on the LAN/intranet can access the Internet through this gateway. (Given the costs charged by the ISP, the organization may of course restrict this to certain users or for certain time limits.) When an individual wishes to be connected directly to the Internet, a computer can be connected to the ISP via a modem, which in turn is plugged into the telephone wall socket. In either case each connected computer needs a Web browser software to be installed. There are two browsers in widespread use: Netscape Navigator and Microsoft Internet Explorer. New versions of each are periodically released.

Searching. The Internet now makes available an enormous and ever-increasing stock of information, but in order to control time and connection costs it is important to be efficient in searching. It is all too easy to just browse from one interesting site or subject to another. Two things help search fast and efficiently: any one of the several powerful search engines now on the Internet, and search skills in selecting and using search terms (keywords).

A search engine consists of a very large database of Websites, with keywords representing the contents of each site, and a search mechanism that regularly scans the World Wide Web for its sites and their contents so as to update the database. It allows users to connect to the database and search this single source for items of interest. The user is provided with the Web address of each such item and can then rapidly connect to each selected address on the Web; the desired items are transmitted for display locally on the user's screen. If desired, items can be saved in part or in whole to the user's computer for closer study off-line (i.e., disconnected from the ISP and the telephone line).

Conducting a search. Search terms are words or phrases that the target documents will contain. Different search engines use different sets of search procedures. It may be best to become accustomed to one or a few search engines and become familiar with their requirements and facilities. The following are examples of well-known search engines:

- Altavista: www.altavista.com
- InfoSeek: www.infoseek.com
- Yahoo: www.yahoo.com
- Lycos: www.lycos.com
- Hotbot: www.hotbot.com

Portals. The sites of some search engines have been developed to be far more than a directory of and pointer to other sites. They provide a range of popular information at their own site. Typically their front page offers a "contents" list with links to the various subject matters covered. Some such "portals" take on a specialist subject area.

7.4 Managing information technology

7.4.1 Procurement

Important aspects of procurement are the detailed specifications, the process of obtaining quotations and choosing one of them, inspection of what has been delivered, and whether equipment should be imported or purchased locally. Each of these will be considered in turn.

Specifications

Guidelines on specifications for the components of a PC have been discussed above in section 7.3.1. A summary is given in table 7.8, which offers an example of a typical stand-alone personal computer in 2000. Bear in mind most of the performance numbers go up each year and that many machines bought in recent years with lower specifications still give good service. It depends on what you need to do with it. For word processing and sending e-mails the older and slower Pentiums and even old AT, 386, and 486 models can be quite sufficient. If the cost of upgrading to a larger machine cannot be met, the next best thing is usually to increase the memory, the cost of which is nowadays quite modest, followed by a larger new or additional hard disk.

Computer magazines offer reviews of current hardware and software. These reviews often include advice on recent innovations and are worth consulting before making any significant purchases.

Purchase

There can be advantages in local rather than overseas procurement. If a problem develops it can be easier to deal with a local supplier, especially if there is a prospect of further sales in the future, than with one who is thousands of kilometers away. It can take a long time for overseas purchases to arrive, which in these times of declining hardware and software prices may prove disadvantageous. Tendering procedures usually take longer for overseas procurement. It may be that by the time an overseas order is delivered the model ordered may be somewhat obsolete and could have been purchased today rather more cheaply than the quotation agreed say six months or more earlier. In some localities local purchase may not be possible because of the lack of local suppliers that can offer a good service.

Overseas procurement comes up against the problem of importation and custom dues. Buying through a local dealer passes the chore of customs clearance to them

Table 7.8: Suggested Specification for a New PC for MIS Use at a Research Institute

Item*	**Specification**
Processor	Pentium 200 MHz
Main memory (RAM)	64–128 MB
Monitor	Color, 17" with 1,024 x 768 pixel resolution
Hard disk	6–12 GB
Diskette drive	1.4 Mb 3.5".
CD-ROM or CD-R or CD-RW	24x or 32x speed (3,600/4,800 KB per second).
Tape backup (optional)	Preferably sufficient capacity to backup the hard disk on one cassette. Not needed with CD-R or CD-RW.
Modem	55 Kbps

* See section 7.3.1 for descriptions of each of these items.

although this extra service will need to be paid for in terms of a slightly higher final bill. Some countries have effective customs clearance agents, and some larger NAROs have an officer delegated to handle this function. Cases are known where consignments have been held up for months due to problems of clearance at customs, and some consignments abandoned because there were no funds to pay the charges. Some donor projects have agreements with the government to have custom duties waived, which is clearly an advantage, provided the procedure is well understood and accepted by all concerned.

If there are promising local suppliers, invite quotations for any substantial order from, say, three suppliers. Cheapest is not necessarily best; consider also the credentials of the supplier: how long have they been supplying IT equipment, can they provide names of any customers that one can call on to assess satisfaction, do they carry any spare parts, and do they have any trained service technicians. Ask to see the workshop and the spares store. These questions also help to deflect tendencies for orders to be placed with persons with greater interest and skill in trading than in information technology!

Donor organizations sometimes insist on international competitive bidding, at least for large orders. The above shows that this may not always be necessary or wise.

Checks and tests on delivery

It is important to unpack and test the equipment promptly after delivery so that any faults can be identified and reported to the supplier without delay. First look for breakages, second that what has arrived is what was ordered, and third that it functions properly. After initial visual inspection, check all items against the invoices, and if in order, assemble and test.

Some faults may derive from the manufacturers. This was demonstrated in a survey of PC reliability and service (PC World, 1998), which recorded the percentage of machines that did not work properly when first turned on. It seems such initial faults are found in about 5–13% of new machines. Fortunately, complete failure—"dead on arrival"—is much less common but does occur occasionally. The survey showed significant differences between manufacturers, so there is merit in checking such magazine-sponsored evaluations prior to making a large computer procurement.

In some countries there is a risk of receiving "recycled" items that may be old (and obsolete), second-hand, and even stolen. Examine the packaging for signs of it having been opened. Sometimes this may be legitimate, for example when the supplier has previously assembled the system for testing. But even in that case one may expect the documentation to be in its original plastic wrapping. Examine the outside of the systems case, monitor, and the keyboard for dirt and scratches. If the systems case is of the "tower" style that stands vertically on the floor, look underneath it for scuff marks. Turn the keyboard upside down over a clean sheet of paper and tap gently to dislodge any accumulated dirt. If this is more than very slight, you have reason to doubt that this new equipment. Examine the contents of the hard disk for the dates of system files and for any obviously user directories and data. Check the version numbers of software. Make a note of the serial numbers of hardware items and check these against delivery notes.

Finally, the PCs should be operated along with the main softwares. Printers should be tested in action, and if modems are included, an e-mail should be sent so that each hardware and software component has been checked.[12]

7.4.2 Maintenance: Information technology in harsh environments

Fortunately the modern PC, at least from the better known manufacturers, is of generally high quality, and in many cases will, with care, run for years with little attention. But it needs only one component to fail to bring the machine to a halt. There are a few rules that will increase the chance of a trouble-free existence. Here we summarize the main threats to PC reliability and preventative measures to reduce problems to a minimum, particularly in developing-country situations, where environmental conditions are often more severe.

Backing up data

Backup copies should be made of all important data at regular intervals—typically daily or weekly depending on how often it is updated-and kept in another place. The writer knows of several cases where, due to the failure to take this simple if irksome precaution, much grief was caused by the loss of the only copy of important data by system failure, theft of the PCs, or their loss in a fire. The commonest medium on which to store backups nowadays are tape cassettes and (re)writable CD-ROMs (see section 7.3.1 "Hardware").

Electricity supply problems

Perhaps the most serious threat to PCs in developing countries is the unreliability of the electricity supply. There are three ways in which the supply can be problematic: complete failure of the supply (blackouts), too low a voltage (brownouts) for the computer to accommodate, and too high a voltage, causing damage to the computer. If blackouts are prolonged they can be solved only with stand-by generators. These may range from a small 4-horsepower petrol generator to serve a single user occasionally, to large diesel-powered units to supply all of a station's computing equipment or even the entire station's electrical requirement. They need to be sited carefully so that the noise and exhaust fumes do not interfere with people.

Variation in the voltage includes surges (increases in voltage) and spikes (large but short-lived voltage increases) and reduced voltage, due for example to overloading of the supply in a locality. A significant spike can burn out the transformer of a PC. The cheapest guard is a spike protector plug, which looks like an ordinary 13 amp plug except that is somewhat deeper in order to hold a condenser, which absorbs small spikes in the electricity supply. Those meeting the USA Underwriters Labora-

12. In a large national agricultural project, the author established the practice of testing all new equipment within 24 hours of delivery, withholding payment until the tests were complete, but then ensuring payment was promptly made on all items passed as satisfactory. The dealers (all procurement was done through local dealers) soon came to accept this procedure, particularly as they learned that payment was made promptly and in full for all correctly supplied equipment.

tory UL1449 standard have to protect against a surge of 6000 volts and a response time of less than 1 nanosecond.

A voltage regulator offers more protection and can allow a PC to continue to run when the mains supply voltage drops, within limits. It comprises a transformer with several coils between which it switches to accommodate to a varying input voltage. These devices provide some protection to the PC but do not save your most recent data, should the voltage drop too far or fail completely. To do that a UPS (uninterruptible power supply) unit is needed.

A UPS unit contains a battery bank as an alternative to the mains electricity supply. When the mains supply fails to provide a sufficient and steady voltage, a UPS switches to the battery supply quickly enough to prevent the computer from being affected. This then provides time for the user to save to disk whatever tasks are being conducted, before switching off. Typical UPS arrangements are designed to provide 20 to 30 minutes of electricity supply from their batteries. The cheapest models provide only an alternative supply on the failure of the mains, but more commonly, a UPS will include a voltage regulator as above.

A small UPS unit of 500 VA can support two PCs and can therefore be located in the users' office. Larger units such as 2000 VA may be used for a computer room, the actual power rating being matched to the number of PCs, printers, etc. being served. Such larger units are best sited outside the room due to the noise of the transformer. In that case provision must be made to warn users that power has switched from mains supply to the UPS and that they should save their work and prepare to switch off. One simple way is to have a light bulb that is always on when the PC is in use, and that runs from the mains supply. When this fails the light noticeably goes out. A small UPS in the room typically emits a warning sound when it switches to its batteries.

A UPS can also fail, so provision should be made for their repair. The life of the lead acid battery usually found in a UPS is shortened by temperatures above around 40 degrees C. Placement of the device, especially the larger models that generate significant heat within their case, can therefore be important. Ensure good ventilation around them and shade from direct sunlight.

Protection against a lightening strike may also be considered. This typically is in the form of a lightening conductor running externally from the highest point of the building to an "earth" in the ground.

Other threats to the personal computer

Switching a personal computer on and off several times a day may cost more in reduced life than the savings in electricity as the system heating up and cooling down affects its components: "Thermal expansion and contraction remain the single largest cause of component failure" (Mueller 1998). Unless electricity is costly, it may be best to turn a PC on at the start of the day and off at the end of the day, or try to arrange tasks so that they are done together in one session, reducing the number of heating and cooling cycles.

Another threat to PCs, serious in some countries, is theft. Rooms have to be fitted with secure burglar bars to the windows and with reinforced doors and locks. Temptation is reduced if the PCs are out of view of those passing by. Important data must be regularly backed up to diskettes, tape cassettes, or other media and stored away from the PC, preferably in another room or building.

Dust can be a factor, although its effect can also be exaggerated. The modern PC seems to be remarkably resilient. The writer encountered very significant dust in West Africa during the annual harmattan, a strong wind that blows in from the Sahara during the dry winter period and brings with it much dust. A layer of dust could be seen on desk tops the next morning. We had covers made up locally for the PCs from soft plastic sheeting, and these were put on at the end of the day to reduce deposit, and kept there whenever the room was dusted and swept out. When a PC was opened to replace parts or for other repairs, we would take the machine outside and use an electric air blower to blow out the dust prior to commencing the repair. Diskettes were always kept in a storage case. With these simple precautions we saw no evidence of dust being a serious problem over three years with some 20 PCs. However, different environments will have different constituents of dust. Mueller (1998) warns of various adverse consequence of both dust and cigarette smoke, and he advises on periodic blowing out with compressed air (e.g., every two years in a reasonably clean environment, more in others).

A survey described in PC World (1998) reported the relative frequency amongst components that gave a problem in normal use (see table 7.9).

Notebook / laptop computers

Notebook computers, which are portable and can run on batteries, are largely immune from electricity supply variations, but they face their own problems. As they are being carried around, they are sometimes dropped or struck against another object. This may stop them completely; the risk is greater if they are moved while in use, and particularly if there is a diskette or CD in place. It is better to remove these to reduce the chance of a head crash (PC World, 1998a). It is better still to switch them off when moving them. Spilling a drink on to a laptop keyboard is potentially much

Table 7.9: Percentage Reported Causes of Main Computer Hardware Problems

Component	Percentage of problems
Hard drive	27
CD-ROM drive	15
Monitor / screen	13
Motherboard	9
Modem / fax board	8
Floppy drive	6
Other	22

Source: PC World (1998)

more serious than doing so on a PC keyboard. In the former case the liquid can quickly seep into the machine itself; in the latter, it is possible to wash out the (disconnected) keyboard in clean water, ideally rinse in distilled water, dry thoroughly and replace. In the worst event only the PC keyboard has to be replaced.

Maintenance contracts

Some suppliers offer a maintenance contract that guarantees a certain level of service against a regular fee. Consider them with caution. The main guarantee is often that the customer pays the supplier a fixed sum in advance and at regular intervals. A better "guarantee" may be found through having a small number of competing agents vying for your custom.

Threats to software: electronic viruses

The biggest software threat in many places, especially where a large population of users share a number of machines, is that of software viruses. These are computer programs written by irresponsible programmers with the intention of interfering with the regular operation of the computer. They can only enter a computer when and if virus-"infected" software is loaded into the machine. This is usually from a diskette but it can also arrive through a modem or via a local area network, either with e-mail attachments or with files downloaded from the Internet.

There are two main lines of defense, and both should be used. The first is to avoid excessive exchange of diskettes, particularly from unreliable sources. The second is to have virus protection software installed on the computer. If computers are networked and access e-mail and the Internet via the network then it makes sense for the network itself also to be fitted with a "fire-wall" against incoming viruses. Anti-virus software needs to be updated periodically to deal with new viruses. All incoming diskettes should be tested for the presence of a virus before any files are transferred to a personal computer. Similarly, files imported on-line (e.g., attached to e-mail messages or downloaded from the Internet) should be checked with the virus protection software before opening or using them. Files sent as e-mail attachments, particularly document files, are liable to host a virus in the form of a macro. It is safer to send such files in rich text format (RTF), which cannot hold a macro. Virus-checking software can be set to monitor the PC continuously, rather than scanning just on booting up.

Health aspects

There are some health risks associated with the heavy use of personal computers. Repetitive strain injury (RSI) is cumulative damage to muscles, tendons and/or nerves of any of the hand, wrist, forearm, shoulder, and neck due to prolonged, repetitive, and forceful or awkward hand movements (Pascarelli and Quilter 1994). The forceful or awkward component of such movements is often caused by poor posture and technique. The adoption of good practices and prompt attention to early symptoms of pain can avoid more serious damage for which the cure may be long and painful and can in some cases even require surgery. Pascarelli and Quilter describe exercises that can help relieve the stress of keyboarding, along with much other practical advice:

One should sit up straight and position the monitor so that the top of the screen is level with or slightly below the height of the eyes. A chair with good lower back support can be helpful. The heights of the chair and the keyboard should be such as to allow the forearms to be approximately horizontal and in line with the hands so that the wrists do not have to bend significantly, in any plane. The forearms and wrists should be free, not resting on the desk surface, so that your whole arm and shoulder are used to move your fingers around the keyboard. Try to maintain a position in which the axis of your forearm, hand and the first joint of the middle finger maintain approximately a straight line. Have the keyboard flat, not tilted towards you as this can force the hands up and the wrist to be bent back. Avoid long periods of continuous computer use; break off regularly, perhaps every 25 minutes, to give the back, arms and hands, and eyes, a rest, by walking around, relaxing the arms and hands, and focussing the eyes on distant objects. This need not be lost time; one can switch to other tasks.

These remarks apply also to the use of a mouse, which is believed to also be a major cause of RSI. Many of the functions of a mouse, such as those of menu selection, can also be effected directly from the keyboard by using "shortcut" keys. Search the Help facility of your software for this term to learn about those available.

Good positioning of the monitor in the room can make a difference too. A monitor is best positioned at right angles to, and some distance away from, the window. A window behind the operator may reflect uncomfortably from the screen, and a window in front of the user may make it harder to read the screen. As with a television there are controls to adjust for brightness and contrast. Those who wear spectacles for reading, with a typical focal length of 40cm, should be aware that an optician can prescribe lenses with a somewhat longer focal length, say of 55–60cm, particularly for viewing a computer monitor.

7.5 Elements of a database system

Introduction

A database is a collection of related data with a minimum of duplication to provide a common pool of information to serve one or more functions. It has a structure that keeps the data separate from the programs that use it and enables different information needs to be served from a single store of data. A database is able to store and manipulate both text and numerical data. In this it overlaps with the functions of word processor and spreadsheet software respectively, but it goes further than both in the degree to which it can store and manipulate data. Over the years several models of database have been developed. Perhaps the simplest is the flat file model.

7.5.1 The flat file database

A flat file database resembles a list or table of items. Each item occupies one line in the table. Attributes of those items occupy the columns. See figure 7.2—this could be the membership list of a club.

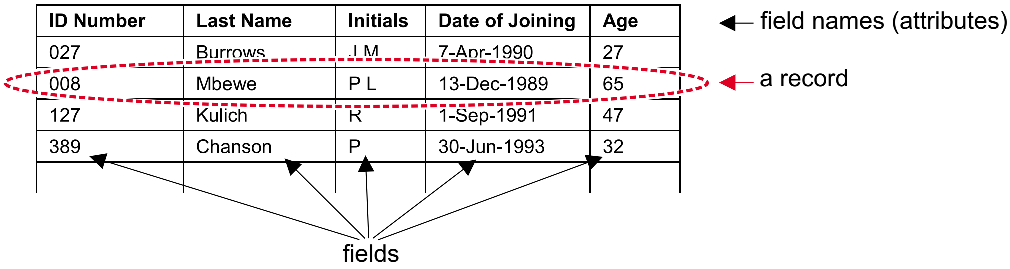


Figure 7.2: Example of a flat file database

Each row in the figure represents a record, in this case a person. Each column represents an attribute of the person and is called a field. For each table the number of fields is usually set at the time the table is established, but records may be added at any time—in this case each time a new member joins the club.

For simple data storage and processing the flat file system is fine. It is quick to set up and easy to understand. It is often the best choice if there is only a single type of object about which data is being stored, e.g., your collection of CDs, a simple membership list of a club, etc.

With greater complexity such as the need to store data on several different objects, or when there are repeating groups of an attribute of an object such as children of members or repeated donations from subscribers, then the flat file system soon runs into difficulty (Date 1990, for example, discusses this further).

Major alternative database structures include hierarchical, network, and relational models. The relational model was the last of these to be developed. Because of its inherent simplicity and functionality it has become the mostly widely used. If an information system is developed in one vendor's relational product, and that product is withdrawn from the market, it is not difficult to move the data content part of the system to another relational product. Work may be needed to reconstruct the user's "front end:" the menu system, and query and report structures. During the 1990s another database model, the object-oriented database, emerged. It was designed to handle complex data types but in doing so it lost the simplicity and powerful querying facility of the relational model. For these reasons its predicted replacement of the relational model did not materialize. The relational model has prevailed though large-scale mini-computer and mainframe systems such as Informix, Oracle, and IBM's DB2 include object features and are sometimes referred to as extended relational or object relational databases.

7.5.2 Text-oriented database management systems

by Hugo Besemer

There is a special category of flat file database systems that have been designed to efficiently retrieve data in datasets that consist mainly of text, such as bibliographic datasets, project descriptions, and full text databases. The many different programs in this category share one feature: for retrieval purposes they build an "inverted file," a permanent index in which (in the case of a bibliographic database) each title word or each keyword may be a separate entry. With each entry, a pointer is stored that specifies the location of the record. These keys enable the system to retrieve records rapidly. They can be combined with the logical or "Boolean" operators AND, OR, and NOT. The inverted file may consist of several physical files.

Text-oriented systems do not impose the restrictions of field length and repeated occurrences of identical data prohibited in the relational database model discussed below. This is very useful for datasets like bibliographic data or project descriptions, where the length of titles or abstracts may vary considerably and where there may be an unknown number of authors or keywords.

This category of database management systems has a number of subcategories:

- "Classical" systems without a predefined data structure and Boolean retrieval facilities. Examples are CDS/ISIS, MINISIS, Asksam, Inmagic's DB/TEXT.
- Reference management systems. These are often popular with scientists, because references from different data sources and output can easily be formatted in different "journal styles" to suit their publications. Example are Reference Manager, ProCite, Endnote.
- Full-text systems, such as FolioViews, which are much used in the legal profession to record legal cases.
- Systems that use other retrieval systems than Boolean logic, such as vector-based or probabilistic retrieval, for example PLS.

The main disadvantage of text-oriented database management systems is their "data redundancy." For example, in a database of project descriptions, the address information of executing or funding organizations is stored with each individual description. In a relational database system there is usually a single table for organizations, in which each piece of information needs to be stored only once. Individual project descriptions will refer to this table. Any change, such as an organization's change of address, therefore needs to be made only once. However, in a text-oriented database management system, such a change of address will have to be made in each project description.

A text-oriented database management system may, because of its good retrieval facilities, be a good way of making a large collection of project descriptions available to the public, but for project management a relational system is a much better solution. One approach might be to manage the data with a relational system and export it regularly to a text-oriented database management system to provide a retrieval interface to the public. Bibliographic data can be handled very well with a text-oriented database as they are "flat" by nature.

7.5.3 The relational database

The relational database model uses rules of data structure and data integrity in order to provide information from otherwise uninformative data.

In the relational model data is stored in two-dimensional tables. Each table typically has data relating only to a single object. An answer to a particular question may require data from a number of such tables, which are then temporarily linked for the purpose of generating that answer. Such links serve no further purpose: the database itself consists of separate tables. Figure 7.3a gives an example of a research information system based on the relational model that includes a table for research activities and a table for scientists.

Each of these tables are straightforward record sets of a particular entity. In practice, of course each would contain a large number of records, hundreds or, in a large NAROs, thousands. A typical MIS would also have many tables.

The relational database model allows us to extract information from two or more such tables. Supposing we wish to know the highest degree of the scientist managing each experiment. We can use the MIS to generate a temporary table that combines those fields from each table to give us this information. The result might look like figure 7.3b.

This example shows how the relational model stores information in simple tables and enables such tables to be combined to generate answers to particular questions.

The main advantage of the relational model over older database models is its inherent simplicity. It is therefore very simple to enter, retrieve (as in the above example), update, and delete data. Its advantage over the flat file model is that is much more flexible for nearly all applications. However, the tables themselves have to conform to certain rules to ensure the smooth operation of processes, such as retrieval, updating, and deletion of data. This does add complexity to operating the relational model, even though the underlying structure is simple. These rules are introduced in the following pages.

Logical, physical, and user's views

Three views of a database are distinguished: the logical, the user's, and the physical views. The user's view is the view that a particular user is presented with for a particular purpose. For example, a user may wish to see just the names and ages of members of a club (see the simple flat file table above), perhaps ranked in order of age. Figure 7.4 shows the resulting user view.

The presentation of this particular view has no effect on the structure of the underlying database as shown in figure 7.2, which shows the logical or conceptual view and embraces all the data elements, whether they are visible to the user or not.

But figure 7.2 is not a good image of the physical layout of the data on a computer disk: the physical view. Data on a disk is stored in concentric circular tracks of magnetic charges. Users have little need to be concerned with this physical view, but database managers may need to think about it. For example in a big database of thousands of records, where several records have been deleted over time (perhaps when members have left an organization), then searching the database can become slow due to all the gaps. This can be remedied by reorganizing the physical storage of the data on the disk. This is a routine task and part of system maintenance.

Research Activities Table

Experiment ID	Experiment Title	Objectives	Scientist
0346	Maize early maturing variety trial		Dhajeel
0298	Plant competition mechanisms		Valkode
0456	Banana diseases survey		Mulenga
0378	Boron deficiency study in cotton		Phiri
0413	Maize / pasture legumes intercropping trial		Valkode
...	...		...

Scientists Table

Scientist	Date of Birth	Highest Degree
Phiri	1-Jan-1956	PhD
Mulenga	23-Sep-1961	BSc
Dhajeel	15-May-1957	MSc
Valkode	11-Nov-1949	PhD
...	...	...

Figure 7.3a: Two tables in a research information system based on the relational database model

Temporary Table

Experiment ID	Experiment Title	Scientist	Highest Degree
0346	Maize early maturing variety trial	Dhajeel	MSc
0298	Plant competition mechanisms	Valkode	PhD
0456	Banana diseases survey	Mulenga	BSc
0378	Boron deficiency study in cotton	Phiri	PhD
0413	Maize / pasture legumes intercropping	Valkode	PhD
...	...	...	...

Figure 7.3b: Derived table of experiments and highest degree of supervising scientist

Last Name	Age
Burrows	27
Chanson	32
Kulich	47
Mbewe	65
...	...

Figure 7.4: A user's view showing some of the data in figure 7.2

In designing a new information system a systems analyst starts with reviewing the users' information needs. This should encompass all of the aggregate information in all user views. From this picture a system designer will move to the conceptual or logical view of a system that provides for all the users needs. This design is then passed to the programmer, who uses database or programming language software to build the system based on the logical design with outputs as requested by the users.

The various features of the relational model described below apply to the external (user) and conceptual views, not to the internal (physical) view implemented on the magnetic medium.

Features of the relational model

There are three important characteristics of the relational database model concerning structure, data integrity, and manipulation. Before giving rules for these three areas, we will define a few terms used in relational modeling.

The structural terms introduced above in the flat file model apply here too: a relation is a table that, like a flat file, has a header of field or attribute names, and a body of data made up of records. Also, the attributes are usually set at the time of designing the system, whereas new records are continually added as part of the normal use of the system. But a relation differs from a flat file in that it is subject to a number of rules or constraints that are necessary to make relations relate to one another consistently.

The primary key of a table is a unique identifier for a record, made up of one or, in the case of a composite primary key, a number of fields. Thus specifying a particular value of the primary key will identify a particular record. In figure 7.3 the scientists' last name could be regarded as a primary key in the scientists table, but it would fail in its function of providing a unique identifier of records if there were two people of the same surname. A special ID (identity) field would be a better choice. In the full relational model each table that represents an object has a primary key.

(It is possible, if unusual, to have more than one candidate primary key, each able to uniquely identify a record. Consider a table of the chemical elements with three fields: name, symbol, and atomic weight e.g., sodium, Na, 23. Each field is able to uniquely identify a record. All fields are therefore candidate keys. One is then chosen as the primary key, and the others are referred to as alternate keys.)

A domain is the pool of all possible values of an attribute. Thus in figure 7.2 the ID Number domain would consist of all numbers from 001 to 999. The term may also be applied to the pool of records defined by a table, query or SQL (structured query language) expression.

Structure

There are three rules relating to the structure of the relational table.

1. The order of records is of no significance or, put another way, records are not ordered. If there were significance in the order, for example if person records were ordered by salary or length of service, this would be undeclared information, which should properly be represented in an additional field for this purpose.

2. The order of attributes or fields within a record is of no significance.
3. Field values are "atomic," i.e., they are not composed of any smaller entities recognized by the database system. Figure 7.5a shows a table that contravenes this rule in the "children" field.

Sometimes a fourth rule is quoted but this applies more to content than structure and is covered by the first rule of integrity below.

4. Records are all unique, i.e., duplicate records are not permitted.

Integrity

There are two rules relating to data integrity in the relational model.

1. **Entity integrity.** Within a table there may be no two records with the same primary key value. No primary key may have a null value, i.e., the primary key field(s) may not be empty.

   A good relational database management system software will not allow entry of a primary key that already exists, as this would produce two records with the same primary key. Nor will it allow a record to be entered and saved while the primary key field(s) remain empty, as such a record would have no means by which it could be identified.

2. **Referential integrity.** To understand this rule we need to explain a new term. A **foreign key** of a table is a field or set of fields that in another table constitute a primary key. For example in figure 7.3a the [scientist] field in the research activities table is a foreign key because in the scientists table it is a primary key.

   The referential integrity rule states that a foreign key of a table (that is any field or set of fields that in another table constitute a primary key) must be either null or have the same value(s) as the same attribute(s) in a record in the table where it is a primary key. Where the foreign key is composite, i.e., made up of two or more fields, all these fields must be either null or all not null (and therefore of the same values as in the table where they comprise the primary key).

   The practical effect of this is to prevent the occurrence of "orphan records" such as for a report for experiment-ID 8500 when the highest experiment-ID issued so far is only 8450. If this were not prevented then when an experiment with ID 8500 is finally approved, then it would be linked to an existing unrelated report.

The better relational database management systems include devices that ensure these rules are maintained. Entering data that would contravene the rule is blocked and automatic cascade delete and cascade update facilities cause all dependent records to follow suite with any change made to the parent record. Sometimes a third rule of integrity is added:

3. Attribute values must be drawn from the underlying domain. This imposes certain database-specific constraints, e.g., all person-ID numbers must be positive.

ID No	Last Name	Inits	Age	Children
027	Burrows	J M	27	Tom Mary
I008	Phiri	P L	65	
127	Kulich	R	47	Jane Lee Harry
...	...	...	...	...

Figure 7.5a: A table with a nonatomic field

ID No	Last Name	Initials	Age	Child1	Child2	Child3
027	Burrows	J M	27	Tom	Mary	
008	Phiri	P L	65			
127	Kulich	R	47	Jane	Lee	Harry
...	...	...	...	...	...	...

Figure 7.5b: A table with repeating fields

ID No	Last Name	Child1	Child 1 gender	Child2	Child 2 gender	Child3	Child 3 gender
027	Burrows	Tom	M	Mary	F		
008	Phiri						
127	Kulich	Jane	F	Lee	M	Harry	M
...	...	...	...	...	...	...	...

Figure 7.5c: A table with repeating groups

Person table

Person ID	Last Name	Initials	Age
027	Burrows	J M	27
008	Phiri	P L	65
127	Kulich	R	47
...	...	...	...

Children table

Person ID	Child No.	Child Name
027	1	Tom
027	2	Mary
127	1	Jane
127	2	Lee
127	3	Mary
...	...	...

Note: Underlined field names constitute the primary key

Figure 7.6: First normal form achieved by separation into two tables

Normalization (dependency theory)

Normalization is a theory and a process applied to data representation to bring a relation (table) more fully into line with the relational model. The stages of the process successively eliminate a particular form of dependency between data items. If left in place such dependencies would interfere with the processing of data needed to produce information to answer queries.

Several "levels" of normalization are recognized. The first one is regarded as particularly important, and a relation conforming to this level (first normal form) is said to be normalized. Second and third normal forms are widely used and fourth, fifth, and sixth normal forms are also recognized.

In the design of a relational database system it is usual to successively alter the system to meet higher levels of normalization. In producing a working system it is then sometimes necessary to abandon some of the moves made in the normalization process as a compromise between the theoretical ideal and a practical system.

First normal form (1NF). A record should not contain repeating groups of attributes. A repeating group can manifest itself in different ways as shown in figures 7.5a, b, and c. The "children" field in figure 7.5a contravenes atomicity: that any one attribute comprises only one data item. Figure 7.5b has repeating fields (child1, child2. child3). Figure 7.5c has repeating groups of fields.

A table with repeating fields or groups is brought into 1NF by either basing it on the lowest level of repeating group, or separating out the repeating unit into another table, along with the key. See figure 7.6. Note that in the new "children" table, two fields are now needed as the primary key (shown with underlined names).

Second normal form (2NF). Full functional dependency: If a table has a multi-attribute[13] primary key then all non-primary key attributes must be dependent on the full primary key, not on just part of it. It follows that a table with a single- attribute primary key is already in 2NF.

Third normal form (3NF). Non-transitive dependency: Non-primary key attributes should depend on the primary key directly, not via another key.

For further explanation of these and other aspects of relational database theory and design, consult one of the many text books available.

Relational operations

Relationships

A relationship between tables establishes how the database selects records from each when building a data set from more than one table, as shown in figures 7.3a and b. This supports the generation of queries, forms, and reports. It works by comparing the contents of one or more relating fields in each table. Usually such fields have the same name(s) in both tables. Usually the primary key of one table will be compared with the corresponding foreign key of the other table. Three types of com-

13. An attribute is equivalent to a field.

parison or relationship are available: one to one, one to many (the most common), and many to many (usually avoided by creating a third table and then two one-to-many relationships).

Queries

Queries provide a powerful method of interrogating a database. They allow individual fields from one or more selected tables to be selected and to have the contents of only those fields displayed. The display can be restricted to just those records that match specified criteria such as "Start Date > 31/Dec/1997" to find only records started in 1998 or later. Queries therefore provide an extension to the standard outputs (reports) built into an MIS. Queries can often help answer ad hoc questions that arise in research management and that are not already provided for in standard outputs.

When a query is likely to be needed again it can be saved. In keeping with the principles of database design, only the search criteria are saved, not the resulting selection of data. When the latter are again needed the query is employed to retrieve it again. In this way it always derives from the latest data in the database.

Modern PC database software usually offers a "query-by-example" facility in which one can choose from lists, first the table or tables to select data from, and then the fields within those tables. A further step allows criteria to be applied to one or more fields to further restrict the selection.

In figure 7.7 a query selects and links two tables to present a list of CARIS 1 descriptors together with their daughter CARIS 2 descriptors. In the second pair of views a criterion has been specified for the CARIS 1 Description, restricting it to cases including the text "animal".[14]

Joins

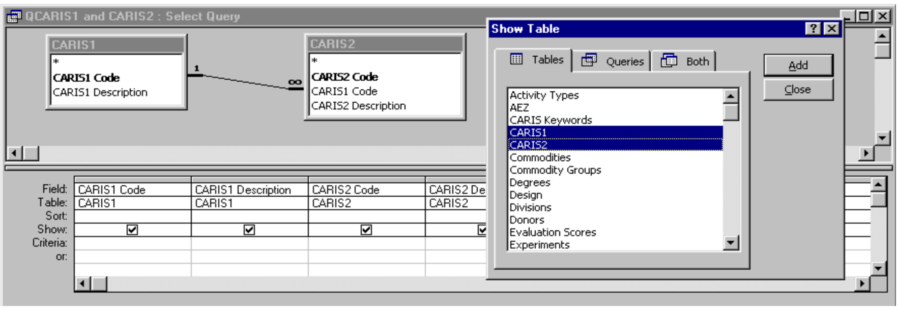
A join is a link between a field in one table and a field of the same data type in another table. A join helps establish a query in which data is retrieved from more than one table.[15]

An **ordinary or inner join** selects records from both tables, where the join fields conform to a specified condition. Usually the condition is that of equality; the two fields contain the same data. The resulting join is sometimes called an equi-join. This is the commonest type (and the first choice in the Access join properties dialog box).

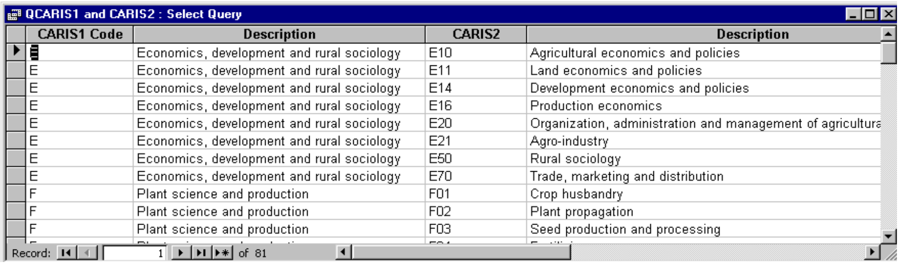
An **outer join** selects all records from one table plus only those records from a second table that have a counterpart in the first table. Records resulting from table 1 records with no counterpart record in table 2 have a null value (empty field) in the table-2 join field. Left and right outer joins are distinguished.

14. CARIS Codes and Descriptors constitute a system of agricultural subject matter classification used by the Food and Agriculture Organization of the United Nations (FAO). ISNAR's INFORM-R MIS uses a subset of these codes, starting with CARIS Code 1 = E, appropriate to agricultural research.
15. In Access, a join is set up just for the current query and is not the same as the relationship set up between tables as part of the whole database structure.

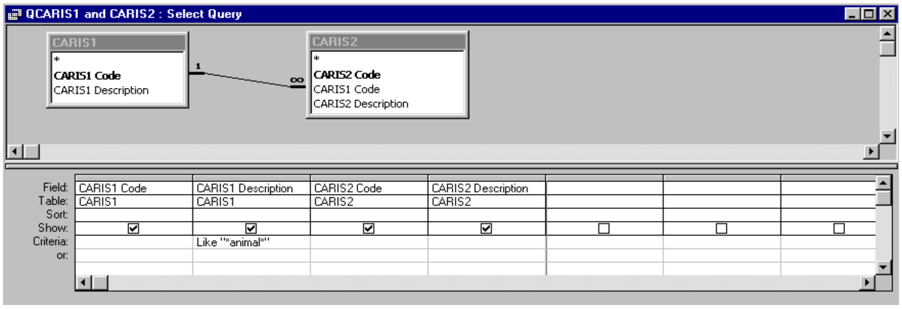
(a) Query design selecting and linking 2 tables.



(b) Data view resulting from the above query design. Note total of 81 records.



(c) Query design restriction of CARIS 1 Description to inclusion of "animal."



(d) Data view resulting from the above query design. Note new total of only 12 records.

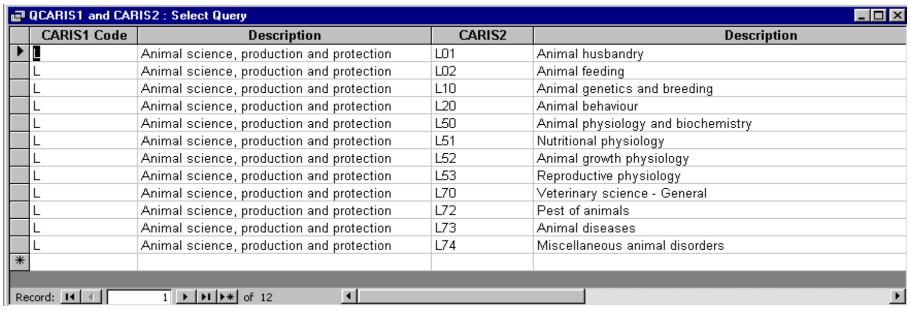


Figure 7.7: Demonstrating the use of the query-by-example facility

A **left outer join** selects all records from the left side, even when there is no matching record on the right, and those records from the right that match with one from the left. A left outer join is indicated by a join line arrow pointing from left to right. The left side usually refers to the "one" table in a 1 to many relationship. (This is the second choice in the Access join properties dialog box.)

A **right outer join** is the converse of a left outer join (and the third choice in the Access join properties dialog box).

A **self join** is a table is joined to itself for some special purpose.

A **Cartesian join** (mathematically the Cartesian product) selects all combinations of records from the two tables. If one table has 50 records and the other 100, the resulting query will have 5,000 records. It can occur in a query by example (QBE) case if fields are included from two tables that have no join between them. It is not often used deliberately but can occur accidentally when designing a query and will then generate a large number of records.

When setting the relationship type between two tables, the type of join to be used between the two can also be set at the same time.

Some advanced database features

Replication. In the context of databases, replication means the duplication of objects (usually data stores such as tables) in more than one location and maintaining consistency between the two. Consistency is achieved by periodically applying the updates of one to the other. Some systems allow this in one direction only, others accept updates at both ends and apply them to the other sites. Heterogeneous replication allows updating across different database software.

The facility is of obvious interest to a NARO when stations maintain their own copy of the national database. From time to time these copies need to be updated with all other stations' data. This would usually be done referring to a single central national version at the research departments HQ. Replication facilitates the merging in a largely automatic way. When stations do not have good telecommunications connections, the stations may send their copy in on diskettes, but good telecommunications should open the prospect of more frequent updating using the database replication facility.

Some PC database software such as Microsoft Access now offer replication, but it is safer to restrict it to the data files and tables rather than the program file(s).

Complex data types. Most personal computer database software can process text, numeric and date data. Although different, these are all alphanumeric. Complex data types refers to images such as photographs and diagrams, and even sounds and video clips (sequences of moving images). The ability to process these data types may have potential for NAROs. Photographic images could include those to assist in the identification of plant pests & diseases, mineral deficiency symptoms in plants, weed species, and health problems in livestock. Database technology is developing to accommodate different complex data types in one database, known as an Object Oriented Database. This stores in one data structure (a) the data and (b) rules about data e.g., how it looks, how it can be processed.

This work is still at an early stage compared with relational database systems but clearly has promise. A significant property of these complex data types is that they require much greater storage capacity and processing power than text. Fortunately the rapid gains in hardware and software, and the development of "multimedia" capabilities, is enabling these new requirements to be met but it will often require the procurement of new more powerful PCs.

7.5.4 Some characteristics of data

Data obsolescence

A limitation of any particular set of data is that it may become obsolete. Data that adequately represents a piece of the real world today may not do so tomorrow. The real world may change; an experiment may move its status from "active" to "completed" or a researcher's highest degree from BSc to MSc. Or we may change our requirements of the part of the real world we are interested in; data collected last year may no longer be of interest today.

There are practical implications for the MIS practitioner. The time between capturing new data and its appearance in outputs needs to be kept short. Failure to do so results in outputs that have drifted from both the real world and users' current needs as described above.

A long lead time from formulation of experimental results to publication can be a problem. Some have begun publishing scientific articles electronically to help reduce the time to publication.

Some data retains value indefinitely as a historical record of events. It is therefore very important that NAROs keep an archive of completed research. Often such research can prove valuable several decades later. There are many instances of research being unwittingly, and expensively, repeated because of lack of information of the earlier work (Vernon 1986).

Data quality

The value of any outputs from an information system depends on the quality of the data. This has to be checked particularly when data is entered and monitored during storage, as several events may affect data. The importance of data quality is reflected in several terms that have come to have specific meaning in this context.

The set of measures that check that incoming data conforms to previously set parameters is referred to as data validation. For example, a date-of-birth field can be set to reject an entry implying an age greater than say 100 or greater than the retirement age. A field for crop species can be set to prevent the accidental entry of a numeral. Data verification is somewhat similar but usually means a test that data being entered is the same as in the source. When data quality is critical the data is sometimes verified by duplicate entry of the same data by different persons: the system accepts only those records for which both entries are the same. When a difference is detected the data is rejected and sent for repeated data entry. Data validation and verification contribute to data integrity, which means freedom from errors. Data security is the defense of the data and the system against events, either deliberate or accidental, that affect its integrity. It may extend to provide privacy, which provides for an individual's right to have their personal information shielded from those not specifically authorized to access it. Passwords can provide privacy within an information system. These can be set to apply selectively to particular parts of the system.

Inconsistency of data refers to two entries of the same data item that are different. Either they were separately and differently entered or one was subsequently updated leaving the other with an older version of the data. This can happen if, for example, the publicity and marketing departments of a company maintain their own lists of customers' addresses. Such duplication of the same stored data is referred to as data redundancy. Database systems are designed to keep such redundancy to a minimum, largely to avoid the risk of inconsistency of data.

Numeric data that is derived by calculation from other data is not normally stored, but is instead calculated afresh from its source data each time it is needed. This avoids using "results" data derived from out-of-date data.

Record locking is a feature of many database management systems. It means that multiple users can access the record simultaneously, but only the first to access the record will be allowed to make any changes. Record locking is a form of concurrency control, which allows several users to access and update the database at the same time without interfering with each other.

The data dictionary

This is a system and accompanying set of procedures for maintaining definitions of the data in a database. It provides a single place for all users to find information on the data in their system. Data about data, such as the field length of and rules applied to each field, is sometimes called metadata. It should be introduced to users at the same time as the system itself. As in actual data in the database, a data dictionary allows data items to be defined just once, irrespective of the number of different applications within a database system that may use it. This provides for consistency and easy updating should this ever be needed. Annex 2 gives a partial data dictionary for the INFORM-R information system; it is generated automatically by the system, reflecting any changes made at the time it is printed.

References

Ahiabenu II, K. 1998. The Ghana All Media and Product Survey (GAMPS) by Research International. *Computers in Africa*, September.

Baker, S. 1997. "Although some aspects of my work have become more efficient with each innovation in PC technology— most notably, data processing and communications—the biggest impact of technological advances in the past ten years has been to improve aesthetics, and to allow me to take work home on a laptop computer. This marginal improvement has come at very high costs in hardware, software and—especially—time spent mastering alleged improvements." Personal communication.

Brynjolfsson, E. and L. Hitt. 1994. Paradox lost? Firm-level evidence of high returns to information systems spending. *Management Science*.

Computers in Africa. 1999. Quoting Mike Jensen. August/September, 8.

Cruz, E.R. da, V. Palma, and A.F. d'Avila. 1982. Rates of return on EMBRAPA's investments: Total investment and physical capital. Brazilia: Empresa Brasileira de Pesquisa Agropecuária.

Date, C.J. 1990. An introduction to database systems. Volume 1. Fifth Edition. Reading, USA: Addison-Wesley

European Commission. 1997. Building the European information society for us all. Final policy report of the high-level expert group. Brussels: EC Directorate-General for Employment, Industrial Relations and Social Affairs.

Fouché, B. 1997. Knowledge networks: A new paradigm for promoting access to information access and exchange. Paper presented at IUCN Workshop: Environmental Documentation and Resource Center Network, November 1997,Windhoek, Namibia.

Hitt, L. and E. Brynjolfsson. 1994. Creating value and destroying profits? Three measures of information technology's contributions. An earlier version of their paper appeared in the Proceedings of the Fifteenth International Conference on Information Systems: The Three Faces of IT Value: Theory and Evidence.

Hochman, Z., C.J. Pearson, and R.W. Litchfield. 1994. Users' attitudes and roles in the development and evaluation of knowledge based decision support systems for agricultural advisers. *Agricultural Systems* 44: 217-235.

IEEE. 1997. Technical information statement on biological and health effects of electric and magnetic fields from video display terminals. Institute of Electrical and Electronics Engineers. Engineering in Medicine and Biology Magazine 16(3): 87-92, 1997. (Also available at the University of Pennsylvania Internet site at http://homepage.seas.upenn.edu/(kfoster/vdt.html.)

Keane, T. 1995. Potato blight warning practice in Ireland. In Phytophthora infestans 150: European Association for Potato Research (EAPR)-Pathology Section Conference, held in Dublin, Ireland, September 1995. Edited by L.J. Dowley et al. Dublin: Boole Press.

Lan, Z. and C.R. Scott. 1996. The relative importance of computer-mediated information versus conventional non-computer-mediated information in public managerial decision making. *Information Resources Management Journal*. 9(1): 27-37.

Laudon, K. and J.P Laudon. 1994. Management information systems. New York: Macmillan Publishing.

Mankin, D., S.G. Cohen, T.K. Bikson. 1996. Teams & technology: Fulfilling the promise of the new organization. Boston, MA., USA: Harvard Business School Press.

Marsh, C. 1998. IT managers turn to ROI to justify IT spend. Results of an online survey. Reported in *Computers in Africa*, May, by Bruce Conradie.

Menou, M.J. 1993. Measuring the impact of information on development. Ottawa: International Development Research Centre.

Mueller, S. 1998. Upgrading and repairing PCs. 10th edition, 1531 pp. USA: Que Corporation.

Nonaka, I., U. Katsubiro, and D. Senoo. 1996. From information processing to knowledge creation: A paradigm shift in business management. *Technology in Society* 18 (2): 203-218.

Pascarelli, E. and D. Quilter. 1994. Repetitive strain injury: A computer user's guide. New York: John Wiley & Sons.

PC World 1998. PC Reliability and Service: Who Can You Trust? PC World. November. 16(11): 97-116.

PC World 1998a. Notebooks for the long haul. PC World. November. 16(11): 155-164.

Philpot, P. 1998. Personal communication.

Tribe, D. 1994. Feeding and greening the world: The role of international agricultural research. Wallingford: CAB International.

Tufte, Edward R. 1983. The visual display of quantitative information. Cheshire, Connecticut: Graphics Press. (The same author has published two further volumes. Envisioning Information (1990) looks at the presentation of complex information in general, not just statistical data, and considers the use of color. These two works provide a valuable guide to anyone working on the presentation of data. A third volume, Visual Explanations (1997), provides some additional material.)

Vernon, R. 1986. A role for information technology in development programmes. *International Agricultural Development*, Jan/Feb, 12-14.

Wang, R.Y. and D.M. Strong. 1996. Beyond accuracy: What data quality means to data customers. *Journal of Management Information Systems* 12(4): 5-34.

Zipp, W. 1994. Improving the transfer and use of agricultural information. A guide to information technology. Washington, DC: World Bank.

Annexes

Annex

A1 Glossary

Many of the following terms have been defined in different ways by different authors. The definitions offered here are to assist readers in using this book rather than to approve any particular definition. Italicized words indicate a cross-reference to a full position in the glossary.

agricultural research intensity ratio (ARI)
Public agricultural research expenditure as a proportion of an agricultural production factor, usually agricultural gross domestic product (AgGDP). In 1980, the World Bank indicated a target of 2.0% by 1990. By 1995, most had not exceeded 1.0%, with African NARS, helped by their relatively small AgGDP and heavy donor support, doing better than those of Asia and Latin America (see table 2.12 "Some indicators of agricultural research context"). Other research intensity ratios are research expenditure per hectare of agricultural land or per person employed in agriculture.

ASCII
Short for American Standard Code for Information Interchange. A code, widely used in computing and communication technology, that represents the alphabet (A-Z, a-z, numerals, punctuation marks and some control characters) entirely with sequences of just two digits, 0 and 1. Eight binary digits (called bits) are needed for each alphabetic character, the whole comprising a byte.

byte (B)
A unit for measuring electronically stored or transmitted data; equivalent to one alphabetic character or two numerical characters; made up of eight binary digits (0 and 1) called bits.

database
An organized collection of related sets of data, managed in such a way to enable the user to view the complete collection, or a logical subset, as a single unit. It typically comprises, in descending hierarchical order, tables, records, and fields.

client-server arrangement
An arrangement of a central "server" computer connected via a network to one or more user's PCs, the "clients." A database that contains data needed by several people resides on the server. When an individual wishes to query the database the query is sent from the client PC to the server, where the database management system interrogates the database, formulates the answer, and transmits it back to the client. Data-processing work is shared between server and client processors so that it minimizes the data traffic on the network. The application on the client PC is known as the front end, the application holding the data on the server as the back end.

empty fields: null or zero length string
An apparently empty field usually means that we do not have information yet to fill it. Microsoft Access enables us to distinguish between two states of "no information." If a field has simply had nothing entered into yet, it is said to be in a null state. If we wish to declare that in a particular case there is no information, as for example the fax no. field for a person whom we know has no fax machine, then we enter a zero length string by typing in two adjacent pairs of quotation marks: "". These are not treated as characters, i.e., the field is treated as if empty. In some systems a null value can be entered into a field, indicating missing or unknown data.

Note that a field with one or more spaces entered with the space bar is not empty: the "spaces" are equivalent to characters. Similarly a numeric zero is a character.

Check the help facility of your software to find out any variations used by that software. In Microsoft Access you need to set the AllowZeroLength property of a field to Yes and the Required property to No to allow the field to be set to zero length.

experiment
The lowest level of research activity for which a research protocol, with title and objectives, is prepared. In some countries a single experiment may be replicated across several sites and repeated over more than one year. The term "trial" is often used to mean the same thing. The term is used here to include surveys and studies.

field
An attribute or feature used to describe an object in a database. Usually the smallest data item or entity of a database, making it "atomic," i.e., not separable into smaller units insofar as the database is concerned. Examples: the start date of an experiment, the highest degree of a researcher.

flat file database
A database in which related data is stored in a single table. Each row represents a record such as a person, and each column (referred to as a field) represents an attribute of the person, e.g., name, age, etc. It is suitable for simple applications of managing data, such as a list of people and their addresses. For more complex applications, particularly where there are several occurrences of one attribute, (e.g., a person's several training events or children) some difficulties can arise, for example when trying to answer certain queries. A r*elational database* is better equipped to handle such circumstances. (See section 7.5.1 "The flat file database.")

front end
A computer system that provides a user with access to data stored in another system, for the purpose of reading, editing or adding new data, or generating reports. (See also *client-server arrangement*.)

Hertz (as in Megahertz)
A unit of frequency, equivalent to cycles per second. It is used to describe the frequency of electromagnetic radiation such as radio waves and of the refresh rate of a computer monitor screen.

INFORM
Information for **R**esearch **M**anagement: a *management information system* that collects and processes selected data to provide research managers with information on the performance of the research program. It was developed by ISNAR in collaboration with some Asian NARS in the late 1980s. The system was kept simple through collecting only the minimum data needed for research management and using scientists' time allocations and costs as a proxy for direct research costs. (See also *INFORM-R*.)

Information and communications technology (ICT)
A collective term covering all those technologies, both hardware and software, dedicated to the capture, storage, processing, transmission, and presentation of information. (See section 7.3: The main technologies.)

information management
All the processes employed within an organization to capture data, process it into information, and distribute this to wherever it is needed. It may include filing, library, accounting, and management information systems. (See also the introduction to chapter 5.)

informatization
The uptake or deployment of information management processes, information services, and information and communication technologies. The term is often applied to society, nations, and sectors.

INFORM-R
The new, "relational" version of *INFORM*, developed in 1995–96 and using new developments in database technology. The main advantages of the *relational database* model used for this MIS are a menu-driven *front end* that enables managers to use the system after only a day or two of training, and its much greater data storage and processing capability. Its main disadvantages are that much data needs to be captured and entered on first implementation within an organization, MIS practitioners need to be trained at each station, and integrating it into the research management cycle takes some time.

INFORM-R Light
A new and simpler version of *INFORM-R*, developed in 1998–99, designed to be implemented quickly.

logical framework
A tool for preparing, monitoring, and evaluating *projects*. It recognizes four levels, activities, outputs, purpose, and goal, linked by hypotheses, e.g., if the activities take place, then the outputs will occur. Measurable indicators are specified for each; the specification includes quantitative, qualitative, and time-related factors (QQT). At each level assumptions are recognized too. A computer-based version is available. Further details of the approach can be found in Horton et al. (1993) pp. 113-119.

management information
Aggregated or summarized information indicating the state of one or more parts or processes of an organization, that is needed by a manager in order to manage those parts or processes. (See also section 2.1 "Introduction.")

management information system (MIS)
A definition for a general MIS is given in section 1.1.5 "Information systems." An MIS for agricultural research is an ongoing data-collection and -analysis system, usually computerized, that provides managers with timely access to information on research inputs, activities, and outputs, in support of effective decision making.

megabyte
A unit for the measurement of electronically stored or transmitted data, equal to 1,048,546 (approximately 1 million) bytes; abbreviated to MB.

NARO
National agricultural research organization: the main national institution established by the government for the pursuit of agricultural research.

NARS
National agricultural research system(s): includes all the public, semi-public, private-sector, and academic agricultural research institutions in the country.

normalization
The process of arranging data within a *database* in order to conform with certain rules that facilitate a range of and guard against some common problems in data processing. It is a feature of *relational database* but not of a *flat-file database*. The first rule forbids the presence of repeating groups such as a person's training events or children within one *table*; this is avoided by splitting into two or more simpler tables. Unnormalized tables can give difficulty, for example when trying to answer certain queries. (See section 7.5.3 "The relational database.")

primary key
A single field or combination of more than one field that uniquely identifies a record in a database, i.e., no two records may have the same primary key.

program
A group of related *experiments* and *projects*, often within a single commodity group or discipline, and usually of sufficient national importance to be expected to continue indefinitely, i.e., without a specific end date.

project
In its stricter sense, a group of related experiments sharing a common objective, with a finite overall budget and time allocation. Each experiment retains its own specific objective. More loosely used to apply to any time-limited research activity.

project budgeting
The process of allocating financial resources to research *projects*. Projects is here used in a general sense to include experiments and programs. (See also section 5.1 "Financial Information in an MIS.")

record
An instance or case of an object about which data is stored in a database. Examples are an experiment, a research scientist, a research station and a job title. Many records of each such object are typically stored in a *table* devoted to that object. A record comprises *fields*, which describe or qualify that record of the object.

Reflex
A *flat-file database* program, designed to run in MS-DOS, which has been used for the great majority of ISNAR's *INFORM* implementations. It was selected for its ease of use, range of features, and graphical outputs. The company owning Reflex stopped maintaining it several years ago.

relational database
A model of database that uses tables conforming more or less to a set of criteria established originally by Codd (1970, 1985). These include, for example, the rule that every record must have a primary key. They also include the rules of *normalization*.

SAB
An acronym for salaries, allowances, and benefits, which is a phrase used in *INFORM* and elsewhere to denote all costs directly attributable to a researcher.

table
A two-dimensional array of rows and columns typical of some database models including the *flat-file* and relational models). Each row represents a single *record* of the database. Columns represent attributes (stored in *fields*) of the object, e.g., a person's name, highest degree, and age.

total factor productivity
A multifactor productivity index expressing total measured output relative to an index of all the measured inputs.

validation
The process of checking that a data item entered into an information system is of the appropriate nature for the *field* it is being put into. Validation rules are usually built into database management systems and operate automatically on data entered into the system. (See also section 7.5.4 "Some characteristics of data.")

verification
The process of checking the accuracy or truth of data. (See also section 7.5.4 "Some characteristics of data.")

Annex

A2 Database Tables and Fields in INFORM-R

The following list of tables and their component fields in INFORM-R shows what the system covers (look at the table names and their descriptions). It also demonstrates the table-based nature of the relational model: look at the several tables serving Experiments, or Persons, for example. It also provides a model from which developers of other agricultural research database systems can borrow tables or check their own tables for fields shown in these tables.

Contact ISNAR for more information on or for a copy of INFORM-R, or consult the ISNAR Website: www.cgiar.org/isnar. Organizations that adopt INFORM-R may find it worthwhile to document changes made to the system to adapt it to local needs. ISNAR will appreciate receiving such information: such changes may have wider applicability and may be added to the generic model.

A2.1 Comments on some particular fields

Chapter 2 covers several fields for the description of research activities and research scientists. The case is made there for a controlled list for scientific disciplines applicable to both scientist and research activity (see table 2.8 "Scientific Disciplines and Their Specialisms for Research Activities and Scientists," page 50).

Annex A3 lists the current data sets offered by INFORM-R on commodities and commodity groups.

Numeric and date fields deserve a comment. As a rough rule, when creating a database, declare a field numeric if its contents are indeed numeric and it is expected to by subjected to any kind of arithmetic. Similarly declare any field with any kind of date, such as a start date or end date, as a date field and not as a text field. Declared date fields can be used in useful calculations such as duration = [End date] - [Start date] or a person's age at any particular date = [Date] - [Date of birth]. A date field will prevent entry of dates that do not exist such as 31st June. Date fields will sort correctly whereas dates in text fields may sort unpredictably depending on how each was entered. A problem arises when the precise date is not known as for example a project's [Expected end date] for which the year and probably the month can be given but not the day. A convention is to simply use the first day of the month.

Activity Types **Key(s): Activity Type**

A look-up list for the Activity Type field in Experiments & Proposal tables, giving types of research activity e.g., experiment, survey, etc. The list can be adapted to the needs of a NARO or country.

Activity Type: One of a look-up list of types of research activity e.g. experiment, survey.

Activity Type Description: Activity Type Description

SYSDB Trigger: This field is used to trigger code to maintain database integrity

AEZ **Key(s): AEZ**
Records the Agro Ecological Zones of each country
AEZ: Agro Ecological Zone code
AEZ Name: Name of Agro Ecological Zone
AEZ Description: Description of Agro Ecological Zone: e.g. rainfall, altitude etc

CARIS Keywords **Key(s): CARIS2 Code; Keyword**
Provides three levels of descriptors to accurately reflect the subject of a research activity. Based on the FAO AGRIS/CARIS Categorization Scheme (1990) and the related FAO AGROVOC agricultural thesaurus (1995).
CARIS2 Code: Three digit number representing subsidiary topics within each CARIS1 code. See Lookup list.
Keyword: One to several terms specifying the detailed nature of a research activity

CARIS1 **Key(s): CARIS1 Code**
CARIS program areas lookup list
CARIS1 Code: One of 12 single letter codes representing major subject areas.
CARIS1 Description: The descriptive title of each of the 12 CARIS 1 codes.

CARIS2 **Key(s): CARIS2 Code**
Subject grouping (CARIS2 codings) lookup list
CARIS2 Code: Three digit number representing subsidiary topics within each CARIS1 code.
CARIS1 Code: One of 12 single letter codes representing major subject areas.
CARIS2 Description: The descriptive title of each CARIS 2 code

Commodities **Key(s): Commodity**
Lookup list of crop and livestock species and non-commodity factors
Commodity: Common name
Commodity Group: One from a standard list of economically important groups of commodities.
Scientific Name: Scientific name of commodity
Commodity Description: Commodity description
National Land Area: Approx area of land under this commodity
Percentage of Agricultural GDP: Percentage contribution to agricultural gross domestic product
Export Value: Export Value
Import Value: Import value
Import Export Value Unit: Usually the local currency or US$
Yield per Ha National: Mean national yield per ha, all farm types
Yield per Ha World: Mean world yield of this commodity
Yield per Ha Research Station: Yield regularly obtained on research stations
Yield Unit: Yield unit: Default is Kg/ha

Commodity Groups **Key(s): Commodity Group**
Lookup list of economically important groups of commodities. Used in management activities such as priority setting and researcher recruitment & training.
Commodity Group: Standard list of economically important groups of commodities.
Commodity Group Description: Description (e.g. cereals, vegetables, livestock)

Degrees **Key(s): Degree**
Lookup list for higher educational qualifications of researchers
Degree: Standard acronym or title of qualification
Degree Description: Degree description
Rank: Ranking order for the purpose of outputs

Design **Key(s): Design**
Look-up list of standard mostly statistical designs of research activities.

Design: One of a list of standard mostly statistical designs of research activities. Others may be added within a NARO.
Design Description:

Divisions **Key(s): Division-ID**
An extra level in the hierarchy of experiment, project, program, where needed.
Division-ID: Division-ID
Division: Division description
Division Objectives: The objectives, strategy or policy of the division

Donors **Key(s): Donor**
Lookup list of donor organizations
Donor: Acronym or short name of donor
Donor Name: Full name of donor
Donor Country: Country of main office of Headquarters
Donor Type: Donor type (local, bilateral, regional, international)

Evaluation Scores **Key(s): Code; Evaluation Indicator Score**
A lookup list for each of the evaluation indicators, and the scores set for each level of each indicator. The score values can be adjusted to suit each country.
Evaluation Indicator: Title of evaluation indicator
Code: Code for use by the system
Evaluation Indicator Score: Levels within each evaluation indicator
Score Value: Value rating applied to each level of each evaluation indicator: can be set for each country
Caption: Caption for use on data capture questionnaire

Experiments **Key(s): Experiment-ID**
Research experiments (Third level of hierarchy of research: Programs, Projects, Experiments) including surveys.
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Experiment Title: Experiment title: should convey the essence of the objectives
Experiment Objectives: A clear statement of what the research is to achieve.
Justification: A statement of why investment in this activity is justified
Start Date: Must be after the start of the research year in which the activity begins; Use 1st of month if actual date not clear.
Planned End Date: Use the 1st of most likely month in the expected end year.
Actual End Date: Applied only on completion of the research activity & used also to identify completed work.
Archive: Archive this experiment in the archive database
Season: E.g. rain, dry, perennial, other
Commodity: The main crop or animal species or non-commodity factor; see Commodities lookup list
Other Crops: Other crops apart from main crop
Main Discipline: Main scientific discipline from a standard list of c. 16; see lookup list.
Specialism: Sub discipline within each Discipline; uncontrolled term but see Specialism lookup list for ideas
Main Pest: Only where it is the subject of research (insect, weed, biological control agent etc)
Other Pests: Pests, other than the main pest given in the Experiments table, that are the subject of this research
Activity Type: Activity type (Experiment, Survey etc. ; see Activity Types lookup list, which can be added to.
Design: Statistical or other design (randomized block, Latin square, factorial, etc)

Replicates: The number of replicates (at most of the sites); must be a number.
Gross Plot Size: e.g. 4 x 6 m
Harvest Plot Size: e.g. 2 x 4 m
Research Station Title: The name of the research station of responsible researcher
Person-ID: Person-ID of the reporting researcher; see Persons lookup list
Dept in Institute: Optional field to show a department or section under which the experiment falls
Project-ID: A unique code for each Project; see Projects lookup list
Approved Comment: Comment or conditions applied when a research activity is approved
Proposal-ID: ID code of the original research activity proposal
CARIS2 Code: Three digit number representing subsidiary topics within each CARIS1 code. See CARIS 2 lookup list.
Keywords: One to several terms specifying the detailed nature of a research activity. Not controlled but see Keywords lookup list.
Updated: Date record was updated
Updated By: Name of person who last updated this record

Experiments and Persons **Key(s): Experiment-ID; Research Year**
Temporary table used for generating report: List of Experiments and Associated Researchers
Experiment-ID: Research Year: Total Research Time: Total Persons: Person-ID1: Research Time1: Person-ID2: Research Time2: Person-ID3: Research Time3: Person-ID4: Research Time4: Person-ID5: Research Time5: Person-ID6: Research Time6: Person-ID7: Research Time7: Person-ID8: Research Time8: Person-ID9: Research Time9: Person-ID10: Research Time10:

Experiments Associated Researchers **Key(s): Experiment-ID; Person-ID**
This table allows a Researchers Time Allocation Questionnaire to include new experiments in which the proposal gave this person as an associated researcher.
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Person-ID: Person-id of associated researcher

Experiments Budget **Key(s): Experiment-ID; Budget Year**
Annual budget allocation for an experiment, and its breakdown. Part of the Project Budgeting System (PBS)
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Budget Year: A 4-digit number representing the year
Govt%: Percentage of the research activity budget to be provided by Government funds.
Donor: Organization that supports this experiment; see Donors lookup list; If none, insert Govt. abbreviation.
Allowances: Allowances
Transport: Transport
Labour: Labour
Materials: Materials
Other Recurrent: Other Recurrent
Capital: Total cost should be depreciated over agreed number of years: Ex: Buildings: 20; Vehicles: machinery, instruments, computers: 5.

Experiments Evaluation **Key(s): Experiment-ID; Evaluation Date**
This table records the evaluation of a research activity after its completion, against its objectives and selected other indicators. Numeric scores can be weighted to reflect relative importance.
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Evaluation Date: Evaluation date

Author: Author of this evaluation, usually the main researcher
Countersigned By: Person countersigning the evaluation; may be team leader, station director
Objectives Achievement Score: Fully realized, Largely realized, Significant problems, Experiment abandoned
Objectives Achievement Comment: Objectives achievement (required if objectives achievement = “Significant problems” or “Experiment abandoned”)
Local Knowledge: Contribution to in-country knowledge; Significant, Some, Insignificant
World Knowledge: Contribution to world knowledge; Significant, Some, Insignificant
Full Report Prepared: Publications: Full experiment / study report prepared ?
Reported in Station or Institute: Publications: Reported in station / institute annual report ?
Reported in Local Journal: Publications: Reported in local journal ?
Reported in International Journal: Publications: Reported in international journal ?
On Farm Results: Results verified in on-farm research ?
Extension Ready: Results fit to pass to Extension service ?
Passed to Extension Service: If results fit to pass extension service and if useful to farmers (or other targeted beneficiaries) : introduced to extension service (or used by other target beneficiary) ?
Impact Review Indicated: Does subsequent IMPACT review look worthwhile ?
Impact Review Years: If IMPACT review worthwhile, in how many years ?
Other Outputs: Other outputs achieved (as listed in research protocol) ?

Experiments Events Dates **Key(s): Experiment-ID; Site Reference; Event; Event Date**
This table stores the dates of significant events such as planting, application of treatments and harvesting.
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Site Reference: A reference character to represent each site at which a research activity is held or repeated e.g. A (main site), B, C etc
Event: The actual event e.g. planting, harvesting
Event Date: Event date (Planting date, Harvest date etc.)
Report: A flag as to whether it would be included in the report printed out
Comments: Any further comment on the event.

Experiments External Support **Key(s): Experiment-ID; Donor; Start Date**
Agreements on support from an outside institution at the Experiment level
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Donor: Organization that supports this experiment ; see Donors Lookup list
Start Date: Start date of support
Support Title: Title of the support agreement
End Date: (Expected) end date of support
Financial Support: Whether financial support is included
Technical Support: Whether technical support is included
Total Value: Total amount of financial support (if Financial support = Y)
Value Unit: Usually local currency or US$
Comment: Comment

Experiments Monitor **Key(s): Experiment-ID; Research Year**
Indicators of the quality of research during its active life, mainly for program leaders and research managers to assess program- or station-wide status of experiments.
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Research Year: A 4-digit number representing a twelve month period defined for each country with regard to the main growing period.

Author: Name of the author of this report (not currently used - author assumed to be the main scientist)
Report Status: Scores: Full report , Adequate report, Partial report, No report yet
Objectives Attainment Score: Scores: As planned, Behind schedule, None yet (=0)
Continued Relevance: Scores: High, Medium, Low
Relevance Comments: Comments on Continued Relevance
Future Status: Scores: Continue unchanged, Modify & Continue, Terminate
Status Comments: Comments on Future Status
Any Extension Messages: Does the research to date justify issuing an extension message?

Experiments Reports **Key(s): Experiment-ID; Research Year; Report Type**
Stores the script part of research activity reports, and some widely applicable indicators of performance. Can be combined with TREATMENT and RESULTS tables to produce a more detailed report.
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Research Year: A 4-digit number representing a twelve month period defined for each country with regard to the main growing period.
Report Type: Type of report (Interim-1,Interim-2 etc., Annual, Final)
Report Date: Report date
Author: Name of the author of this report. Normally the Main Scientist responsible for the experiment.
Report: Short report & discussion of the research and its results
Objectives Attainment: In what ways the current work to date has contributed to the final achievement of the objectives
Remaining Work: What still needs to be done or what new research is needed for the same problem.
Extension Message: Usually only in Final Report
Achievements since Last Report: Any achievements since last report
Comment on Finance: Comments on Budget and Expenditure
Any Difficulties Encountered: Problems or difficulties encountered and suggested solutions

Experiments Results **Key(s): Experiment-ID; Site Reference; Treatment No; Research Year; Result No**
This table stores the results of each treatment of each year of an experiment
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Site Reference: A reference character to represent each site at which a research activity is held or repeated e.g. A (main site), B, C etc
Treatment No: Sequential number of all treatments in this research activity, as given in the Experiments Treatments table
Research Year: The currently selected research year;
Result No: Sequential number of all results from one treatment
Result Numeric: Numeric figure
Result Non-numeric: For non-numeric parameters
LSD Significance: For comparison of each treatment with a control: use * (for 5%), ** (for 1%) or ns (for not significant)
DMRT: Duncan's Multiple Range Test: e.g. aaa, bb, cc etc

Experiments Results Aggregate **Key(s): Experiment-ID; Treatment No; Result No**
This table stores the aggregated results of each treatment for all years of an experiment
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Treatment No: Sequential number of all treatments in this research activity, as given in the Experiments Treatments table
Result No: Sequential number of all results from one treatment

Result Numeric: Numeric figure
Result Non-numeric: For non-numeric parameters

Experiments Results Statistics **Key(s): Experiment-ID; Site Reference; Research Year; Result No**
This table records statistical parameters that apply across all treatments, rather than each treatment, for each result title.
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Site Reference: A reference character to represent each site at which a research activity is held or repeated e.g. A (main site), B, C etc
Research Year: The currently selected research year;
Result No: Identification number for each measurement; Exp Results Titles table gives details.
C of V: Coefficient of Variation : for each Result (where calculated)
LSD: Least Significant Difference (where calculated) e.g. 123 kg/ha, or "ns" if not significant.
LSD level: Usually 0.05 or 0.01 (where calculated)
SE: Standard Error (where calculated)
Other Statistic Title: The Title of any other statistical indicator applying to this result across all treatments.
Other Statistic Value: The value of any other statistical indicator applying to this result across all treatments

Experiments Results Titles **Key(s): Experiment-ID; Result No**
This table provides the titles and units of Results in the Experiments Results table, and whether the result is to be printed in the
Experiment Annual Report.
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Result No: Sequential number of all results; data is in Experiment Results table.
Result Title: Type of result e.g. Grain Yield , Height of lowest pod, No of cobs per plant
Result Unit: e.g. kg grain / ha,
Results to Print: Print results in annual and final reports?

Experiments Sites **Key(s): Experiment-ID; Site Reference**
Sites of an experiment. They inherit all the properties of the parent experiment except for the fields shown here.
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Site Reference: A reference character to represent each site at which a research activity is held or repeated e.g. A (main site), B, C etc
Start Date: The date at which the study started at this site - the month and year are important.
Planned End Date: Planned end date (month and year are important) at this site.
Actual End Date: Actual end date (month and year are important) of the experiment at this site.
Research Station Title: The name of the research station of responsible person at this site.
Location: Location of off-station experiments and replicates
AEZ: Agro Ecological Zone of this site; see AEZ lookup list
Person-ID: Person-ID of person responsible for this site; see Persons lookup list

Experiments Sites Status **Key(s): Experiment-ID; Site Reference; Research Year**
Sites of an experiment. They inherit all the properties of the parent experiment except for the fields shown here.
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities

Site Reference: A reference character to represent each site at which a research activity is held or repeated e.g. A (main site), B, C etc
Research Year: A 4-digit number representing a twelve month period defined for each country with regard to the main growing period.
Current Status: Current Status of the site (related to the Research Year) (Annual flag showing one of: Started, Not started, Canceled)
Current Status Comments: Comments
Site History: Give previous year's crop / fertilizer application / land use etc.

Experiments Sites Treatments **Key(s): Experiment-ID; Site Reference; Treatment No**
Treatments per site of a research activity (normally there is only one treatment per experiment)
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Site Reference: A reference character to represent each site at which a research activity is held or repeated e.g. A (main site), B, C etc
Treatment No: Sequential number of all treatments in this research activity

Experiments Surveys **Key(s): Experiment-ID**
For those research activities that are surveys this table provides fields additional to the Experiments table
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Survey Type: Survey type . See Survey Types Lookup list
Target Population: The group of people that are the target of the survey
Sampling Frame: A description of how the sample to be survey is defined.
Data Collection Method: Data collection method
Sample Size: Sample size
Sample Unit: Sample unit

Experiments Training **Key(s): Experiment-ID; Experiment Training No**
Training events associated with Experiments. Not used at this time.
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Experiment Training No: Sequence Number
Training Objective: Description of the training activity objective
Other Activities Link: Other activities link
No of Trainees: No of trainees
Topics: We may (one day) need to spill this into a separate table
Training Methods: Training methods
Frequency: Note that "Duration" may be covered by Experiments.Start Date & Experiments. End Date

Experiments Treatments **Key(s): Experiment-ID; Treatment No**
Definitions of all treatments within a research activity
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Treatment No: Sequential number of all treatments in this research activity
Main Treatment: Titles of main treatments
Second Level Treatment: Titles of second level treatments - only if needed
Third Level Treatment: Titles of third level treatments - only if needed
Fourth Level Treatment: Titles of fourth level treatments - only if needed
Treatment Comment: Treatment comment

External Factors **Key(s): Sort Class; Item**
This table and the linked Factor Values table describe the 'context' or environment within which research takes place, and which program planning must relate to. A selection of a few of these may suffice for M&E purposes.
Sort Class: Sort Class
Item: Item
Item Title: Item Title
Unit: The unit for numerical items
Explanation: Explanation for this item
Local Target Figure: Local Target Figure for this item
Reference for Local Target: Quote the source and / or the year from which the local target figure derived.
Global Figure: discontinued
Global Reference: Typically a regional average, sometimes an internationally suggested target.
Note: Background information on this factor
Report / Graph: Inset Report and / or Graph if this factor is (to be) used in a report or graph.
Country Comment: Comment to derive management conclusion for the current country.
Source of Factor Values: Source(s) from which data was obtained.

External Factors Candidates for Priorities **Key(s):**
This table and the linked Factor Values table describe the 'context' or environment within which research takes place, and which program planning must relate to. A selection of a few of these may suffice for M&E purposes.
Sort Code: Sort Code
Item: Item
Priority Ranking: 1 = top priority, 2 is next, etc.
Year: Four digits. Year to which the information applies or was updated.
Explanation: Explanation
Objectively Verifiable Indicator: Objectively Verifiable Indicator
Target Statement: Target Statement
Comment: Comment

External Factors Sort Classes **Key(s): Sort Class**
This table and the linked Factor Values table describe the 'context' or environment within which research takes place, and which program planning must relate to. A selection of a few of these may suffice for M&E purposes.
Sort Class: Sort Class
Sort Class Title: Sort Class Title

External Factors Values **Key(s): Sort Class; Item; Year**
This table and the linked Country Factors table describe the 'context' or environment within which research takes place, and which program planning must relate to.
Sort Class: Sort Class
Item: Item
Year: Four digits. Year to which the information applies or was updated.
Numeric Value: The value of numerical items
Non-numeric Value: Statement of status of non-numerical items

Job Titles **Key(s): Job Title**
A lookup list of the job titles found amongst researchers in the current country
Job Title: Job title
Job Title Description: Job title description (not used at present)

Locations **Key(s): Location**
Locations of experiments particularly to identify off-station sites.
Location: The name of a site: may be the name of a farm, or the nearest village. See Locations lookup list for each country.
Research Station Title: The name of the supervising research station.
AEZ: Agro Ecological Zone. See the AEZ Lookup list

Main Disciplines **Key(s): Main Discipline**
A standard short lookup list of around 16 major scientific areas applied to both researchers and experiments. Each may have sub-disciplines (specialisms) and these are uncontrolled though INFORM-R offers a list of commonly used terms.
Main Discipline: A standard lookup list of around 16 major scientific areas applied to researchers and experiments
Main Discipline Description: Main discipline description

Monitor Scores **Key(s): Code; Monitoring Indicator Score**
A lookup list for each of the monitoring indicators, and the scores set for each level of each indicator. The score values can be adjusted to suit each country.
Monitoring Indicator: Title of monitoring indicator
Code: Code for use by the system
Monitoring Indicator Score: Levels within each monitoring indicator
Score Value: Value rating applied to each level of each monitoring indicator: can be set for each country
Description: Explanatory description for each indicator and level

Persons **Key(s): Person-ID**
This being a management information system for the research program, "persons" are researchers. Mostly they are professional scientists but the term "researcher" permits also cases where experienced technicians are managing an experiment.
Person-ID: A unique code for each person, applied and used internally by INFORM-R
Surname: Surname, last or family name of person
Firstname: First name of person
Initials: Initials: only those additional to First name
Birth Date: Date of birth
Gender: Male or female
Is Expatriate: "Yes" Indicates external salary: the Salary Grade is then set to the highest grade for regular scientists
Is Scientist: "Yes" Indicates that this person is regarded as a scientist (independent of his or her highest degree) , Nonscientists are excluded in most of the Inform-R management reports
Research Station Title: Title of the research station to which the person is attached. See Research Station lookup list
Donor: An external organization providing financial or technical support directly to this person
Date of First Appointment: Date of first appointment to the organization
Date of Present Appointment: Date of appointment to the present Job Title position
Date of Discharge: Date of discharge or person leaving the service. The person record is retained within the system but excluded from most outputs.
Main Discipline: Main scientific discipline from a standard list of c. 16; see Main Discipline Lookup list.
Specialism: Sub discipline within each Discipline; uncontrolled term but see Specialism lookup list for ideas (additions allowed)
Commodity: The main plant or animal or other commodity worked on. See Commodities Lookup list for standard names; additions are allowed.
Persons NRN: An optional personal code as used within the country or organization e.g. a national registration number (NRN)

Persons Email: Email address of this person (always precede with “Mailto:”), edit by tab into field, then F2
Updated: Date the record was updated
Updated By: Name of user that updated this record

Persons Costs **Key(s): Person-ID; Research Year**
Records the total fixed costs of each researcher usually comprising salary, allowances and benefits. A new data set is created on rollover to a new research year.
Person-ID: A unique code applied and used by INFORM-R for each researcher
Research Year: A 4-digit number representing a twelve month period defined for each country with regard to the main growing period. Default is the current year.
Job Title: A Lookup list of job titles found amongst researchers in the current country
Rank: A lookup list of standard ranks in the current country. e.g. SRO = Senior Research Officer
Salary Grade: The salary grades of the pay scale. See Salary Grades Lookup list
Grade Level: The levels or steps within each salary grade in the pay scale.
SAB: Salary, Allowances and Benefits (i.e. total direct costs of a staff member)

Persons Degrees **Key(s): Person-ID; Degree**
Records the total fixed costs of each researcher usually comprising salary, allowances and benefits. A new data set is created on rollover to a new research year.
Person-ID: A unique code for each person, applied and used internally by INFORM-R
Degree: The academic qualification. See Degrees Lookup list
Degree Year: The year in which the Degree was awarded
Expected Degree Year: The year in which the Expected Degree is likely to be awarded: has implications for personnel management
Degree University: The university at which the Degree was obtained
Degree Country: The country in which the Degree was obtained
Degree Subject: The scientific subject of the Degree. Uncontrolled: no standard list.
Degree Discipline: The Main Discipline of the Degree. See Main Disciplines Lookup list

Persons Research Time **Key(s): Person-ID; Research Year; Experiment-ID**
Each researcher’s time spent on research, broken down to each experiment he or she is working on.
Person-ID: A unique code applied and used by INFORM-R for each researcher
Research Year: A 4-digit number representing a twelve month period defined for each country with regard to the main growing period. Default is the current year.
Experiment-ID: A unique code for approved experiments, surveys & other base level research activities
Research Time: Time spent on this experiment, as % out of 100% of Total Time which can include Research, Management,
Extension etc .

Persons Short Term Training **Key(s): Person-ID; Training Year; Sequence Number**
Details of each short term training activity for each researcher, each year
Person-ID: A unique code for each person, applied and used internally by INFORM-R
Training Year: Year of short term training
Sequence Number: Sequence-no if more then one training in the same year
Subject: The title or subject of the short term training activity
Duration: Duration of training in weeks
Country: Country in which the training took place
Institution: Name of the institute responsible for the training

Persons Work Time **Key(s): Person-ID; Research Year**
Persons working time split into the main calls on researchers’ time
Person-ID: A unique code for each person, applied and used internally by INFORM-R

Research Year: A 4-digit number representing a twelve month period defined for each country with regard to the main growing period. Default is the current year.
Research Time: Time spent on research as % out of 100% total time, and equal to the sum of times spent on each experiment.
Management Time: Time spent on management tasks as % of total time of 100%
Teaching Time: Time spent on teaching as % of total time of 100%
Consult Time: Time spent on external consultancies as % of total time of 100%
Extension Time: Time spent on extension as % of total time of 100%
Degree Study Time: Time spent on higher degree study leave as % of total time of 100%
Short Term Training Time: Time spent on Short Term Training as % of total time of 100%
Other Time: Time spent on other activities as % of total time of 100%
Other Time Specification: Specification of "other time"

Programs **Key(s): Program-ID**
Details of all programs, which are the top level in the hierarchy of research activities: experiment, project, program.
Program-ID: Program-ID
Program: The title of the Program
Person-ID: Person-ID, as given in the Persons Table, of the person responsible for the program
Program Objectives: The objectives, strategy or policy of the program
Institute: Research Institute-ID (not currently used - dv)
Division-ID: Division-ID
National MIS: Reference to any other relevant information system

Programs External Support **Key(s): Program-ID; Donor; Start Date**
Support from an outside institution at the Program level
Program-ID: Program-ID
Donor: Donor organization that supports this program. See Donors Lookup list.
Start Date: Start date of support
Support Title: Title of the support agreement or arrangement
End Date: End date of support
Financial Support: Financial support
Technical Support: Technical support
Total Value: Total amount of financial support (if Financial support = Y)
Value Unit: Usually local currency or US$
Comment: Any notes on this support agreement

Programs Reports **Key(s): Program-ID; Research Year**
Being the Program Leaders annual report, derived from the constituent project reports, and providing components of the annual report of the research service
Program-ID: Program to which this project belongs
Research Year: Year to which this report applies
Author Surname: Surname of the author: needed only if the author is not the Program Leader & in the Persons table
Author Initials: Initials of the author
Author Title: Title (if any) of the author
Author Job Title: Job title (lookup list: Job Titles)
Program Report: Short Report of the program and its results
Program Comment on Finance: Comments on Budget and Expenditure

Projects **Key(s): Project-ID**
Details of all Projects, which are the second level in the hierarchy of research activities: experiment, project, program.
Project-ID: Project-ID (numbering within Program-ID)

Project: Project title
Project Objectives: Description of the project objectives
Start Date: Must be after the start of the research year in which the project begins; Use 1st of month of actual date not clear.
Planned End Date: Use the 1st of most likely month in the expected end year.
Actual End Date: Applied only on completion of the project & used also to identify completed work.
Program-ID: Program to which this project belongs
Person-ID: Person-ID of the person leading the project

Projects Budget **Key(s): Project-ID; Budget Year**
Details of all Projects, which are the second level in the hierarchy of research activities: experiment, project, program.
Project-ID: Project-ID (numbering within Program-ID)
Budget Year: Budget year
Govt Value: Value of Total Govt funded budget allocation for the year (latest working figure)
Donor Funded Value: Value of Total Donor funded budget allocation for the year (latest working figure)
Value Unit: Usually local currency or US$
Govt %: Percentage of Total Govt funded budget allocation for the year (latest working figure)
Donor Funded %: Percentage of Total Donor funded budget allocation for the year (latest working figure)

Projects External Support **Key(s): Project-ID; Donor; Start Date**
Support from an outside institution arranged at the Project level
Project-ID: Project-ID
Donor: Donor organization that supports this project
Start Date: Start date of support
Support Title: Title of the support agreement or arrangement
End Date: End date of support
Financial Support: Financial support
Technical Support: Technical support
Total Value: Total amount of financial support (if Financial support = Y)
Value Unit: Usually local currency or US$
Comment: Comment

Projects Reports **Key(s): Project-ID; Research Year**
Being the Project Leaders annual report, derived from the constituent experiment reports
Project-ID: Project-ID (numbering within Program-ID)
Research Year: Year in which the project report is produced
Project Report: Short Report of the project and its results
Project Comment on Finance: Comments on Budget and Expenditure

Proposal Appraisal **Key(s): Research Station Title; Proposal-ID**
This table records the review of a Research Proposal against standard quality assurance factors of relevance, quality, benefit and cost.
Research Station Title: The name of the research station of responsible researcher
Proposal-ID: A unique code for research activity proposals.
National Literature Search: Has a search been made in the literature for relevant work done in this country?
No of National Refs: How many literature references did you find on relevant work done in this country?
International Literature Search: Has a search been made in the literature for relevant work done in other countries?

No of International Refs: How many literature references did you find on relevant work done in other countries?
Exp Objectives = Prog Object: Do the objectives of this proposal fit in with those of the parent project and program objectives?
Outputs Measurable?: Are the expected outputs measurable ?
Output Indicators: Please list the indicators by which the outputs can be assessed.
AEZ(s): Which AEZ(s) would the results apply to
Main Beneficiaries: Please identify the main beneficiary (e.g. "Mainly hand hoe cultivators of low rainfall areas"
Good Science?: Is the methodology of this activity based on a rigorous scientific approach?
Peer Review Held?: Has the proposal been reviewed by other scientist(s)?
Date of Peer Review: Date on which the Peer Review held
Peer Review Results: State the main conclusion(s) of the Peer Review
Design Approved?: Has the Biometrics office approved the statistical design?
Chance of Success: 4= excellent chance, 3=good, 2=moderate, 1=small (may be justified if problem is very serious)
Productivity Gain: State type of productivity e.g. yield, better pest control
Productivity Gain Value: Estimated productivity gain per ha
Productivity Gain Unit: Productivity gain unit e.g. kg / ha, % infestation rate
Applicable Crop Area: Estimated total land area to which the results would apply
Estimated Adoption Rate: Estimate adoption rate within the applicable crop area.
Beneficiary Population: Estimate the size of the target beneficiary populations
Typical Farm Size: Typical farm size in hectares
Estimated Time to Achieve Objectives: Estimated Time (years usually) to Achieve Objectives
Estimated Scientist Time Needed: Estimated scientist time needed in person years
Cost Benefit Analysis Done?: Has a Cost Benefit Analysis been done? (Leave blank if cost K3 mil)

Proposed Activity **Key(s): Research Station Title; Proposal-ID**
Details of research proposals: if and when approve, an Experiment-ID is added and the record is moved into the Experiments Table.
Research Station Title: The name of the research station of responsible researcher
Proposal-ID: A unique code (within Research Station) for research activity proposals.
Title: Title of the proposed experiment : should convey the essence of the activity
Proposal Objectives: Proposed experiment objectives
Justification: Justification of the proposed research activity
Planned Start Date: Planned start date of the proposed research activity
Planned End Date: Planned end date of the proposed research activity
Season: Season (rain, dry, perennial, other)
Commodity: The main crop or animal species or non-commodity factor; see Commodities lookup list
Other Crops: Other crop apart from main crop
Main Discipline: Main scientific discipline; see lookup list of c. 16 Main Disciplines
Specialism: Specialism (Sub discipline) within each Discipline; see Specialism lookup list
Main Pest: Main pest (only where it is the subject of research (insect, weed, biological control agent etc)
Other Pests: Pests, other than the main pest given in the Experiments table, that are the subject of this research
Activity Type: Activity type (Experiment, Survey etc. ; see Activity Types lookup list
Design: Statistical or other design (randomized block, Latin square, factorial)
Replicates: The number of replicates (at most of the sites); must be a number.
Gross Plot Size: e.g. 4 x 6 m
Harvest Plot Size: e.g. 2 x 4 m
Person-ID: Person-ID of the reporting researcher (lookup list: Persons)

Advisors: Names of the Advisor(s) participating in the development of the proposal; (semicolon separated)
Dept in Institute: Optional field to show a department or section under which the experiment falls
Project-ID: Parent Project-ID ; see Projects lookup list
Approved: Whether or not the proposed research activity is approved.
Approved Comment: Comment or conditions applied when a research activity is approved
Experiment-ID: A unique code for research activities only needed when the proposed activity is approved.
Updated: Date this record was last updated
Updated By: Name of person who last updated this record

Proposed Activity Associated **Key(s): Research Station Title; Proposal-ID; Person-ID**
Researchers: Researchers that will collaborate with and invest time in this research activity
Research Station Title: The name of the research station of responsible researcher
Proposal-ID: A unique code (within Research Station) for research activity proposals.
Person-ID: Person-id of associated researcher

Proposed Activity Budget **Key(s): Research Station Title; Proposal-ID; Budget Year**
Estimated budget requirements for new proposals
Research Station Title: Research Station title
Proposal-ID: A unique code (within Research Station) for research activity proposals.
Budget Year: Budget Year
Govt %: Percentage of the research activity budget to be provided by Government funds.
Donor: Organization that supports this experiment; see Donors lookup list; If none, insert Govt abbreviation.
Allowances: Allowances
Transport: Transport
Labour: Labour
Materials: Materials
Other Recurrent: Other Recurrent
Capital: Total cost should be depreciated over agreed number of years: Ex: Buildings: 20; Vehicles: machinery, instruments,
computers: 5.

Proposed Activity External Support **Key(s): Research Station Title; Proposal-ID; Donor; Start Date**
Details of any expected donor support for the proposed activity
Research Station Title: The name of the research station of responsible researcher
Proposal-ID: A unique code, within each Research Station, for research activity proposals.
Donor: Organization that supports this experiment ; see Donors lookup list
Start Date: Expected Start date of support
Support Title: Title of the support arrangement or project
End Date: Expected End date of support
Financial Support: Whether financial support is expected
Technical Support: Whether technical support is expected
Total Value: Total amount of financial support expected (if Financial support = Y)
Value Unit: Usually local currency or US$
Comment: Comment

Proposed Activity Sites **Key(s): Research Station Title; Proposal-ID; Site Reference**
Proposed sites of a proposed research activity. They inherit all the properties of the parent except for the fields shown here

Research Station Title: The name of the research station of responsible researcher
Proposal-ID: A unique code (within Research Station) for research activity proposals.
Site Reference: A reference character to represent each site at which a research activity is held or repeated e.g. A (main site), B, C etc
Planned Start Date: Planned end date (month and year are important) at this site.
Planned End Date: Actual end date (month and year are important) of the experiment at this site.
Local Research Station Title: The name of the local research station of responsible person at this site.
Location: Location of planned off-station experiments
AEZ: Agro Ecological Zone of the replicate; see AEZ lookup list
Person-ID: Person-ID of person responsible for this site; see Persons lookup list

Proposed Activity Sites Treatments Key(s): Research Station Title; Proposal-ID; Site Reference; Treatment No
Treatments per site of a proposed research activity (normally there is only one treatment per experiment)
Research Station Title: The name of the research station of responsible researcher
Proposal-ID: A unique code (within Research Station) for research activity proposals.
Site Reference: A reference character to represent each site at which a research activity is held or repeated e.g. A (main site), B, C etc
Treatment No: Sequential number of all treatments in this research proposal

Proposed Activity Surveys Key(s): Research Station Title; Proposal-ID
For those research activities that are surveys this table provides fields additional to the PROPOSED ACTIVITY table
Research Station Title: The name of the research station of responsible researcher
Proposal-ID: A unique code (within Research Station) for research activity proposals.
Survey Type: Survey type. See Survey Types Lookup list
Target Population: The group of people that are the target of the survey
Sampling Frame: A description of how the sample to be surveyed is defined.
Data Collection Method: Data collection method
Sample Size: Sample size
Sample Unit: Sample unit

Proposed Activity Treatments Key(s): Research Station Title; Proposal-ID; Treatment No
Definitions of all treatments in the proposed research activity
Research Station Title: The name of the research station of responsible researcher
Proposal-ID: A unique code (within Research Station) for research activity proposals.
Treatment No: Sequential number of all treatments in this research proposal
Main Treatment: Titles of main treatments
Second Level Treatment: Titles of second level treatments - only if needed
Third Level Treatment: Titles of third level treatments - only if needed
Fourth Level Treatment: Titles of fourth level treatments - only if needed
Treatment Comment: Treatment comment

Ranks Key(s): Rank
A lookup list of standard ranks in the current country. e.g. SRO = Senior Research Officer
Rank: Abbreviation from initials of rank title
Rank Group: Rank group (Scientist, Technical, Director, Expatriate)
Ranking: Relative ranking order
Rank Title: Full title of rank

Research Institutes **Key(s): Institute**

The second in the hierarchy of research station, institute, organization; typically a Research Department of a Ministry, a University Faculty of Agriculture etc.
Institute: Institute title
Organization: (Initials of) Organization to which the institute belongs
Institute Title: Institute title or name in full e.g. Research Branch
Institute Place: Institute place
Institute Mission: Mission statement of the institute (Busch & Binge 1993)

Research Institutes External Support **Key(s): Institute; Donor; Start Date**

Agreements on support from an outside institution to an Institute
Institute: Institute
Donor: Organization that supports this program
Start Date: Start date of support
Support Title: Support title
End Date: End date of support
Financial Support: Financial support
Technical Support: Technical support
Total Value: Total amount of financial support (if Financial support = Y)
Value Unit: Usually local currency or US$
Comment: Comment

Research Organizations **Key(s): Organization**

The top level of the hierarchy of research station, institute, organization;
Organization: Organization: usually a Ministry or University; usually the abbreviation or acronym
Organization Title: Organization title or name in full
Organization Mission: Mission statement of the Organization (Busch & Binge 1993)
Organization HQ Town: Name of the town, city etc where the organization has its HQ
Country: Name of the country in which this organization resides

Research Priorities **Key(s): RP Commodity; Research Year**

List the Commodities (crop or livestock species & non-commodity factors) and in some cases Commodity Groups which are
research priorities in the current county, along with their national and AEZ priority ranking.
RP Commodity: Name of priority commodity or group
Research Year: Priorities in this research year
Commodity Group: Optional, only needed if name of research priority commodity is a commodity group
National Research Priority: National priority ranking for research where 1 = top priority
AEZ1 Score: Priority ranking for Agro-Ecological Zone 1.
AEZ2 Score: Priority ranking for Agro-Ecological Zone 2 .
AEZ3 Score: Priority ranking for Agro-Ecological Zone 3 .

Research Stations **Key(s): Research Station Title**

The base level in the hierarchy of research station, institute, organization; usually the main site of actual field research
Research Station Title: Research station title or name
Research Station Place: Research station place
Research Station Mandate: A statement justifying the existence of the station through a description of its purpose
Research Station Head Name: Name of the current head of the station
Research Station Address 1: Research Station Address 1
Research Station Address 2: Research Station Address 2
Research Station Address PO Box No: Research Station Address PO box no

Research Station Address Town: Research Station Address Town
Research Station Soils Comment: A brief description of the main features of the soils of the station
Year Started: The year that the research station was started
AEZ: Agro Ecological Zone ; see AEZ lookup list
Institute: Institute to which the station belongs e.g. Research Branch
Research Station Abbreviation: Research station abbreviation for use by INFORM-R
Total Area: Total area in hectares
Arable Area: Suitable for cultivation
Pasture Area: Suitable for pasture research
Non Arable Area: Not suitable for cultivation
Latitude: Latitude
Longitude: longitude
Altitude: Altitude in m
Mean Annual Rainfall: Mean annual rainfall in mm

Research Stations Reports **Key(s): Research Station Title; Research Year**
Being the Research Station Directors annual report, derived from the constituent reports, and providing components of the
annual report of the research service
Research Station Title: Research station title or name
Research Year: Year to which this report applies
Research Station Report: Short report of the activities on this research station and its results

Salary Grades **Key(s): Salary Grade; Grade Level; Grade Year**
A lookup list of the salary grades and grade levels of the organization's pay scale, along with the corresponding total salary + allowances + benefits typical of this grade level.
Salary Grade: Salary grade within the organizations pay scale
Grade Level: The levels or steps within each salary grade in the pay scale.
Grade Year: Year in which these grades were introduced.
SAB: Salary, Allowances and Benefits (i.e. total direct costs of a staff member) associated with this salary grade level
Value Unit: Usually local currency or US$
Comments: Comments

Specialism **Key(s): Specialism**
A lookup list of optional sub-disciplines or specialisms within each Main Discipline
Specialism: Suggested specialisms within each discipline
Main Discipline: Main scientific discipline: see standard list of c. 16 in the Main Discipline lookup list
Specialism Description: Specialism description

Survey Types **Key(s): Survey Type**
Lookup list of common Survey types
Survey Type: Survey type
Survey Type Description: Survey type description

Annex

A3 Reference Lists of Agricultural Commodities and Commodity Groups

These lists, taken from the INFORM-R and modified from Gijsbers[1], enable system designers to adopt an existing data set. While any classification scheme is open to debate, this list draws on FAO's Agrovoc index, with adaptations from experience n a number of countries to better serve research management interests. (See section 2.2.2 for further discussion on commodities.)

A3.1 Commodity groups

Commodity group	Examples of commodities
Cereals	Maize, sorghum, millets
Fibres	Cotton, sisal
Fisheries	Tilapia, carp
Forestry	Acacia, Teak
Fruits and nuts	Cashew, citrus
Grain legumes	Beans (Phaseolus), cowpea
Industrial crops	Guar, jojoba, pyrethrum
Livestock	Cattle, goats, bees, pastures
noncommodity	Soils, irrigation, socio-economics
Oilseeds	Groundnuts, soybeans, sunflower
Ornamentals	Flowers
Roots and Tubers	Cassava, sweet potato
Spices	Pepper, vanilla
Stimulants	Tea, coffee, cacao
Sugar	Sugar cane, sugar beet
Tree crops	Rubber, coconut
Vegetables	
Vegetables, exotic	Cabbage, tomato, lettuce
Vegetables, indigenous	Amaranthus, Corchorus

1. Gijsbers, G. 1991. Methods and procedures for the development of INFORM. INFORM Guidelines, Part 2. The Hague: International Service for National Agricultural Research.

A3.2 Commodities

Some noncommodity items and some group items are included in this table to facilitate its use within the INFORM-R MIS in the comparison of agreed research priorities with resource allocations.

Commodity	Commodity group	Scientific name
Acacia sp.	Forestry	Acacia
Amaranthus	Vegetables, indigenous	Amaranthus
Apples	Fruits and nuts	Malus
Apricots	Fruits and nuts	Prunus armenciaca
Arrowroot	Roots and Tubers	Maranta arundinacea
Asparagus	Vegetables, exotic	Asparagus officinalis
Avocados	Fruits and nuts	Persea americana
Bambara groundbeans	Grain legumes	Vigna subterranea
Bamboos	Fibers	Bambusaceae
Bananas and Plantains	Fruits and nuts	Musa sp.
Barley	Cereals	Hordeum sp.
Beans	Grain legumes	Phaseolus vulgaris
Bees	Livestock	Apis
Betel	Stimulants	Piper betle
Broad Beans	Grain legumes	Vicia faba
Buffalo	Livestock	Bubalus bubalis
Cabbage	Vegetables, exotic	Brassica oleracear var. capitata
Cacao	Stimulants	Theombroma cacao
Calliandra	Grain legumes	Calliandra spp
Camels	Livestock	Camelus
Canes and Rattans	Fibers	Arundinaria spp.; Calamus rotang
Cardamom	Spices	Elettaria cardamomum
Carp	Fisheries	
Carrots	Vegetables, exotic	Daucus carota
Cashews	Fruits and nuts	Anacardium occidentale
Cassava	Roots and Tubers	Manihot esculenta
Castor Beans	Oilseeds	Ricinus communis
Catfish	Fisheries	
Cattle	Livestock	Bos
Cattle, beef	Livestock	Bos
Cattle, dairy	Livestock	Bos taurus
Cauliflower	Vegetables, exotic	Brassica oleraceae botrytis
Celery	Vegetables, exotic	Apium graveolens
Chickens	Livestock	Gallus gallus
Chickpeas	Grain legumes	Cicer arietinum
Chillies	Spices	Capsicum annuum
Chingofwa	Vegetables, indigenous	Commelina
Cinnamon	Spices	Cinnamomum zeylanicum
Citrus	Fruits and nuts	Other
Cleome	Vegetables, indigenous	Cleome
Cloves	Spices	Eugenia caryophyllus
Coconut	Tree crops	Cocus nucifera
Coffee	Stimulants	Coffea sp.

Commodity	Commodity group	Scientific name
Corchorus	Vegetables, indigenous	Corchorus spp
Coriander	Spices	Coriandrum sativum
Cotton	Fibers	Gossypium sp.
Cowpeas	Grain legumes	Vigna unguiculata
Crabs	Fisheries	
Currants	Fruits and nuts	Ribes sp.
Donkeys	Livestock	Equus asinus
Ducks	Livestock	Anas platyrhynchos (mallard), Cairina moschata (Muscovy)
Durians	Fruits and nuts	Durio zibethinus
Eggplant	Vegetables, indigenous	Solanum melongena
Engineering	noncommodity	
Equines	Livestock	Equus
Eucalyptus	Forestry	Eucalyptus sp.
Farming Systems	noncommodity	
Feed Crops	Livestock	
Flowers	Ornamentals	
Frogs	Fisheries	Rana
Fruits	Fruits and nuts	
Garlic	Vegetables, exotic	Allium sativum
Ginger	Spices	Zingiber officinale
Gliricidia	Forestry	Gliricidia sepium
Goats	Livestock	Capra
Grapefruit	Fruits and nuts	Citrus paradisi
Grapes	Fruits and nuts	Vitis
Grass pea	Livestock	Lathyrus sativus
Groundbeans	Grain Legumes	Vigna subterranea
Groundnuts	Oilseeds	Arachis hypogaea
Guar	Industrial crops	Cyamopsis tetragonoloba
Guavas	Fruits and nuts	Psidium guajava
Hemp	Fibers	Cannabis sativa
Hevea Rubber	Tree crops	Hevea sp.
Information	noncommodity	
Irrigation	noncommodity	
Jojoba	Industrial crops	Simmondsia chinensis
Jute	Fibers	Corchorus sp.
Kapok	Fibers	Ceiba pentandra
Kenaf	Fibers	Hibiscus cannabinus
Lemon Grass	Spices	Cymbopogon sp.
Lemons	Fruits and nuts	Citrus limon
Lentils	Grain legumes	Lens esculenta
Lettuce	Vegetables, exotic	Lactuca sativa
Leucaena	Forestry	L. leuocephala
Llamas and Alpacas	Livestock	Lama glama, L. pacos
Lima Beans	Grain legumes	Phaseolus lunatus
Limes	Fruits and nuts	Citrus aurantiifolia
Litchis	Fruits and nuts	Litchi chinensis
Livingstone potato	Roots and Tubers	Plectranthus esculentus
Llamas and Alpacas	Livestock	Lama glama
Lobsters and Crayfish	Fisheries	
Lucerne	Livestock	Medicago sativa
Macadamia	Fruits and nuts	Macadamia

Commodity	Commodity group	Scientific name
Macroptilium	Livestock	Macroptilium spp.
Maize	Cereals	Zea mays
Mangoes	Fruits and nuts	Mangifera indica
Many	noncommodity	
Melons	Fruits and nuts	Cucumis melo
Milkfish	Fisheries	
Millet	Cereals	
Millet, finger	Cereals	Eleusine coracana
Millet, pearl	Cereals	Pennisetum americanum
Mulberry	Livestock	Morus spp.
Mullet	Fisheries	
Multipurpose trees	Forestry	
Mung Beans	Grain legumes	Vigna mungo, V. radiata
Mussels	Fisheries	
Mustard	Oilseeds	Brassica spp
Ninde	Industrial crops	?Cassia senna
None	noncommodity	
Nutmeg	Spices	Myristica fragrans
Oats	Cereals	Avena fatua
Octopuses and Squids	Fisheries	Octopus vulgaris
Oil palms	Tree crops	Elaeis guineensis
Okra	Vegetables, indigenous	Hibiscus esculentus
Onions	Vegetables, exotic	Allium sepa
Oranges	Fruits and nuts	Citrus sinensis
Oregano	Spices	Origanum sp.
Oysters	Fisheries	
Papayas	Fruits and nuts	Carica papaya
Passion fruit	Fruits and nuts	Passiflora sp.
Pastures	Livestock	
Peaches	Fruits and nuts	Prunus persica
Pears	Fruits and nuts	Pyrus communis
Peas	Grain legumes	Pisum sativum
Pepper	Spices	Piper nigrum, Capsicum annuum
Pigeon Peas	Grain legumes	Cajanas cajan
Pigeons	Livestock	
Pine	Forestry	Pinus sp.
Pineapples	Fruits and nuts	Ananas comosus
Post-Harvest	noncommodity	
Potatoes	Roots and Tubers	Solanum tuberosum
Prawns	Fisheries	
Pumpkins	Vegetables, indigenous	Curcurbita sp.
Pyrethrum	Industrial crops	Chrsyanthemum cinerariaefolium
Rabbits	Livestock	Oryctolagus cuniculus
Ramie	Fibers	Boehmeria nivea
Rape	Vegetables	
Rice	Cereals	Oryza sativa
Rice Beans	Grain legumes	Vigna umbelleta
Rye	Cereals	Secale cereale
Safflower	Oilseeds	Carthamus tinctorius
Safflowers	Oilseeds	Carthamus tinctorius
Saffron	Spices	Crocus sativus

Commodity	Commodity group	Scientific name
Sago Palms	Roots and Tubers	Metroxylon laeve
Senna	Grain legumes	Cassia senna
Sesame	Oilseeds	Sesamum indicum
Sesbania	Vegetables, exotic	Sesbania
Sharks	Fisheries	
Sheep	Livestock	Ovis aries
Shrimps	Fisheries	
Silkworms	Livestock	
Siratro	Livestock	Macroptilium atropurpureum
Sisal	Fibers	Agave sisalana
Socio-economics	Noncommodity	
Soils	Noncommodity	
Sorghum	Cereals	Sorghum sp.
Soybeans	Oilseeds	Glycine max
Spices	Spices	
Spinach	Vegetables, indigenous	Spinacia oleracea
Squash and Marrow	Vegetables, indigenous	Cucurbita maxima ,C. pepo vars.
Strawberry	Fruits and nuts	Frageria spp.
Sturgeon	Fisheries	
Sugar cane	Sugar	Saccharum officinarum
Sunflowers	Oilseeds	Helianthus annuus
Sunnhemp	Noncommodity	Crotalaria juncea
Sweet Peppers	Vegetables, indigenous	Capsicum annuum
Sweet Potatoes	Roots and Tubers	Ipomoea batatus
Swine	Livestock	Sus
Tea	Stimulants	Camellia sinensis
Teak	Forestry	Tectona grandis
Tephrosia	Industrial crops	Tephrosia vogelii
Tilapia	Fisheries	Tilapia
Tobacco	Stimulants	Nicotiana sp.
Tomatoes	Vegetables, exotic	Lycopersicum esculentum
Triticale	Cereals	Triticum x Secale
Tronchuda	Vegetables	Brassica oleracea
Tuna	Fisheries	
Turkeys	Livestock	
Turmeric	Spices	Curcuma domestica
Turtles	Fisheries	
Unknown	Noncommodity	
Urena lobata	Fibres	Urena lobata
Vanilla	Spices	Vanilla fragrans
Vegetables	Vegetables	
Vegetables, exotic	Vegetables, exotic	
Vegetables, indigenous	Vegetables, indigenous	
Velvet Beans	Grain Legumes	Mucuna deeringiana
Vernonia	Vegetables, exotic	Vernonia
Vetiver grass	Livestock	Vetiveria zizanioides
Watermelons	Fruits and nuts	Citrullus lanatus
Watershed management	Noncommodity	
Wheat	Cereals	Triticum sp.
Winged Beans	Grain legumes	Psophocarpus tetragonolobus
Yams	Roots and Tubers	Dioscorea

Annex

A4 Worldwide Agricultural Research Information Systems and Sources

This annex provides agricultural scientists with examples of sources of agricultural research information.

Abstracts on Tropical Agriculture
See TROPAG

AfricaLink
AfricaLink supports programs of the US Agency for International Development (USAID) in agriculture, natural resource management, and the environment. It assists USAID's partner networks in Africa with access to information technologies and with effective information management.

AfricaLink targets the end users of information technologies, focusing in particular on the scientists and policy makers who are members of USAID's partner networks in the agricultural, environmental, and natural resource management sectors.

AfricaLink is essentially a support activity, and therefore works under the direction of USAID staff that already have a programmatic relationship with African networks. USAID staff request assistance from AfricaLink for their partner networks in order to further their strategic objectives.

For example a member of USAID staff might have a close working relationship with a regional network of maize research scientists, and will request AfricaLink assistance in order to improve the capacity of that network to collaborate and to share information through the use of information and communication technologies, including but not limited to the Internet.

Website: www.info.usaid.gov/regions/afr/alnk/

ASARECA (Association for Strengthening Agricultural Research in East and Central Africa)
ASARECA was established by the 10 National Agricultural Research Institutes of Burundi, the Democratic Republic of Congo, Eritrea, Ethiopia, Kenya, Madagascar, Rwanda, Sudan, Tanzania, and Uganda, to promote cooperation between the regional research systems towards greater effectiveness and efficiency. The association has a number of research networks, e.g., agroforestry, root crops, beans, potato and sweetpotato, and banana. In 1997 it established three additonal networks: sorghum and millet, to be "backstopped" or provided with an advisory service in the

field by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT); maize and wheat, backstopped by the International Maize and Wheat Improvement Center (CIMMYT); and dairy, beef and small ruminants.

A quarterly newsletter, AgriForum, provides information about the regional research networks and programs as well as other activities related to cooperative work in the region.

ASARECA Secretariat, P.O. Box 765, Entebbe, Uganda
Fax: +256 42 21126 / 21070
Email: asareca@imul.com asareca@imul.com
Website: www.asareca.org

AGLINET (Agricultural Libraries Network)
AGLINET is operated by the Library and Documentation Systems Division (GIL) of the Food and Agriculture Organization of the United Nations (FAO).

Website: www.fao.org/library/dlubin/aglincen/aglinlst.htm

AgNIC (Agriculture Network Information Center)
AgNIC provides access and links to agriculture-related information, subject-area experts, and other resources. It was established by an alliance of the US National Agricultural Library, land-grant universities, and other organizations committed to facilitating public access to agricultural and related information.

The AgNIC Alliance is an informal association of US land-grant university libraries and other agricultural libraries, extension services, and other organizations. It provides access to high-quality agricultural information available on the Internet. Members take responsibility for small vertical segments of agricultural information (such as basic, applied, and developmental research; extension; and teaching activities in areas such as food, agriculture, renewable natural resources, forestry, and physical and social sciences).They develop Websites in their specific area of responsibility.

Website: www.agnic.org

AgREN (Agricultural Research and Extension Network)
The AgREN network of the Overseas Development Institute (ODI) was established in the mid-1980s to link policymakers, practitioners, and researchers in the agriculture sector of developing countries. AgREN was founded on a strong belief in the importance of information exchange and learning from both positive and negative experience. It aims to provide its members with up-to-date information and the opportunity to maintain a dialog with others who have similar professional interests.

Agricultural Research and Extension Network
Overseas Development Institute
Portland House, Stag Place
London SW1E 5DP

United Kingdom
Telephone: +44 171 393 1600
Fax: +44 171 393 1699
Internet: www.odi.org.uk/agren

AGRICOLA
AGRICOLA is a bibliographic database covering most aspects of agriculture and food. Managed by the US National Agricultural Library (NAL), Maryland, USA, it represents the catalogs of NAL and several other libraries. There over 3 million records dating back to 1970, of which some 90% are journal articles. There is an increasing focus on publications from the USA to avoid overlap with AGRIS. The collection is accessible on-line and on CD-ROMs.

Website: www.nal.usda.gov/ag98/

AGRIS
Started in 1975, AGRIS is a bibliographic database of FAO covering all aspects of agriculture and food. It comprises contributions from 159 national centers and some regional and international centers, and is approaching 3 million records (1998 data). Most refer to journal articles but about 16% refers to "gray" literature (literature that is not commercially available either on its own or in a journal on the market). The inclusion of gray literature provides a potentially valuable service as much of this would otherwise be unknown outside the organization of its origin. One has to apply to individual authors or their institute to obtain a copy of the item. About a fifth of AGRIS records include an abstract.

A monthly report of new entries is produced as AGRINDEX, and much of the system is available on line and on CD-ROM. There is an associated thesaurus of agricultural terms, in English, French, and Spanish, called AGROVOC.

See also FAO; WAICENT

Website: www.fao.org/agris

AROW (Agricultural Research Organizations on the Web)
Initiated and managed by ISNAR, the AROW site contains a worldwide directory of over 1,100 organizations and universities working in agricultural research that have a home page on the Web.

Internet: www.cgiar.org/isnar/arow/index.htm

Bibliography of Agricultural Information in Africa
This chapter by Ikhizama in Aino (1995) gives a useful list under the following headings: general agricultural information (including short country lists); bibliographic sources; communications / media; education and training; extension; information

technology; libraries and documentation centers; management; networks; researchers; retrieval; rural development; user studies.

Aina, L.O., A.M. Kaniki, and J.B Ojiambo. 1995. Agricultural Information in Africa. Ibadan, Nigeria: Third World Information Services

CARIS (Current Agricultural Research Information System)
Started in 1975, CARIS is a database of current agricultural research projects being carried out in or on behalf of developing countries. Detailed project data is supplied by 132 national and 19 international and 6 regional centers. Some 30,000 currently active projects are recorded. Details include title, objectives, start and expected end dates, and identification of the station and researchers responsible. Subject coverage is wide: all aspects of agriculture and food. Outputs consist of magnetic tape and diskettes with data from the whole collection or from a single country or region.

See also FAO; WAICENT

Website: www4.fao.org/caris

CAB ABSTRACTS
UK-based CAB International publishes a database with abstracts from worldwide sources, including some 12,000 journals and about 3 million records from 1973 to date. The abstracts are a major advantage to scientists in remote locations, where obtaining the original article is often not feasible. Output is through the publication of some 50 subject-matter monthly abstract journals, on-line databases, and magnetic tape and CD-ROMs.

Website: www.cabi.org/publishing/products/database/abstracts/index.asp

CGIAR (Consultative Group on International Agricultural Research)
The CGIAR, established in 1971, is an informal association of some 60 public- and private-sector members that supports a network of 16 international agricultural research centers. CGIAR's mission is to contribute to food security and poverty eradication in developing countries through research, partnership, capacity building, and policy support. The CGIAR promotes sustainable agricultural development based on the environmentally sound management of natural resources. The CGIAR Website provides links to the sites of each of the 16 international agricultural research centers.

Website: www.cgiar.org/

CRIS (Current Research Information System)
CRIS is the documentation and reporting system of the US Department of Agriculture (USDA) for ongoing and recently completed research projects in agriculture, food and nutrition, and forestry. It contains details of over 30,000 current publicly sup-

ported agricultural and forestry research projects of the USDA agencies, the State Agricultural Experiment Stations, the state university land-grant system, and other cooperating state institutions. Approximately 4,000 new project descriptions and about 20,000 progress and publication reports updating existing projects are entered in CRIS annually. The database is updated weekly.

It covers nearly all fields of agriculture including crops, livestock, poultry, fish, forestry and natural resource management. It includes as a subfile the Inventory of Canadian Agri-Food Research (ICAR). This is an up-to-date database for agricultural and food research in Canada and describes over 4,000 projects from industry, universities, and provincial and federal establishments and is produced by the Canadian Agri-Food Research Council.

The CRIS database may be accessed directly over the Internet, on an AGRISEARCH CD-ROM produced by SilverPlatter Information and through DIALOG, the commercially available on-line retrieval system of Knight-Ridder Information.

Others who would like a search to be conducted on their behalf may submit requests, preferably by letter, specifying: name, mailing address, and telephone number, date information needed, topic or subject area to be searched (within 100 words) and the purpose of the request, to:

Current Research Information System, Science and Education Resources Development
Cooperative State Research, Education, and Extension Service/USDA
National Agricultural Library Bldg, 5th Floor
Beltsville, MD 20705-2351, USA.
Fax: (301) 504-6272
E-mail: cris@cris.nal.usda.gov
Website: http://cris.csrees.usda.gov/menu.html

CTA (Technical Centre for Agricultural and Rural Cooperation)

CTA was established in 1983 under the Lomé Convention between the African, Caribbean and Pacific (ACP) States and the European Union Member States. It provides services to improve access to information for agricultural and rural development for ACP countries, and strengthens these countries' capacity to produce, acquire, exchange and utilise information in these areas. (See also the case study by Niang in section 6.3 of this volume.)

CTA
P.O. Box 380
6700 AJ Wageningen
The Netherlands
E-mail: cta@cta.nl
Website: www.agricta.org

EFITA (European Federation for Information Technology in Agriculture, Food and the Environment)
EFITA's mission is to facilitate the exchange of information and experience, to develop knowledge in the area of ICT in agriculture in order to enhance the competitiveness of Europe, and to promote the awareness of ICT in agriculture. To accomplish these missions EFITA enriches the activities and harmonize the efforts of its member organisations and all working units. It initiates, proposes, participates in, and contributes to new activities and projects. It also initiates, participates in, and contributes to projects at European level. It organizes international congresses, seminars and fora. It promotes the publication of a leading journal, reports and recommendations. It initiates and maintains contacts with other organizations on an international level. It aims to be at the leading edge of ICT developments and the application of ICT in agriculture.

Website: www.efita.org

ELDIS (Electronic Development and Environment Information System)
ELDIS provides access and links to a wide range of information on development and environmental issues. Its Website is concerned particularly with agriculture. It offers a data and database hosting service.

Tel: (01273) 606261
Fax: (01273) 621202 or 691647
Website: http://nt1.ids.ac.uk/eldis/agric/agric.htm

FAO (Food and Agriculture Organization of the United Nations)
FAO was founded in 1945 to raise levels of nutrition and standards of living, to improve agricultural productivity, and to better the condition of rural populations. Today, FAO is the largest autonomous agency within the United Nations system with 175 member nations plus the EC and more than 1,500 professional staff.

See also AGRIS; CARIS; WAICENT

Website: www.fao.org

IAALD (International Association of Agricultural Information Specialists)
IAALD facilitates professional development of and communication among members of the agricultural information community worldwide. IAALD's goal is to enhance access to and use of agriculture-related information resources. IAALD promotes the agricultural information profession, supports professional development activities, fosters collaboration, and provides a platform for information exchange.

Website: www.lib.montana.edu/~alijk/iaald.html

Leland Initiative

The Leland Initiative is a five-year, US$15 million project of the US government to extend full Internet connectivity and other global information infrastructure technologies to about 20 African countries in order to promote sustainable development. Most of the activities are geared towards environment, democracy, health and agricultural development, connectivity for NGOs, training, Internet gateway and Internet awareness. The other activities include global trading, which is intended to bring computers to African schools.

USAID - Leland Initiative
2201 "C" Street, NW
Room 2744/NS
AFR/SD/SA
Washington, DC 20523-0089
Fax: 202-216-3373
Email: leland.initiative@usaid.gov
Website: www.info.usaid.gov/regions/afr/leland

Overseas Development Institute

See AgREN

PICTA (Partnership for Information and Communication Technologies in Africa)

PICTA is an informal group of donors and executing agencies committed to improving information exchange and collaboration around ICT activities in Africa. It builds on the work of the African Networking Initiative (ANI) and the African Internet Forum (AIF). This web site, its resources, and the PICTA-CL discussion forum act as a common reference point for ICT-related development activities on the continent. As a resource to its members and to the public at large, PICTA maintains the African ICT Activity Information Management System (AI-AIMS), which is a set of information resources designed to exchange knowledge on ICT activities in Africa.

E-mail: picta@bellanet.org picta@bellanet.org
Website: www.bellanet.org/partners/picta/

SAAINET (Southern African Agricultural Information Network)

SAAINET's main objective is to promote sharing, exchange, and dissemination of information to effectively meet the agricultural information needs in the Southern African Development Community region. The project is funded by the International Development Research Centre (IDRC) and CTA.

SITIA (Service for Information Technology in International Agriculture)

SITIA is a USA-based NGO focusing on the development of information technologies that lead towards sustainable agriculture, natural resource conservation, and rural economic development.

SITIA
1040 Hamilton Court
Menlo Park, California 94025 USA
Telephone: 1-650-833-6035
Fax: 1-650-325-2313
E-mail: m.perry@cgnet.com
Website: www.sitia.org

SilverPlatter Information
SilverPlatter produces a wide range of CD-ROMs including those with the agricultural databases: AGRISEARCH, AGRICOLA, and AGRIS.

SilverPlatter Information, Inc.
100 River Ridge Drive
Norwood, MA 02062-5026, USA.
Website: www.asog.co.at/silverplatter/index1.htm

SINGER (System-wide Information Network for Genetic Resources)
The CGIAR's genetic resources information exchange network SINGER provides common access to information concerning the collections of plant genetic resources held by the 11 CGIAR centers that are active in the study of plant and animal genetic material. Together, these collections comprise over half a million samples of plant and animal (including fish) germplasm of major importance for food and agriculture.

SINGER links the genetic resources databases of the CGIAR centers and allows searches for information relating to the identity, origin, characteristics, and distribution of the genetic resources in the individual center collections and access to further specific data on the collections, e.g., crop characterization data. The infrastructure of SINGER ensures that the central network database is kept current relative to the individual center databases. The system is managed by the International Plant Genetic Resources Institute (IPGRI).

The SINGER database is a relational model with some 21 tables, built in Microsoft SQL Server software running on Microsoft NT Server. The central database is managed on contract by CGNET in the USA. Each of the centers maintains a replicate copy of the SINGER database, with a link to their own existing genetic resources database. Updating from the research centers to the network operations center is through IVDN lines or via high-capacity magnetic media. Updating is done by the largely automatic process of replication, a facility built into Microsoft SQL Server software.

Secretariat of SINGER, c/o IPGRI
Via delle Sette Chiese 142
00145 Rome
Italy
Tel: 39-6-51892-225
Fax: 39-6-575-0309
E-mail: singer@cgiar.org

Website: http:/singer2.cgiar.org

TROPAG (Abstracts on Tropical Agriculture)

TROPAG is a database of the Royal Tropical Institute, Amsterdam, Netherlands, with abstracts of articles and other publications on tropical and subtropical agriculture with an emphasis on items of practical value to extension and development project workers. The source includes some 6,000 journals. Abstracts on Tropical Agriculture is printed as a monthly journal and the database is available on-line, on magnetic tape and CD-ROM. There is a useful geographic index, in additional to the usual subject matter index.

Website: www.kit.nl/ibd/html/body%5Ftropag.htm

WAICENT (World Agricultural Information Center)

FAO's WAICENT is a collection of information services that enables remote users to gain access to much of FAO's large collection of information on agriculture, fisheries, and forestry. Of particular interest to us are FAOSTAT, AGRIS, and CARIS. FAOSTAT provides on-line access to a database of over one million time series records covering international statistics in trade, agricultural production, food balance sheets, fertilizers, pesticides, and others. It is also available on CD-ROM, diskettes, and as printed annual yearbooks.

Website: www.fao.org/waicent/search/default.htm

Annex

A5 A Syllabus for Training in the Implementation of a Management Information System

Training is crucial to the successful implementation of an MIS. An outline syllabus is offered here, based on ISNAR experience in several countries. The approach recognizes three groups of users, each with different training needs. Research managers, at both station and HQ levels, need to be able to access all the built-in reports of the system. This assumes each has a personal computer on their desk. MIS practitioners, and their clerical support staff where available, need additionally to manage data entry, and some processing. Interested researchers should be allowed to follow either of these two training schedules. The MIS national coordinators will need a much deeper insight into the system such that they can competently modify it in response to the user community's needs. This call for competence also in the underlying software.

1. Research managers (1–2 days)

Objectives: *To provide research managers, with no assumed computer competence, with an overview of the MIS's main facilities, and the ability to find management information in all the areas that the system covers.*

Topics:

The menu system:
- A conducted tour (modify to match the menu of the MIS being implemented)
- Difference between reports and data management

Built-in reports:
- On-screen v. print to paper
- Human resources and research activities
- The national directories
- Aggregate data
- Detail: person and research activities
- Research reports: research results; monitoring & evaluation

Features of the report window:
- The buttons; record and report selectors; special messages
- Access toolbar buttons: e.g., "Find" and "Sort" buttons

Structure of data:
- Concept of tables – one for each entity
- Main entities: experiments, persons,
- Look-up tables: e.g., commodities, stations, disciplines

Metadata:
- List of outputs, and their management notes
- Reports questionnaire: User comments form

Special topics:
- The research year
- Controlled and uncontrolled terms; the discipline field

Management Issues:
- Modifying the Ssstem:
  - how reports can be modified using devices within the output
  - how new outputs can be obtained through the INFORM practitioners
  - how the database can be modified, usually through national consultation and through the national research information officer

2. MIS pactitioners at each station, and supporting clerical staff (1–2 weeks)

Objectives: *To provide the practitioner with the ability to manage the system on a day-to-day basis, including routine data capture, data entry, and report generation, and annual dispatch of the station data set to national HQ.*

Knowledge assumed

- Content of the syllabus for research managers
- Familiarity with basic Windows features:
  - use of "scroll bar" to scroll down a list; how this compares with the page-down keyboard key
  - move from one tiled window to another
  - change the position, width, height of a window on the screen
  - maximize, minimize and standard size of a window
  - exit to previous screen, "open door" icon
  - "Tab" rather than Enter to move to next item
  - toolbar and menu choices: faint means not available from here.
  - field and character-within-field views; F2; Shift-F2
  - use of the mouse
  - use of Windows keyboard shortcut Win-E for Windows Explorer

For a short on-screen Windows 95/98/2000 tutorial, click the Start button in Windows, Help, "Ten minutes to using Windows" or "Introducing Windows." Then see the several other items in Start, Help, Contents.

Familiarity with the database objects:
- tables, forms, queries, and reports
- ability to create and modify these

- tables:
  - add a new record
  - copy a (changed) table (to send to another site for example):
  - Alternative using a text delimited structure
- Dynaset: A dynamic subset, or selection, of data in selected fields in selected records of one or more selected table(s). A dynaset is generated through a query.
- List box: list of items is fixed, and is always displayed in full.
- Queries: types: select, update, delete
- Combo box: addition of new items can be allowed; list is not displayed until selected.
- In both list and combo box: the list may be typed in as a *value list*, or may come from an existing table or query.
- Toolbars: the main ones used and their functions, how to open, close, move.
- Error messages
- Graph generation (e.g., use of MicroSoft Graph)

System features:
- Installation
- Need for and methods of making backups
- The research year
- Controlled and uncontrolled terms; the discipline field
- Data capture
  - printing data capture sheets
  - entering data
- Look-up tables
- Data consistency check
- Printing reports: (check number of pages first.)
- INFORM-R configuration: the ZZSysConfig Form

Metadata:
- (List of outputs, and their management notes)
- (Reports questionnaire: User comments form)
- List of tables and fields and their descriptions
- The database window – F11

3. MIS national coordinators

Objectives: *To provide the ability to manage the national data set, including the (annual) merging of the station data sets, checking and correcting the data, and modifying the system to meet the changing needs of the organization.*

A national coordinator will need:
- an in-depth knowledge of the underlying database software (e.g. Access, Dbase, Paradox, etc.), in the version in use in that country.

- good programming ability where the software allows: e.g., for Microsoft Access database, programming in Visual Basic may be advisable or, if already in use in the system, essential.
- a good knowledge of the essentials of database theory
- in-depth knowledge of the MIS:
  - familiarity with all components covered by the menu system
  - a full understanding of the support to the experiment life cycle (including handling of proposals, proposal approval, stage reports e.g. interim, monitoring, annual, evaluation, completion and the difference between data and metadata reports)
  - the mechanisms for handling move to new year
  - the mechanisms for generating all reports (including the underlying queries, record selection criteria e.g. experiment v. replicate,)
  - full understanding of the processes of station to national aggregation and distribution back to station
  - familiarity with all data tables
  - familiarity with all system tables
  - familiarity with the underlying Visual Basic code wherever this is employed
- in-depth knowledge of the contents of the Research Managers and MIS Practitioners syllabuses.

References from the Annexes

Codd, E.F. 1970. A relational model of data for large shared data banks. CACM 13 (6), June 1970. [An abstract is given in Date (1990)].

Codd. E.F. 1985. Is your DBMS really relational? *Computerworld*, October 14, and Does your DBMS run by the rules? October 21.

Date, C.J. 1990. An introduction to database systems. Volume 1. Fifth Edition. Reading, USA: Addison-Wesley

Gijsbers, G. 1991a. Methods and procedures for the development of INFORM. INFORM Guidelines Part 2. The Hague: ISNAR.

Horton, D., P. Ballantyne, W. Peterson, B. Uribe, D. Gapasin, and K. Sheridan. 1993. Monitoring and evaluating agricultural research. A sourcebook. Wallingford: CAB International, and The Hague: ISNAR

Index

A key to acronyms and abbreviations is included in the index; another list is available on page xxiii at the beginning of the book; a glossary is shown in Annex A1 on pages 289–293.

W

Z

Covers: Chlorine-free paper
Inside pages: Recycled paper

Produced by ISNAR Publication Services

Text editing:	Jan van Dongen
Index:	Indexing Specialists, Hove, UK
Design:	Richard Claase and Jan van Dongen
Layout:	Jan van Dongen
Cover:	Richard Claase
Printer:	Rapporten Service Drukkerij bv, Rijswijk, The Netherlands