EGYPT DEED REGISTRATION SYSTEM PROPERTY INDEX MAPS – QUALITY ASSURANCE / QUALITY CONTROL

EGYPT FINANCIAL SERVICES PROJECT
TECHNICAL REPORT #92

December 18, 2006

This publication was produced for review by the United States Agency for International Development. It was prepared by Chemonics International Inc.

DATA PAGE

Submitted by:	Raymond Struyk, Chief of Party Egypt Financial Services (EFS) Project 4 Hayet El Tadrees Square Dokki, Cairo, Egypt Tel: (20) 2 762-6140 Fax: (20) 2 762-6150 www.egyptfs.com Contract No. 263-C-00-05-00003-00
Submitted to:	EFS CTO: Paul Bruning EFS DCTO: Ingi Lotfi Private Sector Programs Office of Policy and Private Sector USAID Mission to Egypt
Task:	Quality Assurance / Quality Control Specifications
KRA:	KRA 2.12 Implementation Plan for Improved Registry Office and Related Infrastructure
Activity:	Activity 2.12.3 Detailed Plan of Operation and Work Schedule to Establish Registry Office in Mokattam
Author:	Sean Johnson, Land Formalization Specialist, ILS Inc.
Managed by:	R. Dougal Menelaws
Reviewed by:	Eng. Mosaad Ibrahim and Eng. Fatma Abdel Kader, Task 2 Team Members
Date:	December 18, 2006
Version:	1.0 (first issue)
Keywords:	Property Index Maps, quality assurance, quality control, quality evaluation, accuracy, specifications, standards, cadastral maps, metadata

This publication was made possible through support provided by the Office of Financial and Information Technology, U.S. Agency for International Development, under the terms of Contract No. 263-C-00-05-00003-00.

TABLE OF CONTENTS

Acronyms

ESA	Egyptian Survey Authority
GIS	Geographic Information System
GPS	Global Positioning System
IRO	Improved Registry Office
QA	Quality Assurance
QC	Quality Control
PID	Property Identification number
PIM	Property Index Map
REPD	Real Estate Publicity Department
RMSE	Root Mean Square Error
RO	Registry Office

Executive Summary

The purpose of this specification is to provide assurance to the Egyptian Survey Authority and the administrators and users of the (personal) deeds registration system that Property Index Maps are created to a certain standard and quality.

The Property Index Map (PIM) identifies and locates property objects – parcels and buildings – and provides an index to more detailed and precise information, such as deeds and cadastral data. Because the PIM is not used to determine the area of a property or the location and dimensions of boundaries, it does not need to be as accurate as a cadastral map. This primary use – as an index to more details and accurate data – is reflected in the specification for map content and accuracy.

The management of quality requires that a product must first have a specification, against which quality measures and controls can be applied and quality evaluation conducted. A technical specification is documented elsewhere and summarized here. The quality evaluation measures will be applied to the following geographic data elements defined in the specifications: - lineage, currency, positional accuracy, attribute accuracy, and logical consistency.

Production of the PIM comprises to a large degree the compilation of existing geographic data, and the quality evaluation must reflect that the contractor is not always the primary source of the data. Moreover, as the contractor is responsible for all production processes – from compilation to final delivery – the management and control of quality is responsibility of the contractor. Quality is assured by the evaluation of a product created in conformance with a specification and reported in metadata.

User requirements for the PIM specifies that it is a feature based geographic data model composed of two key spatial entities: parcels and buildings. These will be represented by points at the end of lines that form closed polygons. As objects, these features require logical consistency: that is, lines must join lines, polygons must be closed, and with no gaps or overlaps.

To achieve the specified accuracy standards, quality control in production is required. This includes simple measures of supervision and automated validation, and will apply to all production steps such as scanning, geo-referencing, field surveying etc.

Quality evaluation will be a two-stage process, duplicated by both the supplier and the buyer. Prior to delivery of the PIM the supplier is expected to have undertaken a raft of quality checks, automatic and manual, covering geometric integrity, topological consistency, accuracy, and completeness. The second stage repeats the same quality evaluation check, but this time undertaken by the buyer, at the time of delivery but prior to acceptance. If all quality checks are passed the product is accepted; if any fails the product is reject and returned to the supplier.

Quality control checks comprise the following:

1. Geometric integrity (no line duplication, overlaps, overshoots, undershoots, gaps, or artifacts)
2. Topological integrity (polygons are closed, defined as area objects, no gaps, no overlaps, no slivers)
3. Parcels are not split by buildings, and buildings are not split by parcels (buildings must be completely within or wholly or partially coincident with parcels unless the property is a "villa" that spans two parcels)
4. All parcels are within block polygons; all blocks are within section polygons; all section polygons are within city polygons; and, all city polygons are within governorate polygons.
5. All parcels have internal PID points and text; all blocks have internal identification number; all sections have internal identification number; all cities have internal identification number; all governorates have internal identification number.
6. All lines are associated with polygons
7. All polygons are labeled
8. Visual spatial comparison of PIM vector dataset with source vector data, either in a software or using hardcopy check plots. If at any point the vector difference between the captured and source material exceeds 2.5 (two point five) meters, the error must be flagged (reported), investigated and corrected if appropriate. Note: large discrepancies may arise due to source material and differences between planned and actual plot layouts, which is not a quality control issue, but still needs to be reported.
9. The number of parcels on input materials matches the number of parcels in the PIM
10. All roads have annotation text
11. Positional accuracy is checked by:
    a. Ground 'truthing' 9 (nine) well-defined feature points per square kilometer[1], either parcel or building corners, distributed uniformly as possible, by comparing vector

[1] This represents, approximately, a 0.5% sample of independent (feature) points.

map data coordinates with real world coordinates. The dataset will pass the absolute positional accuracy quality check if less that 5% of the absolute position comparative results are greater than 2.5 (two point five) meters.

b. Comparative measurements of 9 (nine) well-defined feature sides per square kilometer, both parcel boundary and building lines, on the ground and the corresponding line on the map, distributed throughout the project area, and if less than 5% of the relative position comparative tests are greater than the ratio of 1:200, which is error related to ground distance (e.g. +/-0.5m for 100 meters on the ground) the dataset will pass the relative positional accuracy test.

Note: a small sample size is justified because (i) the absolute positional accuracy is also checked indirectly (below); and (ii) most of the data points are derived and compiled from existing maps and geo-referenced in blocks.

12. Visual spatial comparison in software of PIM vector dataset with very high resolution remotely sensed and rectified raster satellite images (ground resolution of equal or better than 1.0 meters). If at any point the vector difference between the captured material and the image exceeds 2.95 (two point nine five) meters, the error must be corrected.

13. All attribute table record fields are populated (those that require to be populated) with correctly formatted data. A software routine should count the number of records and the report the number fields that are empty; if more than 5% of fields that must contain data do not contain data then the data set is rejected.

14. All attribute table records with primary keys (parcel PID) are related to parcels in the PIM; all child records in linked tables (unit PID) have parent records in primary table. Any orphan records will result in rejection of the dataset (i.e. compliance must be 100%).

15. To check attribute accuracy a sample of 16 properties (parcels and buildings) well distributed per square kilometer throughout a project area or zone, where data is independently recollected and compared against the originally collected data. If more than 5% of the sample records contain different data from the original records, the dataset will be rejected and the errors corrected.

16. If an attribute dataset has been rejected because of inaccuracy, and after errors have been corrected, another sample check will be undertaken, with the same number but different sample properties. If more than 5% of the sample records contain different data from the original records, the dataset will be rejected and the errors corrected.

17. The quality checks will be continued, with different sample data each time, until the dataset passes the quality evaluation measure. Note: Although a strictly random sample is not required, the sample should be well distributed and representative; sample points cannot be repeated in a later sample.

A quality control and evaluation report, along with a metadata file, must document, as a minimum: lineage, currency and quality; quality control measures; quality evaluation checks applied; name of person who performed the internal quality control; time spent on quality control; number, type and size of errors found in the internal quality control, and the measures taken to correct the errors.

In summary:

Characteristic	***Measure(s)***	***Count***	***Indicator***
Data format and structure	Conformance with specification	1	Yes / No
Completeness	Spatial coverage; duplicate data; missing data	All	95% of all feature objects (parcels and buildings)
Logical consistency (data integrity)	Entity relationships; links; value range; gaps; overlaps; undershoots; overshoots	All	95% of all entities
Positional accuracy	Absolute domain; relative domain	9/sq.km; 9/sq.km	95% within 2.5m; 95% within 1:200
Attribute accuracy	Correctness, character length, format and values	All	95%

QUALITY ASSURANCE AND CONTROL FOR PROPERTY INDEX MAPPING

SECTION 1 – INTRODUCTION

A. Scope

This Quality Assurance (QA) and Quality Control (QC) document applies to cadastral operations to support the establishment of an improved deeds registration (Siguel El-Shaksi) system in Egypt. Specifically, these cadastral operations are applicable for the production of property index maps by compilation of existing geographical datasets and by mapping techniques. These specifications do **not** cover techniques of cadastral surveying for boundary demarcation, subdivision, and the production of cadastral diagrams, plans and mapping.

This document on quality assurance and control is designed to provide the Egyptian Survey Authority (ESA) with the means to assess the quality of a data set (property index map) created and delivered by an external source or contractor. The assessment of data quality is made against predefined specifications or standards, which are documented elsewhere.

The Egyptian Survey Authority is both a data user and supplier. In the context of land registration, ESA is a data supplier. The end users of property index maps are the property owing public (customers), through an intermediary, the Improved Registration Office (IRO). The IRO is the local point of customer contact for the Real Estate Publicity Department (REPD) of the Ministry of Justice in any particular administrative area.

B. Background

The purpose of property index maps is twofold:

a. To enable identification of properties through location and position with respect to administrative areas (governorates, districts, municipalities) and to adjoining properties (parcels and buildings).

b. To assign and administer unique property identification numbers that will serve as an index to all deeds attesting to title and documents supporting registration and processing of transactions.

From a data user's perspective, data quality is the applicability of a dataset for a particular use – its fitness for purpose, in other words. From a data provider's perspective, data quality is the degree to which data in a dataset conforms to product specifications.

It is important that the data provider, in this case ESA, understands as clearly as possible, customer's uses for property index maps. This is required in order to interpret and implement the map specifications. It is also needed for understanding and evaluating information about the data (metadata) provided to customers on data quality (specifications) and history (lineage) to help customers decide on fitness of the dataset for their purpose. These user and provider perspectives may be referred to as 'external quality' and 'internal quality', respectively, and the overall quality of a geographical information product, such as a property index map, is dependent on the standard of both external and internal quality. This quality assurance specification covers both the internal and the external aspects of quality assurance: product specifications, quality evaluation and reporting (metadata).

C. Definitions

In this document, the following definition of terms apply:

Quality assurance:	The means by which an institution satisfies itself that the standards and the quality of its products and services can be maintained and enhanced.
Quality management:	That aspect of the overall management function that determines and implements quality policy.
Quality control:	The operational techniques and activities that are used to fulfill requirements of quality. Quality control verifies that those systems used to monitor the delivery of products and services are being carried out satisfactorily.
Quality assessment:	The identification of those issues, or problems, which are attributable to the influence or impact of any scheme for the assessment of quality. The emphasis is upon "measurability" against some framework, which represents dimensions of quality – a standard.

SECTION 2 – QUALITY ELEMENTS

A quality element is a measurable, quantitative component of information about data quality. Generally, there are five data quality elements: currency (temporal validity), positional accuracy, attribute accuracy (thematic accuracy), logical consistency and completeness. In addition, the non-quantitative component, lineage, provides further information about data quality.

A. Lineage

Lineage describes the history of a dataset – the sources, methods and processes used in its creation and maintenance, from real-world abstraction and capture through to representation in spatial digital data. However, full lineage reporting for the points, lines and areas is not recorded within the spatial dataset but as an attribute to the feature object that is defined by the spatial data. In the case of property index maps, attributes of the real property feature object (parcels and buildings) would provide information about the date the feature geometry or attribution was last updated, update status (indicating the type of change applied to a feature), change history, and the acquisition method for coordinates of data points.

B. Currency

Currency describes how up to date the data are. This is a parameter of the update policy applicable to the dataset. For some datasets, such as topographic mapping, the update policy may be determined by balancing user requirements with economic resources. For example, a map sheet is not brought up to date until there is a sufficient number of feature changes that justifies revising and re-publishing (distributing) a new version. For property index mapping, the update policy requires that the map is continuously revised; that is, changes taking place to feature geometry are immediately reflected in the spatial dataset.

The currency attribute for property index map is set at the date for its initial production (delivery). Thereafter, continuous revision is effected by implementation (and enforcement) of the procedures for property mutation that ensure real time consistency between the map and registry components of the deeds registration system. This real time relationship is two-way: the property index map is not updated until a transaction takes place in the registry, and a transaction in the registry is not committed (finalized and a property created or changed) until the index map is updated.

C. Positional Accuracy

Positional accuracy can be defined as the measure of the digital representation of a feature point with respect to its position on the ground in both absolute and relative terms. Absolute accuracy is measured against actual position of the feature point on the surface of the earth with respect to a coordinate reference system; relative positional consistency of a feature point is in relation to other local points.

Positional accuracy is measured using predefined limits within which check comparisons must fall. Both absolute and relative positional accuracy limits are usually based on the scale of the representation of the spatial data set or map. The absolute accuracy of large-scale vector data can be represented by root mean square error[2] values in relation to map scale, which in turn dictates the methods and resolution of data collection. Table 1 indicates typical accuracy assessments of large-scale topographic mapping produced from aerial photographs (by photogrammetry).

Mapping scale	***Root mean square error***	***95% confidence level***	***99% confidence level***
1:1,000	+/-0.50m	+/-0.86m	+/-1.07m

Table 1. Absolute accuracy of large scale topographic map data (estimated) in meters

Measures of relative positional accuracy are both subjective and objective. A feature that is straight on the ground should be represented as a straight line on the map, and visual interpretation is an effective, if subjective, way of making this comparison. The objective measure for relative positional accuracy is also dependent on map scale, and a distance measured on the ground should be within predefined limits for the same distance derived from the map. For example, the estimated standard deviation between an actual measurement and a derived distance of the same feature on 1:1,000 scale mapping should not exceed 1:300. A ratio is the most appropriate way to express relative accuracy; for example, a measurement on the ground of 100 meters should be represented on the map by a derived distance that is within +/-0.33 meters.

A factor that compromises positional accuracy, particularly absolute accuracy, is the lineage of the map. A property index map that has been prepared, or partly prepared, from existing material, which has been scanned and geo-referenced (transformed), will be less accurate than a map compiled directly by feature measurement. Depending on the lineage of the source material used in the compilation of the property index map and the derivation of

[2] The root mean square error value represents here the discrepancy between the captured position of feature point and its true position in relation to the National coordinate reference system.

transformation parameters, an absolute RMSE error for a data point can be multiplied by a factor of three. That is, an error of +/-0.86m for topographic maps will equate to +/-2.5m for property index maps derived from secondary sources (other maps and plans).

D. Attribute Accuracy

Attribute accuracy is the accuracy of information about the characteristics of features represented in the property index map. These characteristics include the feature representation (classification), text information (feature descriptions), and change history attributes at feature level.

E. Logical Consistency

Consistency is a measure of the degree to which the data logic complies with data structures defined in specifications. Quality control for logical consistency of property index maps (parcels and building objects) comprised of points, lines and polygons, is mainly comprised of checks for geometric and topological consistency, validity of record structure and validity of values. For example, connectivity is checked to make sure polygons are ‘closed’ properly. Checks are typically completed automatically by software, ensuring that values of feature classifications and attributes, geometry and topology, database schema and file formats are complete and valid in accordance with the data specification.

SECTION 3 – QUALITY MANAGEMENT

Quality Management (QM) is a process that leads to the delivery of quality-compliant products and services. Ultimately the quality of supplied data is, in most part, only as good as the quality of data capture processes, or the quality of source data in the case of data compilation. This section provides an overview of the dataset supply process (capture, compilation etc), highlighting key data quality considerations, which collectively are described as 'quality management'.

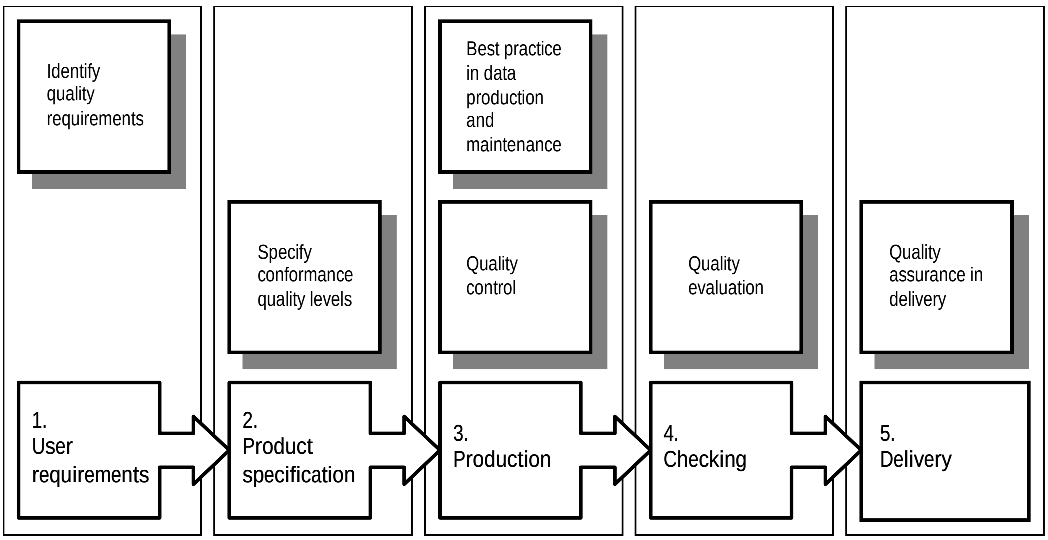


Figure 1. The quality management process

A. Identifying Quality (User) Requirements

It is important to identify quality requirements by first identifying use and user requirements for the property index map. This prerequisite includes defining the supporting attribute data for property index maps, and defining the product specifications and hence quality levels. However, given a potentially large number of possible users, it can be a difficult task to fully understand, identify and describe all uses in sufficient detail to be able to define a complete and comprehensive specification. Moreover, at this early stage in the development and use of property index maps, it is important to focus on priority needs, and thus sufficient to categorize uses and users into two general groups – primary and secondary.

	Use	***Users***
Primary	Property identification (location and relative position)	Registry office and customers for registration services
Secondary	Integration and provision of land information	Government agencies concerned with planning, land tax, service delivery etc.

Table 2 – Primary and secondary use and users of property index map

ESA are primarily a provider of property index maps and data; they are not a primary user. The primary users are the customers and the registration system itself, and ESA are providing a service to them by maintaining the property index map. The specific purpose and narrow use of the property index map means that the specification and quality measures will be different from general-purpose topographic maps or cadastral maps. This distinction must be appreciated.

A1 Primary Use

The property index map provides a spatial representation of part of the real world; the part specifically related to separately held (owned or occupied) properties. This representation is achieved by geometrical arrangements of points, lines and polygons. The object of registration (land parcel or building) is defined by a series of interconnecting lines that form a closed polygon.

The most important quality of this spatial representation is the definition of the object itself – the fact that it exists, and in a particular absolute location (registration area, district, section and block) and a relative location (next to adjacent properties). These attributes of existence and location are the primary uses of the index map: “does the property exists, and where is it?” The quality of spatial accuracy (the polygon area and lengths of polygon sides) is of less importance.

In addition to the qualities of existence and location, a property index map must be complete; in other words, all registered properties must be indicated. Completeness refers to both coverage and currency, and any changes or mutations to registered properties must be reflected immediately. The temporal quality of the index map must be classed as ‘continually kept up to date’.

A2 Secondary Use

With respect to secondary uses, the property index map provides a valuable spatial data framework for broader land-related information. The integration and exchange of spatial data is achieved through both the map and attributes of the map, specifically the Property Identification number (PID). In this respect, the quality of spatial accuracy (size and geometry) becomes relatively more important.

A balance is therefore required between the more abstract primary uses of “does the property exists, and where is it?” and the more concrete requirements of “what shape and size is the property?” As a general rule, land information held by other agencies is spatially referenced according to the needs and resources of the agency concerned. For example, utility and commercial distribution companies often use street addressing for location-based services. Agencies responsible for land and environmental management typically use base mapping – topographic maps if available; high resolution remotely sensed imagery when suitable maps are unavailable.

Satisfying the general needs of most secondary users, and balancing these against the specific primary need of property registration (registry organization and services delivered to registry customers) requires, therefore, that the spatial accuracy of property index maps should be approximately consistent with general-purpose base mapping (the most common form of spatial representation). In an urban context this would be between maps scales of 1:1,000 and 1:2,500.

A3 Other Uses

It is common practice throughout the world for agencies to request data based on their perceived needs rather than real needs. This is more so where the data is provided at less than the cost of production. In instances where users are required to pay the full economic cost for data, there may be a reassessment of the cost/benefit equation, and less comprehensive and less accurate data is often then found to be adequate for satisfying business needs. A good example is utility companies who can demand accurate cadastral maps for managing distribution networks; a need satisfied in many other countries by base topographic mapping without reference to the cadaster.

The primary function of accurate cadastral survey data, in either plan, map or numerical form, is to provide evidence of the location of boundary marks and lines, and to aid the re-establishment of such points and lines if they are moved or lost or for the resolution of a dispute. Such accuracy is not necessarily a requirement for registering the property (unless otherwise explicitly specified in legislation) as demonstrated by the land law and registration systems that operate in many common law jurisdictions with non-numerical property descriptions, such as India. Accurate measurement is, however, highly desirable in places lacking substantial physical linear boundary demarcations (such as walls, hedgerows, ditches, fences) and where marks (such as stones or iron rods) are easily lost or moved. Although the accurate survey to provide evidence for the location and relocation of boundary marks is also used for describing the parcel (e.g. Australia) this need not be the case, and different methodologies can be used for describing properties and for defining boundaries. Many jurisdictions utilize mapping techniques such as photogrammetry[3] to compile less accurate property index maps especially when registering for the first time properties that already exist; they may also employ more accurate cadastral surveys for defining, demarcating and documenting new parcels created by subdivision. Two survey methodologies and standards can coexist: one for property identification and first registration of existing properties, and another for subdivision and providing better evidence of ill-defined boundary lines.

User requirements for Property Index Maps

1. Identification and location of registered and registrable properties
2. Allocation and administration of Property Identification Numbers
3. Spatial framework for parcel-based land information systems

B. Product Specification

A product specification document defines an abstraction of the real world based on user needs (defined above). Important components in the specification should include: data content and structure; feature classifications and attribute definition; consistency rules; reference system information; data capture information; data quality standards; and, metadata.

A simple definition of quality is that a dataset is of the required quality if it satisfies the requirements expressed in a particular specification based on primary user requirements. This specification defines, and is not defined by, data capture specification. Quality control procedures are, therefore, required in the capture or compilation process to verify that the data conforms to the product specification.

[3] A technique for producing maps for a series of vertical aerial photographs

C. Production (quality control)

Quality control is the checks that are applied to all parts of the data production process (capture, compilation, storage, editing, manipulation). These are designed to prevent errors from occurring, or identifying errors early in the process, that may not be detected in the subsequent evaluation process, particularly if the checks are based on data sampling.

Various tools can be utilized during data production, such as control variables, which enforce data consistency and content. Random checking applied throughout the process is another often used quality control process. This can be a simple as having a more experienced person supervising and checking the work of less experienced people.

D. Checking (quality evaluation)

The quality evaluation process will comprise five different process steps: (1) identifying the data quality elements and scope; (2) identify the data quality measures; (3) select and apply an appropriate evaluation method; (4) determine the data quality result; and, (5) determine and report conformance.

The data quality elements of the property index map that will be evaluated include: the property index map spatial characteristics and associated attribute tables. Elements related to the production process and delivery of other data sets needed for improved deeds registration, such as scanning of deeds and other documents, are not a part of this quality assurance specification.

The data quality elements of the property index map that will be tested or measured include:

Characteristic	***Measure(s)***	***Indicator***
Data format and structure	Conformance with specification	Yes / No
Completeness	Spatial coverage; duplicate data; missing data	Percentage of universe
Logical consistency (data integrity)	Entity relationships; links; value range; gaps; overlaps; undershoots; overshoots	Percentage of universe
Positional accuracy	Absolute domain; relative domain	Percentage of universe
Attribute accuracy	Character length, format and values	Percentage of universe

How each measure is tested can be classified into two methods: direct and indirect. Direct evaluation can be further divided into full inspection and sampling.

	Full	***Sample***
Direct	Data format and structure; completeness (some elements); logical consistency	Positional accuracy; attribute accuracy
Indirect	Completeness (some elements)	Attribute accuracy

Direct quality evaluation provides objective quality results, which better enables the data producer/supplier (contractor) and the user/buyer (client) to confirm that conformance quality levels are met. Indirect evaluation is valuable but being more subjective is less reliable for setting conformance levels.

The means of conducting direct evaluation can be automatic (in software), which is most appropriate for full inspection, and non-automatic approaches appropriate for inspection by sampling. For example, completeness is a measure suited for direct automatic evaluation of positive entries in a database table. Indirect non-automatic evaluation using, for example, existing maps or imagery in a visual comparison with the property index map can test completeness with respect to spatial coverage.

Sampling of geographic datasets, including property index maps, can be problematic. Ideally, sampling should be random, but a stratified non-random sampling procedure may be used; for example, by using a grid to identify sample points. However, a more random sample can be used to evaluate attribute accuracy. A sampling ratio or size is suggested in section 4.

The outcome of the quality evaluation will be either acceptance or rejection of the dataset. If the dataset is rejected, then after the data are corrected, a new evaluation (checks) will be conducted before final and full acceptance that everything is in conformance with the specification.

E. Delivery (quality assurance and metadata)

The value of geographic data, including property index maps, depends less on its cost and more on its fitness for a particular purpose. A critical measure of fitness is data quality, but more specially the assurance that the data set is fit for a particular customer's use. Elements of the quality control process (assurance) have been described above. Demonstrating the conformance to data specifications, and documenting the lineage, control checks, and evaluation results when delivering data – in other words, assuring the quality of data and its fitness for purpose – is the role of metadata (data about data).

The delivery of property index maps for a particular geographic area will be accompanied by a metadata file and quality evaluation report that documents, amongst other things, the quality evaluation measures applied before the data was delivered. Immediately prior to, or upon delivery, but before final acceptance, the recipient will also evaluate data quality. This quality evaluation by the recipient will substantially repeat the evaluation conducted by the data producer. In essence, the quality evaluation is a full inspection of the metadata and quality report statements.

SECTION 4 – QUALITY MANAGEMENT (QA/QC) OF PROPERTY INDEX MAPS

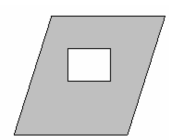
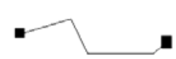
A. User Requirements

The user requirements for property index maps to support the improved deeds registration system were discussed in section 3 above. Briefly, the intended uses of the property index map may include: property identification; property identification number allocation; regional and large-scale urban planning; asset management; property transaction and valuation analysis; graphical property index. The property index map is not intended for use in legal descriptions of properties or parcel boundaries, or where sub-meter accuracy for property or boundary location is a requirement.

Fitness for purpose of the property index maps is determined by the content and accuracy of the compilation materials, which varies according to map type and scale. Parts are compiled from 1:1,000 scale as-built surveys (+/-0.25m accuracy) and other parts equivalent to 1:2,500 scale (+/-2.5m accuracy) map and imagery data. Use of the data is therefore suited to applications at the same or smaller scale to that of the source (1:2,500 scale and +/-2.5m accuracy), and any enlargement or extrapolation of the data will result in proportionally increased errors.

B. Product Specifications

The property index map is a feature based geographic data model that is composed of two key spatial features: parcels and buildings. Supplementary data includes property identification numbers, street names, and block and section polygons and numbers. The spatial data structure schema of the property index map can be summarized as:

Polygons:	Used to represent area features of: section, block, parcel, building and servitudes (rights of way, etc.)	
Lines:	Used to construct area feature polygons	
Points:	Used to locate polygon attribute text, annotation text, and to define ends (nodes) of lines	
Attributes:	Text and tags (links to external attribute data and text descriptions)	

Product specifications, format and content of the PIM are documented fully elsewhere. Briefly, these are:

B1 Spatial Data Format

The data shall be: vector point, line and area format. Points are represented by an [X and Y] coordinate pair, being an easting and northing referenced to the ESA mapping grid, plus a text string identifier. (Note: a [Z] elevation value is not required). A vector(s) between two or more points defines a line. A line has a tag for an attribute record that indicates its source, date of capture and accuracy. Areas are represented by a closed series of lines, and the property identification number provides the tag (link) to attribute data for each area (enclosed polygon).

Spatial data tables for points, lines, and areas, will be in a format documented in technical specifications (see separate technical specifications and guidelines for property index mapping).

B2 Spatial Data Integrity

Spatial data features will not overlap, overshoot, undershoot, contain gaps, or contain artifacts, as illustrated in figure 2. Features of the same type will not be coincident; i.e. they will not be duplicated (one feature exactly on top of another feature). Adjoining polygons may share the same line.

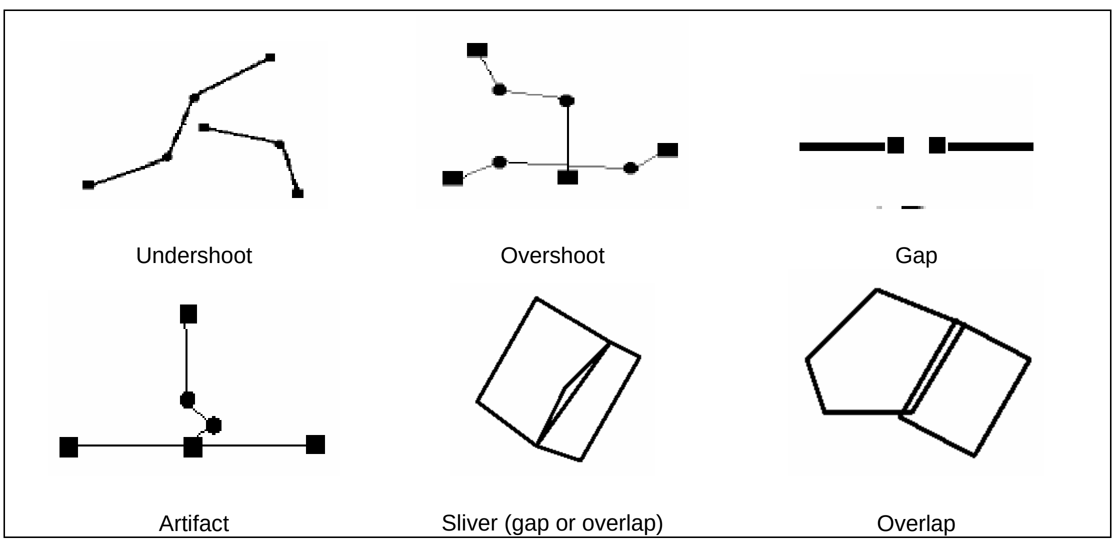


Figure 2. Examples of lack of data integrity

B3 Numerical characteristics

All coordinates shall be held or stored in double precision format, in a coordinate reference system (datum and projection) consistent with 1:1,000 scale base topographic mapping for the same area. Coordinate reference system parameters are provided by ESA.

B4 Positional accuracy

With the property index map being compiled from a variety of spatial data sources, its absolute accuracy must be considered as commensurate with the lineage of source materials and base mapping (1:2,500 scale and/or rectified very high resolution remotely sensed satellite imagery). This means that the position of features on the map must be within 1mm of their true, absolute, position. This equates to +/-2.5 meters permissible error, the 95% confidence level.

For relative accuracy, between adjacent features, this is also dependent on map scale and compilation material. The features on the map must be within 0.2mm (scaling resolution) of their true relative position, and this can be expressed as a ratio (depending on distance) of 1:200 where a distance of 100m must be accurate to +/-0.50 meters.

Accuracy	***Value***
Absolute accuracy standard	+/-2.5 meters (95% confidence)
Relative accuracy standard	1:200 (+/-0.5m over 100 meters)

C. Production (quality control)

The data supplier should undertake the following procedures during the production process:

1. Procedure manual including input check-list to guide operator;
2. Use of operator ‘pick-lists’ for standard inputs;
3. Software application to check for data integrity;
4. Validation of attribute tables for accuracy, correctness and completeness types according to approved reference tables

These procedures, which help to ensure conformity of the data to the technical specification, will apply to the following production steps (and quality measures):

1. Scanning of cadastral maps and plans (resolution)

2. Determination of coordinates for geo-referencing purposes (accuracy)
3. Map and plan geo-referencing and integration (transformation results)
4. Scanning of deeds, forms and documents (resolution and completeness)
5. Map, plan, deed, form and document attribute data entry (accuracy and completeness)
6. Field verification, data collection and data entry (procedures, accuracy and completeness)
7. Property identification number assignment to index map, database records and image files (accuracy and completeness)
8. Database fields, records and relations (integrity and completeness)
9. Metadata capture and records (completeness)
10. Index map vectorization and integration (consistency, accuracy and completeness)

In the vectorization process, where an operator has discovered irregularities in the source material or made an assumption in the interpretation of the source material, an error flag will be placed. A list will be made of these flags for later investigation and the action taken reported.

The buyer or client for the PIM should not actively participate in the enforcement or oversight of the above production quality control measures. This will be the responsibility of the supplier (contractor). The client's input and review of quality will be at the delivery stage (E).

D. Checking (quality evaluation)

Data quality evaluation will be a two-stage process. The first stage will be done by the data creator/compiler (supplier) before delivery (quality evaluation); the second stage by the data recipient (client/buyer) before acceptance of delivery (quality assurance).

D1 Property Index Map Spatial Characteristics

These are expected to include the following **direct** and **automated** (in software) pre-delivery checks:

1. Geometric integrity (no line duplication, overlaps, overshoots, undershoots, gaps, or artifacts, as illustrated in figure 2, above)
2. Topological integrity (polygons are closed, defined as area objects, no gaps, no overlaps, no slivers)

3. Parcels are not split by buildings, and buildings are not split by parcels (buildings must be completely within or wholly or partially coincident with parcels unless the property is a “villa” that spans two parcels)
4. All parcels are within block polygons; all blocks are within section polygons; all section polygons are within city polygons; and, all city polygons are within governorate polygons.
5. All parcels have internal PID points and text; all blocks have internal identification number; all sections have internal identification number; all cities have internal identification number; all governorates have internal identification number.
6. All lines are associated with polygons
7. All polygons are labeled

The following **direct** and **manual** pre-delivery checks for completeness and accuracy:

1. Visual spatial comparison of PIM vector dataset with source vector data and raster images, either in a software or using hardcopy check plots. If at any point the vector difference between the captured and source material exceeds 2.5 (two point five) meters, the error must be flagged (reported), investigated and corrected if appropriate. Note: large discrepancies may arise due to source material and differences between planned and actual plot layouts, which is not a quality control issue, but still needs to be reported.
2. The number of parcels on input materials matches the number of parcels in the PIM
3. All roads have annotation text
4. Positional accuracy is checked by:
    c. Ground ‘truthing’ 9 (nine) well-defined feature points per square kilometer[4], either parcel or building corners, distributed uniformly as possible, by comparing vector map data coordinates with real world coordinates. The dataset will pass the absolute positional accuracy quality check if less that 5% of the absolute position comparative results are greater than 2.5 (two point five) meters.
    d. Comparative measurements of 9 (nine) well-defined feature sides per square kilometer, both parcel boundary and building lines, on the ground and the corresponding line on the map, distributed throughout the project area, and if less than 5% of the relative position comparative tests are greater than the ratio of 1:200, which is error related to ground distance (e.g. +/-0.5m for 100 meters on the ground) the dataset will pass the relative positional accuracy test.

[4] This represents, approximately, a 0.5% sample of independent (feature) points.

Note: a small sample size is justified because (i) the absolute positional accuracy is also checked indirectly (below); and (ii) most of the data points are derived and compiled from existing maps and geo-referenced in blocks.

The following **indirect** and **manual** pre-delivery checks:

1. Visual spatial comparison in software of PIM vector dataset with very high resolution remotely sensed and rectified raster satellite images (ground resolution of equal or better than 1.0 meters). If at any point the vector difference between the captured material and the image exceeds 2.95 (two point nine five) meters, the error must be corrected.

D2 Property Index Map Attribute Tables

The tables associated with the A vector data set are expected to include the following **direct** and **automated** (in software) pre-delivery checks:

1. All record fields are populated (those that require to be populated) with correctly formatted data. A software routine should count the number of records and the report the number fields that are empty; if more than 5% of fields that must contain data do not contain data then the data set is rejected.
2. All records with primary keys (parcel PID) are related to parcels in the PIM; all child records in linked tables (unit PID) have parent records in primary table. Any orphan records will result in rejection of the dataset (i.e. compliance must be 100%).

Note: cross check between parcel polygon area calculated from the property index map, and the area given in a document, such as a mutation form, is not a requirement. However, such a check may be useful to highlight possible issues, and only large (> 10%) area differences discovered need reporting. Large differences discovered that are due to compilation, digitizing, geo-referencing errors must be fixed; large discrepancies due to plots not laid out in conformance with the development plan must be flagged and reported.

Attribute accuracy is assured by undertaking the following **direct** and **manual** pre-delivery checks:

1. A sample of 16 properties (parcels and buildings) well distributed per square kilometer throughout a project area or zone, where data is independently recollected and compared against the originally collected data. If more than 5% of the sample

records contain different data from the original records, the dataset will be rejected and the errors corrected.

2. If a dataset has been rejected because of attribute inaccuracy, and after errors have been corrected, another sample check will be undertaken, with the same number but different sample properties. If more than 5% of the sample records contain different data from the original records, the dataset will be rejected and the errors corrected.

3. The quality checks will be continued, with different sample data each time, until the dataset passes the quality evaluation measure. Note: Although a strictly random sample is not required, the sample should be well distributed and representative; sample points cannot be repeated in a later sample.

D3 Quality Reporting

A quality control report must document, as a minimum, the following: lineage, currency and positional and attribute quality; quality control measures implemented; quality evaluation checks applied; name of person who performed the internal quality control; time spent on quality control; number, type and size of errors found in the internal quality control, and the measures taken to correct the errors.

E. Delivery (quality assurance and metadata)

Immediately prior to, or just after, but before acceptance of delivery, the recipient (buyer) will perform another series of quality evaluation tests; these will be substantially the same as those conducted by the supplier, and listed above. However, the recipient will conduct one further evaluation: of the supplier's quality evaluation report and the dataset metadata record. This evaluation will, in effect, be a check of the metadata statements that the data passed all required quality control measures. It will also check the completeness of the metadata record.

If any of the evaluation measures specified at **Section D** above are not passed, then the data set, in its entirety, is rejected and returned to the supplier for correction. Upon re-submission of the corrected dataset(s) the evaluation will be repeated.

Specifications for format and content of the metadata file are given in annex B.

ANNEX A – PROPERTY INDEX MAP SPECIFICATIONS

A1 Property Index Map Data Content

Entity	***Typology***	***Description***
Points	Attribute text	Property identification number
		Street name
Lines	Parcel boundary	Straight line between boundary corner points
	Building line	Straight line between building corners
	Servitude line	Straight line between servitude corners
Polygons	Parcel	Closed loop of linked lines
	Building	Closed loop of linked lines within or wholly or partly coincident with parcel polygon
	Servitude	Closed loop of linked lines within or wholly or partly coincident with parcel and/or building polygon

ANNEX B – QUALITY REPORTING AND METADATA REQUIREMENTS

B1 Quality Evaluation Report

A report that documents all the quality controls put in place during property index map production and the results of the quality evaluation prior to delivery. All checks must be reported, including checks where the data failed to pass.

B2 Metadata

Metadata is data about data. In other words, it is a structured summary of information that describes the data. Metadata includes, but is not restricted to, characteristics such as the content, quality, currency, and availability of the data.

Core metadata elements are as follows:

Category	***Element***	***Definition***	***Occurrences***	***Field***
Dataset	Title	The name of the dataset	1	Text (50)
Custodian	Owner	Organisation responsible	1	Text (50)
Description	Abstract	Brief description	1	Text (150)
	Key words	For searching	N	Text (50)
	Geographic extent name	Common name of the area that the data set covers	N	Text (100)
	Geographic region	Common name of the region where the extent is located	N	Text (50)
	Geographic extent	Latitude and longitude (decimal degrees) of E, S, W, N extents	4	Number
	Format	Description of what format data is stored	1	Text (25)
Data currency	Begin date	First data added	1	Date
	End date	Last data added	1	Date
Data quality	Lineage	A brief history of the source and processing steps	1	Text (200)
	Positional accuracy	A brief assessment of the data set as a whole and the checks carried out	1	Text (200)
	Attribute accuracy	A brief assessment of the reliability and completeness of associated data and the checks carried out	1	Text (200)

Category	***Element***	***Definition***	***Occurrences***	***Field***
	Logical consistency	A brief assessment of the geometric and topology data rules	1	Text (200)
	Completeness	A brief assessment of the extent and coverage of data	1	Text (200)
Contact	Information	Organisation, person, address, telephone, etc	1	Text (200)