CapacityProject
compartiendo conocimientos
Julio de 2008
resumen técnico 10
the Capacity Project

Consideraciones sobre la calidad de los datos en el fortalecimiento de los sistemas de información de recursos humanos (SIRH)

Samwel Wakibi, IntraHealth International

Introducción

Hoy en día, las organizaciones se valen de datos para incrementar su eficacia y mejorar la planificación del futuro (Rochnik y Dijcks, 2006). Cuando la calidad de los datos es deficiente se pierde tiempo, dinero y la confianza de los clientes y esto puede poner en apuros a las organizaciones. Se estima que en el ámbito industrial el índice típico de error en la calidad de los datos, que es del 1% al 5%, puede provocar una pérdida del 10% de los ingresos (Redman, 1996). Conforme los datos se recopilan, se analizan y se traducen en informes que son de utilidad para la planificación y la toma de decisiones, pueden ocurrir problemas con la calidad de los datos a medida que la información pasa a través de límites organizacionales y sistemas.

Los temas relacionados con la calidad de los datos han sido de fundamental importancia en la experiencia de ejecución de programas de Capacity Project, un proyecto de alcance mundial financiado por USAID que ayuda a los países en desarrollo a fortalecer los recursos humanos para la salud (RHS) para enfrentar mejor los desafíos que conlleva la implementación y mantenimiento de programas de salud de calidad superior. El Proyecto ha adquirido experiencia en temas relacionados con la calidad de los datos, particularmente en el campo del fortalecimiento de los sistemas de información de recursos humanos (SIRH), para brindar apoyo en la planificación y gestión de los recursos humanos para la salud.

Este resumen técnico presenta algunos conceptos de calidad de datos e incluye ejemplos que ilustran la importancia del manejo de datos, específicamente en el campo de los recursos humanos para la salud, según la experiencia de Capacity Project en el fortalecimiento de los sistemas de información de recursos humanos en los países en desarrollo.

Conceptos básicos de calidad de los datos

Una definición del término *calidad de los datos* que es ampliamente reconocida en los campos de economía, negocios y medicina es la "aptitud para el uso" actual y futuro; en otras palabras, el grado en que los datos satisfacen las necesidades y expectativas de los usuarios (Chapman, 2005; Carson, 2000). La calidad de los datos describe la condición de los datos, la serie de procesos requeridos para alcanzar esta condición y la precisión de los datos. Para que los datos sean aptos para el uso, deben estar libres de duplicaciones, faltas de ortografía, omisiones y variaciones innecesarias y deberán ajustarse a una estructura definida (Chapman, 2005; Carson, 2000; Brown, Stouffer y Hardee, 2007).

La calidad de los datos se refiere a los siguientes aspectos:

- Exactitud
- Oportunidad
- Congruencia
- Precisión
- Integridad
- Relevancia

- La *exactitud* se refiere a la cercanía a los valores medidos, observaciones o estimaciones del valor real o verdadero, sin manipulación ni sesgos políticos o personales; en otras palabras, la exactitud es una medida del grado en que los datos reflejan la realidad. Las preguntas orientadoras para evaluar la exactitud de los datos se relacionan con la confiabilidad de las fuentes de datos y el proceso de recopilación de datos.
- La *oportunidad* se refiere a la disponibilidad de datos cuando se requieren. Algunos factores relacionados con esto son: conocer el período cuando los datos se recopilaron, cuándo se realizó la actualización más reciente de datos, por cuánto tiempo permanecerán vigentes los datos y si son procesados con el fin de brindar información oportuna para realizar las actividades cotidianas o tomar decisiones.
- La *congruencia* describe la ausencia de contradicciones evidentes y es una medida de validez y confiabilidad internas. Las preguntas orientadoras para evaluar la congruencia de los datos se relacionan con la medida en que se aplican las mismas definiciones, códigos y formatos para el mismo tipo de datos provenientes de diferentes fuentes.
- La *precisión* se refiere al uso consistente de un indicador dado al generar los mismos resultados. Por ejemplo, por medio del uso de un formulario de recopilación de datos con un alto grado de precisión se generarán las mismas respuestas si el formulario se utiliza repetidas veces para un tema dado. La precisión y la exactitud difieren en cuanto que una medida puede ser precisa sin ser exacta; por ejemplo, una medida puede generar los mismos resultados incorrectos repetidas veces.
- La *integridad* se refiere a la falta de errores de omisión, como la omisión de registros en un conjunto de datos, o una variable sin datos. Para evaluar lo completo, la pregunta es si se han incluido todos los datos elegibles.
- La *relevancia* se refiere a la disponibilidad de los detalles o datos requeridos. Ayuda a responder las preguntas relacionadas con la estructura de la base de datos o el diseño del formulario de recopilación de datos. Por ejemplo: ¿Se recopilan los datos apropiados y se almacenan para utilizarse en las iniciativas actuales y futuras? ¿Los registros de datos recopilados son suficientemente detallados como para responder preguntas relacionadas con las actividades previstas? La regla por defecto es que si se desconoce cómo se aplicará la información, entonces será necesario proporcionar datos con el mayor nivel de detalle que sea posible. Esto brinda cierta confiabilidad adicional a los datos, ya que es más fácil desechar las partes redundantes o innecesarias que buscar la información faltante (New Zealand E-government Unit, 2004).

USAID
DEL PUEBLO DE LOS ESTADOS UNIDOS DE AMÉRICA

Experiencia de campo

El fortalecimiento de los sistemas de información de recursos humanos abarca cinco campos de acción: desarrollo participativo del liderazgo de los grupos interesados; fortalecimiento de la infraestructura de tecnología de información y comunicación; desarrollo y utilización de software y capacitación relacionada; mejoramiento continuo de la iniciativa; y desarrollo de destrezas en los campos relacionados, como la toma de decisiones basadas en datos.

Capacity Project ha facilitado actividades de fortalecimiento de los sistemas de información de recursos humanos en Uganda, Ruanda, Tanzania, Namibia, Sudán del Sur, Suazilandia, Sudáfrica y Lesoto y además, inició la implementación de actividades similares en otros países en desarrollo. En este proceso, Capacity Project ha abordado temas relacionados con la calidad de los datos que abarcan desde la exactitud, integridad y congruencia hasta la oportunidad de los datos y la información generada.

Las fallas que se identificaron en la calidad de los datos se han atribuido al uso de tecnología y procedimientos de operación inapropiados, la falta de destrezas en el manejo de datos, falta de conciencia respecto a la calidad de los datos y falta de comprensión de la importancia de los datos en los procesos de toma de decisiones. Adicionalmente, han surgido otros temas relacionados con la calidad de los datos a raíz de influencias ambientales, tales como acontecimientos vitales que no fueron documentados en el momento de ocurrir, documentos perdidos en guerras, conflictos civiles e incendios y documentos dañados a causa de un almacenamiento inadecuado y las inclemencias del tiempo.

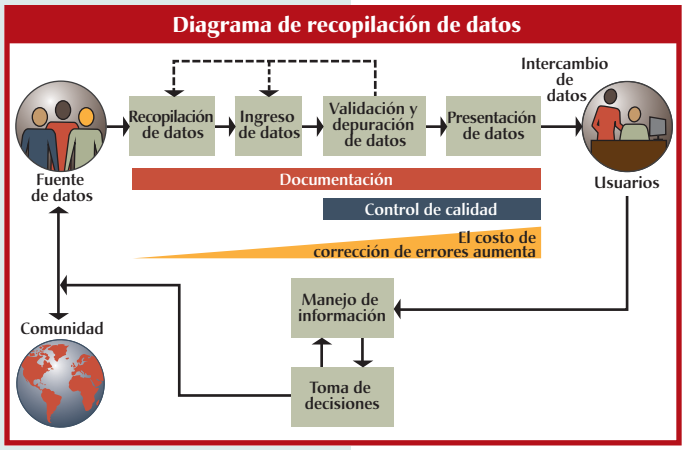

Figura 1: Diagrama de recopilación de datos

(Adaptado de Chapman, 2005)

A medida que los datos van pasando por la cadena de manejo de información, los problemas relacionados con su calidad se magnifican y el costo que implica la corrección de errores se incrementa (ver la Figura 1). Al final, si los gerentes no cuentan con información de alta calidad acaban por tomar decisiones orientadas por señales falsas, y esto conduce a ineficacias como una distribución desigual de los recursos humanos para la salud, los "trabajadores fantasma" (empleados anteriores que permanecen en planilla aunque ya no estén empleados) y atrasos en el pago de salarios, como se ha observado en algunos sistemas de recursos humanos en los países donde trabaja Capacity Project.

Estudios de caso

Los siguientes ejemplos provienen de la experiencia de Capacity Project en el trabajo de campo.

Registros duplicados y datos faltantes: el caso de Uganda

El registro del Uganda Nurses and Midwives Council – UNMC (Consejo de enfermeras y parteras de Uganda) mantiene datos sobre las enfermeras, parteras y estudiantes de enfermería y partería que trabajan en Uganda. UNMC lleva un registro de datos relacionados con la participación en programas de capacitación previa al servicio (denominado *elaboración de fichas*), los resultados de los exámenes de las estudiantes de enfermería y partería, así como el registro, otorgamiento de licencias y renovación de licencias para enfermeras y parteras profesionales. Gracias al apoyo de Capacity Project, el registro se mejoró y pasó de ser un sistema de archivo de documentos impresos a un sistema computarizado.

Al desarrollar la base de datos se encontró que algunos identificadores únicos (como los índices) se repetían múltiples veces. Anteriormente, los índices se generaban manualmente y eran asignados por diferentes personas en un proceso que se coordinaba en forma deficiente. Como medida de corrección, se modificó el sistema computarizado para permitir la creación de un número de identificación personal interno y único para identificar a las estudiantes y las enfermeras calificadas formadas en Uganda. Otros desafíos relacionados con la calidad de los datos incluían los casos de registros faltantes, registros incompletos y archivos perdidos o faltantes.

Actualmente, a través del uso de la nueva base de datos UNMC tiene la capacidad de reportar el número de enfermeras capacitadas en el país y su situación de retención, además del número total de estudiantes de enfermería por distrito de domicilio.

Incongruencias en los datos: el caso de Lesoto

Capacity Project ayudó al Ministerio de Salud y Bienestar Social (MOHSS, por sus siglas en inglés) a desarrollar una base de datos para gestionar los recursos humanos para la salud y actualmente está en proceso de implementarla. Los datos a manejar a través de la base de datos se comparten entre el Ministerio de Finanzas (salarios de los empleados), el Ministerio de Salud y Bienestar Social (registro y utilización de empleados) y el Ministerio de Servicios Públicos, que es responsable de la contratación.

Durante el proceso de desarrollo de la base de datos el equipo se dio cuenta que estos tres ministerios mantenían sistemas de codificación paralelos para las plazas establecidas. Además, los ministerios clasificaban las unidades del Ministerio de Salud y Bienestar Social en forma diferente: lo que se clasificaba como "programa" o "departamento" en un ministerio podía denominarse "subprograma" en otro ministerio.

A pesar de haber resuelto el problema de los tres sistemas de codificación paralelos, esto ya había provocado redundancias en la base de datos. Para eliminar las redundancias la computadora combinó los tres conjuntos de códigos internamente y asignó un número único a cada ítem.

Oportunidad de la información y los datos: el caso de Suazilandia

Para Suazilandia, Capacity Project desarrolló e implementó una solución "intermedia" (una base de datos menos elaborada para satisfacer las necesidades de corto plazo). Ahora el Ministerio de Salud y Bienestar Social tiene la capacidad de actualizar los registros de los recursos humanos para la salud con regularidad.

El proceso de desarrollo de la base de datos reveló que con el sistema manual ocurrían demoras, lo que significaba que los jubilados y los empleados que dejaban sus plazas permanecían en planilla por mucho tiempo, los empleados nuevos tenían que esperar mucho antes de percibir su salario, los contratos de los empleados no se monitoreaban y los pagos no se realizaban con prontitud. Como consecuencia de ello, algunos empleados nuevos y contratados iniciaron juicios contra el gobierno; y para empeorar las cosas,

debido la incapacidad del sistema de reconocer las plazas vacantes y responder rápidamente estas plazas permanecían sin estar disponibles, por lo que no era posible contratar a nuevos empleados y por consiguiente, se redujo significativamente el número de empleados disponibles para proveer servicios de salud.

La nueva base de datos se ha utilizado para validar los datos de planilla y manejar la rotación y escasez de personal en forma eficaz. Adicionalmente, hay más información disponible para facilitar la toma de decisiones relacionadas con la contratación y traslado de personal.

Buenas prácticas en el fortalecimiento de los sistemas de información de recursos humanos

Las buenas prácticas en el fortalecimiento de los sistemas de información se evalúan en términos de la calidad agregada, el ahorro de tiempo y la reducción de costos de manejo de datos (United States Department of Defense, 1999; NHS Connecting for Health, 2007). Adicionalmente, Capacity Project aborda temas de calidad de los datos y al mismo tiempo, brinda asistencia en el fortalecimiento de sistemas de información de recursos humanos. El enfoque de Capacity Project incorpora un protocolo de aseguramiento de calidad que vela por la exactitud, lo oportuno, la congruencia, precisión, lo completo y la relevancia de los datos con relación a las necesidades especiales de los clientes.

Estas buenas prácticas se basan en la experiencia de campo adquirida en el trabajo con datos de diversos sistemas de información de recursos humanos en países del África Oriental, Central y del Sur. Entre los ejemplos de los principales aspectos abordados y enfoques adoptados se pueden incluir los siguientes:

Desarrollo de sistemas

Las bases de datos que se implementaron son longitudinales; registran los datos personales, profesionales y laborales de los trabajadores de salud y los cambios relacionados. Abarcan el período desde la capacitación hasta el retiro para asegurar que los datos que se requieren para responder preguntas relacionadas con políticas y gestión de recursos humanos para la salud estén completos y sean exactos.

Capacity Project emplea tecnología eficiente de bajo costo para apoyar el desarrollo de los sistemas de información de recursos humanos. Se prefiere el software "Open Source" (software de código abierto), ya que está disponible en forma gratuita y tiene la capacidad de manejar el volumen y procesamiento de datos previstos en forma eficaz. El software "Open Source" se distribuye en forma gratuita bajo una licencia que permite que cualquier persona que así lo desee pueda estudiar, copiar, utilizar y modificar el código de origen y redistribuirlo en forma modificada o sin modificar, sin restricciones. Se puede obtener acceso por Internet a las bases de datos de código abierto que se desarrollan para manejar los datos de recursos humanos; esto permite depurar e ingresar datos fuera de las oficinas, lo que reduce al mínimo las demoras en la actualización de datos.

Los sistemas de información se fortalecen por etapas planificadas para obtener resultados rápidos sin sobrecargar la infraestructura de apoyo. Los sistemas de codificación de las diferentes organizaciones se armonizan para asegurar que las bases de datos o módulos relacionados sean consistentes unos con otros, y se asigna un identificador único a cada registro. Además, este sistema de codificación sincronizada asegura que las búsquedas en todas las bases de datos generen los mismos resultados.

Para reducir al mínimo los errores en el ingreso de datos se utilizan dos métodos de ingreso de datos: ingreso doble y pruebas al azar. Los sistemas de ingreso doble constituyen una función importante que se ha integrado en el desarrollo de bases de datos.

- El *ingreso de datos doble* se refiere a un proceso que estipula que los datos deben ser ingresados en dos momentos distintos, idealmente por dos empleados diferentes; y la segunda persona que ingrese los mismos datos debe revisar si hay discrepancias entre ambos conjuntos de datos ingresados y seleccionar el ingreso exacto con base en los datos originales. Éste es el método más exacto de los métodos de aseguramiento de calidad de los datos pero implica una duplicación de costos y tiempo, en comparación con el proceso de ingresar datos una sola vez.
- Las *pruebas al azar* se refieren al proceso de generar una lista al azar de los datos ingresados que un segundo empleado o un supervisor debe comparar con los datos originales. Este método necesariamente es menos exacto que el anterior, pero se evita la duplicación de costos y se reduce el tiempo requerido para ingresar los datos.

Ambos métodos permiten llevar un registro de los errores en el ingreso de datos; el análisis de este registro se puede utilizar para mejorar la capacitación del personal responsable del ingreso de datos, realizar cambios en los procedimientos de operación o modificar el software o la interfaz del usuario.

Seguridad de datos e infraestructura

Para asegurar la seguridad de los datos se utilizan contraseñas y acceso restringido, según las funciones de los empleados, para limitar el acceso únicamente a los usuarios autorizados. Por ejemplo, los empleados responsables del ingreso de datos tienen acceso a un conjunto de datos muy limitado (lo que se requiere para mantener la calidad) mientras que los analistas o administradores tienen un acceso mucho mayor. Además, el acceso al espacio físico donde los servidores albergan las bases de datos es controlado, y se instituye un sistema de respaldo (backup).

Manejo

En los casos cuando se requiere, se realizan cambios en los formularios de recopilación de datos para asegurar que los datos sean precisos y que estén completos. Esta tarea se lleva a cabo en colaboración con los equipos locales para cumplir con los requerimientos de reporte de datos a nivel local y reducir al mínimo la carga de recopilación de datos (que constituye una importante fuente indirecta de errores).

Las personas encargadas de la recopilación de datos y el equipo de manejo de datos trabajan en estrecha colaboración para reforzar la comprensión mutua del trabajo de los demás. Mediante este proceso de retroalimentación se mejora la calidad de los datos en el momento de recopilarlos, con lo que se reduce al mínimo la necesidad de realizar correcciones más adelante en la cadena de manejo de información.

La recopilación de datos se mantiene actualizada, se asignan códigos únicos, los datos se procesan con regularidad y se colocan en una red para ponerlos a disposición de todos aquellos que los necesiten; sin embargo, sólo las personas autorizadas para ello tienen acceso para garantizar la confidencialidad y seguridad de los datos. Esta práctica de disponibilidad segura permite asegurar que:

- La base de datos se mantenga actualizada, ya que el ingreso de datos es descentralizado

Visite el Centro Global de Recursos Humanos para la Salud (RHS) para encontrar, compartir y aportar conocimientos y herramientas de recursos humanos para la salud. Para las personas que trabajan a nivel nacional o mundial, el Centro Global de Recursos Humanos para la Salud proporciona información para:

- Mejorar la planificación estratégica y la toma de decisiones
- Fortalecer la elaboración de informes y presentaciones
- Apoyar el cabildeo por los recursos humanos para la salud (RHS)
- Mejorar el desarrollo profesional
- Ahorrar tiempo

- Se elimine la práctica ineficaz de crear bases de datos disipadas y carentes de estructura
- Los datos se compartan y se utilicen dentro de las organizaciones y entre diversas organizaciones para la toma de decisiones y planificación.

Depuración de datos

Los datos provenientes de otras fuentes, como encuestas y otras bases de datos, se utilizan para detectar datos corruptos o inexactos. Los datos se corrigen o se eliminan del registro o la base de datos para que sean consistentes con otros conjuntos de datos similares en el sistema. Los procesos de depuración de datos y estrategias de prevención de errores se implementan en forma simultánea con la recopilación, ingreso y procesamiento de datos para reforzar el aseguramiento de la calidad de los datos.

Ingreso de datos

Para reducir al mínimo los errores al ingresar datos en las bases de datos se utilizan los documentos originales como fuentes principales. La necesidad de mecanografiar se reduce al mínimo a través del uso de menús emergentes, con lo que además se acelera el ingreso de datos y se mejora la consistencia de los datos. De esta manera se descentraliza el ingreso de datos y se lleva a cabo más cerca del punto donde se generan los datos; esta tarea se convierte en una parte del trabajo rutinario de los empleados que recopilan los datos.

Capacitación y adquisición de destrezas

Se imparten cursos de capacitación para los empleados y el desempeño de éstos se monitorea estrechamente. Los cursos se relacionan con el trabajo y están dirigidos principalmente a las personas encargadas de la recopilación de datos, los operadores que ingresan los datos y los gerentes que guían la elaboración y uso de los informes. Además, los empleados participan activamente en el desarrollo e implementación de las bases de datos.

Conclusión

La mejor definición de la calidad de los datos es su "aptitud para el uso" en la toma de decisiones y formulación de políticas. La exactitud, oportunidad, consistencia, precisión, integridad y relevancia son medidas de la calidad de los datos.

Capacity Project ha investigado sobre las mejores prácticas y las ha adoptado para asegurar la calidad de los datos en su labor de fortalecimiento de los sistemas de información de recursos humanos. Además, Capacity Project ofrece cursos de capacitación para que los grupos interesados adquieran destrezas de desarrollo de sistemas y manejo de datos, con el fin de asegurar la calidad, sostenibilidad y mejoramiento continuo de los sistemas que se implementan. Adicionalmente, los gerentes de salud reciben capacitación en el uso de datos para tomar decisiones informadas. Se espera que al tener acceso a datos de calidad superior, sistemas eficaces y oportunos y capacitación integral en el uso de datos, pronto los países tendrán la capacidad de utilizar la evidencia regular y formular políticas apropiadas para abordar los desafíos que enfrentan en el sector de la salud.

Referencias

Carson, C. What is data quality? a distillation of experience. Washington, DC: Departamento de Estadística, Fondo Monetario Internacional, 2000. Consultado el 15 de junio de 2007 en: *http://www.thecre.com/pdf/imf pdf*

Chapman, A. Principles of data quality, version 1.0. Copenhague, Dinamarca: Global Biodiversity Information Facility, 2005. Consultado el 7 de julio de 2008 en: *http://circa.gbif net/irc/DownLoad/kjeYAKJSmRGFqwAaUY4x8KZ1jH4pYxtv/F37w1fUI4R0AgTiySEZttf0yRVSbNGn/Data%20Quality pdf*

Eckerson, W. Data quality and the bottom line: achieving business success through a commitment to high quality data. TDWI Report Series. Chatsworth, CA: The Data Warehousing Institute, 101 Communications, 2002. Consultado el 7 de julio de 2008 en: *http://download.101com.com/pub/tdwi/Files/DQReport pdf*

New Zealand E-government Unit. New Zealand E-government Interoperability Framework (NZ e-GIF), Version 2.1. Wellington, Nueva Zelanda: New Zealand E-government Unit, 2004. Consultado el 6 de agosto de 2007 en: *http://www.e.govt nz/standards/e-gif/e-gif-v-2-1/e-gif-v2-1 pdf*

NHS Connecting for Health, Department of Health (UK). Best practice guidelines. Leeds, Inglaterra: NHS Connecting for Health, 2007. Consultado el 15 de junio de 2007 en: *http://www.connectingforhealth nhs.uk/systems andservices/data/clinicalcoding/data quality/best_practice* Redman, T. Data quality for the information age. Boston, MA: Artech House Publisher, 1996.

Rochnik, N., Dijcks, J. Oracle Warehouse Builder 10gR2: transforming data into quality information. An Oracle whitepaper. Redwood Shores, CA: Oracle, 2006. Consultado el 6 de agosto de 2007 en: *http://www.oracle.com/technology/products/warehouse/pdf/ transforming%20data%20into%20quality%20information pdf*

United States Department of Defense. Best practices for data quality oversight of environmental sampling and testing activities. Informe Final. United States Department of Defense, 1999. Consultado el 6 de agosto de 2007 en: *http://www.navylabs navy mil/Archive/BP-1999 pdf*

Brown, W., Stouffer, R., Hardee, K. Data quality assurance tool for program-level indicators. Chapel Hill, NC: MEASURE/Evaluation Project, University of North Carolina at Chapel Hill, 2007. Consultado el 7 de julio de 2008 en: *http://www.pepfar.gov/documents/organization/79628 pdf*

the Capacity Project

The Capacity Project
IntraHealth International, Inc.
6340 Quadrangle Drive
Suite 200
Chapel Hill, NC 27517
Tel. (919) 313-9100
Fax (919) 313-9108
info@capacityproject.org
www.capacityproject.org

Esta publicación fue posible gracias al apoyo del pueblo de los Estados Unidos, a través de la Agencia de los Estados Unidos para el Desarrollo Internacional (USAID). El contenido de la publicación es responsabilidad de Capacity Project y no necesariamente refleja las opiniones de USAID o el Gobierno de los Estados Unidos.

El asocio de Capacity Project

IntraHealth International

innovating to save lives
Jhpiego
an affiliate of Johns Hopkins University

PATH
A catalyst for global health

IMA World Health
Advancing health & healing the world over

LATH
Liverpool Associates in Tropical Health

TRG

MSH Management Sciences for Health
a nonprofit organization strengthening health programs worldwide