

Revista Calidad en la Educación Superior
Programa de Autoevaluación Académica
Universidad Estatal a Distancia
ISSN 1659-4703
Costa Rica
revistacalidad@uned.ac.cr

CALIDAD DE DATOS Y SU EVALUACIÓN: UN CASO DE ESTUDIO

DATA QUALITY AND ITS EVALUATION: A STUDY CASE

Lorena Zúñiga Segura¹
lzuniga@uned.ac.cr

Elisa Sánchez Godínez²
elsanchez@uned.ac.cr

Centro de Investigación y Evaluación Institucional (CIEI)
Universidad Estatal a Distancia, Costa Rica

VI Edición
Volumen III, Número 2
Noviembre 2012
pp. 33 - 49

Recibido: abril, 2012
Aprobado: agosto, 2012

¹ Investigadora del Centro de Investigación y Evaluación Institucional (CIEI)

² Coordinadora de la Unidad de Información y Análisis Estadístico del Centro de Investigación y Evaluación Institucional (CIEI)

Resumen

En los diferentes procesos que se llevan a cabo en una institución o empresa se utiliza información, es por ello que desde la recolección o captura de los datos debe contemplarse la calidad de los mismos.

El presente artículo describe una metodología para la evaluación de la calidad de los datos propuesta por el autor Arkady Maydanchik y su aplicación a una base de datos determinada.

Los resultados generales que se obtuvieron permiten señalar oportunidades de mejora, que contribuirán con la calidad de los datos, y por ende, en todos los procesos que hacen uso de ellos para la toma de decisiones. El análisis fue realizado en el año 2011, se evaluó información que pertenece a la Universidad Estatal a Distancia (UNED) de Costa Rica, la cual fue registrada entre los años 1980 al 2011.

Palabras claves: evaluación de la calidad de datos, información, bases de datos.

Abstract

Organizations perform different kinds of processes that use information. For this reason, the data quality must be taken into account from the data entry activities through organizational information systems.

This article describes the data quality assessment methodology proposed by Arkady Maydanchick, which was applied in order to assess the data quality of an organizational database.

The findings show that there are opportunities for improvements, which will contribute to data quality and therefore all processes that use these data for decision making support. The analysis was conducted in 2011; information from the Universidad Estatal a Distancia (UNED) of Costa Rica was evaluated, the database assessed contains information between the years 1980 to 2011.

Keywords: data quality assessment, information, databases

Introducción

En general, el término calidad es comúnmente asociado con productos terminados (vehículos, ropa, materiales de construcción, entre otros), o bien con servicios, por ejemplo calidad de la educación, de servicios médicos, en el servicio al cliente, sólo por mencionar algunos. Sin embargo, es menos frecuente escuchar este término asociado a los datos y la información, a pesar de que es precisamente esta última uno de los elementos centrales en la toma de decisiones en toda empresa o institución.

Muchos de los procesos de una organización generan información que es almacenada en bases de datos, con ella se pueden realizar estudios para mejorar los mismos procesos y, de esta manera, avanzar hacia el logro de la calidad. En una organización de servicio, como un centro médico, por ejemplo, la información sobre el tiempo que dura un paciente, desde el momento en que llega al centro, hasta que sale con sus medicamentos, es de vital importancia, ya que le permite al tomador de decisiones conocer la eficiencia en el servicio y buscar la mejora continua mediante el monitoreo de un indicador de calidad, que sería el tiempo promedio de duración de la consulta médica, incluyendo por ejemplo, la entrega de los medicamentos.

Sin embargo, para llegar a confiar en un indicador de calidad es menester confiar en la información que sirvió como base para generarlo, de manera que se debe tener especial cuidado con los datos que se introducen o registran. Los datos están en constante cambio, son los procesos propios de la organización los que determinan la periodicidad de esos cambios. Así, por ejemplo, la frecuencia de los nacimientos en un país marca un ritmo importante en la periodicidad de las actualizaciones que se deben realizar a la base de datos que almacenan estos hechos, de igual manera ocurre con los registros de las defunciones. Esta periodicidad es diferente en las instituciones de educación, en donde las matrículas de sus estudiantes se realiza una vez al año y, en algunos casos, como en las universidades, dos o más veces.

En este último caso, la información relacionada con la entrega de la docencia, sufre cambios en el año, dependiendo del proceso de matrícula de los estudiantes, graduaciones, registros de notas, actualizaciones de datos personales del estudiantado, entre otros. Históricamente, las actualizaciones han sido realizadas utilizando los servicios de digitadores, los cuales toman los datos de un formulario y la introducen en la base de datos mediante sistemas informáticos, o bien, el mismo estudiante, a través del portal web de la universidad, realiza la matrícula de los cursos a los cuales tiene posibilidad, según el plan de estudios de la carrera de su interés.

Por el hecho de ser personas quienes introducen los datos, se corre siempre el riesgo de que contengan algún error y queden así almacenados, por tal motivo, este aspecto debe planificarse para minimizar el registro de datos incorrectos. Una forma de hacerlo es capacitando a los digitadores, cuidando la claridad en las instrucciones dadas y verificando que la información que se introduce sea la correcta. Si no se tiene especial cuidado en estos aspectos, la base de datos se puede estar llenando con información sin utilidad, que no puede ser usada por las altas autoridades o gerentes de instituciones y empresas, puesto que “datos con errores conducen a decisiones mal informadas” (Whiting, 2006).

Si el sistema al cual se accede no tiene buenos controles de entrada de datos, tales como la detección de valores nulos o no permitidos en variables, consideradas como importantes, cuando éstas sean necesarias para tomar una decisión o para realizar un estudio, no se podrá contar con la información completa. Si hay errores en los datos almacenados, se podrían presentar comportamientos abruptos de la variable o un comportamiento lineal, sin variaciones en el tiempo.

Las revisiones o evaluaciones sistemáticas de la calidad de las bases de datos son necesarias para poder detectar problemas que pudieran existir, y en caso de requerirse, dar inicio a procesos de limpieza o depuración de los datos almacenados. En el caso de las instituciones de educación superior, contar con información de calidad es importante, ya que esta es utilizada en los procesos de toma de decisiones

sobre diversos temas institucionales, uno de los cuales es la entrega de la docencia. Por citar solamente algunos ejemplos, se puede decir que es vital conocer la cantidad de estudiantes que se matriculan, esto con el propósito de mejorar la atención que se les brinda. La información sobre sus condiciones socioeconómicas es necesaria en la asignación de becas, la relativa a su rendimiento académico es útil durante la revisión de procesos de formación, entre otros.

Pero, ¿cómo saber si las bases de datos que se utilizan para diferentes fines, contienen información de calidad?, y si es así, ¿qué metodología seguir para darle la calidad deseada? Dar respuesta a estas y otras preguntas, es el objetivo de este trabajo.

Como un primer paso, es necesario conocer algunos aspectos relacionados con la calidad, los elementos básicos que debe cumplir una base de datos para tenerla, así como los pasos para identificar posibles errores en su contenido.

Este trabajo inicia con el abordaje de esos aspectos, posteriormente, se muestran algunos ejemplos de problemas de calidad de datos muy posiblemente producto de una digitación incorrecta, también se señala la necesidad de establecer controles estrictos de entrada de datos en los sistemas de información, incluyendo los controles cruzados con otras variables, así como la definición de campos de llenado obligatorio, lo anterior con el fin de minimizar errores y problemas posteriores al utilizar la información.

Calidad de datos

En términos generales, la calidad de datos “examina si los datos de una organización son confiables, consistentes, actualizados, están libres de duplicidades y si son apropiados para sus objetivos”. (Fisher, 2009, p.xvi).

Algunos de los elementos que Adelman (2009) señala como indicadores de calidad de datos se explican a continuación:

Tabla 1. Indicadores de calidad de datos

Indicador	Descripción
Integridad	Se refiere a que la información no debe ser alterada ni destruida o borrada, dejando datos “huérfanos”. Por ejemplo, que por error se eliminen datos personales de un estudiante pero permanezca en la base de datos información sobre cursos matriculados. En este caso, no habría forma de relacionar las asignaturas con el estudiante que las cursó.
Consistencia	Tanto el formato, como los datos en sí son consistentes en toda la base de datos. Por ejemplo, si el formato para los códigos de carrera es ‘0000’, es decir, cuatro dígitos, entonces todos los códigos de carrera deben cumplir con tal formato. De forma que no se encuentre la carrera 0001 y en otra tabla aparezca como 01
No redundancia	Un mismo dato, por ejemplo, el año de ingreso a la institución, no aparece innecesariamente en más de una tabla de la base de datos
Buen diseño de las bases de datos	En el diseño de las bases de datos se deben contemplar las reglas de normalización para bases de datos, a fin de lograr un diseño que, entre otras cosas, minimice la redundancia de información

Fuente: Elaboración propia basado en Adelman (2009)

Cuando se piensa en datos que son introducidos por personas haciendo uso de sistemas informáticos, se debe tener claro que existe el riesgo de que los datos, si no se toman las medidas adecuadas, queden registrados de manera incorrecta, incompleta o del todo no se registren.

Según Maydanchik (2007), pueden haber numerosas causas que generan posteriormente problemas de calidad de datos, entre éstas se pueden mencionar: a) la entrada o registro de datos de forma manual combinado con formularios o pantallas mal diseñadas o poco amigables con el usuario; b) el procesamiento de los datos en sistemas transaccionales, los que usualmente son modificados para incorporar nuevos requerimientos. Estos cambios en los sistemas y las bases de datos no siempre son probados antes de ponerlos en funcionamiento. c) cambios no capturados, los datos cambian con el tiempo y pueden volverse obsoletos si el objeto o elemento que

describen cambia. Por ejemplo, cierta información que durante un tiempo se capturó porque era importante en esos años, posteriormente, no se volvió a registrar debido a que ya no era necesaria ni relevante.

Un ejemplo del primer punto ocurre en las instituciones educativas cuando se digitan los datos de un estudiante de primer ingreso, el segundo caso se puede dar cuando se procesa información de una pregunta que recientemente ha sufrido cambios, por ejemplo la consideración de una nueva característica o categoría en un producto como los celulares en el comercio y en el tercer punto, la no utilización de información que con el tiempo es innecesaria, tal como la tenencia de televisores en blanco y negro en un hogar.

Un elemento importante a tomar en cuenta, y que justifica en buena medida llevar a cabo una iniciativa de calidad de datos, es el costo que representa para una organización mantener y utilizar datos duplicados, inexactos o con otro tipo de errores. Según estimaciones realizadas en los Estados Unidos, el uso de este tipo de información le ha costado a las empresas de ese país cerca de \$600 mil millones al año. (Adelman, 2009).

Tener datos de alta calidad implica que se puede generar información confiable para análisis y toma de decisiones, así por ejemplo, en el campo de la educación superior, se podría determinar que es mejor ofrecer tal o cual carrera en una región determinada, o que ciertos servicios (y por lo tanto los recursos necesarios para ofrecerlos) pueden ser mejor orientados a poblaciones con características específicas, o bien, que ciertos medios de comunicación masiva son de mayor penetración entre la población estudiantil, entre otras muchas posibilidades.

El análisis de calidad de datos planteado en este trabajo, se llevó a cabo en la Universidad Estatal a Distancia, sin embargo, es aplicable perfectamente a cualquier tipo de organización, incluyendo las instancias de educación, tanto primaria como secundaria, ya que en general, estas instituciones registran y almacenan información

personal del estudiante, su nivel y rendimiento académico, el centro de estudios del que proviene, entre otros aspectos. Con ella se pueden realizar estudios sobre su perfil de entrada, el rendimiento académico durante los estudios, su perfil de salida, la percepción sobre los servicios recibidos, así como otros temas de interés para la instancia educativa.

Esta información es valiosa para las autoridades porque les permite analizar aspectos transversales y longitudinales de los estudiantes, conocimiento que pueden utilizar para mejorar o ampliar los servicios ofrecidos a esta población. Este es un motivo suficientemente poderoso para pensar que los datos deben ser de calidad. Otro elemento de consideración es el costo monetario de mantener datos que presentan problemas, ya que “el tiempo invertido diagnosticando y corrigiendo los errores es tiempo gastado de manera no productiva” (Loshin, 2001, p.3).

La evaluación planteada en este trabajo si se lleva a cabo en coordinación con las instancias respectivas, puede mejorar de una manera significativa la información sobre la población meta, esto se traduce en una herramienta con suficiente potencial para las autoridades.

Metodología

Se utilizó para este análisis una base de datos institucional con información que es consultada por varias instancias en diferentes procesos. El análisis fue realizado en el año 2011. Se evaluó información comprendida entre los años 1980 al 2011. En este caso particular, se decidió seguir en buena parte la metodología de evaluación de calidad de datos que propone el autor Arkady Maydanchik (2007), misma que se describe brevemente en la siguiente tabla:

Tabla 2

Metodología para la evaluación de la calidad de los datos

Paso	Descripción
Recopilación de metadata general	Listados de campos, descripciones, valores válidos, modelos relacionales.
Análisis de datos y perfilado de datos	Es el proceso de examinar los datos y entender su contenido y estructura actual.
Poblar el área de pruebas (base de datos local)	Se debe crear un área de trabajo o pruebas en la cual replicar la base de datos real.
Diseño de reglas de calidad de datos	Estas reglas son restricciones sobre valores válidos de los datos y relaciones que pueden ser revisados mediante software.
Implementación de las reglas de calidad de datos	Este proceso implica la codificación o programación de las reglas utilizando una combinación de herramientas: consultas SQL, procedimientos almacenados, programas.
Afinamiento de las reglas de calidad de datos	Se basa ampliamente en la verificación manual de una muestra de datos y la comparación de resultados con los errores identificados por las reglas de calidad de datos.
Tabulación de resultados agregados	Cada resultado agrega errores identificados por las reglas de calidad de datos en un único valor – un porcentaje de los registros con datos buenos entre todos los registros considerados.
Cuadro de mando de calidad de datos	Es el producto central del proyecto de evaluación de calidad de datos. Es un programa que brinda información sobre la calidad de los datos y permite análisis agregado o detallado, según se requiera.

Fuente: elaboración propia basado en Maydanchik (2007), 2012

Cabe aclarar que en este primer acercamiento, y por limitaciones de tiempo, no se llegó al diseño e implementación de un software que automatice totalmente la evaluación que se lleva a cabo, por lo que el cuadro de mando de calidad de datos está aún por ser diseñado e implementado. Tampoco se utilizaron herramientas de software especializado para calidad de datos, solamente se empleó el SQL Server Management Studio 2008 para la construcción de la base de datos, la implementación y ejecución de las reglas necesarias.

El objetivo general de esta primera evaluación fue conocer el estado actual del contenido de la base de datos en cuestión, específicamente, de aquellas tablas

utilizadas para la generación de estadísticas periódicas. De esto, se deduce que no se llevó a cabo una revisión de la base de datos en su totalidad.

La figura 1 muestra los pasos ejecutados para realizar la evaluación de la base de datos seleccionada. Estos se describen posteriormente.

Figura 1. Evaluación de la calidad de los datos seleccionados



Fuente: Elaboración propia basado en Maydanchik (2007), 2012

- 1. Recopilación de la metadata general:** aunque se trataba de una base de datos conocida por quien realizó la evaluación, fue necesario efectuar algunas consultas relacionadas con el significado de algunos códigos, por ejemplo, códigos de estado civil, que no se encuentran almacenados en una tabla separada en la base de datos real.
- 2. Selección de datos:** fue necesario identificar previamente las tablas que serían evaluadas, así como los campos dentro de éstas, que serían relevantes para la evaluación. También, se requirió identificar las relaciones existentes entre tablas, con el fin de determinar el estado actual de los datos involucrados en dichas relaciones.
- 3. Diseñar y poblar el área de pruebas:** se diseñó y construyó una base de datos local para contener los que iban a ser evaluados. Por facilidad, se utilizaron los mismos nombres (tanto en tablas como en los campos o variables) que en la base de datos real, se respetó también el tipo de datos de cada campo, así por ejemplo,

si el nombre de la persona se almacena en un espacio de 20 caracteres, en el área de trabajo, se hizo de igual manera.

Una vez creada esta base de datos, se hizo la transferencia y copia del contenido tabla por tabla, de la base de datos real hacia la de evaluación. En total, fue necesario copiar el contenido total de 23 tablas aproximadamente, entre las que iban a ser evaluadas y otras llamadas maestras, que si bien no fueron evaluadas, contenían información esencial especialmente en la fase de evaluación de relaciones.

- 4. Perfil de los campos:** a cada uno de los campos seleccionados como relevantes se le aplicó la técnica de perfil de datos, para esto se elaboraron tablas de frecuencias (absolutas y relativas) que contienen cada valor encontrado en el campo y la cantidad de veces que aparece. En total, se realizó un perfil para alrededor de 22 campos o variables. A continuación se muestra un ejemplo de lo que contenía cada tabla de frecuencias o perfil de los campos, cabe aclarar que no se muestran los resultados reales:

Tabla 3
Perfil del campo, Código de provincia

Valor del campo	Cantidad	Porcentaje
01	100	50%
02	10	5%
03	57	28.5%
Nulos	33	16.5%

Fuente: elaboración propia

- 5. Diseño de reglas de calidad de datos:** se identificaron las relaciones existentes entre dos o más tablas, con el fin de diseñar las reglas de calidad de datos que permitieran evaluar el estado de los datos de esas relaciones.
- 6. Implementación de las reglas de calidad de datos:** una vez diseñadas las reglas de calidad de datos, éstas fueron implementadas utilizando el SQL (Structured

Query Lenguaje, por sus siglas en inglés). Las reglas diseñadas e implementadas incluyen validaciones para:

- a. **Formato del campo:** por ejemplo validación del formato de las fechas de nacimiento de la persona, cantidad de dígitos de sus números telefónicos y formato de las direcciones de correo electrónico.
- b. **Valores válidos:** que no existan valores fuera de rangos o conjuntos de valores permitidos. Se usan en conjunto con el perfil del campo.
- c. **Restricciones de precisión:** permitieron evaluar que los valores de un campo tengan la misma precisión.

7. **Perfil de las relaciones:** al igual que con los campos, se hizo un perfil de las relaciones entre tablas. Por ejemplo, se esperaría que todo registro sobre una persona específica en la tabla A deba tener al menos una referencia en la tabla B. Se identificaron alrededor de 12 relaciones.

Resultados

Una vez aplicadas todas las reglas de calidad de datos que se diseñaron, partiendo de lo más simple, como las frecuencias de los datos hasta lo más complejo como la relación entre dos o más tablas, se encontraron diversos problemas en algunos de los campos o variables evaluadas, éstos se pueden agrupar en las siguientes categorías:

- Valores nulos: variables cuyo valor no está asignado (valores nulos) o contiene espacios en blanco. Este aspecto debe cuidarse, debido a que es información que no se obtuvo en su momento, y para obtenerla, se requiere contactar a la persona que puede brindar el dato, o bien, realizar un control cruzado con otras bases de datos cuyo contenido sea oficial.
- Valores no válidos: Son campos con valores fuera de rangos o del conjunto de valores permitidos. En este caso, se puede introducir información que no corresponde con los rangos válidos de la variable en análisis, por ejemplo, la

digitación en puntos o caracteres como paréntesis o comillas, que no pueden ser asociados a los campos establecidos a priori.

En estos casos, la solución es incrementar los controles de entrada de datos al sistema, así como mejorar los ya existentes.

- Duplicidad de valores: este problema se presenta en aquellos campos cuyo valor, debiendo ser único, se repite en diferentes registros. En otras palabras, varios registros de la base de datos comparten información que, por sus características, es única, como por ejemplo, número de teléfono celular o la dirección de correo electrónico. En este caso, se requiere también de mejoras en el sistema de información para que se controle y evite el ingreso de datos duplicados.
- Formatos incorrectos: se refiere a variables que no cumplen con el formato establecido, ya sea porque no se utilizaron todos los caracteres con los que contaba el formato, o porque se incluyeron caracteres no contemplados en el mismo. Este tipo de problema se presenta en variables que almacenan fechas, en algunos casos, se encuentran barras inclinadas o menos dígitos de los requeridos.
- La aplicación de las reglas de calidad de datos para el perfil de las relaciones también evidenció la existencia de problemas, por ejemplo, al relacionar dos tablas, se encontraron casos donde los registros de la tabla A no tenían correspondencia con los de la tabla B, lo que da como resultado información incompleta o registros huérfanos. Este tipo de problemas tiene repercusiones serias en las organizaciones, por ejemplo, si se piensa en información financiera relativa a la recolección de impuestos, al no existir una correspondencia entre el contribuyente y la recaudación en un año determinado, o bien, los saldos de una cuenta de un cliente de un banco y sus datos personales.

Es importante mencionar que buena parte de los problemas de calidad se presentan en los grupos de registros más antiguos, y que en el transcurso del período al que pertenecen los datos evaluados (1980 – 2011), el sistema de información que permite

su registro ha tenido mejoras en sus controles de entrada. Sin embargo, hay otros problemas de calidad que se dan en registros mucho más recientes.

Posterior al diseño y aplicación de las reglas que permitieron hacer la evaluación, se ha coordinado con las instancias respectivas para analizar el porqué de los problemas encontrados y tomar las acciones correspondientes, con el fin de mejorar los controles y corregir los errores que se detectaron. Para tal propósito, se ha hecho uso de otras fuentes de información alternativa, tales como los expedientes de los estudiantes. No obstante, los datos que no pueden ser recuperados podrían ser considerados como valores perdidos por parte de los usuarios de los datos y ser desestimados a la hora de generar estadísticas.

Un aspecto importante de señalar es que la metodología utilizada identifica problemas en la calidad de los datos, mediante la revisión de las restricciones de las variables y los cruces o relaciones entre tablas de la misma base de datos, sin embargo, la actualización es un elemento que está fuera del alcance de la metodología seguida.

Conclusiones

Velar por la calidad de los datos que se generan en una organización es un proceso tan importante como los propios de la organización. Actualmente, vivimos en una sociedad en la que el conocimiento, generado mediante el análisis de datos y de información, constituye una herramienta poderosa para apoyar la toma de decisiones, por tal motivo estos deben cumplir esta condición.

La calidad en los datos de un sistema de información garantiza que las consultas que se hagan para tomar decisiones, como por ejemplo, definir la contratación de tutores o profesores, tamaño de los grupos, apertura de cursos, entre otros, sean confiables.

Por otro lado, pueden ser utilizadas para realizar autoevaluaciones de las carreras y posteriormente acreditarlas, lo que redundaría en una mejor educación para los estudiantes.

Las revisiones de la calidad deben ser sistemáticas, por lo que se hacen algunas recomendaciones basadas en el ejercicio que se realizó en la Universidad Estatal a Distancia de Costa Rica, las cuales se presentan a continuación.

Debido a restricciones de tiempo no fue posible llegar a la etapa de implementación del software, para la evaluación de la calidad de los datos, que sugiere la metodología. Este vendría a generar indicadores que permitirían cuantificar la calidad de los datos evaluados, así como la comparación del porcentaje de calidad de esos datos a través de los años.

Con la información que se obtiene en estas evaluaciones, se pueden fijar metas, con el fin de mejorar los indicadores y cumplir con los estándares de calidad establecidos por la organización.

Como resultado de la evaluación, se puede decir que es de suma importancia mejorar controles de entrada de datos, tales como el establecimiento de campos de llenado obligatorio, en los diferentes sistemas de información. Esto con el fin de minimizar el registro de datos erróneos o la no digitación de ellos.

Tomando en cuenta que aunque existan controles de entrada de datos en los sistemas de información, siempre existe la posibilidad de registrar datos con algún grado de error, en los casos que se requiera, será también necesario instaurar procesos de control de calidad de los datos, mediante los cuales éstos sean revisados y corregidos posteriormente a su almacenamiento.

En aquellos sistemas de información, donde el usuario no es quien registra sus propios datos, se vuelve de gran importancia la capacitación del personal encargado de digitalizarlos.

Una vez realizada la primera evaluación, se debe planificar en el tiempo, las subsiguientes evaluaciones con el propósito de monitorear tanto la eficacia de los

controles de entrada de los datos, como el nivel de calidad que va alcanzando la información registrada.

Es de suma importancia que la calidad de los datos se convierta en una iniciativa apoyada por quienes dirigen la organización, de forma tal que se conformen equipos de trabajo, para que se dediquen tiempo completo a depurar las bases de datos y a monitorear la calidad de su contenido.

Si bien es cierto, una iniciativa de calidad de datos puede representar inicialmente un costo adicional para la organización, por el hecho de que las personas que conformarían el equipo deben capacitarse, y probablemente ser sustituidas temporalmente en sus puestos originales, además de la compra de equipo y software para realizar el trabajo, el costo de no solucionar los problemas de calidad de datos podría ser mucho mayor, con el paso de los años.

Debe complementarse la metodología utilizada en este análisis, con la revisión periódica y la actualización de aquellos datos que no presentaron problemas en la revisión.

Finalmente, es de gran importancia que exista coordinación entre todas las instancias que hacen uso del sistema, cuya base de datos se evalúa, ya que esto contribuye a crear condiciones para elevar la calidad de los datos, así como para evitar el no registro de información relevante para los investigadores, autoridades y ejecutivos.

Bibliografía

Adelman, S. (2009). Yet Another Piece On Data Quality. Enterprise Information Management Institute.org. In: *Eiminsight Magazine Current Issue*, volumen 3(6). Available : <http://www.eiminstitute.org/library/eimi-archives/volume-3-issue-6-june-2009/yet-another-piece-on-data-quality>

Fisher, T. (2009). *The Data Asset. How Smart Companies Govern Their Data for Business Success. Estados Unidos: John Wiley & Sons, Inc.*

Loshin, D. (2001). *Enterprise Knowledge Management. The data quality approach. Estados Unidos. Morgan Kaufmann Publishers.*

Maydanchik, A. (2007). *Data Quality Assessment. United States of America: Technics Publications.*

Whiting, R. (2006). Hamstrung by defective data. Information Week. Available: <http://www.informationweek.com/news/187200771>