

# Evaluación de la calidad de la información en las organizaciones, a través de herramientas tecnológicas<sup>1</sup>

Evaluation of the Quality of Data and Information in Organizations through Technological Tools

Avaliação da qualidade dos dados e informações nas organizações através de ferramentas tecnológicas

<https://doi.org/10.15332/24631140.7785>

Artículos de investigación

Alejandra Piedrahita Mazo<sup>2</sup>

Diana Montoya Quintero<sup>3</sup>

Recibido: 12 de noviembre del 2021

Revisado: 20 de abril del 2022

Aceptado: 31 de mayo del 2022

Citar como:

Piedrahita Mazo, A. y Montoya Quintero, D. (2022). Evaluación de la calidad de la información en las organizaciones, a través de herramientas tecnológicas.

*SIGNOS, Investigación en Sistemas de Gestión*, 14(2).

<https://doi.org/10.15332/24631140.7785>



## Resumen

La información constituye un activo estratégico para las diferentes prácticas organizacionales. Por esa razón, la calidad de los datos es fundamental para el buen funcionamiento de todo negocio, ya sea grande, mediano o pequeño. Los sistemas informáticos son un punto clave para la toma de decisiones en

---

<sup>1</sup> Artículo de investigación

<sup>2</sup> Estudiante de Ingeniería de la Calidad y tecnóloga en Calidad, Instituto Tecnológico Metropolitano. Correo electrónico: [alejandrapiedrahita209288@correo.itm.edu.co](mailto:alejandrapiedrahita209288@correo.itm.edu.co); ORCID: <https://orcid.org/0000-0003-4581-8466>; CvLAC: [https://scienti.minciencias.gov.co/cvlac/visualizador/generarCurriculoCv.do?cod\\_rh=0000050936&lang=es](https://scienti.minciencias.gov.co/cvlac/visualizador/generarCurriculoCv.do?cod_rh=0000050936&lang=es)

<sup>3</sup> Doctora en Ingeniería de Sistemas e Informática, Instituto Tecnológico Metropolitano. Correo electrónico: [dianamontoya@itm.edu.co](mailto:dianamontoya@itm.edu.co); ORCID: <https://orcid.org/0000-0003-0761-4067>; CvLAC: [https://scienti.minciencias.gov.co/cvlac/visualizador/generarCurriculoCv.do?cod\\_rh=0000478342&lang=es](https://scienti.minciencias.gov.co/cvlac/visualizador/generarCurriculoCv.do?cod_rh=0000478342&lang=es)

cualquier empresa que pretenda ser competitiva, debido a que gestionan datos útiles en diferentes ámbitos. En ese sentido, es necesario que la información tenga la calidad adecuada, con la integridad como uno de los atributos más relevantes. Diversas investigaciones orientadas por enfoques y métodos específicos han buscado identificar problemas en la calidad de los datos y la información, así como definir métodos y mecanismos para garantizarla y medirla. Disponer de datos e información de alta calidad, en definitiva, permite responder mejor a las prioridades de las organizaciones, interactuar de forma inteligente con los clientes, y desarrollar la operación de la empresa de manera eficiente. De acuerdo con lo anterior, este artículo tiene el objetivo de analizar la evaluación de la calidad de los datos y la información, a través de herramientas tecnológicas con base en la revisión de la literatura. Los resultados indican que una adecuada gestión de la calidad de los datos y la información depende, en gran medida, de la medición. El establecimiento de indicadores y la recogida de métricas son fundamentales para comprender cada componente del ciclo de calidad en los procesos singulares de cada organización.

**Palabras clave:** calidad, evaluación, datos, información, herramientas tecnológicas.

## Abstract

Data and information are a strategic asset for the different organizational practices, therefore the quality of data is essential for the proper functioning of any business, whether large, medium or small. Information systems are a key point for decision making in any company that intends to be competitive. This is due to the fact that these systems are based on data management for the generation of useful information for decision making. In this context, it is necessary that the information generated within the different organizations comes from data with the appropriate quality, integrity being one of the most relevant attributes. Various approaches and methods have been studied by different researchers in the area, most of which have as their main objectives to identify the scenarios where data quality and information quality problems can occur, as well as to define methods and mechanisms to ensure and measure them. It is important to emphasize that the quality of data and information is constantly being studied to provide a better understanding of them, providing reliable data, which will allow to make more accurate and safer decisions within the different organizations. Having high quality data and information available will ultimately allow to better respond to key organizational priorities and interact intelligently with customers, as well as achieve greater operational efficiency. Therefore, this article aims at characterizing the evaluation of data and information quality through technological tools according to the literature review. The different results indicate that the proper management of data and information quality will depend to a large extent on knowing how to measure it.

Defining indicators and collecting metrics will allow to gain an understanding of each component of the data quality cycle, as each organization is unique.

**Keywords:** Quality, evaluation, data, information, technological tools.

## Resumo

Dados e informações são um ativo estratégico para as diferentes práticas organizacionais, portanto, a qualidade dos dados é essencial para o bom funcionamento de qualquer empresa, seja ela grande, média ou pequena. Os sistemas de informação são um ponto-chave para a tomada de decisões em qualquer empresa que pretenda ser competitiva. Isto se deve ao fato de que estes sistemas são baseados no gerenciamento de dados para a geração de informações úteis para a tomada de decisões. Neste contexto, é necessário que a informação gerada dentro das diferentes organizações decorra de dados com a qualidade apropriada, sendo a integridade um dos atributos mais relevantes. Diversas abordagens e métodos têm sido estudados por diferentes pesquisadores da área, tendo a maioria deles como principais objetivos a identificação dos cenários onde a qualidade dos dados e os problemas de qualidade da informação podem ocorrer, assim como a definição de métodos e mecanismos para garanti-la e mensurá-la. É importante salientar que a qualidade dos dados e informações é constantemente estudada para proporcionar uma melhor compreensão deles, fornecendo dados confiáveis, que permitirão tomar decisões de forma muito mais segura e precisa dentro das diferentes organizações. Ter dados e informações de alta qualidade à sua disposição irá permitir responder melhor às principais prioridades organizacionais e interagir inteligentemente com os clientes, assim como alcançar maior eficiência operacional. Portanto, o objetivo deste artigo é caracterizar a avaliação da qualidade dos dados e das informações através de ferramentas tecnológicas de acordo com a revisão bibliográfica. Os diferentes resultados indicam que a gestão adequada da qualidade dos dados e das informações dependerá, em grande parte, de saber como medi-la. O estabelecimento de indicadores e a coleta de métricas permitirá a compreensão de cada componente do ciclo de qualidade dos dados, uma vez que cada organização é única.

**Palavras-chave:** Qualidade, avaliação, dados, informações, ferramentas tecnológicas.

## Introducción

Actualmente, todo tipo de datos e información<sup>4</sup> es susceptible de ser evaluado, sobre todo si se requiere recolectar una cantidad específica. Existe un corpus teórico contrastado relativo a los criterios de evaluación del almacenamiento de datos e información en los soportes tradicionales, incluso en los electrónicos. Sin embargo, los datos y la información telemática, especialmente la accesible a través de Internet, aún es objeto de reflexión e investigación. Se plantea la necesidad de definir parámetros y procedimientos para analizar la calidad de la información y los datos accesibles en línea (Merlo Vega, 2003).

Actualmente, las organizaciones consideran los datos y la información como recursos valiosos para sus procesos internos y su interrelación con el entorno, lo cual justifica el estudio de su calidad y los esfuerzos por mejorarla. Para evaluar la calidad de los datos y la información en diferentes organizaciones, es necesario contar con un conjunto de atributos que permita medirla y conocer la opinión del cliente. Este último componente facilita el diseño de estrategias de excelencia en la prestación del servicio, y genera un auténtico valor dentro del proceso de gestión de las diferentes industrias (Torres y Rojas, 2008).

Con el rápido desarrollo de la tecnología de recopilación y almacenamiento, la capacidad de absorber la información útil oculta detrás de los datos se hace cada vez más importante (Pan et ál., 2018). En este sentido, es indispensable evaluar la calidad de los datos brutos y la información procesada como base para la toma de decisiones científicas y el desarrollo de proyectos de intercambio de información. Esto evita diversos problemas para la gerencia empresarial (Chen et ál., 2009).

Se estima que el 80 % de todos los datos que se producen en la actualidad no están estructurados, lo cual muestra la importancia de la generación de conocimiento sobre evaluación de la calidad, mediante criterios de integridad y precisión (Taleb et ál., 2019). Para consolidar la comprensión y medición de la calidad, las empresas se están convirtiendo en "sistemas abiertos", al incorporar diversas técnicas y herramientas tecnológicas que garantizan la calidad y estabilidad de los datos en un sistema de gestión de información computarizado (Mandke y Nayar, 2004).

---

<sup>4</sup> Los datos son los elementos primarios o valores básicos que, en sí mismos, no dicen nada. La información resulta de la sistematización y procedimiento de datos. La información es fundamental para la toma de decisiones.

## Marco referencial

En los últimos años se ha enfatizado en la importancia de los datos y la información, ya que se han popularizado expresiones como: “los datos son la nueva moneda”, “los datos son el nuevo petróleo”, “los datos son la mina oculta”, entre otras (Gómez y Piattini, 2018). Muchos de los procesos de una organización generan información que es almacenada en bases de datos. Estas son insumo para realizar estudios para mejorar la calidad en diferentes ámbitos organizacionales. En una empresa de servicio, por ejemplo, un centro médico, los datos y la información sobre el tiempo que dura un paciente, desde su llegada hasta el momento de salida con sus medicamentos, son de vital importancia. Estos permiten al tomador de decisiones conocer la eficiencia del servicio y buscar la mejora continua mediante el monitoreo de un indicador de calidad, que, en este caso, sería el tiempo promedio de duración de la consulta médica, incluyendo la entrega de medicamentos (Zúñiga Segura y Sánchez Godínez, 2012)

La medición de la calidad de los datos y la información implica una mirada contextual y multidimensional, lo que significa que se mide según diversas dimensiones, definidas de acuerdo con el enfoque de investigación y los tipos de datos. Por ejemplo, Panahy et ál. (2014) señala la relevancia de “tener los datos correctos, en el formato correcto, en el lugar correcto, en el momento correcto”. Por lo tanto, la evaluación de la calidad de los datos se vincula a un enfoque según el ámbito o servicios de información en que se aplique, que van desde redes de sensores hasta sistemas de ciudades inteligentes. A menudo las evaluaciones se derivan de las características intrínsecas de los datos, y suelen estar desconectadas de contextos de aplicación específicos o no son aplicables o eficientes para grandes conjuntos de datos (Li et ál., 2019).

A lo largo de la historia de las organizaciones, la complejidad del reprocesamiento de datos e información, así como el tamaño y el volumen de estos han limitado mejoras en la calidad en sí misma y en la evaluación (Kumar et ál., 2019). La gestión de la calidad de los datos está implicada en la medición, dado que la primera contempla actividades de planificación, implementación y control de los datos, mediante diferentes metodologías bien estructuradas. Por su parte, la medición de la calidad permite hallar diferencias entre diversos grupos de datos, a partir de procedimientos y prácticas del marco utilizado. Con lo anterior, se concluye que el resultado de la calidad de los datos y la información afecta los beneficios o costos dentro de cualquier organización (Bowo et ál., 2019).

A continuación, se presentan diferentes enfoques frente a las metodologías y métricas planteadas por diversos autores acerca de la calidad de la información.

### **Métricas aplicadas a la evaluación de la calidad de los datos**

En la actualidad, hay un problema de escasez de métricas consistentes para la evaluación de la calidad de bases de datos. En todos los casos estudiados, las métricas se encuentran sometidas a la subjetividad, tanto del punto de vista del autor que las propone, como de las personas que las aplican. Más aún, existe un aparente desconocimiento por parte de los autores que proponen sus métricas con respecto a su significado y los objetivos de implementación (Brinck, 2002).

También es importante hablar de las diferentes dimensiones de la calidad de datos. Estas representan un conjunto de atributos, factores o valores que califican un producto o servicio. Un factor representa un aspecto particular de una dimensión de calidad. Cada uno puede medirse con diferentes métodos, que implementan métricas específicas. De esta forma, una métrica puede ser medida por varios métodos. La métrica se define a través de la semántica, la unidad de medición y la granularidad de la medida (Clerici y Fernández, 2019). La figura 1 representa los diferentes conceptos y su jerarquía dentro de la calidad de los datos.

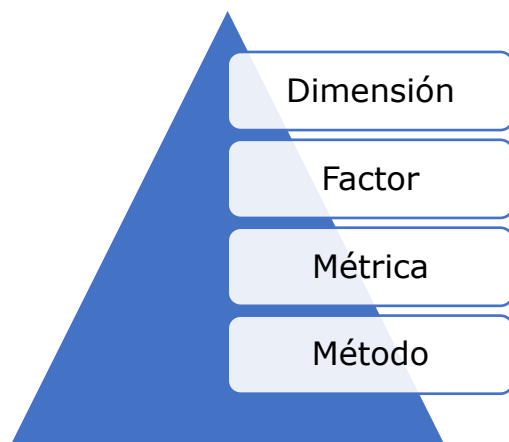


Figura 1. Conceptos y jerarquía dentro de la calidad de los datos

Fuente: Clerici y Fernández (2019).

### **Metodologías planteadas por otros autores**

Existen diversas metodologías para la evaluación de datos e información en los contextos empresariales. En esta sección se exponen algunas metodologías relevantes.

### ***Estudio de caso de PT JAS con el marco TDQM***

Esta metodología utiliza el método de entrevistas y observaciones al jefe de personal operativo y de tecnologías de información del negocio. Además de probar la credibilidad de los datos y documentos, Bowo et ál. (2019) acuden a la triangulación, que consiste en obtener datos de la misma fuente con diferentes técnicas de recopilación de datos. El objetivo es definir las dimensiones de la calidad de los datos, medir una métrica, identificar fuentes de errores de datos, y crear parámetros de mejora para evaluar de forma correcta la calidad de los datos y la información.

### ***Metodología para pequeñas y medianas empresas***

Esta metodología consiste en la validación de datos mediante la planificación de recursos empresariales (ERP) y el sistema de ejecución de fabricación (MES). Los resultados de la evaluación se resumen en dimensiones de calidad y pueden utilizarse en estudios comparativos, en cualquier tipo de industria. El marco analítico propuesto muestra la frecuencia con que las dimensiones de calidad se ven afectadas por observaciones erróneas. Las dependencias entre las medidas de calidad se calculan a través de ecuaciones y valores entrópicos (Pontevedra et ál., 2019).

### ***Metodología con métricas ponderadas***

En este caso, se proponen nuevas métricas basadas en ponderaciones de datos para hacerlas más prácticas en cuanto a la ponderación de los datos, con el fin de organizar la medición y mejorar la calidad. Se incluyen diferentes dimensiones de calidad como: precisión, integridad, puntualidad y relevancia (Vaziri et ál., 2019).

### ***Identificación de herramientas tecnológicas de evaluación***

En este diseño metodológico se realiza un relevamiento de las herramientas existentes de medición de calidad de datos, en el mercado y los servicios disponibles en la web. A partir de lo anterior, se propone el diseño, la implementación y documentación de un modelo para la evaluación de la calidad de bases de datos relacionales. También, permite su ejecución y visualización de resultados (Clerici y Fernández, 2019).

### ***Calidad de datos y su evaluación: un caso de estudio***

El autor Arkady Maydanchik (citado en Zúñiga Segura y Sánchez Godínez, 2012) describe una serie de pasos para la evaluación y aplicación de la calidad en una base de datos determinada. Los resultados obtenidos a partir de esta metodología permiten determinar diferentes oportunidades de mejora en la calidad de los datos. Por lo tanto, aportan gran valor en los procesos de toma de decisiones.

## Metodología

El diseño de investigación propuesto se basa en las metodologías planteadas por Clerici y Fernández (2019) y Zúñiga Segura y Sánchez Godínez (2012). De estas, se rescatan diferentes métricas y pasos necesarios para evaluar de forma correcta la calidad de los datos y la información. En este caso, se enfatiza en la evaluación de las métricas de calidad de los datos.

Los pasos para evaluar la calidad de los datos y la información son:

- *Determinar la exactitud de los datos e información:* este primer paso debe indicar qué tan precisos, válidos y libre de errores están los datos representados en el sistema de información empresarial. Se definen factores como: correctitud, correctitud sintáctica y precisión, para definir la exactitud.
- *Establecer la completitud de los datos:* se debe establecer si el sistema de información contiene toda la información de interés, y si este cubre aspectos extensionales (cantidad de entidades/estados de la realidad representados) e intencionales (cantidad de datos sobre cada entidad/estado). Para la completitud, se definen los siguientes factores: cobertura y densidad.
- *Realizar análisis y perfilado de datos:* se examinan los datos para entender su contenido y estructura actual en el interior de la organización.
- *Verificar la frescura de los datos y la información:* se debe involucrar una perspectiva temporal de los datos del sistema de información organizacional, es decir, que tan viejos son. Los factores que se definen para la frescura son: actualidad, oportunidad y volatilidad.
- *Determinar la consistencia de los datos y la información:* en este paso la consistencia captura la satisfacción de reglas semánticas definidas sobre los datos, que pueden ser de integridad para una base de datos o de usuarios. Los factores para este paso son: integridad de dominio e integridad-interrelación. En todos los casos se mide qué tan bien se satisfacen las reglas de integridad.
- *Determinar la unicidad de los datos y la información:* se identifica el nivel de duplicación de los datos registrado en el sistema de información organizacional, lo cual ocurre cuando una misma entidad está representada dos o más veces. Los factores que se definen son: no-duplicación y no-contradicción, en ambos casos se debe medir la cantidad de datos repetidos.
- *Diseñar de reglas de calidad de los datos:* estas se refieren a restricciones sobre valores válidos y relaciones que pueden ser revisadas mediante un *software* seleccionado por cada organización.

A continuación, se representa la metodología propuesta para este artículo.



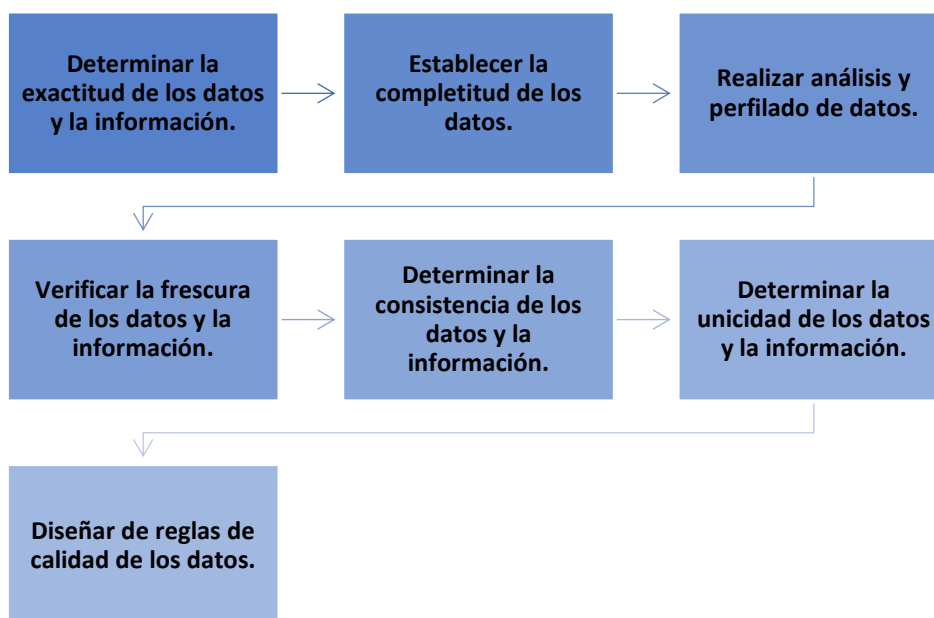


Figura 2. Metodología para la evaluación de la calidad de los datos y la información

Fuente: elaboración propia.

## Resultados

Tabla 1. Criterios para la evaluación de la calidad de datos e información, a través de herramientas tecnológicas

Herramienta		Dimensión		Metodologías	
Datacleaner	X	Exactitud.	X	Estudio de caso de PT JAS con el marco TDQM.	
Data Quality Services		Consistencia.	X	Metodología para pequeñas y medianas empresas.	
Microsoft Power Bi	X	Completitud.	X	Metodología con métricas ponderadas.	
Programación en R	X	Unicidad.	X	Identificación de herramientas tecnológicas de evaluación.	X
Open Studio for Data Quality de Talen		Frescura.	X	Calidad de datos y su evaluación: un caso de estudio.	X
Excel	X	Integridad.			
Tableau Public		Conformidad.			

Fuente: elaboración propia.

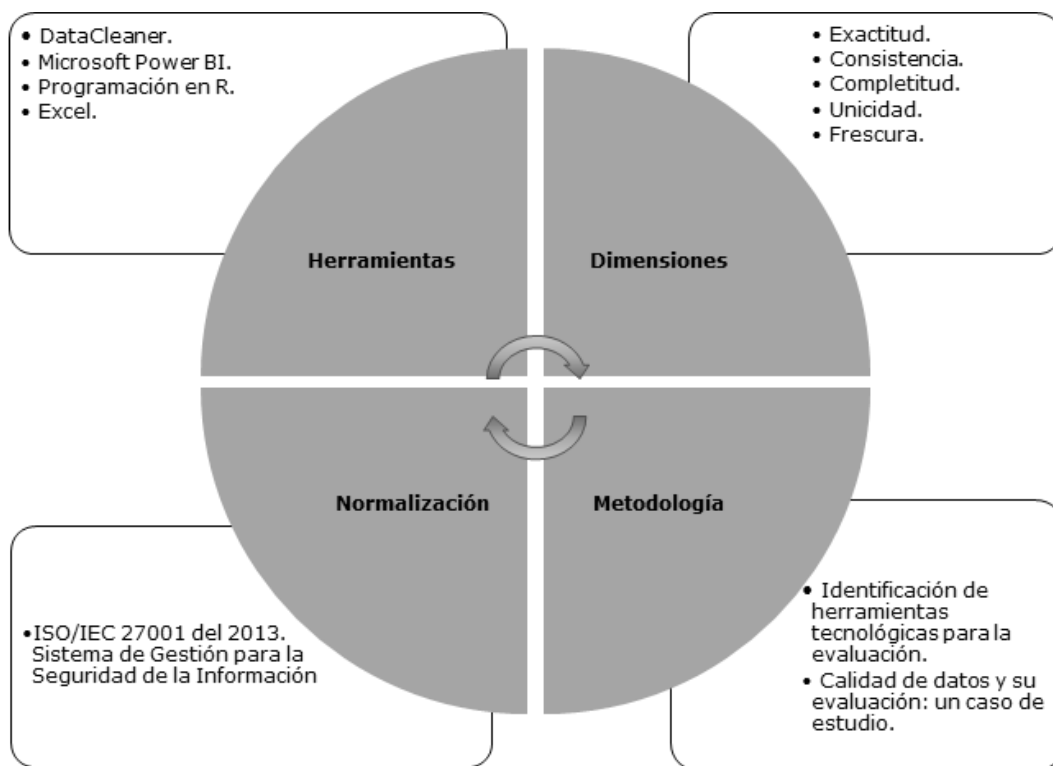


Figura 3. Componentes y elementos para evaluar la calidad de los datos y de la información

Fuente: elaboración propia.

## Discusión y conclusiones

Con la evaluación de la calidad de los datos y la información en el interior de las organizaciones, es posible identificar de forma temprana los riesgos, y así lograr disminuirlos. Cada vez es más importante velar por la calidad de los datos y la información que se generan en las diferentes organizaciones. La evaluación de la calidad de los datos y la información es un proceso central para la toma de decisiones administrativas y es tan importante como los propios de cada empresa.

La metodología planteada puede utilizarse en pequeñas, medianas y grandes empresas, ya que permite evaluar diferentes métricas de calidad, al proporcionar acciones coherentes y necesarias para realizar un diagnóstico eficaz de las causas que originan los problemas en los datos y la información.

La calidad de los datos y la información no solo se vincula con los repositorios empresariales, sino que también escalan a nivel tecnológico y proporcionan diferentes alternativas para su ejecución. Es de suma importancia que la calidad

de los datos se convierta en una iniciativa apoyada por quienes dirigen la organización, de tal forma que se conformen equipos de trabajo dedicados a depurar las diferentes bases de datos. Así, es posible monitorear en detalle la calidad del dato y la información de una forma correcta.

Finalmente, se concluye que para que la evaluación de la calidad de los datos y la información sea efectiva es importante invertir en la capacitación y concientización de todos los colaboradores de una organización. La concientización de las empresas en torno al manejo adecuado de los datos y la información es la única vía para obtener grandes ventajas y beneficios a nivel organizacional.

## Referencias

- Bowo, W. A., Suhanto, A., Naisuty, M., Ma'mun, S., Hidayanto, A. N. y Habsari, I. C. (2019). Data quality assessment: A case study of PT JAS using TDQM Framework. *Proceedings of 2019 4th International Conference on Informatics and Computing, ICIC 2019*. pp. 1-6. <https://doi.org/10.1109/ICIC47613.2019.8985896>
- Brinck, V. S. (2002). *Análisis de métricas de calidad para esquemas conceptuales de bases de datos* (pp. 1-10). <http://inf.udec.cl/~revista/ediciones/edicion8/Vsanjuan.pdf>
- Clerici, F. y Fernández, B. (2019). *Herramienta para la evaluación de calidad de datos* [Informe final de proyecto de grado]. Instituto de Computación Facultad de Ingeniería, Universidad de la República, Montevideo. <https://www.colibri.udelar.edu.uy/jspui/bitstream/20.500.12008/20268/1/tg-clerici-fernandez.pdf>
- Chen, B., Wang, B., Zheng, C. y Hu, X. (2009). Research and implementation of information quality improvement bing. *4th International Conference on Cooperation and Promotion of Information Resources in Science and Technology, Coinfo 2009* (pp. 225-229). <https://doi.org/10.1109/COININFO.2009.17>
- Gómez, A. y Piattini, M. (2018). Importancia de la calidad de los datos en la transformación digital. *Ruidera*, (13), 1-15. <https://ruidera.uclm.es/xmlui/handle/10578/18798>
- Kumar, J., Crow, M. C., Devarakonda, R., Giansiracusa, M., Guntupally, K., Olatt, J. V., Price, Z., Shanafield, H. A. y Singh, A. (2019). Provenance-aware workflow for data quality management and improvement for large continuous scientific data streams. *Proceedings - 2019 IEEE International Conference on Big Data, Big Data 2019* (pp. 3260-3266). <https://doi.org/10.1109/BigData47090.2019.9006358>
- Li, A., Zhang, L., Qian, J., Xiao, X., Li, X. y Xie, Y. (2019). Todqa : Efficient Task-Oriented Data Quality Assessment. *IEEE*. <https://doi.org/10.1109/MSN48538.2019.00028>
- Mandke, V. V. y Nayar, M. K. (2004). Beyond quality: The information integrity imperative. *Total Quality Management and Business Excellence*, 15(5-6), 645-654. <https://doi.org/10.1080/14783360410001680134>
- Torres, M. y Rojas, D. (2008). Modelo de evaluación de la calidad de la información corporativa en los servicios médicos. *Enlace*, 5(3). [http://ve.scielo.org/scielo.php?script=sci\\_arttext&pid=S1690-75152008000300003](http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1690-75152008000300003)

- Merlo Vega, J. A. (2003). La evaluación de la calidad de la información web : aportaciones teóricas y experiencias prácticas. *Sociedad de la Información*, (8), 101-110.
- Pan, X., Zhang, M. y Chen, X. (2018). A method of quality improvement based on big quality warranty data analysis. *IEEE International Conference on Software Quality, Reliability and Security Companion*, 643-644. <https://doi.org/10.1109/QRS-C.2018.00115>
- Panahy, P. H. S., Sidi, F., Affendey, L. S. y Jabar, M. A. (2014). The impact of data quality dimensions on business process improvement. *2014 4th World Congress on Information and Communication Technologies, WICT 2014* (pp. 70-73). <https://doi.org/10.1109/WICT.2014.7077304>
- Pontevedra, V., Colangelo, E., Günther, L. C., Colangelo, E., Wiendahl, H., Bauer, C., Santana, A., Afonso, P., Zanin, A. y Wernke, R. (2019). Data quality assessment for improved decision-making: A methodology for small and medium-sized enterprises. *Procedia Manufacturing*, 29 (pp. 583-591). <https://doi.org/10.1016/j.promfg.2019.02.114>
- Taleb, I., Serhani, M. A. y Dssouli, R. (2019). Big data quality assessment model for unstructured data. *Proceedings of the 2018 13th International Conference on Innovations in Information Technology, IIT 2018* (pp. 69-74). <https://doi.org/10.1109/INNOVATIONS.2018.8605945>
- Vaziri, R., Mohsenzadeh, M. y Habibi, J. (2019). Measuring data quality with weighted metrics. *Total Quality Management and Business Excellence*, 30(5-6), 708-720. <https://doi.org/10.1080/14783363.2017.1332954>
- Zúñiga Segura, L. y Sánchez Godínez, E. (2012). Calidad de datos y su evaluación: un caso de estudio. *Calidad en la Educación Superior*, 3(2), 33-49. <https://doi.org/10.22458/caes.v3i2.444>