



Framework B-DataCom: métricas para la comprensibilidad de los datos en entornos de analítica de negocio

B-DataCom framework: metrics for data comprehensibility in business analytics environments

Silvia Alejandra Molina España ✉

Universidad del Cauca, Popayán, Colombia

silviamolina@unicauca.edu.co

ORCID: 0009-0002-8318-533X

Sandra Lorena Buitrón Ruiz

Universidad del Cauca, Popayán, Colombia

sandrabr@unicauca.edu.co

ORCID: 0000-0001-6127-1849

Francisco José Pino Correa

Universidad del Cauca, Popayán, Colombia

fjpino@unicauca.edu.co

ORCID: 0000-0003-0668-4485



<https://doi.org/10.36825/RITI.14.35.004>

Recibido: Junio 14, 2026

Aceptado: Octubre 01, 2026

Resumen: En las organizaciones desarrolladoras de software existen procesos técnicos misionales enfocados en la construcción del producto de software, pero también existen otros procesos de negocio, no menos relevantes, que soportan dicha construcción como la toma de decisiones organizacionales. En este contexto, la calidad de los datos de negocio se ha vuelto cada vez más relevante, sin embargo, existe una brecha sobre la comprensibilidad de estos datos, debido a la ausencia tanto de su conceptualización, como de mecanismos formales para la medición de su calidad; limitando de esta manera su aprovechamiento para los procesos de análisis y toma de decisiones organizacionales. Este trabajo presenta el *framework* B-DataCom, el cual propone métricas para la medición de la comprensibilidad de los datos de negocio en organizaciones desarrolladoras de software. Se realiza un mapeo sistemático de la literatura para identificar elementos viables de medición con respecto a la calidad del dato en un contexto empresarial; lo que fundamentó el *framework*, considerando además una extensión en la Ontología de la medición para su conceptualización integrada de la semántica de la comprensibilidad. B-DataCom define las características descriptivas de la comprensibilidad del dato, un modelo conceptual que representa la teoría valorada en la investigación y un conjunto de 11 métricas GQM (*Goal-Question-Metric*). Luego de un ejercicio de instanciación de las métricas, sobre el proceso MPS (Mejora de Procesos) referenciado en COMPETISOFT, se concluye que las métricas propuestas pueden ser escaladas a través de futuros estudios de caso.

Palabras clave: *Comprensibilidad de los Datos, Goal-Question-Metric (GQM), Toma de Decisiones, Calidad de los Datos.*

Abstract: In software development organizations, there are core technical processes focused on the construction of the software product, but there are also other business processes, equally relevant, that support this construction, such as organizational decision-making. In this context, the quality of business data has become increasingly relevant; however, there remains a gap regarding data understandability, due to the absence of both its conceptualization and formal mechanisms for measuring its quality. This limits its effective use for organizational data analysis and decision-making processes. This work presents the B-DataCom framework, which proposes metrics for measuring the understandability of business data in software development organizations. A systematic mapping of the literature was conducted to identify viable measurement elements regarding data quality in a business context, which grounded the Framework, also considering an extension of the Measurement Ontology for its integrated conceptualization of the semantics of understandability. B-DataCom defines descriptive aspects of data understandability, a conceptual model representing the theory evaluated in the research, and a set of 11 GQM (Goal-Question-Metric) metrics. After an instantiation exercise of these metrics in the MPS (Process Improvement) process referenced in COMPETISOFT, it is concluded that the proposed metrics can be scaled through future case studies.

Keywords: *Understandability, Goal-Question-Metric (GQM), Decision-Making, Data Quality.*

1. Introducción

En la industria del software, la calidad de los datos es fundamental para respaldar tanto procesos analíticos como la toma de decisiones empresariales [1]. En este contexto, al gestionar y analizar grandes volúmenes de información, la integración de la ciencia de datos, el Big Data y el proceso exploratorio de la analítica de datos (EDA) fortalece significativamente la calidad de la información [2], [3], [4]. Por otra parte, sobre la calidad de los datos, existen varias características que pueden considerarse para su medición, como las que ofrece el estándar ISO/IEC 25012 [5]; sin embargo, persiste un vacío investigativo relacionado con la conceptualización y medición de la comprensibilidad de los datos para respaldar estrategias competitivas en la toma de decisiones organizacionales frente a desafíos como la heterogeneidad de formatos y la información inconsistente o errónea [6], [7]. En consecuencia, surge la necesidad de disponer de mecanismos que ayuden a determinar el grado de comprensibilidad de los datos para fortalecer la calidad de la información en procesos organizacionales [8].

Este artículo deriva de la investigación de maestría titulada “*Framework* para la medición de la comprensibilidad de los datos de negocio en organizaciones desarrolladoras de software”. El *framework* B-DataCom, que integra la Ontología de la Medición y el modelo de medición *Goal-Question-Metric* (GQM), definiendo el conjunto inicial de once métricas teóricas orientadas a evaluar esta característica de calidad en pro de fortalecer procesos analíticos para una mejor toma de decisiones organizacionales. El documento se organiza de la siguiente manera: en la sección 2 se presenta la metodología del mapeo sistemático de la literatura [9]; la sección 3 expone los resultados y la discusión que sustentan la propuesta de investigación; la sección 4 presenta el modelo conceptual del *framework*, B-DataCom y las métricas definidas; la sección 5 presenta los resultados de una instanciación de las métricas sobre el proceso de negocio MPS referenciado en COMPETISOFT [10] que permiten demostrar la viabilidad de la propuesta; finalmente, en la sección 6 se presenta las conclusiones del avance investigativo.

2. Metodología

El desarrollo del mapeo sistemático sobre la literatura de la comprensibilidad de los datos en el contexto del desarrollo de software, se toma el protocolo propuesto por [11] y cuyos resultados iniciales se reportan en [9]. En la Tabla 1 se resume los principales elementos metodológicos considerados durante la ejecución del protocolo, de ella se obtiene el siguiente análisis: los estudios primarios cubren algunos aspectos de calidad de los datos; sin embargo, no incorporan técnicas explícitas para su medición. También, los estudios resaltan la importancia que tienen los datos en la analítica en ámbito organizacional proporcionando contribuciones en la mejora de la calidad de datos, lo que ayuda a la toma de decisiones organizacionales. A pesar de esto, aún se evidencia también la necesidad de encontrar características concretas que permitan describir la comprensibilidad de los datos en el contexto del desarrollo de software, ya que, los artículos integran características, tales como fácil, que puede ser asociada con la comprensibilidad, pero no como un atributo directo.

Tabla 1. Resumen de la metodología del mapeo sistemático.

Proceso	Hallazgos identificados en la literatura
Pregunta de investigación	¿Qué propuestas relacionadas con la comprensibilidad de los datos se han planteado para apoyar la toma de decisiones en el contexto de organizaciones desarrolladoras de software?
Cadena de búsqueda de información	(<i>Data Analytics</i>) OR (Understandability) AND (<i>Software Development</i>)
Base de datos	Scopus
Período de búsqueda	2016-2025
Proceso de selección	Se aplicaron criterios de inclusión y exclusión orientados a identificar estudios relacionados con la comprensibilidad de los datos, la calidad de los datos y la analítica de datos en el contexto del desarrollo de software
Resultado del proceso	263 artículos recuperados, 20 estudios relevantes y 12 estudios primarios seleccionados para el análisis.
Calidad del aporte de las propuestas primarias	Se hace una evaluación de aporte de calidad a las propuestas primarias en donde se aplicaron tres preguntas de profundización, y para una se evaluó el grado de aporte de calidad (GAC) por medio de la escala de <i>Likert</i> [12], donde <i>1 aporte Bajo</i> , <i>2 aporte Medio</i> y <i>3 aporte Alto</i> . De acuerdo con los resultados obtenidos, se decide continuar con el total de 12 trabajos de investigación ya que todos proporcionan información vital y pertinente para el mapeo sistemático.

Fuente: Elaboración propia.

3. Resultados y discusión del mapeo sistemático

En la organización de la información sobre la comprensibilidad de los datos, se clasifica en dos dimensiones de análisis.

3.1. Trabajos relevantes

El análisis de los estudios primarios permitió identificar tres núcleos temáticos: los estudios [8] y [13] abordan la *comprensibilidad y calidad de los datos*, analizando cómo estos ayudan a los modelos de calidad y generan utilidad a los servicios y productos. Estos estudios abordan indirectamente la comprensibilidad, pero no la presentan como un atributo medible, lo que justifica la necesidad de un *framework* para respaldar e impulsar la toma de decisiones empresariales. Por otra parte, los estudios [14], [15], [16], [17], [18] y [19] *abordan el uso e integración de metadatos*, analizando cómo este ayuda a la comprensión de archivos de datos públicos, flujos de trabajo computacionales, proyectos de software, búsqueda de servicios y repositorios, computación reproducible (RCS) y mejora de requisitos. Estos estudios identifican el rol de los metadatos para mejorar la comprensibilidad de los datos, pero no proporcionan métricas específicas para evaluar esta característica. Finalmente, los estudios [3], [4], [6], [20], [21], [22], [23], [24], [25], [26], [27] y [28] abordan la *analítica de datos, Big Data y toma de decisiones*, analizando cómo estos impactan y ayudan a las organizaciones en procesos de digitalización, estrategias de negocios de proyectos de desarrollo de software, modelos de construcción (BIM), optimización de análisis de datos en contextos empresariales, gemelos digitales y *frameworks*. Estos estudios analizan, en contextos organizacionales, el rol del análisis de datos, el Big Data y la toma de decisiones en el desarrollo de software, pero no consideran la comprensibilidad como un posible elemento importante dentro de estos mismos procesos, dejando a un lado la calidad semántica de los datos.

3.2. Dimensiones de análisis del mapeo sistemático

El análisis de los estudios primarios permitió organizar la información en dos dimensiones de análisis: Metodología y Comprensibilidad de los datos. En la dimensión de la metodología se identifican, en los estudios primarios, los componentes: tipos de contribución y tipos de validación. Por otra parte, en la dimensión de la

comprensibilidad de los datos se identificaron dos tópicos que permiten entender el estado propio de ésta en el contexto del desarrollo de software: (i) Elementos de Medición de Calidad del Dato: Técnicas, métricas y métodos de medición de calidad de datos; (ii) Características Descriptivas de la Comprensibilidad de los datos encontradas en la literatura que asocien o mencionen la comprensibilidad. A continuación, la Tabla 2 presenta el consolidado de ambas dimensiones.

Tabla 2. Consolidado de las dimensiones de análisis identificadas en los estudios primarios.

Dimensión	Componente/Tópico	Hallazgos identificados en la literatura
Metodológica	Tipo de contribución	Estudios, Enfoque, Propuesta metodológica y Herramienta.
	Tipo de validación	Teórica, Encuestas, Estudios de caso y Experimentos.
Comprensibilidad	Elementos de medición de calidad del dato	Escala de Likert (escala de percepciones), Número de problemas completamente especificados, Modelo de calidad Q-Rapids, KPI (<i>Key Performance Indicator</i>), MAGE-TAB <i>Proteomics</i> , <i>Reliable Matching of Multi-Cultural Names</i> (enfoque), <i>Standardization</i> (técnica), <i>Cleansing</i> (técnica), <i>Data Profiling Tools</i> (herramienta), <i>Matching</i> (técnica), <i>Survivorship Techniques</i> (técnica). Dimensión de calidad: Precisión, Exactitud, Completitud, Consistencia, Oportunidad, Conformidad, Unicidad, Legibilidad y Asimetría. Métricas: métrica de Distribución Única y Métrica de Dos Distribuciones.
	Características descriptivas de la comprensibilidad	Fácil, Claro, Evidente, Inteligible.

Fuente: Elaboración propia

A partir del análisis presentado en la Tabla 2 y de los resultados del mapeo sistemático, se identifican dos hallazgos principales. En primer lugar, los estudios primarios utilizan para la validación de sus propuestas, *Estudios de Caso* y contribuciones orientadas a *Herramientas*. Estos resultados reflejan soluciones aplicables sobre un contexto real; permitiendo de esta manera mejoras en investigaciones empíricas y bajo diferentes usos de herramientas. En segundo lugar, aunque la mayoría de los estudios primarios caracteriza la comprensibilidad con el concepto de *Fácil* y emplea *indicadores de calidad de datos*, especialmente indicadores clave de desempeño (KPI), ninguno establece una conceptualización precisa ni mediciones orientadas a la comprensibilidad de los datos. Con base a lo anterior, se establece la necesidad de contar con mecanismos que ayuden a evaluar la comprensibilidad para contextos organizacionales del desarrollo de software.

4. Framework B-DataCom

El *framework* B-DataCom (*Business Data Comprehensibility Framework*), está conformado por tres componentes: (i) *Modelo Conceptual*, (ii) *la Extensión de la Ontología de la Medición* y (iii) *el conjunto de Métricas definidas mediante el modelo Goal-Question-Metric (GQM)* [29]. El resultado de esta integración proporciona una base inicial para la medición estructurada de la comprensibilidad de los datos de negocio en procesos organizacionales.

4.1. Modelo conceptual del framework B-DataCom

A partir de los resultados del mapeo sistemático presentado en [9], se retoma el Modelo Conceptual del *framework* B-DataCom, como una solución preliminar al problema abordado en esta investigación, permitiendo establecer una medición estructurada sobre la comprensibilidad de los datos de negocio. Este primer componente se fundamenta e integra la *Ontología de la Medición* [30], centrándose en la *Semántica de la Comprensibilidad* (características descriptivas y los factores que la impactan), y el apoyo del *Modelo de Medición GQM*; delimitando de esta manera el alcance de la propuesta, como se muestra en la Figura 1.

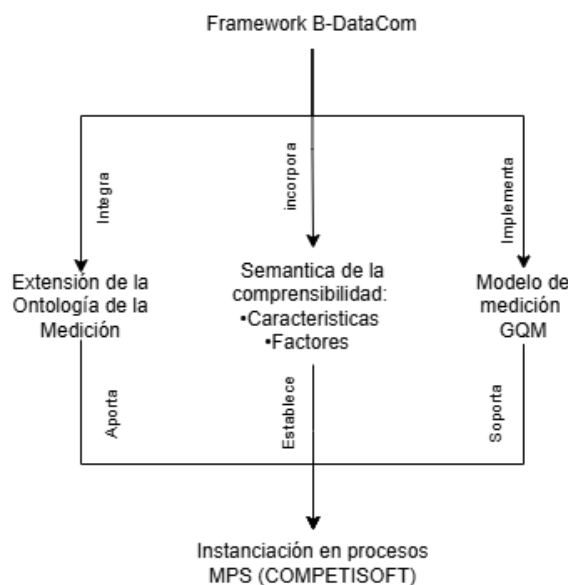


Figura 1. Modelo conceptual del *framework* B-DataCom. Fuente: adaptado de [9].

Dentro del modelo conceptual, la Semántica de la Comprensibilidad establece y define el significado de la comprensibilidad, sirviendo de esta manera como insumo para el modelo de medición GQM. Con el objetivo de definir atributos medibles y factores que impacten la comprensibilidad en el contexto organizacional del desarrollo del software, se establecieron los siguientes elementos:

- **Características descriptivas:** La comprensibilidad puede describirse a través de aspectos que permitan su entendimiento. Con base en esto, se decide evaluar los 15 sinónimos de la palabra *comprensibilidad* según la RAE [31]. Se descartaron los que no aportan algún elemento adicional en su semántica. En total quedaron siete aspectos seleccionados para la caracterización de la comprensibilidad de los datos. De los cuales, cuatro hacen parte de los que se encontraron en la literatura: *fácil, claro, evidente e inteligible* y, tres pertenecen a la RAE: *manifiesto, cognoscible y accesible*. Este conjunto de siete aspectos descriptivos será utilizado como *atributos medibles* de la comprensibilidad.
- **Factores:** La comprensibilidad puede verse impactada por factores que afecten su uso e interpretación. Con base en esto, se eligieron los que se identificaron en la literatura que presentan algún nivel de impacto en la comprensibilidad de los datos. En total, se obtuvieron cuatro y son los siguientes: *capacidad tecnológica, estandarización, gobernanza en el ciclo de vida del dato, interoperabilidad*.

4.2. Extensión de la Ontología de la Medición

El segundo componente del *framework* B-DataCom, está compuesta por cuatro sub-ontologías: Caracterización y Objetivos, Medidas Software, Formas de Medir y Acción de Medir, las cuales permiten responder a la necesidad de medición de la *comprensibilidad* de los datos en el contexto del desarrollo de software. De esta manera, se integra la Semántica de la Comprensibilidad, mediante la inclusión de las *características descriptivas* y los *factores* que impactan la comprensibilidad, extendiendo así la Ontología de la Medición hacia este dominio específico e incluyendo la teoría del Modelo de Medición GQM.

4.3. Modelo de Medición de GQM (Goal-Question-Metric)

El tercer componente del *framework* B-DataCom corresponde al conjunto de métricas definido mediante el modelo *Goal-Question-Metric* (GQM) que consiste en obtener indicadores en términos del grado de comprensibilidad de los datos, y se constituye por: el Objetivo de la Medición (*Goal*) que evalúa la comprensibilidad en el contexto de una organización desarrolladora de software; Preguntas de Medición (*Questions*) que permitan identificar datos directos e indirectos referentes a la comprensibilidad; Métricas (*Metrics*) que respondan a las preguntas de medición generando de esta manera indicadores en términos del grado de comprensibilidad de los datos. Con base en esto, a continuación, en la Tabla 3, se describe la definición de las preguntas (Q) y métricas (M) para los

atributos (que se agruparon de acuerdo con su definición), factores y consolidados de ambos para medir el grado de comprensibilidad.

Tabla 3. Preguntas y Métricas relacionados a los atributos/factores de la comprensibilidad.

ID del atributo/factor/consolidado	Pregunta (Q)	Métrica (M)
Atributo 1 (A1): Inteligible/Accesible/Claro	¿Cuál es el grado de accesibilidad que se tiene sobre los datos de negocio para el proceso de la organización?	Grado de accesibilidad (GA) = (NDV) número de datos de negocio a los cuales se les conoce sus posibles valores en el proceso / (NTP) número total de datos de negocio identificados en el proceso. $GA = \frac{NDV}{NTP} \quad (1)$
Atributo 2 (A2): Evidente/Manifiesto	¿Qué nivel de evidencia (presencia) explícita se tiene sobre los datos de negocio para el proceso de la organización?	Nivel de evidencia (NE) = (NDE) número de datos de negocio que se encuentran explícitamente documentados / (NTP) registrados en el proceso/ número total de datos de negocio identificados el proceso. $NE = \frac{NDE}{NTP} \quad (2)$
Atributo (A3): Fácil	¿Cuál es el porcentaje de facilidad para identificar el propósito de los datos de negocio para el proceso de la organización?	Porcentaje de facilidad (PF) = (NDP) número de datos de negocio cuyo propósito está descrito de forma explícita en el proceso / (NTP) número total de datos de negocio identificados en el proceso. $PF = \frac{NDP}{NTP} \quad (3)$
Atributo (A4): Cognoscible	¿Cuál es el grado de información que poseen los datos de negocio para el proceso de la organización?	Grado de información (GI) = (NIM) número de datos de negocio que poseen información descriptiva en términos de metadato / (NTP) número total de datos de negocio identificados en el proceso. $GI = \frac{NIM}{NTP} \quad (4)$
Factor 1 (F1): Capacidad tecnológica	¿Cuál es el grado de soporte tecnológico disponible para la gestión y comprensión de los datos de negocio dentro el proceso de la organización?	Grado de capacidad tecnológica (GCT) = (NHT) número de datos de negocio gestionados mediante herramientas tecnológicas, automatizadas o digitalizadas / (NTP) número total de datos de negocio identificados en el proceso. $GCT = \frac{NHT}{NTP} \quad (5)$
Factor 2 (F2): Estandarización del dato	¿Cuál es el grado de estandarización de los datos del proceso de la organización?	Grado de estandarización de los datos (GED) = (NPR) Número de datos propios del proceso replicados / (NTP) número total de datos propios de negocio identificados en el proceso. $GED = \frac{NPR}{NTP} \quad (6)$

Factor 3 (F3): Gobernanza de la información	¿Cuál es el grado de control y gestión definido sobre los datos de negocio dentro del proceso de la organización?	Grado de gobernanza en el ciclo de vida del dato (GGI) = (NDR) número de datos de negocio que poseen responsables definidos dentro del proceso/ (NTP) número total de datos de negocio identificados en el proceso. $GGI = \frac{NDR}{NTP} \quad (7)$
Factor 4 (F4): Interoperabilidad	¿Cuál es el grado de intercambio entre procesos de los datos de negocio?	Grado de interoperabilidad (GIO) = (NDI) número de datos de negocio que son intercambiados entre procesos organizacionales / (NTP) número total de datos de negocio identificados en el proceso. $GIO = \frac{NDI}{NTP} \quad (8)$
Consolidado de atributos (CA)	¿Cuál es el grado consolidado de atributos de la comprensibilidad de los datos dentro del proceso de la organización?	Grado consolidado de atributos (GCA)% = $(GA * F_{GA}) + (NE * F_{NE}) + (PF * F_{PE}) + (GI * F_{GI})$ (9) Los factores (F) tiene un valor equitativo del 25%
Consolidado de factores (CF)	¿Cuál es el grado consolidado de los factores que impactan la comprensibilidad de los datos dentro del proceso de la organización?	Grado consolidado de factores (GCF)% = $(GCT * F_{GCT}) + (GED * F_{GED}) + (GGI * F_{GGI}) + (GIO * F_{GIO})$ (10) Los cuatro factores (F) tiene un valor equitativo del 25% para dar un total del 100%
Consolidado de factores y atributos (CAF)	¿Cuál es el grado de comprensibilidad de los datos considerando tanto los atributos como los factores?	Grado general de comprensibilidad (GGC)% = $(GCA * F_{GCA}) + (GCF * F_{GCF})$ (11) Los dos factores (F) tiene un valor equitativo del 50% para dar un total del 100%

Fuente: Elaboración propia.

Las métricas propuestas se estructuran en cinco grupos. El primer grupo corresponde a las métricas orientadas a los *Atributos* (A), que se encargan de evaluar las propiedades semánticas que caracterizan la comprensibilidad de los datos descritas anteriormente, con el fin de determinar qué tan comprensibles son los datos dentro de un proceso organizacional. Estas métricas se agruparon en cuatro atributos, A1, A2, A3 y A4; donde cada uno se compone por una o varias características según las definiciones similares entre ellas. El segundo grupo corresponde a las métricas de los *Factores* (F), las cuales evalúan qué condiciones, del entorno organizacional, favorecen o limitan dicha comprensibilidad. El tercer y cuarto grupo corresponde al *Consolidados de los Atributos* (CA) y al *Consolidado de los Factores* (CF), que integra los resultados de los Atributos y Factores, permitiendo obtener tanto una valoración global de cada dimensión como una visión integrada de las características propias de los datos y el contexto en el que los factores son gestionados. Finalmente, el quinto grupo corresponde al grado *General de Comprensibilidad* (GGC), identificado como CAF en la Tabla 3, que integra ambos consolidados (CA y CF), constituyendo de esta manera el indicador global sobre la comprensibilidad de los datos de negocio; sirviendo como apoyo para la identificación de oportunidades de mejora en el contexto de la toma de decisiones en organizaciones desarrolladoras de software.

Dentro de la definición de un modelo de medición, como sistema, es importante determinar umbrales de interpretación de los resultados cuantitativos que dichas medidas arrojen. Para ello, el *framework* B-DataCom sugiere una escala de interpretación inicial genérica y homogénea, conformada por tres niveles de calidad: *Alto*, *Medio* y *Bajo*. Esta escala se aplica a cada una de las métricas propuestas para evaluar tanto los cuatro grupos de Atributos como los cuatro Factores identificados en el proceso evaluado. Cada uno de estos niveles podría establecerse de manera homogénea en un rango de 0 a 33 % para una Baja Calidad, 34 a 66 % para una Calidad Media y 67 a 100 % para una Alta Calidad, lo que permitiría interpretar el grado de comprensibilidad de los datos del proceso de negocio evaluado con respecto a cada uno de estos atributos. De igual manera, esta misma escala de umbrales podría aplicarse para la interpretación de los Factores que impactan la comprensibilidad, indicando qué tan preparada está la organización frente a la implementación de los mecanismos representados por el *framework*. Finalmente, con respecto a las métricas CA, CF y CAF, como métricas compuestas y finales para la determinación del grado de comprensibilidad de los datos en un proceso organizacional evaluado, los umbrales también indicarán un grado Alto, Medio o Bajo, de manera consistente con las demás métricas propuestas.

5. Evaluación inicial del Framework a través de instanciación

Para la evaluación inicial del *framework* B-DataCom, se realizó una instanciación sobre el Modelo de Referencia de Procesos de Mejora (MPS) de COMPETISOFT. Para ello, se seleccionó un solo subproceso representativo de cada una de sus tres categorías organizacionales: Alta Dirección (DIR), Gerencia (GER) y Operación (OPE), con sus procesos correspondientes. De estos subprocesos, se identificaron 200 datos de negocio que sirven como unidad de análisis para la aplicación de las once métricas propuestas descritas anteriormente. A partir de estos datos de negocio se calculó el valor de las once métricas, correspondientes a los atributos, factores y consolidados de la comprensibilidad. Los resultados obtenidos se presentan de manera resumida en la Tabla 4, con el propósito de evidenciar y validar el procedimiento de aplicación del *framework* B-DataCom en un contexto organizacional de desarrollo de software.

Tabla 4. Resultados de la instanciación del *framework* B-DataCom sobre el modelo de referencia de procesos COMPETISOFT (MPS).

Categoría	Proceso	Subproceso	GQM Atributos GCA (Factor del 25% para cada métrica)		GQM Factores GCF (Factor del 25% para cada métrica)		GQM Consolidado GGC (Factor del 50% C/U)
Alta Dirección (DIR)	Gestión de Negocio (GN)	Planificación estratégica	(1)	GA= 60%	(5)	GCT = 70%	(9) GCA=60%
			(2)	NE= 70%	(6)	GED = 60%	(10) GCF=60%
			(3)	PF= 60%	(7)	GGI = 50%	(11) GGC=60%
			(4)	GI= 50%	(8)	GIO = 60%	
Gerencia (GER)	Gestión de Procesos (GP)	Planificación de procesos	(1)	GA= 75%	(5)	GCT = 80%	(9) GCA=71.25%
			(2)	NE= 70%	(6)	GED = 70%	(10) GCF=71.25%
			(3)	PF= 75%	(7)	GGI = 65%	(11) GGC=71.25%
			(4)	GI= 65%	(8)	GIO = 70%	
Operación (OPE)	Administración de proyecto (AP)	Planificación de proyecto	(1)	GA= 75%	(5)	GCT = 82.1%	(9) GCA=72.3%
			(2)	NE= 71.4%	(6)	GED = 71.4%	(10) GCF=74.07%
			(3)	PF= 75%	(7)	GGI = 67.8%	(11) GGC=73.18%
			(4)	GI= 67.8%	(8)	GIO = 75%	

Fuente: Elaboración propia.

Los resultados obtenidos se interpretan de acuerdo con la escala propuesta por el *framework* B-DataCom: Baja Calidad (0 %–33 %), Calidad Media (34 %–66 %) y Alta Calidad (67 %–100 %). Bajo esta interpretación, Alta Dirección obtuvo un GGC del 60 %, correspondiente a una Calidad Media, lo que sugiere oportunidades de mejora en la definición, documentación y estandarización de los datos de negocio que respaldan la planificación estratégica. Por su parte, Gerencia (71,25 %) y Operación (73,18 %) alcanzaron una Alta Calidad, evidenciando una mayor organización y comprensibilidad de los datos utilizados para la gestión y ejecución de los procesos. En

conjunto, estos resultados muestran que el *framework* B-DataCom permite identificar cuantitativamente el grado de comprensibilidad de los datos de negocio y priorizar acciones de mejora sobre aquellos procesos que requieren fortalecer la calidad semántica de sus datos, contribuyendo así a una toma de decisiones organizacionales más consistente.

6. Conclusiones

A partir del análisis de la literatura, se logra evidenciar una ausencia de medición estructurada para la característica de la comprensibilidad de los datos, incluyendo el estudio de su semántica en contextos de organizaciones desarrolladoras de software. También, se comprueba que no existe una conceptualización y caracterización específica para esta característica, lo que limita el aprovechamiento de los datos de negocio en la operacionalización de los procesos organizacionales, dificultando su completa integración como activo de valor para la toma de decisiones. En otras palabras, se concluye que el no abordar este tipo de comprensibilidad en contextos organizacionales no permite reconocer, ni trabajar sobre el impacto de esta ausencia en procesos como la analítica de datos.

Esta investigación, por tanto, presenta el *framework* B-DataCom con el cual se busca aportar en la disminución de la brecha relacionada a la conceptualización y medición del aspecto de comprensibilidad de los datos de negocio como un elemento esencial para el uso efectivo de los mismos en las organizaciones, por medio de once métricas GQM que sirvan de base para establecer indicadores sobre dicho grado de comprensibilidad. El *framework* fue construido bajo los lineamientos de fundamentos teóricos como lo es el dominio de la Ontología de la medición, dentro de la cual se realizó una extensión cubierta en términos de características descriptivas de la comprensibilidad y factores que la impactan.

La evaluación inicial de la propuesta consiste en instanciar las métricas para el proceso MPS de manera que se pudiera identificar limitaciones de aplicación del conjunto de mediciones y de esta manera ajustar lo necesario antes de ejecutar un segundo ciclo de evaluación a través de métodos empíricos como lo son los estudios de caso. Como trabajos futuros se espera diseñar y lograr resultados de evaluación con respecto a la idoneidad del *framework* y de esta manera identificar mejoras en su definición, estructura y la semántica de las métricas en contextos reales de manera que se pueda fortalecer la aplicabilidad del *framework* basada en la complejidad empresarial.

7. Agradecimientos

Esta investigación está financiada por COLFUTURO por medio de la convocatoria “Formación de capital humano de alto nivel para promover el liderazgo social y comunitario y afrontar la Cuarta Revolución Industrial en el departamento del Cauca (4RI)” del 2024.

8. Referencias

- [1] Merino, J., Caballero, I., Rivas, B., Serrano, M., Piattini, M. (2016). A data quality in use model for Big Data. *Future Generation Computer Systems*, 63, 123–130. <https://doi.org/10.1016/j.future.2015.11.024>
- [2] Badshah, A., Daud, A., Alharbey, R., Banjar, A., Bukhari, A., Alshemaimri, B. (2024). Big data applications: Overview, challenges and future. *Artificial Intelligence Review*, 57 (11), 1-49. <https://doi.org/10.1007/s10462-024-10938-5>
- [3] Yeh, Y.-T., Eden, R., Fieft, E., Syed, R. (2025). The role of use for the business value of big data analytics. *The Journal of Strategic Information Systems*, 34 (2), 1-36. <https://doi.org/10.1016/j.jsis.2025.101888>
- [4] Willetts, M., Atkins, A. S. (2024). Evaluation of a software positioning tool to support SMEs in adoption of big data analytics. *Journal of Electronic Science and Technology*, 22 (1), 1-11. <https://doi.org/10.1016/j.jnlest.2023.100229>
- [5] International Organization for Standardization. (2008). *Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model* (ISO/IEC Standard No. 25012:2008). <https://www.iso.org/standard/35736.html>

- [6] López, L., Mazano, M., Gómez, C., Oriol, M., Franch, X., Martínez-Fernández, S., Vollmer, A. M. (2020). QaSD: A quality-aware strategic dashboard for supporting decision makers in Agile software development. *Science of Computer Programming*, 202, 1-10. <https://doi.org/10.1016/j.scico.2020.102568>
- [7] Putrama, I. M., Martinek, P. (2024). Heterogeneous data integration: Challenges and opportunities. *Data in Brief*, 56, 1-23. <https://doi.org/10.1016/j.dib.2024.110853>
- [8] Martínez-Fernández, S., Vollmer, A. M., Jedlitschka, A., Franch, X., López, L., Ram, P. (2019). Continuously assessing and improving software quality with software analytics tools: A case study. *IEEE Access*, 7, 68219–68239. <https://doi.org/10.1109/ACCESS.2019.2917403>
- [9] Molina España, S. A., Buitrón Ruiz, S. L., Pino Correa, F. J. (2026). Medición de la comprensibilidad de datos en analítica de negocio: El framework B-DataCom. *Ingeniería y Competitividad*, 28 (2), 1-19. https://revistaingenieria.univalle.edu.co/index.php/ingenieria_y_competitividad/article/view/15823
- [10] Grupo Alarcos. (s. f.). *COMPETISOFT*. Universidad de Castilla-La Mancha. <https://alarcos.esi.uclm.es/competisoft/Framework/>
- [11] Petersen, K., Feldt, R., Mujtaba, S., Mattsson, M. (2008). *Systematic mapping studies in software engineering*. 12th International Conference on Evaluation and Assessment in Software Engineering, Italy. <https://doi.org/10.14236/ewic/EASE2008.8>
- [12] Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22 (140), 1–55. https://legacy.voteview.com/pdf/Likert_1932.pdf
- [13] Lee, B., Ahmed-Kristensen, S. (2025). D3 framework: An evidence-based data-driven design framework for new product service development. *Computers in Industry*, 164, 1-14. <https://doi.org/10.1016/j.compind.2024.104206>
- [14] Dai, C., Füllgrabe, A., Pfeuffer, J., Solovyeva, E. M., Deng, J., Moreno, P.... Perez-Riverol, Y. (2021). A proteomics sample metadata representation for multiomics integration and big data analysis. *Nature Communications*, 12 (1), 5854. <https://doi.org/10.1038/s41467-021-26111-3>
- [15] Goble, C., Cohen-Boulakia, S., Soiland-Reyes, S., Garijo, D., Gil, Y., Crusoe, M. R., Peters, K., Schober, D., (2024). FAIR computational workflows. *Data Intelligence*, 2 (1–2), 108–121. https://doi.org/10.1162/dint_a_00033
- [16] Kim, Y., Kim, J., Jeon, H., Kim, Y., Song, H., Kim, B. (2021). Githru: Visual analytics for understanding software development history through Git metadata analysis. *IEEE Transactions on Visualization and Computer Graphics*, 27 (2), 656–666. <https://doi.org/10.1109/TVCG.2020.3030414>
- [17] Řezník, T., Raes, L., Stott, A., De Lathouwer, B., Perego, A., Charvát, K., Kafka, Š. (2022). Improving the documentation and findability of data services and repositories: A review of (meta)data management approaches. *Computers & Geosciences*, 169, 1-28. <https://doi.org/10.1016/j.cageo.2022.105194>
- [18] Leipzig, J., Nüst, D., Hoyt, C. T., Ram, K., Greenberg, J. (2021). The role of metadata in reproducible computational research. *Patterns*, 2 (9), 1-28. <https://doi.org/10.1016/j.patter.2021.100322>
- [19] Bandara, M., Rabhi, F. A., Bano, M. (2023). A knowledge-driven approach for designing data analytics platforms. *Requirements Engineering*, 28, 195–212. <https://doi.org/10.1007/s00766-022-00385-5>
- [20] Fannouch, A., Gharib, J., Gahi, Y. (2025). Enhancing DataOps practices through innovative collaborative models: A systematic review. *International Journal of Information Management Data Insights*, 5 (1), 1-18. <https://doi.org/10.1016/j.jjime.2025.100321>
- [21] Kozik, R., Choraś, M., Puchalski, D., Renk, R. (2019). Q-Rapids framework for advanced data analysis to improve rapid software development. *Journal of Ambient Intelligence and Humanized Computing*, 10, 1927–1936. <https://doi.org/10.1007/s12652-018-0784-5>
- [22] Kusumawardhani, F. K., Ratmono, D., Wibowo, S. T., Darsono, D., Widyatmoko, S., Rokhman, N. (2024). The impact of digitalization in accounting systems on information quality, cost reduction and decision making: Evidence from SMEs. *International Journal of Data and Network Science*, 8, 1111–1116. <https://doi.org/10.5267/j.ijdns.2023.11.023>
- [23] Liberatore, M. J., Pollack-Johnson, B., Clain, S. H. (2017). Analytics capabilities and the decision to invest in analytics. *Journal of Computer Information Systems*, 57 (4), 364–373. <https://doi.org/10.1080/08874417.2016.1232995>
- [24] Ahmad, M. O., Ahmad, I., Rana, N. P., Khan, I. S. (2023). An empirical investigation on business analytics in software and systems development projects. *Information Systems Frontiers*, 25, 917–927. <https://doi.org/10.1007/s10796-022-10253-w>

- [25]Sompolgrunk, A., Banihashemi, S., Golzad, H., Le Nguyen, K. (2024). Strategic alignment of BIM and big data through systematic analysis and model development. *Automation in Construction*, 168, 1-22. <https://doi.org/10.1016/j.autcon.2024.105801>
- [26]Ryzhakova, G., Malykhina, O., Pokolenko, V., Rubtsova, O., Homenko, O., Nesterenko, I., Honcharenko, T. (2022). Construction project management with Digital Twin Information System. *International Journal of Emerging Technology and Advanced Engineering*, 12 (10), 19–28. https://doi.org/10.46338/ijetae1022_03
- [27]Rique, T., Perkusich, M., Dantas, E., Albuquerque, D., Gorgônio, K., Almeida, H. (2023). On adopting software analytics for managerial decision-making: A practitioner’s perspective. *IEEE Access*, 11, 73145–73163. <https://doi.org/10.1109/ACCESS.2023.3294823>
- [28]Zhu, J., Zhuang, E., Fu, J., Baranowski, J., Ford, A., Shen, J. (2016). A framework-based approach to utility big data analytics. *IEEE Transactions on Power Systems*, 31 (3), 2455–2462. <https://doi.org/10.1109/TPWRS.2015.2462775>
- [29]Basili, V., Trendowicz, A., Kowalczyk, M., Heidrich, J., Seaman, C., Münch, J., Rombach, D. (2014). *Aligning Organizations Through Measurement: The GQM+Strategies Approach*. Springer. <https://doi.org/10.1007/978-3-319-05047-8>
- [30]Ferreira, M., García, F., Ruiz, F., Bertoa, M. F., Calero, C., Vallecillo, A., Piattini, M., Mora, B. (2006). *Medición del Software Ontología y Metamodelo* (Informe Técnico UCLM-TSI-001). Universidad de Castilla-La Mancha. <https://n9.cl/izkla>
- [31]Real Academia Española. (s. f.). Comprensible. En *Diccionario de la lengua española* (Edición del Tricentenario). <https://dle.rae.es/comprendible>