

Principios, lineamientos, dimensiones y atributos para la evaluación de calidad de Datos Abiertos de Gobierno.

Principles, guidelines, dimensions, and attributes for the quality evaluation of Open Government Data.

Adriana Milena Rangel-Carrillo¹, Gina Paola Maestre-Góngora², Mariutsi Alexandra Osorio-Sanabria³

^{1,2}Universidad Cooperativa de Colombia, Bogota - Colombia, ³Universidad Pontificia Bolivariana, Medellín - Colombia

ORCID: ¹[0000-0002-2206-748X](https://orcid.org/0000-0002-2206-748X), ²[0000-0002-2880-9245](https://orcid.org/0000-0002-2880-9245), ³[0000-0001-7375-3052](https://orcid.org/0000-0001-7375-3052)

Recibido: 19 de agosto de 2020.

Aprobado: 11 de noviembre de 2020.

Resumen— El propósito de esta investigación fue establecer una base conceptual para la evaluación de calidad de los datos abiertos de gobierno, al delimitar los principios, lineamientos, dimensiones y atributos de calidad requeridos para implementar una operación efectiva de las directrices dadas por el Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia. Se realizó una revisión de literatura, en la cual se definió el objetivo de la investigación, se realizó una consulta en diferentes fuentes, se prepararon los datos y se clasificó la literatura de acuerdo con los criterios de selección definidos y así alinear los componentes seleccionados en un esquema estratégico que prioriza la explotación de los datos como fuente de tendencias e innovaciones. Se identificaron tres principios esenciales: Aptitud para el Uso, Orientación a Procesos y Mejora Continua de los cuales se desprenden tres dimensiones para considerar el contexto, las propiedades y representación de valor de los datos, con 6 lineamientos y 12 atributos que aportan de manera gradual con el nivel de calidad requerido según los documentos técnicos de referente para el contexto colombiano. En conclusión, el modelo propuesto permite desarrollar una evaluación de calidad de los datos abiertos de gobierno para habilitar, a través de su publicación y reutilización, los beneficios de la conversión de la materia prima de los datos en productos que se materializan en información relevante y coherente para el sector público, privado y la ciudadanía en general.

Palabras Claves: Datos Abiertos de Gobierno, Atributos de calidad de datos, Dimensiones de Calidad de datos, Lineamientos de Calidad de datos, Principios de Calidad de datos.

Abstract— The purpose of this research was to establish a conceptual basis for the quality assessment of open government data, by delimiting the principles, guidelines, dimensions and quality attributes required to implement an effective operation of the guidelines given by the Ministry of Technologies of Information and Communications of Colombia. A literature review was carried out, in which the objective of the research was defined, a query was made in different sources, the data was prepared and the literature was classified according to the defined selection criteria and thus align the selected components in a strategic scheme that prioritizes the exploitation of data as a source of trends and innovations. Three essential principles were identified: Fitness for Use, Process Orientation and Continuous Improvement, from which three dimensions emerge to consider the context, properties and value representation of the data, with 6 guidelines and 12 attributes that gradually contribute with the level of quality required according to the technical reference documents for the Colombian context. In conclusion, the proposed model makes it possible to develop a quality assessment of open government data to enable, through its publication and reuse, the benefits of converting the raw material of the data into products that are materialized in relevant information and coherent for the public and private sectors and the general public.

Keywords: Open Government Data, Data quality attributes, Data quality dimensions, Data quality guidelines, Data quality principles.

*Autor para correspondencia.

Correo electrónico: jpalacios@udistrital.edu.co (Jose Ignacio Palacios Osma).

La revisión por pares es responsabilidad de la Universidad de Santander.

Este es un artículo bajo la licencia CC BY (<https://creativecommons.org/licenses/by/4.0/>).

Forma de citar: A. M. Rangel-Carrillo, G. P. Maestre-Góngora y M. A. Osorio-Sanabria, "Principios, lineamientos, dimensiones y atributos para la evaluación de calidad de Datos Abiertos de Gobierno", Aibi revista de investigación, administración e ingeniería, vol. 8, no. S1, pp. 54-65, 2020, doi: [10.15649/2346030X.950](https://doi.org/10.15649/2346030X.950)

I. INTRODUCCIÓN

Las instituciones de carácter público producen y gestionan una gran cantidad de datos, en diversos formatos, a gran velocidad y variedad, que son publicados para facilitar el acceso a la ciudadanía en general y su reutilización en diferentes contextos; dichos datos se denominan datos abiertos de gobierno (OGD, por sus siglas en inglés) [1]. En los últimos años han tomado fuerza varias iniciativas de OGD, teniendo como principal objetivo la transparencia y la generación de valor público a través de estos.

El valor de los OGD se enmarca en su publicación y posterior reutilización, lo cual demanda contar con un nivel de calidad adecuado, que genere confianza en los consumidores de los datos. En la actualidad un alto porcentaje de datos puede ser imprecisos, errados o incompletos. El concepto corresponde a “Bad data” [2], con el cual se exponen las falencias en la calidad de los datos que dificulta el proceso de análisis, al ser considerados no aptos para su uso. Algunos efectos de este problema se ven reflejados en: Toma de decisiones no acertadas, dificultad para aprovechar su valor predictivo, interferencia en el proceso de obtención de nuevos conocimientos [3], el 75% del presupuesto destinado para el Análisis de los datos se utiliza en la adaptación de los datos, se requiere utilizar el 60% del tiempo para evaluar los datos, limpiar y organizar [4].

Dichos problemas de calidad en los datos deterioran su idoneidad y adecuación a su propósito, por lo tanto, el proceso de evaluación de calidad debe ser iterativo e incremental. Sin embargo, a pesar de la evidente crisis relacionada a la calidad de los datos abiertos de gobierno, se evidencia una renuencia a abordarla por parte de los involucrados, debido al costo que dicho proceso representa.

Debido a lo anterior, se requiere delimitar los principios, lineamientos, dimensiones y atributos para la estandarización y aseguramiento de calidad de los OGD, alineando este esquema con los documentos técnicos establecidos para el contexto colombiano desarrolladas por Ministerio de Tecnologías de la Información y Comunicación de Colombia (MinTIC), teniendo como propósito facultar los datos para su uso con un enfoque en los consumidores de los datos [5].

Este trabajo se centra en establecer las bases conceptuales requeridas para evaluar la calidad de los OGD, y así responder a: ¿Cuáles son los principios de calidad necesarios para OGD?, ¿Cuáles lineamientos permiten alcanzar los principios de calidad seleccionados, aplicables a OGD?, ¿Qué dimensiones y atributos de los datos son requeridas para la evaluación de calidad de OGD?, a través de los cuales se puedan aplicar las guías y manuales establecidas por MinTIC. Posteriormente se presentan los resultados, análisis e interpretación y las conclusiones donde se resalta el proceso de calidad como un cimiento fundamental para estructurar los datos, promover su uso y fortalecer la generación de valor, con el fin de ofrecer una nueva visión de diferentes fenómenos de la realidad [6].

II. MARCO TEÓRICO

Las oportunidades que se han generado con la era de la información y el internet ha desencadenado que se aumente de manera exponencial la cantidad de datos, la diversidad y estructura. Lo anterior ha incentivado que en el sector público y privado los datos sean considerados activos de información que pueden ser aprovechados por tecnologías avanzadas como la inteligencia artificial, el internet de las cosas, la robótica, la nanotecnología, la biotecnología y la computación en la nube, entre otras.

a. *Los datos*

Los datos son la materia prima requerida para el desarrollo de diferentes actividades, su uso y aplicación se ha diversificado en varias áreas y ciencias, para llegar a ser de beneficio en la vida cotidiana de la humanidad. Se pueden observar dichos beneficios en pronósticos del clima, flujo vehicular, hasta llegar a resolver desafíos más importantes tales como [7]:

- El Metro de Quito: A través de los datos se establecieron puntos estratégicos de origen y destino del metro.
- Factura Electrónica en Colombia: La formalización de los datos requeridos y actualizados ha permitido mitigar el gran problema de evasión de impuestos.
- El teleférico como transporte masivo en Bolivia: Se evidenció la seguridad que provee el teleférico, como un hito en la innovación para el transporte masivo, gracias a los datos recolectados de los diferentes medios de transporte masivo sobre accidentes de tránsito en las calles y vías de todo el país, desde el año 2014 hasta el año 2018.
- Inteligencia Artificial para mejorar la productividad en granjas de trucha en Perú: A través de los datos tomados de cámaras, sensores y satélites, se aplicó Inteligencia Artificial para determinar la cantidad y momento óptimo de alimentar a las truchas y así maximizar la productividad en dichas granjas.
- Blockchain para la titulación de tierras en Bolivia: Los datos son almacenados en copias idénticas, por diferentes actores del proceso, para descentralizar su custodia y evitar la corrupción de estos.
- La Metodología BIM (Building Information Modeling): A través de los datos se imita el proceso real de construcción y se pronostican diferentes etapas de una obra, además de simular cambios sobre el uso de una obra específica y los efectos de dichos cambios.

b. *Datos Abiertos de Gobierno*

Entre la clasificación de los datos, se encuentran aquellos que son considerados de carácter público, los cuales pueden ser usados y reutilizados por cualquier persona de manera libre; estos han sido denominados OGD y se estiman como factores claves que proveen una plataforma de transformación global para que los gobiernos [8], empresas privadas y públicas, y los ciudadanos en general, tengan bases claras e información real para tomar decisiones que permitan el desarrollo de sus áreas de desempeño laboral, profesional e intelectual [9].

Los OGD tienen como base un conjunto de principios que soportan el uso, acceso y la reutilización de dichos datos, tales como: Deben ser procesables por máquinas, Accesibles, Completos, deben obtenerse de la fuente de origen, no propietarios, oportunos, actualizados, con

licenciamiento abierto y deben estar disponibles para cualquier persona sin discriminación [10], [11]. Estos principios trazan la ruta para que los datos puedan generar valor de las siguientes maneras [9]:

- Promueve la transparencia y el control social gracias a la visibilidad de la gestión pública.
- Ayuda a que las personas puedan asumir un rol más activo en la sociedad al tomar sus decisiones individuales y comunitarias con información real.
- Facilita la innovación de productos y la creación de nuevos modelos de negocio.
- Hace énfasis en el mejoramiento de la gestión gubernamental y la eficiencia en la prestación de sus servicios.
- Permite la planeación y prospectiva de escenarios futuros para la sociedad causados por decisiones, programas y políticas definidas.
- Fortalece la explotación del potencial de uso de los datos al visualizar tendencias y nuevas perspectivas.

c. *Calidad de datos*

El concepto de calidad ha evolucionado, desde la Inspección aplicada a Productos (1.450 A.C.), el control de calidad para los procesos (1.700 y 1.800 D.C.), el aseguramiento de la calidad en los sistemas (1.950 D.C.), la calidad total (1.990 D.C.), hasta la Gestión de la Calidad Total o Excelencia (2.000 D.C.) [12]; donde la calidad trasciende de ser considerada una prioridad competitiva a convertirse en un requisito imprescindible para que las organizaciones y la sociedad puedan aprovechar los datos que se producen a nivel interno y externo en la generación de valor.

De acuerdo con lo anterior, la calidad de los datos enfrenta múltiples retos, entre ellos se pueden encontrar:

- Retos relacionados al volumen y la velocidad: [13].
- Retos relacionados a la variedad: [14].
- Retos relacionados a la veracidad y valor: [14].
- Alta tasa de reutilización: [15].
- supervisar el equilibrio [15].
- “Rejuvenecimiento” de los datos: [15].

Las características expuestas aportan a que un alto porcentaje de datos pueden ser imprecisos, errados o incompletos [16]. El concepto corresponde a “Bad data” [2], con el cual se exponen las falencias en la calidad de los datos que dificulta el proceso de análisis de estos, al ser considerados no aptos para su uso.

Para garantizar la calidad de los datos, se toman como base los siguientes principios [14]:

1. Los datos representan una unidad de valor y afectación en costos para los clientes interesados, quienes ven a dichos datos como un producto.
2. Al igual que un producto, los datos cuentan con un nivel de calidad, el cual es el resultado del proceso de generación y procesamiento aplicado.
3. Para determinar la calidad de los datos, influyen varios factores, tales como el propósito para el cual fueron generados, su uso potencial, los usuarios interesados, el tiempo de oportunidad y vigencia de dichos datos.

Para el cumplimiento de los anteriores principios, se han desarrollado estándares que involucran la gestión y aseguramiento de la calidad de los datos, tal como la norma ISO/IEC 25012 [17], mediante la cual se identifican y definen los criterios y requisitos de calidad de los datos y se evalúa su conformidad, además del cumplimiento de reglas de negocio y legislación aplicable; todo esto con el objetivo de responder a las necesidades de obtener datos aptos para el uso en las siguientes circunstancias:

- Organizaciones que no cuentan con procesos de calidad de datos o dicho proceso es débil.
- Datos incompletos que generan información insuficiente, con afectación negativa en resultados y satisfacción de los clientes.
- Falta de coherencia e integración de los datos.
- Alineación y uso de datos antiguos con nuevos.
- Datos cambiantes en periodos cortos de tiempo.

El modelo propuesto por la norma ISO/IEC 25012 [17] define dos puntos de vista de la calidad de los datos:

- La calidad inherente, que destaca el potencial intrínseco que tienen los datos para satisfacer las necesidades implícitas y explícitas de su propósito de creación, el cual determina el valor de dichos datos cuando sus características o dimensiones de calidad cumplen las restricciones definidas.
- La calidad dependiente del sistema, que corresponde a la dependencia que tienen los datos del dominio tecnológico en donde se usan.

En la **Tabla 1** se presenta con una “X” la clasificación de las características de calidad para los puntos de vista definidos por [17]: inherente y dependiente del sistema.

Tabla 1: Modelo de Calidad de Datos ISO/IEC 25012.

Características	Puntos de vista de la Calidad de Datos	
	Inherente	Dependiente del sistema
Exactitud	X	
Complejidad	X	
Consistencia	X	
Credibilidad	X	
Actualidad	X	
Accesibilidad	X	X

Características	Puntos de vista de la Calidad de Datos	
	Inherente	Dependiente del sistema
Conformidad	X	X
Confidencialidad	X	X
Eficiencia	X	X
Precisión	X	X
Trazabilidad	X	X
Comprensibilidad	X	X
Disponibilidad		X
Portabilidad		X
Recuperabilidad		X

Fuente: Elaboración propia.

d. Calidad de Datos para Big Data

Las características o dimensiones de calidad descritas con anterioridad contemplan factores adicionales de calidad de Big Data, relacionados con sus criterios de procesamiento, denominados las “V”: Volumen, Velocidad, Variedad, Valor, Variabilidad, Veracidad y Visibilidad [18], los cuales aportan características especiales, tales como el tamaño y la estructura que generalmente exceden las capacidades de adquisición, almacenamiento y tratamiento de los datos de manera oportuna, a través de métodos y herramientas tradicionales; además sobrepasan la capacidad de percepción de los seres humanos [19].

En Big Data se encuentran diferentes tipos de datos, tal como los datos estructurados que tienen longitud, formato y tamaño definido; los datos semiestructurados, que hace referencia a aquellos que no presentan una estructura definida de forma clara, pero cuentan con una organización en sus metadatos y, por último, los datos no estructurados, que se caracterizan por no tener un formato específico, lo que hace muy difícil su gestión de manera tradicional. De manera regular en Big Data el 80% de los datos son no estructurados y el 20% restante se divide en estructurados y semiestructurados [13]. UNECE en [20] identifica un componente adicional en la calidad de Big Data, el cual define “Hiperdimensiones” de calidad, las cuales corresponden a una agrupación de dimensiones de calidad de los datos como se presenta en la **Tabla 2**.

Tabla 2: Marco de Trabajo sugerido para la calidad de Big Data.

Hiperdimensión	Dimensión	Factores para considerar
Fuente: Factores asociados a los tipos de datos, características de su origen y gobernabilidad.	Ambiente institucional o de negocio	Sostenibilidad del proveedor de datos de la entidad. Fiabilidad Transparencia e interpretabilidad
	Privacidad y seguridad	Legislación Almacenamiento de datos vs. Proveedor de los datos. Restricciones Percepción
Metadatos: Factores asociados a los tipos de datos, características de su origen y gobernabilidad.	Complejidad	Limitaciones técnicas Tipo de dato: Estructurado o no estructurado Legibilidad Jerarquías y Anidamiento
	Complejidad	Si los metadatos están disponibles, son interpretables y están completos.
	Usabilidad	Recursos necesarios para importar y analizar Análisis de riesgos
	Factores relacionados con el tiempo	Oportunidad o puntualidad. Periodicidad Variación
	Enlaces	Presencia y calidad de las variables de enlace. Niveles de enlaces.
	Coherencia / Consistencia	Estandarización Variables claves disponibles (variables de clasificación, construcción de métricas).
	Validez	Transparencia de los métodos y procesos. Solidez de los métodos y procesos.
Datos: Factores relacionados con la calidad de los datos en sí	Precisión y Selectividad	Evaluación de errores en los datos. Conjunto de datos de referencia. Selectividad.
	Enlaces	Calidad de las variables de enlace.
	Coherencia / Consistencia	Coherencia entre la descripción del metadato y los valores de los datos.
	Validez	Coherencia entre los procesos y métodos, y los valores de los datos.

Fuente: Elaboración propia.

El aseguramiento de las dimensiones de calidad es un factor fundamental para aprovechar al máximo el valor de Big Data, pues se requiere contar con información confiable que permita tomar decisiones oportunas, a través de procesos eficientes que faciliten la conversión de los grandes volúmenes de datos de manera ágil. Dentro de los procesos principales de Big Data se encuentran la gestión de datos y el análisis de datos. El primero implica una serie de pasos e involucra tecnologías para adquirir, almacenar, preparar y recuperar los datos, como insumo para realizar el análisis de datos, que consiste en aplicar técnicas para analizar y adquirir inteligencia a partir del Big Data, por lo cual se ubica el análisis de datos en el proceso general de la “percepción interna” de Big Data [21], la regla es que cuanto mayor es la muestra de datos, más precisas serán las estadísticas y demás técnicas de análisis [21].

e. Calidad de datos para OGD

La calidad de los OGD se aborda desde un concepto interdisciplinario, el cual se fundamenta en la aptitud para el uso [22]. El uso a su vez se diversifica en varias posibilidades y contextos, bajo una perspectiva objetiva, por medio de atributos acordados, y una perspectiva subjetiva que está ligada a los requisitos y experiencia de los consumidores de los datos [23].

La idoneidad de los OGD para el propósito de la reutilización es un pilar de la calidad de los datos [24], [25]; sin embargo, la calidad puede variar ampliamente generando inconsistencias en los términos utilizados en los conjuntos de datos y falta de granularidad [26], ocasionando de esta manera, una barrera en el uso de los datos y los portales donde se publican [27].

f. Calidad de OGD en Colombia

A través de la guía para el uso y aprovechamiento de los datos abiertos propuesta por el gobierno colombiano y actualizada en septiembre de 2019 [28], se definen procesos para la gestión de los datos abiertos, los cuales hacen énfasis en el monitoreo de la calidad por medio de mediciones de calidad, uso e impacto.

Por otra parte, el potencial de los OGD es amplio y genera grandes aportes para la ciudadanía en general. Para diciembre de 2019 MinTIC, actualizó los requisitos de calidad de OGD para obtener el sello de calidad en las instituciones públicas, alineados con el ciclo de vida de los datos, a través del cual se establecen cinco etapas necesarias para el aprovechamiento del valor que llevan en sí los datos abiertos: i) preparación, ii) acceso, iii) uso, iv) evaluación y v) mejora, [29]. A través de cada etapa se verifican seis requisitos de calidad que los OGD deben cumplir para obtener el “Sello de Excelencia”. Los requisitos definidos son [29]:

1. Identificar el conjunto de datos a publicar
2. Priorizar los datos a publicar
3. Asegurar la completitud de los datos
4. Asegurar la disponibilidad de los datos
5. Garantizar la actualidad de los datos
6. Documentar los datos a través de metadatos

Para garantizar la calidad de los OGD en Colombia, el MinTIC ha establecido como estándar 16 criterios de calidad que representan los pilares de verificación y cumplimiento. Estos requisitos se describen en la **Tabla 3**[30].

Tabla 3: Estándar de Criterios de Calidad en Colombia.

Criterio	Descripción
Confidencialidad	Los datos solo deben ser accedidos por personas autorizadas para proteger la información reservada y clasificada.
Relevancia	Los datos publicados deben ser de utilidad para los usuarios. En este concepto los datos tienen más o menos relevancia de acuerdo con el caso de análisis.
Actualidad	Vigencia y actualidad de los datos publicados.
Trazabilidad	Histórico del conjunto de datos disponible: fechas de creación, publicación y actualizaciones.
Conformidad	Cumplimiento de lineamientos y estándares vigentes como DCAT (Data Catalog), para la descripción de la Metada.
Exactitud	Datos diligenciados de manera correcta.
Completitud	Información completa, datos completos para todas las columnas.
Consistencia	Datos coherentes y libres de contradicción.
Precisión	Nivel de desagregación en que están publicados los datos con respecto al nivel de desagregación en que fueron generados.
Portabilidad	Formatos sin restricciones para la reutilización de los datos.
Credibilidad	Información veraz y confiable para los usuarios.
Comprensibilidad	Características que permiten al usuario leer e interpretar los datos.
Accesibilidad	Herramientas tecnológicas que garanticen el acceso al usuario que lo requiera.

Fuente: Elaboración propia.

III. METODOLOGÍA O PROCEDIMIENTOS

La metodología utilizada en este estudio corresponde a una revisión de literatura basada en [31], la cual sigue cuatro pasos tal como se muestra en la Figura 1.

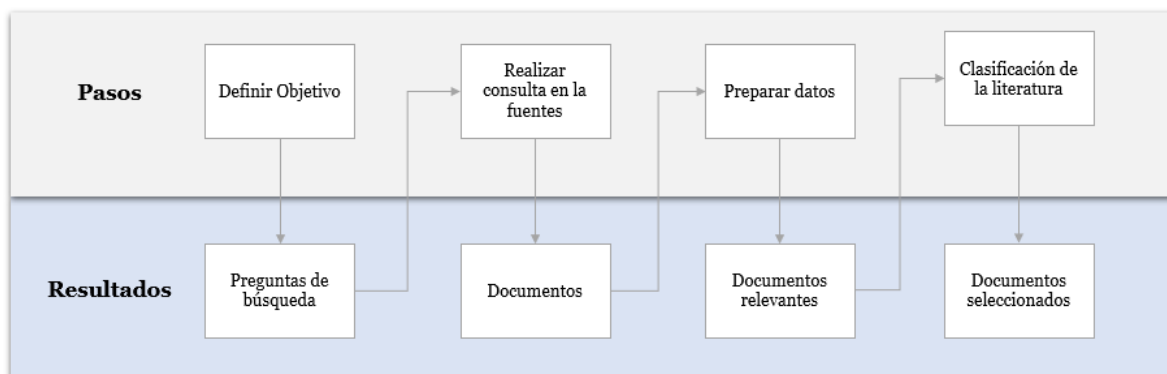


Figura 1: Metodología.

Fuente: Elaboración propia.

a. *Definir Objetivo*

Se propuso resolver las siguientes preguntas con relación al objeto de estudio:

- P1. ¿Cuáles son los principios de calidad necesarios para OGD?
- P2. ¿Cuáles lineamientos permiten alcanzar los principios de calidad seleccionados, aplicables a OGD?
- P3. ¿Qué dimensiones y atributos de los datos son requeridas para la evaluación de calidad de OGD?

b. *Consulta de fuentes*

Con el fin de realizar una síntesis que integre diferentes investigaciones, de los cuales se pueda obtener un panorama del estado actual del objeto de estudio, se realizó la búsqueda en las siguientes fuentes: Biblioteca digital IEEE Xplore y Scopus haciendo uso de las palabras clave: Open Government Data Quality, Big Data Quality, Open Government Data Quality Methodology y Big Data Quality Methodology.

c. *Preparar datos*

Con base en las palabras claves nombradas anteriormente se aplicó una búsqueda sobre los metadatos de “Resumen” y “Título” en las fuentes seleccionadas. Se prepararon los documentos excluyendo duplicados y revisando el título y resumen para identificar los documentos que realizan un aporte determinante en la investigación y así evitar referencias irrelevantes. Posteriormente, se revisó el tipo de documento (artículos y documentos técnicos) y el rango de fecha de publicación: 2010 a 2020. Finalmente, se seleccionaron 48 documentos de la siguiente manera:

- Todos los documentos: Aplicar términos de búsqueda: 724 resultados.
- Documentos relevantes: Preparar los documentos al eliminar duplicados, revisar título y resumen: 62 resultados.
- Documentos seleccionados: Verificar tipo de documento, pertinencia y año: 48 resultados.

d. *Clasificación de la literatura*

Los 48 documentos relevantes seleccionados se clasificaron teniendo en cuenta los siguientes criterios:

- Enfoque en OGD.
- Preservar el propósito principal de cualquier tipo de dato.
- Cubrir aspectos intrínsecos de los datos y su potencial.
- Abordar las características más relevantes que deben ser preservadas y adecuadas en los OGD para su evaluación de calidad.

IV. RESULTADOS ANÁLISIS E INTERPRETACIÓN

En [32] se expone que los datos abiertos son un subdominio de Big Data, y que los problemas de calidad de los datos se agudizan en Big Data; motivo por el cual se vuelve fundamental enfocar los esfuerzos en evaluar y garantizar la calidad de este subdominio representativo. Es por esta razón, que acoger los grandes datos como un recurso renovable y asegurar su calidad, converge en beneficios de impacto económico y social, donde estos datos requieren un foco de atención y procesamiento para asegurar la capacidad de diversificar y ampliar su significado, según el contexto de estudio, en este caso el gobierno, con el fin de obtener resultados más deliberados, reutilizables y generar respuestas que no podían preverse al momento de su creación [3].

A continuación, se presentan los principios, lineamientos, dimensiones y atributos identificados en la revisión de literatura, los cuales se organizan de manera sistemática y así, establecer un modelo conceptual en árbol para la evaluación de calidad de OGD en el contexto colombiano. Dicho modelo aporta una guía a los productores de OGD (entidades del gobierno colombiano) para asegurar el nivel de calidad de los datos publicados y se conforma como una plataforma que facilita la aplicación de los documentos técnicos de MinTIC fortaleciendo el énfasis que estos documentos tienen en la promoción del uso y aprovechamiento de los OGD para generar valor público [28].

a. *Principios para la evaluación de calidad de OGD*

Se identifican tres principios fundamentales para la evaluación de calidad de OGD:

- **Aptitud para el uso:** Asegurar la capacidad del dato para servir a un propósito definido [5], [6], [14], [20], [23], [24], [28], [33]–[37].
- **Orientado a procesos:** Gestionar la calidad de los datos, a través de procesos que permitan implementar la estrategia de calidad de forma efectiva, con actividades interrelacionadas para una ejecución sistemática [14], [23], [33], [36]–[38].
- **Mejora continua:** Garantizar la calidad de los datos a través de la planeación, ejecución de acciones concretas y verificación periódica [14], [20], [23], [24], [33], [35]–[38].

La calidad es un concepto interdisciplinario que se construye con diferentes perspectivas y mediciones, dentro de ella se encuentra la calidad de los datos con un enfoque multidimensional centrado en el aseguramiento de la aptitud para su uso [23]. De acuerdo con dicho concepto, los movimientos e iniciativas gubernamentales de datos abiertos promueven su publicación y uso a través de estándares de calidad [5] para obtener valor [22] y garantizar un impacto positivo en los consumidores potenciales [6].

En concordancia, es requerido un proceso definido y sistemático que aporte en la producción de OGD con alta calidad [39], a través de herramientas que faciliten su transformación en información comprensible [40] y ayude a gestionar adecuadamente los riesgos relacionados con su publicación [38]. Dicho proceso requiere:

- Una interacción constante entre los productores y los consumidores de los OGD [9].
- Una estrategia clara para aplicar antes y después de la publicación [9].
- Acciones concretas para realizar seguimiento e identificar el impacto y valor generado [9].
- Buena implementación y monitoreo para generar usos de impacto [28].

El rol del gobierno con iniciativas de OGD en progreso, no sólo reside en publicar los datos para la ciudadanía, sino que está compuesto por otras actividades tales como aumentar la variedad de los datos y mejorar continuamente su calidad [41]. Para validar el conjunto de OGD se requiere un mecanismo de mejora continua antes y después de realizar la publicación [42], con el fin de transformar los datos en información y conocimiento que represente un recurso potencial para la innovación [38].

La creación de soluciones innovadoras, a través de la apertura y el uso de los datos, habilita la generación de valor público y requiere el mejoramiento de la calidad de los OGD para que las empresas pequeñas y medianas puedan utilizar dichos datos, crear nuevos productos, fortalecer los negocios y conocer su mercado [40].

b. Lineamientos para la evaluación de calidad de OGD

Los OGD se pueden considerar un producto que genera valor, estimula la competitividad, el crecimiento económico y son reutilizables [22]. Dentro de los objetivos de las iniciativas de datos abiertos en el mundo, la reutilización [23] es un motivador para maximizar la generación de valor público, que se habilita al publicar datos con calidad y garantizar su transformación en productos, servicios o nuevas aplicaciones [5].

La publicación con calidad de los OGD facilita la explotación de los datos por parte de sus consumidores, proceso que genera impacto a través del uso [22], tanto a nivel individual, como colectivo y organizacional [6]. Estos consumidores juegan un papel muy importante en la reutilización de los datos y en la valoración del impacto generado por los mismos. No tiene sentido tener muchos datos disponibles si no se exploran, es decir, se descubren y se usan, además de poder analizar y descubrir el valor inmerso en ellos, a través de su explotación, labor que recae sobre los consumidores [22]. De esta manera, se identifican tres lineamientos para el principio de “Aptitud para el uso”:

- **Reutilización:** Habilitar el uso de los datos en varios contextos.
- **Impacto:** Generar valor para los consumidores, según el uso de los datos en un contexto específico.
- **Orientado al consumidor:** Facilitar al consumidor del dato el conocimiento contextual y acceso a los datos.

Los gobiernos con iniciativas activas de OGD, deben garantizar la confiabilidad de los datos, promover la colaboración para generar valor público [40], [43], asegurar su disponibilidad y uso [30], lo cual sólo se logra manteniendo un nivel adecuado de calidad. Este proceso de adecuación tiene como objetivo preparar los datos para su procesamiento y reutilización [9].

Uno de los desafíos a los cuales se enfrentan los productores de datos con la publicación, es mantener los conjuntos de datos adecuados y actualizados [40] para facilitar su interpretación en el contexto correspondiente [39] y permitir una comprensión adecuada de los mismos [38].

Bajo el enfoque anterior se identifican dos lineamientos para el principio de “Orientado a Procesos”:

- **Aseguramiento:** Realizar actividades planificadas y sistemáticas con el fin de alcanzar requisitos de calidad para las propiedades inherentes de los datos.
- **Adecuación:** Acondicionar los datos para cumplir con los requisitos de calidad relacionados a sus propiedades inherentes.

Un factor fundamental en los OGD es que el uso de los datos está ligado a su calidad [28]. De acuerdo con la capacidad de asegurar la calidad de los datos, es posible obtener valor a partir de su reutilización [38], generando a su vez:

- Valor económico [3], [32], [39], [44]–[46].
- Valor comercial [35]–[37], [47].
- Valor social [3], [14], [44], [45].
- Valor público [38], [40], [44]–[46], [48].

Se identifica el lineamiento de valor para el principio de “Mejora continua”, a través del cual se busca garantizar la representación de los datos como una unidad de valor para sus consumidores.

c. Dimensiones para la evaluación de calidad de OGD

La evaluación de calidad de los OGD requiere del componente contextual del dato [37] para habilitar su comprensión y hacerlo entendible para los consumidores [33]. También es importante asegurar las cualidades o atributos intrínsecos de los datos, con los cuales se logra habilitar el uso [20] y lograr percepciones claras desde su valor inherente [3].

La calidad inherente, que destaca el potencial intrínseco que tienen los datos para satisfacer las necesidades implícitas y explícitas de su propósito de creación, el cual determina su valor cuando sus características o dimensiones de calidad cumplen las restricciones definidas [17].

A su vez, los datos en su componente figurativo son representaciones del impacto y valor generado [28] al facilitar la extracción de nuevas conexiones de objetos del sector relacionado y crear nuevas predicciones respecto a problemas o fenómenos del entorno [3]. Se identifican tres dimensiones:

- **Dimensión del contexto del dato:** Corresponde al contexto del dato [33], [36].
- **Dimensión de las características del dato:** Propiedades inherentes del dato [20], [33], [36].
- **Dimensión de representación del dato:** Corresponde a la representación del dato en su contexto o en sus usos potenciales [33], [36].

Garantizar las dimensiones de los OGD impulsa la liberación de respuestas del interior de los datos, convergiendo en líneas de valor público con impactos diferenciados en el área financiera, política, social, estratégica para la innovación, ideológica, en el respeto hacia las entidades públicas y confianza en los datos publicados por estas [5].

d. Atributos para la evaluación de calidad de OGD

La capacidad de los datos para prestar de manera correcta un servicio se logra con el cumplimiento de criterios definidos para estandarizar su calidad [2], los cuales apuntan a algún atributo de los OGD. Para este estudio se identificaron y seleccionaron 12 atributos de calidad para OGD:

1. **Accesibilidad:** Facilitar el acceso a la información (datos y metadatos) en varios contextos [5], [14], [20], [23], [33], [35]–[37].
2. **Actualización:** Vigencia del dato en su contexto [5], [24].
3. **Complejidad:** Los valores del dato y metadatos relacionado están completos [14], [20], [33], [35]–[37]
4. **Comprensible:** Permitir al consumidor leer e interpretar lo que representa el dato [5], [24], [33], [36]
5. **Conformidad:** Los datos cumplen con estándares y lineamientos [5], [24].
6. **Consistencia:** Datos coherentes y libres de contradicción [6], [14], [33], [36], [37].
7. **Exactitud:** Datos diligenciados correctamente [5], [14], [20], [23], [24], [33], [35], [36].
8. **Integridad:** Mantener y asegurar la consistencia de los datos [6], [14], [24], [37].
9. **Precisión:** Grado en que un dato es correcto y conciso [35], [37].
10. **Relevancia:** Utilidad del dato para el consumidor, según el contexto [5], [6], [14], [20], [33], [35], [36]
11. **Trazabilidad:** Capacidad de seguir el proceso de evolución del dato y su representación en cada etapa de su ciclo de vida [5], [24].
12. **Uso:** Capacidad del dato para servir a un fin determinado [28].

A continuación, en la Figura 2 se propone un modelo conceptual que describe los principios, lineamientos, dimensiones y atributos relevantes, de acuerdo con los criterios de selección definidos, para finalizar con la alineación de estos componentes en un esquema estratégico que prioriza la explotación de los datos como fuente de tendencias e innovaciones. Se presentan los lineamientos que declaran los principios seleccionados y se establece la jerarquía de dependencia para detallar las dimensiones y los atributos que en conjunto proveen una visión de los objetivos requeridos que facilita la evaluación de los OGD.

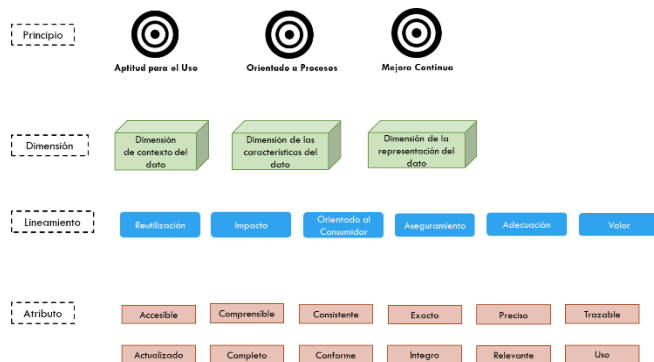


Figura 2: Modelo propuesto.
Fuente: Elaboración propia.

A continuación, se describe cada ítem que conforma el modelo conceptual propuesto (ver 4).

Tabla 4: Modelo Propuesto.

Item	Descripción
Principios	
Aptitud para el uso	Asegurar la capacidad del dato para servir a un propósito definido.
Orientado a procesos	Gestionar la calidad de los datos, a través de procesos que permitan implementar la estrategia de calidad de forma efectiva, con actividades interrelacionadas para una ejecución sistemática.
Mejora continua	Garantizar la calidad de los datos a través de la planeación, ejecución de acciones concretas y verificación periódica.
Dimensiones	
Dimensión del contexto del dato	Corresponde a las características del contexto de origen y uso del dato para facilitar su comprensión y aplicación.
Dimensión de las características del dato	Hace referencia a las propiedades inherentes del dato, a través de las cuales se define, interpreta y proyecta su uso.
Dimensión de representación del dato.	Corresponde a la representación del dato en su contexto de origen o en los posibles contextos de reutilización.
Lineamientos	

Item	Descripción
Aseguramiento	Realizar actividades planificadas y sistemáticas con el fin de alcanzar requisitos de calidad para las propiedades inherentes de los datos.
Adecuación	Acondicionar los datos para cumplir con los requisitos de calidad relacionados a sus propiedades inherentes.
Reutilización	Habilitar el uso de los datos en varios contextos.
Impacto	Generar valor para los consumidores, según el uso de los datos en un contexto específico.
Orientado al consumidor	Facilitar al consumidor del dato el conocimiento contextual y acceso a los datos.
Valor	Garantizar la representación de los datos como una unidad de valor para sus consumidores.
Atributos	
Accesibilidad	Facilitar el acceso a la información (datos y metadatos) en varios contextos.
Actualización	Vigencia del dato en su contexto.
Complejidad	Los valores del dato y metadatos relacionado están completos.
Comprensible	Permitir al consumidor leer e interpretar lo que representa el dato.
Conformidad	Los datos cumplen con estándares y lineamientos.
Consistencia	Datos coherentes y libres de contradicción.
Exactitud	Datos diligenciados correctamente.
Integridad	Mantener y asegurar la consistencia de los datos.
Precisión	Grado en que un dato es correcto y conciso
Relevancia	Utilidad del dato para el consumidor, según el contexto.
Trazabilidad	Capacidad de seguir el proceso de evolución del dato y su representación en cada etapa de su ciclo de vida.
Uso	Capacidad del dato para servir a un fin determinado.

Fuente: Elaboración propia

El modelo conceptual propuesto se ajustó para su posible aplicación como caso de estudio en el Ministerio de Hacienda y Crédito Público (MHCP) de Colombia, con el objetivo de mejorar la publicación de datos abiertos con un nivel de calidad adecuado. Este caso de estudio se definió teniendo en cuenta que el modelo se diseñó considerando los documentos y lineamientos técnicos emitidos por el MinTIC. En la Figura 3 se presenta el modelo ajustado, así como la relación de los atributos de calidad de los OGD con los lineamientos y las dimensiones que en conjunto proveen una visión de los principios orientadores requeridos para facilitar la evaluación de calidad de los OGD.

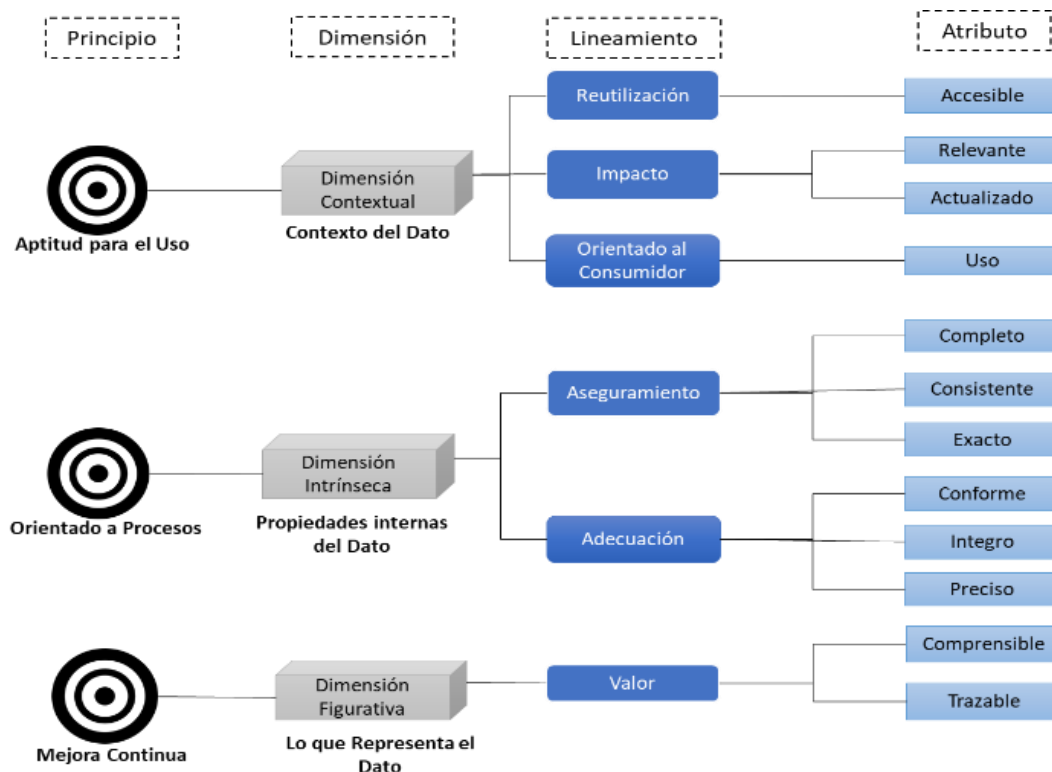


Figura 3: Modelo propuesto aplicado a MHCP.

Fuente: Elaboración propia.

Una limitación identificada en este estudio corresponde a que la publicación de OGD para el MHCP es un tema aún en exploración. Sin embargo, esta propuesta puede representar un hito de interés debido a las políticas actuales y el trabajo que se quiere adelantar en Colombia respecto a la gestión y aprovechamiento de datos públicos con Big Data. Por otra parte, la alineación de los principios, lineamientos, dimensiones y atributos podría variar según el caso de estudio, lo que se espera es que la coherencia en esta alineación sea acorde con los objetivos del negocio y las necesidades de evaluación, así mismo se podrían establecer otros criterios o asignar prioridades en la alineación de los elementos por lo cual el modelo podría expandirse o contraerse.

Por otra parte, este modelo se encuentra en construcción, como trabajo futuro se espera validar su contenido con al menos cinco entidades públicas del Estado colombiano para identificar fortalezas y debilidades. Además, se plantea realizar la propuesta metodológica modelada en BPMN (Business Process Model and Notation) a partir de los elementos identificados en este trabajo para su aplicación al caso de estudio del MHCP.

V. CONCLUSIONES

Este trabajo propone un modelo conceptual soportado en principios, dimensiones, lineamientos y atributos de calidad de OGD, los cuales se identificaron a partir de la revisión de literatura. El modelo presentado corresponde a una versión inicial, sujeta a cambios, de acuerdo con una discusión posterior con expertos para validar su pertinencia y de esta manera, dar una solución más adecuada para su futura adopción e implementación.

Como resultado de la revisión de literatura se encontró que, los OGD con calidad, representan un desafío que debe ser considerado y tratado para fomentar la colaboración y participación ciudadana. Así como para posibilitar la reutilización de los datos en procura de la innovación pública y privada. Los consumidores de estos datos requieren una base sólida de confianza que active el uso de la información pública, la cual sólo se puede adquirir luego de la implementación de un proceso de aseguramiento de calidad de datos que fortalezca e intensifique el valor inherente de los datos y active su potencial en diferentes contextos.

Se observó que el desarrollo de estrategias de apertura de datos genera valor en el descubrimiento de tendencias y en la identificación de nuevas perspectivas de un problema para obtener patrones que apoyen la innovación. Además, las líneas conceptuales que definen la estructura de soporte para que los OGD se transformen en valor público, centran su motivación en promover la reutilización de los datos en diversos dominios y así, facilitar las implementaciones innovadoras construidas a partir de nuevo conocimiento.

De otro lado, se determinó que la calidad de los OGD es un componente que ha evolucionado de ser un requerimiento para la publicación y acondicionamiento de la aptitud para su uso, hasta consolidarse como un proceso inherente a los datos, a través del cual se logran habilitar los beneficios de la explotación de sus propiedades para usos de impacto que suman valor a los consumidores.

En este sentido, los principios orientadores seleccionados para fundamentar la calidad de los OGD se complementan entre sí para buscar la conversión de la materia prima de los datos en productos que se materializan en información relevante y coherente para el sector público, privado y la ciudadanía en general. Mientras que, las dimensiones y atributos de calidad de los OGD son factores determinantes en las tendencias de reutilización de los datos, debido a que, a través de su aseguramiento, proveen una visión más amplia de los recursos inherentes a ellos y facilitan la proyección de su potencial en la toma de decisiones.

Se proyecta que a mediano plazo los OGD presentarán las características de Big Data. Debido a esto, se identifica la necesidad de definir una metodología de evaluación de calidad de los OGD, teniendo como propósito facultar los datos para su uso con un enfoque en los consumidores de los datos. Por lo anterior, la calidad de los OGD debe darse como un proceso previo a la apertura de los datos, para aportar valor en la creación y generación de nuevos modelos de negocio en el ámbito empresarial y a su vez, facilitar la innovación en el sector público, a partir de ideas y mejores prácticas del sector privado.

VI. REFERENCIAS

- [1] J. F. Mahecha, N. E. López, and J. A. Velandia, "Assessing data quality in open data: A case study," 2017 Congr. Int. Innov. y Tendencias en Ing. CONIITI 2017 - Conf. Proc., vol. 2018-Janua, pp. 1–5, 2018, doi: 10.1109/CONIITI.2017.8273343.
- [2] TodoBI, "11 Consejos sobre Bad Data: el enemigo silencioso en Business Intelligence y Big Data," TodoBI, 2019. <https://www.todobi.com/11-consejos-sobre-bad-data-el-enemigo/> (accessed Jul. 14, 2020).
- [3] BSA The Software Alliance, "¿Por qué son tan importantes los datos?," 2017. Accessed: May 22, 2020. [Online]. Available: https://data.bsa.org/wp-content/uploads/2015/10/BSADataStudy_es.pdf.
- [4] T. Redman, "Bad Data Costs the U.S. \$3 Trillion Per Year," Harvard Business Review Web site, 2016. <https://hbr.org/2016/09/bad-data-costs-the-u-s-3-trillion-per-year> (accessed Jul. 14, 2020).
- [5] Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia, "Guía de estándares de Calidad e interoperabilidad de los Datos Abiertos del Gobierno de Colombia," 2016. Accessed: May 04, 2020. [Online]. Available: https://herramientas.datos.gov.co/sites/default/files/A_guia_de_estandares_final_0.pdf.
- [6] M. S. Talukder, L. Shen, M. F. Hossain-Talukder, and Y. Bao, "Determinants of user acceptance and use of open government data (OGD): An empirical investigation in Bangladesh," Technol. Soc., vol. 56, pp. 147–156, Feb. 2019, doi: 10.1016/j.techsoc.2018.09.013.
- [7] A. Gómez, A. Funez, C. Robert, D. Zepeda, y S. Juárez, "Cómo innovar en proyectos de desarrollo: 13 casos de éxito en Latinoamérica | Publications," BID, 2019. https://publications.iadb.org/publications/spanish/document/Cómo_innovar_en_proyectos_de_desarrollo_Trece_casos_de_éxito_en_Latinoamérica.pdf (accessed Jul. 14, 2020).
- [8] Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia, "Colombia es el tercer país con mejor progreso en datos abiertos: OCDE - Ministerio de Tecnologías de la Información y las Comunicaciones," 2020. <https://www.mintic.gov.co/portal/inicio/Sala-de-Prensa/Noticias/116115:Colombia-es-el-tercer-pais-con-mejor-progreso-en-datos-abiertos-OCDE> (accessed Sep. 04, 2020).
- [9] Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia, "Guía de Datos Abiertos en Colombia," 2016. Accessed: May 04, 2020. [Online]. Available: <https://herramientas.datos.gov.co/es/blog/conoce-la-guia-de-datos-abiertos>.
- [10] J. Kucera, "Open Government Data Publication Methodology," J. Syst. Integr., vol. 6, no. 2, pp. 52–61, 2015, doi: <https://doi.org/10.20470/jsi.v6i2.231>.
- [11] Cumbre Global de la Alianza para el Gobierno Abierto, "Carta Internacional de Datos Abiertos (2015)," México, 2015. Accessed: May 27, 2020. [Online]. Available: http://www.datosabiertos.gob.ec/wp-content/uploads/downloads/2016/08/Carta_Internacional_de_Datos_Abiertos2015.pdf.
- [12] K. Torres-Saumeth, T. Ruiz Afanador, L. Solís Ospino, and F. Martínez Barraza, "Calidad y su evolución: una revisión* Quality and its evolution: a review," Universidad Autónoma del Caribe, 2012. Accessed: Jul. 14, 2020. [Online]. Available: <https://dialnet.unirioja.es/servlet/articulo?codigo=4400435&info=resumen&idioma=ENG>.
- [13] Power Data, "Big Data: ¿En qué consiste? Su importancia, desafíos y gobernabilidad," Power Data Especialista en Gestión de Datos, 2019. <https://www.powerdata.es/big-data> (accessed Jul. 14, 2020).

- [14] L. Cai and Y. Zhu, "The Challenges of Data Quality and Data Quality Assessment in the Big Data Era," *Data Sci. J.*, vol. 14, no. 0, p. 2, May 2015, doi: 10.5334/dsj-2015-002.
- [15] D. Loshin, "Understanding Big Data Quality for Maximum Information Usability," 2014. Accessed: Jul. 14, 2020. [Online]. Available: www.dataqualitybook.com.
- [16] Gobierno de España, "Manual práctico para mejorar la calidad de los datos abiertos," Madrid, Dec. 2017. Accessed: Jun. 02, 2020. [Online]. Available: https://datos.gob.es/sites/default/files/doc/file/manual_practico_para_mejorar_la_calidad_de_los_datos_abiertos_1.pdf.
- [17] ISO/IEC 25012, "ISO 25012," ISO/IEC 25012, 2019. <https://iso25000.com/index.php/normas-iso-25000/iso-25012> (accessed Jul. 14, 2020).
- [18] S. Mukherjee and R. Shaw, "Big Data-Concepts, Applications, Challenges and Future Scope," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 5, no. 2, 2016, doi: 10.17148/IJARCCCE.2016.5215.
- [19] N. Miloslavskaya and A. Tolstoy, "Big Data, Fast Data and Data Lake Concepts," in *Procedia Computer Science*, Jan. 2016, vol. 88, pp. 300–305, doi: 10.1016/j.procs.2016.07.439.
- [20] Europe United Nations Economic Commission, "A Suggested Framework for the Quality of Big Data Deliverables," 2014. [Online]. Available: https://statswiki.unece.org/download/attachments/108102944/Big_Data_Quality_Framework_-_final_Jan08-2015.pdf?version=1&modificationDate=1420725063663&api=v2.
- [21] P. Russom and T. Org, "BIG DATA ANALYTICS FOURTH QUARTER 2011 TDWI RE SE A RCH Co-sponsored by BIG DATA A N A LY TIC S FOURTH QUARTER 2011 TDWI BEST PRACTICES REPORT Introduction to Big Data Analytics," 2011.
- [22] J. Attard, F. Orlandi, and S. Auer, "Value Creation on Open Government Data," in *2016 49th Hawaii International Conference on System Sciences (HICSS)*, Jan. 2016, pp. 2605–2614, doi: 10.1109/HICSS.2016.326.
- [23] J. Attard, F. Orlandi, S. Scerri, and S. Auer, "A systematic review of open government data initiatives," *Gov. Inf. Q.*, vol. 32, no. 4, pp. 399–418, Oct. 2015, doi: 10.1016/j.giq.2015.07.006.
- [24] M. Yi, "Exploring the quality of government open data Comparison study of the UK, the USA and Korea," *Electron. Libr.*, vol. 37, no. 1, pp. 35–48, 2019, doi: 10.1108/EL-06-2018-0124.
- [25] Ministerio de Ciencia Tecnología Innovación y Comunicaciones - Brasil, "Plan de Datos Abiertos 2020-2021," Brasil, 2020. Accessed: May 27, 2020. [Online]. Available: <http://www.mctic.gov.br/>.
- [26] A. Zuiderwijk, M. Janssen, and I. Susha, "Improving the speed and ease of open data use through metadata, interaction mechanisms, and quality indicators," *J. Organ. Comput. Electron. Commer.*, vol. 26, no. 1–2, pp. 116–146, Apr. 2016, doi: 10.1080/10919392.2015.1125180.
- [27] N. F. Warraich, T. Rasool, and M. Sajid, "Challenges to Use Open Government Data through the Citizens' Lens: A Systematic Review," *J. Polit. Stud.*, vol. 26, no. 2, p. 14, 2019.
- [28] Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia, "Guía para el uso y aprovechamiento de Datos Abiertos en Colombia," 2019. Accessed: May 04, 2020. [Online]. Available: https://herramientas.datos.gov.co/sites/default/files/Guia_de_Datos_Abiertos_de_Colombia.pdf.
- [29] Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia, "Requisitos de calidad para Datos Abiertos," 2019. Accessed: May 04, 2020. [Online]. Available: https://sellodeexcelencia.gov.co/documents/UTSF_SDE_Requisitos_de_calidad_para_datos_abiertos_2019_12_02_v_2_0.pdf.
- [30] Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia, "Datos Abiertos Colombia," 2019. <https://herramientas.datos.gov.co/es/fichatecnica/calidad> (accessed Apr. 29, 2020).
- [31] E. Manrique, G. Maestre-Gongora, y M. Osorio-Sanabria, "Caracterización de buenas prácticas para Gestión de Sistemas de Información en Entidades Públicas Colombianas," *Ingeniere*, vol. 14, no. 24, pp. 65–76, 2018.
- [32] P. Ciancarini, F. Poggi, y D. Russo, "Big Data Quality: A Roadmap for Open Data," in *Proceedings - 2016 IEEE 2nd International Conference on Big Data Computing Service and Applications, BigDataService 2016*, May 2016, pp. 210–215, doi: 10.1109/BigDataService.2016.37.
- [33] T. Koltay, "Quality of Open Research Data: Values, Convergences and Governance," *Information*, vol. 11, no. 4, p. 175, Mar. 2020, doi: 10.3390/info11040175.
- [34] Congreso de la República de Colombia, "LEY 1712 DE 2014," Imprenta Nacional de Colombia, 2014. <http://suin.gov.co/viewDocument.aspx?ruta=Leyes/1687091> (accessed May 04, 2020).
- [35] J. Merino, I. Caballero, B. Rivas, M. Serrano, and M. Piattini, "A Data Quality in Use model for Big Data," *Futur. Gener. Comput. Syst.*, vol. 63, pp. 123–130, Oct. 2016, doi: 10.1016/j.future.2015.11.024.
- [36] A. Wahyudi, G. Kuk, and M. Janssen, "A Process Pattern Model for Tackling and Improving Big Data Quality," vol. 20, no. 3, pp. 457–469, 2018, doi: <https://doi.org/10.1007/s10796-017-9822-7>.
- [37] D. Williams and H. Tang, "Data Quality Management for Industry 4.0: A Survey," 2020. [Online]. Available: www.asq.org.
- [38] J. Kucera and D. Chlappek, "Benefits and Risks of Open Government Data," *J. Syst. Integr.*, vol. 5, no. 1, pp. 30–41, 2014, doi: <http://dx.doi.org/10.20470/jsi.v5i1.185>.
- [39] S. Martin, M. Foulonneau, S. Turki, and M. Ihdjadene, "Open Data: Barriers, Risks and Opportunities," Luxemburgo, 2012. Accessed: Apr. 28, 2020. [Online]. Available: <http://www.data.rennes-metropole.fr/>.
- [40] A. Munte-Kunigami and F. Serale, "Los datos abiertos en América Latina y el Caribe," *Los datos abiertos en América Lat. y el Caribe*, 2018, doi: 10.18235/0001202.
- [41] Z. Yang and A. Kankanhalli, "Innovation in government services: The case of open data," in *IFIP Advances in Information and Communication Technology*, 2013, vol. 402, pp. 644–651, doi: 10.1007/978-3-642-38862-0_47.
- [42] S. Cadena-Vela, A. Fuster-Guilló, and J. N. Mazón, "Publicando datos abiertos considerando criterios de calidad," 2019.
- [43] Diario Oficial de la Federación, "GUÍA de Implementación de la Política de Datos Abiertos - México," pp. 1–13, 2015, [Online]. Available: http://www.dof.gob.mx/nota_detalle.php?codigo=5397117&fecha=18/06/2015&print=true.
- [44] E. Kalampokis, E. Tambouris, and K. Tarabanis, "Open government data: A stage model," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011, vol. 6846 LNCS, pp. 235–246, doi: 10.1007/978-3-642-22878-0_20.
- [45] Consejo Nacional de Política Económica y Social de la República de Colombia, "Política Nacional de Explotación de Datos (Big Data)," 2018. Accessed: Apr. 25, 2020. [Online]. Available: <https://colaboracion.dnp.gov.co/CDT/Conpes/Económicos/3920.pdf>.

- [46] N. A. M. Sabri, N. A. Emran, and N. Harum, "Government open data portals: A measurement of data veracity coverage," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 12, pp. 1975–1983, Oct. 2019, doi: 10.35940/ijitee.L2908.1081219.
- [47] J. Wiczorkowski, "Barriers to using open government data," in *ACM International Conference Proceeding Series*, Jun. 2019, pp. 15–20, doi: 10.1145/3340017.3340022.
- [48] S. Goldsmith y S. Crawford, "El análisis predictivo: impulsando la mejora a partir de los datos," 2016, [Online]. Available: <https://publications.iadb.org/publications/spanish/document/Innovaciones-en-la-prestación-de-servicios-públicos-Número-4-El-análisis-predictivo-Impulsando-la-mejora-a-partir-de-los-datos.pdf>.