

Calidad de datos aplicada a la base de datos abierta de casos registrados de COVID-19

Ariel Pasini^{ORCID}, Juan Ignacio Torres^{ORCID}, Silvia Esponda^{ORCID}, Patricia Pesado^{ORCID}

Instituto de Investigación en Informática LIDI (III-LIDI)*

Facultad de Informática – Universidad Nacional de La Plata 50 y 120 La Plata Buenos Aires

*Centro Asociado Comisión de Investigaciones Científicas de la Pcia. de Bs. As. (CIC)

{apasini, jitorres, sesponda, ppesado}@lidi.info.unlp.edu.ar

Abstract. Para lograr una mayor transparencia y fomentar la participación de los ciudadanos en la toma de decisiones, y responder de manera eficiente, los gobiernos de diferentes países ponen una gran cantidad de información a disposición de su comunidad. Esta información, tomó más relevancia durante la pandemia del COVID-19. Sin embargo, si estos datos no poseen un buen nivel de calidad, la información pierde confiabilidad. Se presenta una evaluación de la calidad de datos del archivo público “COVID-19. Casos registrados en la República Argentina” utilizando el modelo que provee la norma ISO/IEC 25012 y el proceso de evaluación definido por la ISO/IEC 25040.

Keywords: Calidad de datos, ISO/IEC 25000, Gobierno abierto, Datos abiertos

1 Introducción

Desde el primer caso detectado en Wuhan (Hubei, China), el COVID-19 se ha expandido rápidamente a través de aproximadamente 200 países/regiones amenazando con la vida de las personas e irrumpiendo significativamente en la economía y la sociedad mundialmente.

La publicación de datos abiertos es crucial durante un período de pandemia, ya que, a través del análisis y procesamiento de estos, los ciudadanos ayudan en el proceso de toma de decisiones para responder de manera inteligente [1]. Es fundamental que los datos públicos en formatos abiertos (ya sea a través de datasets o reportes diarios) sean de alta calidad. Una baja calidad de datos afecta de manera negativa tanto al proceso de creación de valor como a la transparencia y la confianza en el gobierno [2].

El modelo general definido en la norma ISO/IEC 25012 – “Data Quality Model” [3] permite medir la calidad de los datos almacenados en un sistema informático. Es utilizado junto con otras normas para planificar y llevar a cabo evaluaciones de

calidad de datos. Este modelo, clasifica los atributos de calidad en quince características de acuerdo con dos puntos de vista que no son excluyentes: inherente y dependiente del sistema. Las características alcanzan diferente importancia y prioridad según las necesidades de las partes interesadas.

Por otro lado, la norma ISO/IEC 25040 – “Evaluation Process” [4] proporciona un modelo de referencia general para evaluar, que tiene en consideración las entradas al proceso de evaluación, restricciones y aquellos recursos que sean necesarios para obtener las salidas que correspondan. El proceso consta de un total de cinco actividades.

El presente artículo propone realizar la medición de la calidad de los datos obtenidos desde el archivo público “COVID-19. Casos registrados en la República Argentina” [5], mantenido por la Dirección de Epidemiología y Análisis de Situación de Salud, tomando ciertas características de calidad definidas por la norma ISO/IEC 25012 y siguiendo las actividades definidas en el modelo de evaluación establecido por la norma ISO/IEC 25040.

En la segunda sección se introducen las normas y las métricas a utilizar para la evaluación de la calidad de los datos. En la sección tres se introduce el concepto de Gobierno Abierto, se mencionan ventajas y desventajas de la publicación de datos abiertos y se resalta la importancia de la participación de los ciudadanos en el proceso de generación de valor a través del uso de estos datos. Luego, en el capítulo 4 se introduce detalladamente el recurso a evaluar y se realiza la evaluación de la calidad de la base de datos COVID-19. Finalmente, se presentan las conclusiones del artículo.

2 Calidad de Datos

La familia de normas ISO/IEC 25000, conocida como SQuARE (System and Software Requirements and Evaluation) propone evaluar la calidad de un producto de software bajo un marco de trabajo común. En el contexto del artículo, se utilizarán dos normas de la familia, la ISO/IEC 25012 y la ISO/IEC 25040.

2.1 Modelo de calidad de datos ISO/IEC 25012

Innegablemente, la cantidad de datos manejada por sistemas informáticos se encuentra en aumento en todas partes del mundo. Es necesario en todo proyecto de tecnología de la información maximizar la calidad de los datos que se intercambian, procesan y utilizan entre sistemas. Una baja calidad de datos puede generar información insatisfactoria y resultados que no son de utilidad.

La norma define un modelo de calidad de datos en un formato estructurado dentro de un sistema informático. Este modelo está compuesto por quince características divididas en dos puntos de vista: inherente y dependiente del sistema. Cabe destacar que algunas características son pertinentes desde los dos puntos de vista.

Calidad de los datos inherente:

Este punto de vista se refiere al grado en que las características tienen el potencial de satisfacer las necesidades establecidas e implícitas cuando se utilizan los datos bajo condiciones especificadas. Las características pertinentes desde el punto de vista inherente son: **Exactitud**: hace referencia a que los datos poseen atributos representados de manera correcta (puede ser sintáctica o semántica). **Compleitud**: los datos poseen valores para todos los atributos esperados. **Consistencia**: los datos no son contradictorios y son consistentes unos con otros. **Credibilidad**: los datos poseen atributos ciertos y creíbles (incluye concepto de autenticidad). **Actualidad**: los datos poseen atributos que son actualizados adecuadamente.

Calidad de los datos dependientes del sistema:

Hace referencia al grado en que se alcanza y preserva la calidad de los datos en un sistema informático cuando los datos se utilizan bajo condiciones específicas. Existe una dependencia entre la calidad de los datos y el dominio tecnológico en el cual los mismos se utilizan. Las características pertinentes desde el punto de vista dependiente del sistema son: **Disponibilidad**: define el grado en que los atributos pueden ser recuperados por los usuarios y/o aplicaciones que tengan acceso. **Portabilidad**: grado en que los datos pueden ser instalados, reemplazados o copiados de un sistema a otro manteniendo la calidad. **Recuperabilidad**: los datos deben mantener y preservar un nivel de operaciones y calidad ante fallos.

Calidad de los datos inherente y dependientes del sistema:

Las características que son pertinentes desde ambos puntos de vista (inherente y dependiente del sistema) son: **Accesibilidad**: grado en que se puede acceder a los datos. **Cumplimiento**: los datos cumplen con normas, convenciones o regulaciones en relación a la calidad de datos en un contexto específico. **Confidencialidad**: los datos tienen atributos que deben ser accedidos e interpretados solamente por usuarios autorizados. Es parte de la seguridad de la información, junto con la integridad y la disponibilidad. **Eficiencia**: los atributos se pueden procesar y proporcionar niveles esperados de rendimiento utilizando tipos y cantidades de recursos correspondientes. **Precisión**: los atributos pertenecientes a los datos presentan valores exactos o que permiten discriminación. **Trazabilidad**: se analiza si se proporciona un registro para la auditoría del acceso a los datos y de cualquier modificación respecto a los mismos. **Comprensibilidad**: grado en que los usuarios pueden leer e interpretar los datos, que deben estar expresados de manera apropiada.

2.2 Definición de métricas según ISO/IEC 25012

La ISO/IEC 25012 provee métricas de ejemplo para cada característica. Para cada característica, se utilizarán las funciones de medición contenidas en la norma. Se presenta una a modo de ejemplo:

- Tipo (inherente o dependiente del sistema)
- Nombre de la medida de calidad

- Función de medición
- Variables

En la tabla 1 se presenta como ejemplo la métrica para la característica Completitud.

Tabla 1. Definición de métrica para la Completitud

Tipo	Inherente
Nombre de la medida de calidad	Completitud de los datos en un archivo
Función de medición	Valor = A/B
Variables	A = número de datos requeridos para el contexto particular en el archivo B= número de datos en el contexto particular de uso previsto

2.3 Modelo de evaluación

La norma ISO/IEC 25040 define un modelo de referencia teniendo en cuenta entradas, restricciones y recursos necesarios para obtener las salidas deseadas. Para llevar a cabo la evaluación de un producto de software, se define un proceso de cinco actividades:

1. **Establecer los requisitos de la evaluación:** establecer el propósito de la evaluación y los requisitos de calidad del software a evaluar, identificando las partes interesadas en el producto, los riesgos (si existen) y el modelo de calidad a utilizar.
2. **Especificar la evaluación:** dentro de esta actividad se seleccionan los módulos de evaluación y se definen los criterios de decisión tanto para las métricas como para la evaluación.
3. **Diseñar la evaluación:** se define el plan con las actividades de evaluación que se deben realizar, teniendo en cuenta la disponibilidad de recursos.
4. **Ejecutar la evaluación:** en esta cuarta actividad se realizan las mediciones y se aplican los criterios de decisión tanto para las métricas como para la evaluación.
5. **Concluir la evaluación:** se cierra la evaluación de la calidad, realizando un informe con los resultados finales y las conclusiones de la evaluación en base a lo obtenido.

3 Gobierno y datos abiertos

El término de gobierno abierto se fue construyendo a través del tiempo. Sin embargo, fue popularizado recién en el 2009 por la Administración de Barack Obama en el “Memorandum on Transparency and Open Government” [6] en donde se proponen tres principios:

1. Un gobierno debe ser transparente. La transparencia en un gobierno puede ser alcanzada proveyendo a los ciudadanos con información acerca de lo que está haciendo, lo cual promueve la rendición de cuentas.

2. Un gobierno debe ser participativo. La participación pública en la formación de políticas y aportando conocimiento, ideas y experiencia mejora tanto la efectividad del gobierno como la calidad de sus decisiones.
3. Un gobierno debe ser colaborativo. La colaboración involucra activamente a los ciudadanos en el trabajo de su gobierno. Este principio requiere de cooperación entre individuos, empresas, asociaciones y agentes en todos los niveles de gobierno.

Se puede entender al gobierno abierto como una plataforma tecnológica ya que utiliza la tecnología de información y comunicación a su alcance para lograr su propósito. Esto incluye tanto sitios web como las redes sociales. [7]

3.1 Participación de los ciudadanos en el análisis de los datos abiertos

La participación de la ciudadanía es crucial en el concepto de gobierno abierto. Los gobiernos alrededor del mundo buscan aumentar la participación de los ciudadanos debido a su importancia para la toma de decisiones tanto en procesos políticos como administrativos. Para ello, es fundamental poner al alcance de la sociedad datos públicos en formatos abiertos. La revolución en las TIC (tecnologías de información y comunicación) cambió radicalmente la interacción entre gobiernos y ciudadanos.

La Web social o Web 2.0, que fomenta la participación de los individuos, permite alcanzar fácilmente el cumplimiento de los tres propósitos mencionados en el apartado anterior. Este paradigma se refiere al contenido que es creado, compartido y procesado por usuarios a través de medios de comunicación social además de la formación de una red social en la cual los mismos se conectan entre sí. [8]

Los gobiernos pueden expandir el alcance de los datos abiertos que publican a través del uso de estas redes sociales electrónicas. Para impulsar esto, es necesario crear comunidades en torno a los datos, que los consulten y analicen además de compartir nuevos datos. Además, un punto clave a tener en cuenta, es la redistribución de los mismos, a mayor velocidad de redistribución, mayor será el valor adquirido por los datos.

3.2 Datos abiertos: ventajas y desventajas

La publicación de datos abiertos en sus diferentes formas, como se mencionó previamente, permite a los ciudadanos usar y crear información a través de una red colaborativa. Con la misma surge una inteligencia colectiva en la que el público ayuda a mejorar la toma de decisiones participando activamente.

Además, al fomentar la publicación de datos abiertos, se aumenta la satisfacción del pueblo, ya que promueve igualdad de acceso a los datos, y se obtiene mayor confianza y transparencia. [9]

Sin embargo, hay ciertos factores que hacen que el uso de los datos y la participación ciudadana se vea perjudicada, debido a que no promueven la generación

de valor agregado. Otro punto para tener en cuenta es la falta de un análisis sistemático sobre qué debería ser publicado y qué esperan los usuarios de los datos abiertos, en muchos casos esa falta de análisis conlleva a publicar información incompleta, o incluso excesiva, lo que impide concentrarse en lo relevante. Por lo tanto, es importante asegurar la calidad de los datos para garantizar un análisis acertado.

4 Base de datos de COVID-19

La publicación de datos epidemiológicos abiertos junto con su análisis es un elemento fundamental ante el COVID-19. Es por lo que, para profundizar la transparencia durante esta pandemia, el ministerio de Salud de la Nación publicó distintos datasets relacionados a las tareas de lucha contra el coronavirus. En particular, en el siguiente trabajo se hizo foco en el recurso “*COVID-19. Casos registrados en la República Argentina*”, de actualización diaria, que se comenzó a publicar el 15 de mayo de 2020 y es mantenido por la Dirección Nacional de Epidemiología y Análisis de Situación de Salud.

El mismo cuenta con veinticinco campos, que se mencionan a continuación junto con sus tipos: número de caso (número entero), sexo (texto), edad (número entero), edad indicada en meses o años (texto), país de residencia (texto), provincia de residencia (texto), departamento de residencia (texto), provincia de establecimiento de carga (texto), fecha de inicio de síntomas (fecha según ISO-8601 - “Data elements and interchange formats — Information interchange — Representation of dates and times”), fecha de apertura del caso (fecha ISO-8601), semana epidemiológica de fecha de apertura (número entero), fecha de internación (fecha ISO-8601), indicación si estuvo en cuidado intensivo (texto), fecha de ingreso a cuidado intensivo en caso de corresponder (fecha ISO-8601), indicación de fallecido (texto), fecha de fallecimiento en caso de corresponder (tiempo ISO-8601), indicación si requirió asistencia respiratoria mecánica (texto), código de provincia de carga (número entero), origen de financiamiento (texto), clasificación manual del registro (texto), clasificación del caso (texto), código de provincia de residencia (número entero), fecha de diagnóstico (tiempo ISO-8601), código de departamento de residencia (número entero) y última actualización (fecha ISO-8601).

4.1 Evaluación de la base de datos COVID-19

Se realizará la evaluación teniendo en cuenta las actividades propuestas por la norma ISO/IEC 25040.

4.1.1 Propósito de la evaluación.

El propósito de la evaluación es: *analizar si los datos presentes en el recurso considerados como obligatorios están presentes, qué tan actualizada está la clasificación de los casos, si los campos cumplen con los formatos esperados, si están expresados correctamente y si los datos son consistentes entre sí.*

Para lograrlo, se evaluarán las características: *Exactitud*, *Compleitud*, *Consistencia*, *Actualidad* y *Comprensibilidad*.

4.1.2 Especificación de la evaluación.

En esta actividad, además de definir los módulos de evaluación (presentado en la sección 2.2), se definen los criterios de decisión tanto para las características como para la evaluación final.

En la tabla 2 se presentan los rangos de valores para determinar el nivel de aceptación para las características: *exactitud*, *consistencia*, *completitud* y *comprensibilidad*.

En el caso de la característica *actualidad*, para determinar el nivel obtenido, se deben verificar tres campos en la base de datos: *confirmados*, *fallecidos*, *descartados* en la tabla 3 se puede ver los valores para determinar el nivel de cada uno de ellos y en la tabla 4 la combinación de estos tres campos para establecer el nivel alcanzado por la característica.

Tabla 2. Criterios de decisión para las características Exactitud, Consistencia, Compleitud y Comprensibilidad.

Características	Nivel	Rango de valores
Exactitud	Inaceptable	Si Valor ≥ 0 y Valor $< 0,3$
Consistencia	Mínimamente aceptable	Si Valor $\geq 0,3$ y Valor $< 0,7$
Compleitud	Rango objetivo	Si Valor $\geq 0,7$ y Valor $< 0,9$
Comprensibilidad	Excede los requerimientos	Si Valor $\geq 0,9$

Tabla 3. Criterios de decisión para los campos del criterio actualidad.

Campos	Nivel	Rango de valores
Confirmados (C)	Inaceptable	Si Valor ≥ 0 y Valor $< 0,3$
Fallecidos (F)	Mínimamente aceptable	Si Valor $\geq 0,3$ y Valor $< 0,7$
Descartados (D)	Rango objetivo	Si Valor $\geq 0,7$ y Valor $< 0,9$
	Excede los requerimientos	Si Valor $\geq 0,9$

Tabla 4. Criterios de decisión para la característica actualidad

Características	Nivel	Rango de valores
Actualidad	Inaceptable	C, D y F: Inaceptable
	Mínimamente aceptable	C: D y F: Mínimamente aceptable
	Rango objetivo	C, D y F: Rango Objetivo
	Excede los requerimientos	C, D y F: Excede los requerimientos

En la tabla 5 se muestran los criterios de decisión para la evaluación final, teniendo en cuenta los criterios de decisión para cada característica.

Tabla 5. Criterios de decisión para la Evaluación final.

Nivel Final	Valor de las características
Inaceptable	E, C, A y M: Inaceptable
	O: Mínimamente Aceptable
Mínimamente aceptable	E: Mínimamente aceptable
	C y M: Rango objetivo
	O y A: Mínimamente Aceptable
Rango Objetivo	E, O, A y M: Rango Objetivo
	C: Excede los requerimientos
Excede los requerimientos	E, C, O, A y M: Excede los requerimientos

Exactitud (E) Completitud (C) Consistencia (O) Actualidad (A) Comprensibilidad (M)

4.1.3 Diseño de la evaluación.

Se propone realizar una evaluación de la calidad de los datos desde la publicación del recurso (15 de mayo de 2020) hasta el 17 de febrero de 2021. Se evalúan las características definidas en la ISO/IEC 25012 acordes al propósito de la evaluación basada en la guía propuesta por la norma ISO/IEC 25040 y luego se calcula el promedio para determinar el valor final. Para realizar la evaluación se tendrán en cuenta las especificaciones que figuran en la tabla 6.

Tabla 6. Especificaciones de las características a evaluar

Característica	Especificaciones
Exactitud (Sintáctica + Semántica)	Para cada campo se tendrá en cuenta que coincida con el tipo de dato definido por el mantenedor de la base de datos, en el contexto acorde a la descripción.
	Número de Caso -> Número entero
	Sexo -> Texto (M o F)
	Edad -> Número Entero entre 0 y 116
	Edad indicada en meses o años -> Texto (“años” o “meses”)
	País de Residencia -> Texto
	Provincia de Residencia -> Texto
	Departamento de Residencia -> Texto
	Provincia de establecimiento de carga -> Texto
	Fecha de inicio de síntomas -> Fecha ISO-8601 (date)
	Fecha de apertura del caso -> Fecha ISO-8601 (date)
	Semana Epidemiológica de fecha de apertura -> Número Entero
	Fecha de internación -> Fecha ISO-8601 (date) – si corresponde
	Indicación si estuvo en cuidado intensivo -> Texto (SI/NO)
	Fecha de ingreso a cuidado intensivo -> Fecha ISO-8601 (date) – si corresponde
	Indicación de fallecido -> Texto (SI/NO)
	Fecha de fallecimiento -> Tiempo ISO-8601 (time) – si corresponde
Indicación si requirió asistencia respiratoria mecánica -> Texto	

	(SI/NO) Código de Provincia de carga -> Número Entero Origen de financiamiento -> Texto (Público/Privado) Clasificación manual del registro -> Texto Clasificación del caso -> Texto (Descartado, confirmado o sospechoso). Código de Provincia de residencia -> Número Entero Fecha de diagnóstico -> Tiempo ISO-8601 (time) Código de Departamento de residencia -> Número Entero Última actualización -> Fecha ISO-8601 (date)
Completitud	En cada registro se evalúa que cada campo obligatorio esté presente
Consistencia	Los datos deben ser consistentes entre sí. Un ejemplo de inconsistencia a tener en cuenta es que una persona que no ha fallecido tenga fecha de fallecimiento.
Actualidad	La actualidad se mide en base a la diferencia entre el día de publicación del dataset y la fecha de diagnóstico (descartado/confirmado) o fecha de fallecimiento (fallecido). Para realizar la evaluación se considera actual aquella fecha de diagnóstico o de fallecimiento que tenga una demora de máximo 2 días con respecto a la fecha de emisión del recurso.
Comprensibilidad	Se evalúa que los campos estén expresados con una unidad o forma apropiada y deben ser legibles por el usuario.

4.1.4 Ejecución de la evaluación.

Al realizarse las mediciones se obtuvieron los valores para cada característica, y al aplicarse los criterios de decisión se obtuvieron los resultados reflejados en la tabla 7.

Tabla 7. Valores y resultados para cada característica

Característica	Valor	Resultado
Exactitud	1	Excede los requerimientos
Completitud	0.82	Rango objetivo
Consistencia	1	Excede los requerimientos
Actualidad	Confirmados = 0.75	Rango objetivo
	Descartados = 0.77	Rango objetivo
	Fallecidos = 0.33	Mínimamente aceptable
Comprensibilidad	0.92	Excede los requerimientos

El resultado más bajo, se relaciona con el campo *fallecidos* de la característica de *actualidad*, el promedio de cantidad de días de demora fue de 2.84 días para los *confirmados*, 9.65 días para los *descartados* y de 13.23 días para los *fallecidos*.

4.1.5 Finalización de la evaluación.

En la tabla 8 se presentan los resultados obtenidos para cada característica, aplicando el criterio definido para la evaluación final (tabla 5), si bien tres de las cinco características analizadas exceden los requerimientos, la actualidad de los fallecidos determinó que la característica *actualidad* sea mínimamente aceptable, por lo tanto, el mayor nivel que se puede otorgar a la evaluación final del propósito analizado es: **Mínimamente Aceptable**. Siendo este, un resultado satisfactorio por lo definido en la norma ISO/IEC 25040.

Tabla 8. Resultado final de la evaluación

Característica	Nivel Obtenido
Exactitud	Excede los requerimientos
Complejidad	Rango Objetivo
Consistencia	Excede los requerimientos
Actualidad	Mínimamente aceptable
Comprensibilidad	Excede los requerimientos

En cuanto a la *comprensibilidad*, si bien excede los requerimientos, se notó que no existió diferencia en el uso de los tipos Fecha ISO-8601 (date) y Tiempo ISO-8601 (time), los campos con estos tipos mostraban datos con el mismo formato.

5 Conclusiones

Se presentaron dos normas de la familia ISO/IEC 25000 (SQuaRE). Se mostraron las quince características presentes en la ISO/IEC 25012, de las cuales se seleccionaron cinco para un contexto específico de uso y para las cuales se definieron las métricas basadas en las funciones de medición definidas en la norma.

Se definió el concepto de gobierno abierto, haciendo énfasis en los beneficios y las problemáticas de publicar datos abiertos además de resaltar la importancia de la participación de los habitantes en el uso de estos para la generación de valor y ayudando a tomar decisiones de manera eficaz.

Se realizó la evaluación de la calidad de datos a la base de datos “COVID-19. Casos registrados en la República Argentina”, siguiendo el modelo de evaluación propuesto por la norma ISO/IEC 25040. El resultado final fue “mínimamente aceptable” debido a que el valor obtenido para la actualidad de los fallecidos se encontraba dentro de este rango. Esto no quita que el resultado haya sido satisfactorio.

Cabe destacar, que, en una evaluación realizada diariamente bajo las mismas condiciones durante el mes de junio de 2021, los valores se mantuvieron en los mismos rangos, lo cual marca una estabilidad en la calidad de los datos.

6 Agradecimientos

Esta publicación fue realizada en el contexto del Proyecto CAP4CITY – “Strengthening Governance Capacity for Smart Sustainable Cities” (www.cap4city.eu) co-financiado por el Programa Erasmus+ de la Unión Europea. Acuerdo Número 598273-EPP-1- 2018-1-AT-EPPKA2-CBHE-JP. Número de proyecto: 598273



7 Referencias

- [1] O. Morgan, “How decision makers can use quantitative approaches to guide outbreak responses,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 374, no. 1776, 2019, doi: 10.1098/rstb.2018.0365.
- [2] H. Chen, D. Hailey, N. Wang, and P. Yu, “A review of data quality assessment methods for public health information systems,” *Int. J. Environ. Res. Public Health*, vol. 11, no. 5, pp. 5170–5207, 2014, doi: 10.3390/ijerph110505170.
- [3] ISO/IEC 25012:2008, “Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model.”
- [4] ISO/IEC 25040:2011, “Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Evaluation process.”
- [5] “COVID-19. Casos registrados en la República Argentina” [Online]. Available: <https://datos.gob.ar/>. - <http://datos.salud.gob.ar/>
- [6] O. Government, “Memorandum on Transparency and Open Government,” *Fed. Regist.*, pp. 21–22, 2009, [Online]. Available: <https://www.archives.gov/files/cui/documents/2009-WH-memo-on-transparency-and-open-government.pdf>.
- [7] R. Sandoval-Almazán, “Gobierno abierto y transparencia: Construyendo un marco conceptual,” *Convergencia*, vol. 22, no. 68, pp. 203–227, 2015, doi: 10.29101/crcs.v0i68.2958.
- [8] S. A. Chun, S. Shulman, R. Sandoval, and E. Hovy, “Government 2.0: Making connections between citizens, data and government,” *Inf. Polity*, vol. 15, no. 1–2, pp. 1–9, 2010, doi: 10.3233/IP-2010-0205.
- [9] M. Janssen, Y. Charalabidis, and A. Zuiderwijk, “Benefits, Adoption Barriers and Myths of Open Data and Open Government,” *Inf. Syst. Manag.*, vol. 29, no. 4, pp. 258–268, 2012.