

# Evaluación del tratamiento de los datos abiertos geolocalizados

---

## » Agustín Simón

CAETI - Facultad de Tecnología Informática - Universidad Abierta Interamericana, Ciudad de Buenos Aires, Argentina.

## Roxana Martínez

Unidad de Información Financiera – UIF. Ciudad de Buenos Aires, Argentina.  
CAETI - Facultad de Tecnología Informática - Universidad Abierta Interamericana.

## Resumen

En los tiempos de hoy en día, los datos abiertos públicos gubernamentales son una herramienta poderosa con el potencial de reducir las asimetrías de información, fomentar la toma de decisiones basada en evidencia, mejorar la rendición de cuentas y generar impacto en la calidad de vida de las personas a través de su uso. Los gobiernos de los diferentes países tienen una oportunidad sin precedentes de generar valor público a través de la apertura y uso de datos. La atención en los datos abiertos se suele centrar en los beneficios que éstos generan en términos de transparencia y rendición de cuentas. Con el fin de promover y mejorar la calidad de los conjuntos de datos abiertos gubernamentales en geolocalización, este trabajo propone el uso de métricas específicas para evaluar su contenido. Como parte de esta propuesta, se ha desarrollado un prototipo utilizando las últimas tecnologías informáticas disponibles en el mercado. Esta herramienta de validación, llamada OpenMapGeoData, permite detectar de manera más sencilla y automática las deficiencias y errores en los conjuntos de datos que podrían dificultar su interoperabilidad con diferentes fuentes de bases de datos y software externo utilizado por otros organismos.

---

**PALABRAS CLAVE:** DATOS ABIERTOS GEOLOCALIZADOS, GOBIERNO ABIERTO

## Evaluation of the processing of geolocated open data

### Abstract

In today's times, open government public data is a powerful tool with the potential to reduce information asymmetries, foster evidence-based decision-making, improve accountability, and generate impact on people's quality of life through its use. Governments in different countries

have an unprecedented opportunity to generate public value through the openness and use of data. The focus on open data is often the benefits it generates in terms of transparency and accountability. In order to promote and improve the quality of open government geolocation datasets, this paper proposes the use of specific metrics to evaluate their content. As part of this proposal, a prototype has been developed using the latest computer technologies available on the market. This validation tool, called OpenMapGeoData, makes it easier and more automatic to detect deficiencies and errors in datasets that could hinder their interoperability with different database sources and external software used by other agencies.

---

**KEYWORDS:** GEOLOCATED OPEN DATA, OPEN GOVERNMENT

## Introducción

Durante la última década, se ha producido un enorme incremento en la disponibilidad de datos gubernamentales abiertos, lo que ha generado oportunidades significativas en la investigación y la toma de decisiones en diversas áreas. Estos conjuntos de datos, accesibles al público y transparentes, contienen valiosa información sobre diversos aspectos de la sociedad, desde indicadores económicos hasta datos medioambientales y demográficos. Uno de los campos en los que los datos gubernamentales abiertos han demostrado su capacidad transformadora es la geolocalización.

La geolocalización implica asignar coordenadas geográficas a elementos o eventos específicos, permitiendo su representación en mapas y el análisis de su ubicación espacial. Los datos gubernamentales abiertos en geolocalización brindan una oportunidad sin precedentes para comprender patrones y fenómenos en un territorio determinado, lo cual facilita la planificación urbana, la gestión de recursos naturales, la toma de decisiones en situaciones de emergencia y muchas otras aplicaciones.

No obstante, a pesar del enorme potencial de los datos gubernamentales abiertos en geolocalización, es crucial asegurar su calidad y utilidad para los investigadores y usuarios finales. La precisión, integridad e interoperabilidad de estos conjuntos de datos son aspectos determinantes que deben abordarse adecuadamente para maximizar su valor y promover su uso efectivo.

Por lo tanto, este trabajo final de posgrado, siendo una línea futura de una tesis doctoral, (Martínez, 2021) se enfoca en analizar y mejorar la calidad de los datos gubernamentales abiertos en geolocalización. El objetivo principal de esta investigación es desarrollar un sólido marco metodológico para evaluar la calidad de estos conjuntos de datos y proponer mejoras en su contenido. Para lograr esto, se aplicarán técnicas avanzadas de validación derivadas de un análisis minucioso, considerando diversas métricas específicas para medir la precisión, integridad, consistencia y actualidad de los datos geolocalizados.

Además, como parte integral de esta investigación, se creará un innovador prototipo de herramienta informática llamada OpenMapGeoData, que facilitará la validación automática de los datos geolocalizados. Esta herramienta aprovechará las últimas tecnologías y algoritmos disponibles en el campo de la informática, y permitirá detectar de manera eficiente y precisa fallos y errores en los conjuntos de datos. Al proporcionar una forma más sencilla y automática de verificar la calidad de los datos geolocalizados, se fomentará su interoperabilidad y se reducirán las barreras para su uso en diversas bases de datos y aplicaciones externas.

Mediante un enfoque empírico respaldado por evidencia, esta investigación tiene como objetivo demostrar la viabilidad y efectividad de OpenMapGeoData como una herramienta de validación de datos gubernamentales abiertos en geolocalización. Se realizarán experimentos y pruebas detalladas utilizando conjuntos de datos reales para evaluar la capacidad de la herramienta para detectar y corregir errores, así como su impacto en la mejora de la calidad de los datos.

Los resultados de este estudio tienen el potencial de contribuir al avance de la investigación en el campo de los datos gubernamentales abiertos en geolocalización, así como proporcionar recomendaciones prácticas para mejorar la calidad y el uso de estos conjuntos de datos. La importancia de este trabajo radica en su capacidad para fortalecer la transparencia, la participación ciudadana y la toma de decisiones basada en evidencia en el ámbito gubernamental y académico.

## Problemas y soluciones

En el contexto de los datos abiertos gubernamentales, la geolocalización se ha convertido en un campo de gran relevancia y potencial. Sin embargo, a pesar de la disponibilidad de conjuntos de datos abiertos en geolocalización, existe un problema fundamental relacionado con la calidad y utilidad de estos datos para los investigadores y usuarios finales.

El problema radica en la falta de garantías en cuanto a la precisión, integridad e interoperabilidad de los datos abiertos gubernamentales en geolocalización. Aunque estos datos son accesibles de manera pública, su contenido puede presentar deficiencias y errores que comprometen su confiabilidad y limitan su aprovechamiento efectivo.

La precisión de los datos geolocalizados es crucial para una correcta interpretación y análisis. Errores en la ubicación espacial de los elementos o eventos registrados pueden llevar a conclusiones erróneas y decisiones equivocadas. Además, la integridad de los conjuntos de datos, es decir, la completitud y consistencia de la información geoespacial proporcionada, es esencial para garantizar una representación fidedigna de la realidad.

Otro desafío importante es la interoperabilidad de los datos abiertos en geolocalización. es fundamental que sean compatibles con diferentes fuentes de datos y sistemas externos. A pesar de ello, la falta de estándares comunes y la heterogeneidad en la estructura y formato de los datos dificultan su integración y uso eficiente.

Uno de los desafíos más importantes que han enfrentado los gobiernos radica en lograr una integración efectiva entre estos cambios y el desarrollo de bases de datos abiertos. Esto implica garantizar que los esfuerzos realizados generen beneficios comunes y se presenten de manera conjunta. Para abordar esta cuestión, se han propuesto diversas soluciones, entre las cuales se destaca el establecimiento de regulaciones administrativas que exigen a las agencias gubernamentales identificar los procesos que facilitan el cambio organizacional. Asimismo, se les insta a incorporar enlaces a las redes sociales con el fin de fomentar el intercambio de ideas y aprovechar el impulso que ha surgido en relación a la alfabetización digital.

No obstante, no todos los países se han encontrado con la posibilidad de desarrollar al máximo las potencialidades que ofrece el Gobierno abierto dadas las diferencias en desarrollo, cultura,

economía e instrumentación de una infraestructura para el aprovechamiento de las tecnologías de información. Así, en Latinoamérica, solo algunos Gobiernos se han dado a la tarea de construir un sistema donde se vinculan los datos abiertos gubernamentales, no sin afrontar diversas problemáticas y buscar soluciones ad hoc. Tal es el caso de Argentina, Brasil y México, los que hasta el momento han logrado avances importantes; aunque es claro que falta mucho por hacer y los retos económicos, políticos y sociales que enfrentan estos países afectan de forma significativa sus esfuerzos.

Además, la actualidad de los datos geolocalizados es un aspecto crítico a considerar. La dinámica de la sociedad y del entorno geográfico implica que los datos se actualicen constantemente. La falta de actualización de los conjuntos de datos abiertos gubernamentales en geolocalización puede llevar a información desactualizada y, por lo tanto, a decisiones incorrectas o ineficientes.

En su tesis doctoral, (Martínez, 2021), se destacan tres conceptos esenciales como pilares para abordar esta problemática:

**Comprensión del contexto de Gobierno Abierto:** implica el respaldo tanto del Estado Nacional como de los ciudadanos. Es fundamental entender y promover la colaboración y la participación activa de ambas partes en la apertura de datos gubernamentales.

**Promoción de datos abiertos:** se propone la apertura de diversos conjuntos de datos a través de plataformas y portales especializados. Estos datos abiertos constituyen una valiosa fuente de información y servicios, y su disponibilidad fomenta el desarrollo de aplicaciones y análisis de datos por parte de la comunidad.

En síntesis, el desafío que enfrentan los datos abiertos gubernamentales en geolocalización radica en su falta de calidad y utilidad. La falta de precisión, integridad, interoperabilidad y actualización comprometen su valor y restringen su impacto potencial en la investigación y la toma de decisiones. Es fundamental abordar esta problemática para maximizar el valor de los datos gubernamentales abiertos en geolocalización y promover su uso efectivo en diversas áreas, abarcando desde la planificación urbana hasta la gestión de recursos naturales y la respuesta a emergencias.

## Propuesta

La propuesta de este trabajo final se basa en la medición de calidad en los datos abiertos geolocalizados. El objetivo principal de esta propuesta es enriquecer la experiencia del lector al comprender y asimilar los conceptos presentados en este estudio.

La medición de calidad y la gestión de datos se refieren a un enfoque riguroso y sistemático para examinar y procesar conjuntos de datos, con el propósito de extraer información relevante y tomar decisiones fundamentadas. En el contexto de esta investigación, se utilizará esta evaluación de manera integral para diagnosticar la calidad de los datos gubernamentales abiertos en geolocalización.

Al aplicar mediciones y la gestión de datos como marco de investigación, se llevará a cabo un examen minucioso de los conjuntos de datos utilizando técnicas estadísticas y herramientas informáticas especializadas. Esto permitirá identificar patrones, tendencias y posibles errores en la información geoespacial proporcionada por los organismos gubernamentales.

Además, se implementarán estrategias eficientes de gestión de datos, que implican establecer procedimientos claros y consistentes para la recopilación, almacenamiento, procesamiento y análisis de los datos geolocalizados. Estas prácticas garantizarán la integridad de los conjuntos de datos y facilitarán su acceso y uso en investigaciones futuras.

Es importante destacar que la aplicación de esta medición de calidad y gestión de datos contribuirá a generar valor para el lector. Esto se logrará al proporcionar una comprensión más profunda y completa de los términos y conceptos presentados en este estudio. Además, el enfoque sistemático y riguroso utilizado en esta investigación permitirá obtener resultados confiables y fundamentados, lo que aumentará la credibilidad y relevancia de los hallazgos.

En resumen, el uso de mediciones de calidad y gestión de datos en este trabajo tiene como objetivo principal agregar valor al lector al facilitar la comprensión y asimilación de los conceptos propuestos. A través de un enfoque riguroso y sistemático, se examinarán los datos gubernamentales abiertos en geolocalización, utilizando técnicas y herramientas especializadas. La implementación de estrategias eficientes de gestión de datos garantizará la integridad y disponibilidad de los conjuntos de datos. En última instancia, se espera que esta propuesta genere un mayor valor al proporcionar información confiable y fundamentada, fortaleciendo así la relevancia y credibilidad de esta investigación.

## Marco teórico

Como indica en su libro, (Muenten-Kunigami, A., & Serale, F., 2018), en esta era digital, los datos se han convertido en el nuevo “oro negro” que impulsa el crecimiento económico y sustenta la llamada “economía de los datos”. En este contexto, los datos masivos y abiertos desempeñan un papel fundamental al mejorar la confianza de los ciudadanos en sus gobiernos y permitirles participar en la acción pública. La apertura de datos públicos ha dado lugar a una nueva concepción del Estado, más receptivo y centrado en las necesidades de los ciudadanos. En este sentido, es crucial que los gobiernos garanticen la confiabilidad de los datos y fomenten la colaboración para generar valor público, al tiempo que salvaguardan la confidencialidad de la información personal. Para que los datos abiertos sean verdaderamente valiosos, es necesario avanzar en la estandarización de la calidad, mejorar la accesibilidad y publicarlos en formatos fáciles de utilizar. De esta manera, se facilitará el acceso a la información y se mejorará la calidad de los servicios públicos.

No obstante, la apertura de los datos gubernamentales se enfrenta a diversos desafíos de naturaleza política, legal y técnica. Estos desafíos incluyen la garantía de la confiabilidad de los datos publicados, la protección adecuada de la privacidad de los individuos, la falta de conexión entre la oferta y la demanda de datos, así como la insuficiente evaluación del impacto de las iniciativas de datos abiertos.

A pesar de estos obstáculos, en América Latina existe un fuerte compromiso con la promoción de la agenda de datos abiertos. La región ha sido pionera en la implementación de iniciativas innovadoras en este ámbito. Varios países latinoamericanos, como Argentina, han logrado establecer regulaciones adecuadas para el acceso a la información pública, han puesto en marcha portales de datos abiertos y están desarrollando legislación en materia de protección de datos personales.

## Gobierno abierto y datos gubernamentales abiertos

Como lo detalla su estudio, (Nicolás, M. A., & Catachura, C. M., 2020), las discusiones relacionadas con el concepto de “gobierno abierto” tienen sus raíces en la década de 1950 en Estados Unidos, cuando se debatía en el Congreso la creación de una ley de acceso a la información abierta. En aquel entonces, la literatura y los estudios en el ámbito de la administración pública en Estados Unidos se centraban en mejorar los procedimientos para el funcionamiento de las políticas gubernamentales. Sin embargo, en tiempos más recientes, el concepto de “gobierno abierto” ha adquirido relevancia al abordar el papel del Estado en la promoción de la transparencia, la rendición de cuentas, la participación ciudadana y la innovación tecnológica en la gestión administrativa.

En su libro (Araujo, 2012), destaca que, si bien los términos ‘gobierno abierto’ y ‘datos gubernamentales abiertos’ pueden considerarse sinónimos, el concepto de gobierno abierto va más allá de la simple disponibilidad de datos. Implica tanto el acceso a “datos abiertos”, es decir, información gubernamental accesible a través de Internet, como diversas acciones que promueven la transparencia gubernamental y, en última instancia, la participación de la sociedad civil. En otras palabras, el concepto de gobierno abierto se materializa cuando los gobiernos se comprometen, ya sea a través de directrices o políticas, a ofrecer datos abiertos de calidad. El término “gobierno abierto” se originó en los debates sobre el acceso a la información pública en Estados Unidos en la década de 1950. Actualmente, el concepto engloba tanto la disponibilidad de datos gubernamentales abiertos como diversas acciones para promover la transparencia y la participación ciudadana. Es fundamental que los gobiernos asuman un compromiso real de ofrecer datos de calidad para materializar el verdadero espíritu del gobierno abierto.

Esos criterios apuntan a la idea de que el acceso a la información gubernamental por parte de los ciudadanos es fundamental para mejorar la democracia. Por lo tanto, se busca que los datos generados y proporcionados por los gobiernos sean accesibles a la población. Los datos de los gobiernos deben cumplir una serie de requisitos para ser considerados abiertos. En última instancia, el objetivo es viabilizar datos que sean de utilidad, es decir, que puedan ser analizados, comparados y procesados por los ciudadanos.

Las discusiones en torno al concepto de “gobierno abierto” han ganado impulso gracias a los debates generados dentro de una organización internacional conocida como “Open Government Partnership” (Asociación para el Gobierno Abierto). Esta iniciativa fue establecida en 2011 y contó con la participación de ocho países fundadores, entre los cuales se encontraba Brasil. Actualmente, la asociación cuenta con la membresía de 70 países.

En su creación, se aprobó la Declaración de Gobierno Abierto, la cual establece el compromiso de elaborar planes de acción nacionales que aseguren la implementación de acuerdos concretos para promover la transparencia, empoderar a los ciudadanos, combatir la corrupción y aprovechar las nuevas tecnologías con el objetivo de fortalecer la gobernanza. Esta declaración representa un hito importante en el avance de la agenda de gobierno abierto a nivel internacional.

La apertura de estos datos fomenta la transparencia al permitir a los ciudadanos examinar y comprender mejor las acciones y políticas gubernamentales. Además, los datos gubernamentales abiertos ofrecen oportunidades para el desarrollo de aplicaciones, herramientas y servicios que mejoran la calidad de vida de los ciudadanos y promueven la innovación.

Se busca fomentar la participación ciudadana y permitir que los ciudadanos puedan generar nueva información o aplicaciones útiles para la comunidad. Para lograr esto, es importante que los datos generados y proporcionados por los gobiernos estén accesibles para la población. Sin embargo, es necesario aclarar en qué sentido se consideran “abiertos” los datos. A continuación, se presentan los criterios para determinar la apertura de los datos:

1. Los datos deben estar completos: Los datos públicos no deben tener restricciones de privacidad u otras limitaciones. Además, deben estar almacenados electrónicamente.
2. Los datos deben ser primarios: Los datos deben estar disponibles en su forma original, sin procesamientos ni agregaciones.
3. Los datos deben ser oportunos: Los datos deben ser oportunos para preservar su valor.
4. Los datos deben ser accesibles: Los datos deben estar ampliamente disponibles para los usuarios y con diversos propósitos.
5. Los datos deben poder ser procesados: Los datos deben estar estructurados de manera adecuada para permitir su automatización mediante diferentes herramientas.
6. Acceso indiscriminado: Los datos deben estar disponibles para cualquier usuario sin necesidad de registro.
7. El formato de los datos debe ser no propietario: Los datos deben estar disponibles en un formato que no sea propiedad de ninguna entidad, evitando el control exclusivo.
8. Los datos deben estar libre de licencias; Según (Naser, 2011) los datos no deben estar sujetos a derechos de autor, patentes ni regulaciones secretas. El gobierno u otros organismos pueden aplicar medidas razonables de privacidad, seguridad y restricciones.

Según, (Nicolás, M. A., & Catachura, C. M., 2020), el “gobierno abierto” tiene potencialidades y limitaciones. En cuanto a las potencialidades, se destaca que el acceso a los datos gubernamentales contribuye a una gestión pública más participativa y al control social por parte de los ciudadanos. El uso de Internet por parte de los gobiernos se enfoca en la publicidad de la gestión y la transparencia de las acciones gubernamentales. Este proceso puede empoderar a los ciudadanos al brindarles acceso a información pública de calidad, lo que les permite ejercer una ciudadanía más activa. Además, se resalta que las potencialidades también pueden conducir a la creación de nuevas herramientas o aplicaciones basadas en datos gubernamentales, fomentando la innovación en la gestión pública. Los servicios pueden ser generados a través de asociaciones público-privadas que permitan a cualquier persona procesar los datos y crear contenido mediante su reutilización.

Por otro lado, (Janssen & Charalabidis, 2012), identifican diferentes limitaciones o barreras asociadas al gobierno abierto y los datos gubernamentales abiertos en geolocalización. Estas limitaciones se agrupan en cinco categorías:

1. Limitaciones institucionales: se refieren a los obstáculos relacionados con el suministro de datos por parte de los proveedores de servicios gubernamentales. Esto incluye la falta de políticas uniformes de divulgación de datos en diferentes niveles de gobierno.
2. Complejidad de uso: se refiere a las dificultades que enfrentan los usuarios al utilizar datos abiertos, como la falta de acceso a los datos, la falta de habilidades técnicas para manipularlos y la falta de información sobre el significado de cada dato.
3. Participación: implica el acceso y uso de datos por parte de los usuarios, como la necesidad de registrarse antes de descargar los datos.



4. Calidad de la información: se relaciona con la calidad de los datos proporcionados a los usuarios, como la presencia de información incompleta, datos obsoletos o no válidos, entre otros.
5. Barreras técnicas: se refieren a obstáculos tecnológicos en la provisión de datos, como la falta de normas, la falta de soporte para hacer que los datos estén disponibles y la ausencia de un software estándar para el procesamiento de datos.

En este trabajo final, se busca explorar y analizar el impacto del gobierno abierto y los datos gubernamentales abiertos en geolocalización en el contexto actual. A través de un estudio exhaustivo, se pretende obtener una comprensión más profunda de los principios, beneficios, desafíos y oportunidades asociados con la apertura de datos gubernamentales. El objetivo es desarrollar estrategias efectivas para implementar y utilizar de manera óptima los datos gubernamentales abiertos en beneficio de la sociedad. Además, se espera que esta investigación contribuya al desarrollo de políticas y prácticas que promuevan la transparencia, la participación ciudadana y la colaboración en el gobierno, mejorando así la calidad de la gobernanza pública en general.

## ¿Por qué es importante abrir los datos en geolocalización?

La apertura de datos en geolocalización se ha convertido en una práctica fundamental en la sociedad moderna. Los datos geolocalizados, que contienen información espacialmente referenciada, ofrecen un gran potencial para impulsar la transparencia, la innovación y la toma de decisiones basadas en evidencia.

En su estudio, (Boulos et al., 2011), remarcan que la apertura de datos geolocalizados genera una mejora de la planificación urbana y la toma de decisiones ya que le permite a los planificadores urbanos y a los responsables de la toma de decisiones acceder a información valiosa sobre la distribución de recursos, la infraestructura existente, el transporte, los servicios públicos y otros aspectos clave. Esto facilita la planificación estratégica, la identificación de necesidades y la implementación de políticas más eficientes y efectivas.

En tanto al estímulo de la innovación y el emprendimiento, (Kitchin, 2014), hace mención a la apertura de datos geolocalizados la cual proporciona una fuente de información rica y diversa para el desarrollo de aplicaciones y servicios innovadores. Al permitir que desarrolladores, emprendedores y empresas accedan y utilicen estos datos, se fomenta la creación de soluciones creativas que mejoran la vida de las personas y generan oportunidades económicas.

Asimismo, es importante tener en cuenta la respuesta a emergencias y gestión de desastres, por ello, (Goodchild et al., 2007), mencionan sobre los datos geolocalizados abiertos los cuales desempeñan un papel crucial en la respuesta a emergencias y la gestión de desastres. Proporcionan información actualizada sobre la ubicación de recursos, áreas afectadas, rutas de evacuación y otras variables importantes para la toma de decisiones en situaciones críticas.

La apertura de datos se ha vuelto cada vez más relevante en nuestra sociedad actual. La disponibilidad y accesibilidad de los datos gubernamentales abiertos tienen un impacto significativo en múltiples áreas, abarcando desde la gobernanza y la participación ciudadana hasta la innovación y el desarrollo económico. En particular, la apertura de datos geolocalizados desempeña un papel fundamental en nuestra sociedad actual. Estos datos nos brindan la



oportunidad de mejorar la planificación urbana, fortalecer la participación ciudadana, estimular la innovación y el emprendimiento, responder eficazmente a situaciones de emergencia y desastres, y fomentar la investigación científica y el desarrollo académico. Al abrir estos datos, se fomenta la transparencia en la gestión pública y se promueve la toma de decisiones basadas en evidencia. Además, se generan oportunidades para mejorar la calidad de vida de las personas al utilizar esta información para abordar problemas y desafíos sociales.

Es fundamental que reconozcamos el valor de los datos y trabajemos en conjunto para promover su apertura. Al hacerlo, podremos aprovechar plenamente su potencial y crear una sociedad más transparente, participativa e innovadora, en la que los datos se conviertan en un recurso valioso para el beneficio de todos.

### ¿Qué son los datos abiertos geolocalizados?

En la Carta Internacional de Datos Abiertos (De Luca, 2019) los datos abiertos son datos digitales que se ofrecen al público con las especificaciones técnicas y legales necesarias para permitir su uso, reutilización y redistribución de forma libre por cualquier individuo, en cualquier momento y en cualquier ubicación. En otras palabras, los datos abiertos son aquellos que están disponibles de manera accesible y sin restricciones, lo que permite a las personas aprovecharlos y aplicarlos a sus necesidades y preferencias, fomentando así la transparencia y la colaboración en el intercambio de información.

Los datos abiertos nos deben permitir a todos:

- Accederlos, a través de formatos digitales
- Utilizarlos, a través de la estructuración en formatos o estándares que faciliten la interoperabilidad.
- Compartirlos, a través de su certeza jurídica e integridad.

Los Datos Abiertos Geolocalizados son conjuntos de información digital que contienen datos específicos relacionados con ubicaciones geográficas y están disponibles de forma abierta y accesible para su uso y reutilización por parte de cualquier persona. Estos datos geolocalizados brindan información precisa sobre la posición geográfica de elementos como lugares, direcciones, puntos de interés, rutas, límites administrativos, entre otros.

Los datos abiertos geolocalizados constituyen una fuente invaluable de información espacial que puede ser utilizada por una amplia gama de actores, desde investigadores hasta desarrolladores de aplicaciones y ciudadanos interesados en explorar su entorno. En un estudio previo, (Moore, 2017), indica que los datos geolocalizados abiertos ofrecen una oportunidad única para mejorar nuestra comprensión de la realidad geográfica y fomentar la toma de decisiones informadas.

Por su parte, (Goodchild et al., 2007), detalla que los datos geolocalizados abiertos son el combustible que alimenta la revolución espacial. Su disponibilidad crea un ecosistema de innovación en el que los desarrolladores pueden crear aplicaciones y servicios que abordan desafíos urbanos, medioambientales y sociales. La disponibilidad de estos datos permite el desarrollo de aplicaciones y servicios que aprovechan el poder de la ubicación para brindar soluciones innovadoras.

La apertura de estos datos implica que están disponibles en formatos y estándares abiertos, lo que facilita su acceso, reutilización y redistribución sin restricciones innecesarias. Esto significa que cualquier individuo, empresa u organización puede utilizar estos datos para diferentes propósitos, como desarrollar aplicaciones móviles, crear mapas interactivos, analizar tendencias geográficas, realizar investigaciones académicas, entre otros. Estos abarcan una amplia variedad de fuentes, incluyendo información recopilada por gobiernos, instituciones públicas, organizaciones no gubernamentales, empresas privadas y la comunidad en general.

Además, los datos abiertos geolocalizados pueden contribuir al análisis espacial y la toma de decisiones basada en evidencia. En un estudio previo (Haklay et al., 2008), señala que la apertura de datos geolocalizados facilita la colaboración y el análisis conjunto de múltiples fuentes de información, lo que permite un mejor entendimiento de los fenómenos espaciales y una toma de decisiones más fundamentada.

La importancia de los Datos Abiertos Geolocalizados radica en su potencial para impulsar la innovación, mejorar la toma de decisiones y promover la transparencia. Al ser accesibles y reutilizables, estos datos permiten a los investigadores, profesionales, emprendedores y ciudadanos en general aprovechar su potencial para generar nuevos conocimientos, impulsar el desarrollo económico y mejorar la calidad de vida de las personas.

Asimismo, los Datos Abiertos Geolocalizados fomentan la colaboración y el intercambio de información entre diferentes actores, lo que puede llevar a soluciones más eficientes y efectivas en áreas como el transporte, la planificación urbana, la gestión de emergencias, el turismo, la agricultura, la salud y el medio ambiente. Al poner estos datos a disposición del público, se fomenta la participación ciudadana, se fortalece la rendición de cuentas y se promueve una mayor comprensión y análisis de los fenómenos geográficos.

En resumen, los Datos Abiertos Geolocalizados representan un valioso recurso para la investigación académica, la toma de decisiones y la innovación en diferentes ámbitos. Su apertura y accesibilidad fomentan la colaboración, la transparencia y el desarrollo de soluciones basadas en evidencias, lo que contribuye a una sociedad más informada y participativa en la que se pueden aprovechar al máximo las ventajas de la información geográfica.

## Calidad de los datos abiertos

La calidad de los datos abiertos geolocalizados es un aspecto fundamental que debe abordarse debido a su impacto en la precisión y confiabilidad de la información geográfica disponible. Es importante destacar este tema por varias razones:

- **Toma de decisiones sólidas:** La calidad de los datos es crucial para tomar decisiones informadas y confiables. Si los datos geolocalizados son imprecisos, incompletos o inconsistentes, las decisiones basadas en ellos pueden ser incorrectas o sesgadas. Por lo tanto, es esencial garantizar la calidad de los datos para asegurar que las decisiones se basen en información confiable.
- **Confianza en la información:** La calidad de los datos abiertos geolocalizados es esencial para generar confianza tanto en los usuarios como en los proveedores de datos. Los usuarios

confían en que la información geográfica proporcionada es precisa y actualizada, mientras que los proveedores de datos se esfuerzan por mantener altos estándares de calidad para asegurar la credibilidad de sus datos.

- **Interoperabilidad y reutilización:** Los datos geolocalizados de alta calidad son más propensos a ser interoperables, lo que significa que pueden integrarse y combinarse fácilmente con otros conjuntos de datos. Esto permite una mayor reutilización de los datos en diferentes aplicaciones y contextos, lo que fomenta la innovación y la generación de valor agregado.
- **Mejora de la experiencia del usuario:** Los datos de calidad mejoran la experiencia del usuario al utilizar aplicaciones o servicios que se basan en información geográfica. Los usuarios pueden confiar en la precisión de los mapas, las direcciones y los puntos de interés, lo que mejora su satisfacción y usabilidad.
- **Evitar sesgos y discriminación:** La calidad de los datos también es relevante para abordar posibles sesgos o discriminación en la información geográfica. Si los datos están incompletos o contienen sesgos implícitos, como la falta de representación equitativa de ciertas áreas o grupos de población, puede haber consecuencias negativas en la toma de decisiones y en la equidad espacial.

## Propuesta de datos abiertos en geolocalización

Según (Datos Gob AR, 2023) para garantizar la coherencia y la interoperabilidad de los datos a nivel internacional, es importante seguir los estándares establecidos para los nombres y códigos de países o territorios. El uso del estándar ISO 3166-1, (ISO 3166-1, 2013) es altamente recomendado y se sugiere incluir un campo con el código alfabético de 3 dígitos del estándar Código alfa-3, (ISO 3166-1, 2013) y otro con el nombre completo del país en español. Para facilitar esto, se pueden utilizar los “Nombres de uso común” de la lista de países y sus códigos alfa-3 publicados por INDEC, (Instituto Nacional de Estadística y Censos, 2023).

Es importante destacar que los nombres oficiales de los países pueden variar según el idioma y la cultura, lo que puede generar confusiones y dificultades en el intercambio de datos. Por ejemplo, en Argentina, nos referimos al país vecino como “Brasil”, mientras que su nombre oficial en portugués es “República Federativa do Brasil” y la traducción oficial en español es “República Federativa del Brasil”. Al usar el código de país según el estándar (ISO 3166-1, 2013), en este caso “BRA”, se resuelve este problema de denominación y se asegura la interoperabilidad.

Para facilitar la comprensión y el uso de los datos, se recomienda utilizar nombres de campos claros y coherentes. En el caso de los códigos de país, se sugiere utilizar el nombre “pais\_id” o, si hay más de un campo de país en el conjunto de datos, se recomienda que cada campo finalice con “pais\_id” (por ejemplo, “pais\_origen\_id”, “pais\_destino\_id”). Además, el campo con el nombre completo del país debe ser llamado “pais\_nombre”. De esta manera, se asegura una fácil interpretación y uso de los datos por parte de los usuarios.

## Métricas

A continuación, se detallan las métricas principales propuestas a utilizarse como referencia para determinar si los conjuntos de datos cumplen con las condiciones necesarias para considerarse

de calidad en geolocalización. Estas métricas incluyen el “Tratamiento de números decimales”, el “Tratamiento de la propiedad Geometry”, el “Tratamiento de la propiedad Coordinates”, la presencia de “Datos faltantes y/o incompletos” y la existencia de “Datos falsos”.

## Tratamiento de números decimales

En este contexto se aborda la necesidad crítica de realizar una evaluación exhaustiva de la propiedad “coordinates” en los datos geoespaciales utilizados. Específicamente, se busca asegurar que los valores numéricos presentes en dicha propiedad cumplan con un requisito de precisión mínima, donde se exija al menos dos dígitos después del punto decimal.

La precisión de los datos geoespaciales es un factor crucial para garantizar la exactitud y confiabilidad de los resultados y análisis que se derivan de ellos. Al tener un número decimal con al menos dos dígitos después del punto decimal en la propiedad “coordinates”, se logra una mayor fidelidad y detalle en la representación de la ubicación geográfica en el mapa.

La evaluación de la propiedad “coordinates” implica un proceso riguroso de verificación y validación de los datos. Se implementarán algoritmos y técnicas especializadas para analizar y examinar cada valor numérico presente en la propiedad. Estos algoritmos realizarán un análisis de patrones y comprobarán si se cumple el requisito mínimo de dos dígitos después del punto decimal. En caso de que algún valor no cumpla con esta especificación, se generarán alertas o se tomarán medidas para corregir o eliminar dichos datos inconsistentes.

La aplicación de esta evaluación de precisión en los datos geoespaciales contribuirá significativamente a la calidad y confiabilidad de los resultados obtenidos en el contexto de esta investigación. Los mapas generados a partir de estos datos contarán con una precisión adecuada y una representación más fiel de la realidad geográfica, lo que mejorará la toma de decisiones informadas y la comprensión de los fenómenos espaciales estudiados.

Además de los beneficios directos en el ámbito de este trabajo final, los resultados y hallazgos obtenidos de esta evaluación podrían tener implicaciones más amplias en el campo de la cartografía y los sistemas de información geográfica. La consideración y aplicación de estándares de precisión en los datos geoespaciales pueden mejorar la interoperabilidad, la comparabilidad y la integración de información geográfica de diferentes fuentes y sistemas.

## Tratamiento de la propiedad geometry

En el marco de este trabajo final, se aborda la necesidad de realizar una evaluación exhaustiva de la propiedad “geometry” que se encuentra presente en el documento de estudio. Esta propiedad desempeña un papel fundamental, ya que contiene las coordenadas geográficas necesarias para identificar y localizar de manera precisa la ubicación geográfica de los datos analizados. Por lo tanto, es de vital importancia llevar a cabo una verificación minuciosa de esta propiedad para asegurar la integridad y fiabilidad de la geolocalización de la información.

La propiedad “geometry” es un componente esencial en el ámbito de los datos geoespaciales, pues proporciona los elementos fundamentales para representar la ubicación geográfica de los objetos o eventos estudiados. Estos datos geoespaciales son ampliamente utilizados en diversas disciplinas, como la geografía, la planificación urbana, la gestión de recursos naturales y el análisis de riesgos, entre otros. La precisión y exactitud de la información geoespacial son factores críticos para obtener resultados confiables y respaldar la toma de decisiones informada.

La verificación minuciosa de la propiedad “geometry” permitirá garantizar la calidad y confiabilidad de los datos geoespaciales utilizados en esta investigación. Asimismo, contribuirá a la generación de resultados más precisos y representativos, lo que fortalecerá la validez y relevancia de los hallazgos obtenidos.

### Tratamiento de propiedad coordinates

En el marco de este trabajo final, se aborda la necesidad de realizar una evaluación rigurosa de la propiedad “coordinates” en los datos geoespaciales utilizados. Esta propiedad desempeña un papel crucial, ya que contiene información esencial sobre la geolocalización de los objetos o eventos estudiados.

La evaluación rigurosa de la propiedad “coordinates” implica llevar a cabo un análisis detallado de los datos geoespaciales. Esto implica verificar la consistencia y exactitud de las coordenadas presentes en la propiedad, asegurando que cumplan con los estándares y las convenciones establecidas en el ámbito de la geolocalización. Además, se examinará la coherencia espacial de las coordenadas, asegurando que reflejen correctamente la ubicación geográfica deseada.

### Datos faltantes o incompletos

En el contexto actual, en el que la disponibilidad y precisión de los datos juegan un papel crucial en la toma de decisiones informadas, es imprescindible llevar a cabo una evaluación de los datos nulos.

A fin de abordar adecuadamente la clasificación de los datos faltantes o incompletos, se estableció una métrica basada en criterios claros y definidos. En este sentido, se consideran casos en los que los valores indicados pueden presentarse como “N/D” (no disponible), “N/A” (no aplicable), “NULL” (vacío), “-” (guion), “- -” (dos guiones) o bien “--” (dos guiones consecutivos). Estos indicadores nos permiten identificar los campos que carecen de información y requieren una verificación detallada.

La identificación y clasificación precisa de los datos nulos resulta crucial para llevar a cabo un análisis completo y confiable. Al abordar los casos mencionados anteriormente, se establece un enfoque riguroso que nos permite comprender la magnitud de los datos faltantes y su posible impacto en los resultados de la investigación. Al examinar detenidamente cada uno de estos casos, podemos determinar si los valores ausentes son el resultado de errores en la recopilación de datos, problemas de transmisión o simplemente la falta de disponibilidad de la información requerida.

Una vez que se han identificado los datos nulos, es necesario emprender medidas correctivas para garantizar la integridad y la completitud de los mismos.

## Datos falsos

Es de vital importancia validar los datos incorrectos en este contexto, ya que estos contienen información necesaria para determinar la ubicación geográfica precisa de los datos.

La identificación y clasificación de los datos incorrectos se aborda mediante criterios claros y bien definidos. En este sentido, se consideran aquellos casos en los que los valores proporcionados difieren de una representación numérica adecuada. Estos casos abarcan cualquier información que no se asemeje a un número válido, como caracteres alfabéticos, símbolos no numéricos o formatos incompatibles con la naturaleza de los datos.

La presencia de datos incorrectos puede tener un impacto significativo en los resultados y conclusiones obtenidos a partir de un análisis. Por lo tanto, es esencial identificar y abordar estos datos de manera exhaustiva. La clasificación y validación de los datos incorrectos nos permiten evaluar la calidad y la confiabilidad de la información utilizada en nuestro estudio, así como determinar la magnitud de su efecto en los resultados finales.

## Implementación

Se describirá el proceso de implementación de estas métricas mediante la creación de una nueva herramienta desarrollada específicamente para tal fin. La herramienta ha sido nombrada "OpenMapGeoData". Esta herramienta utiliza la tecnología Angular en su versión 14 (LTS) para el frontend. Esta aplicación se encarga de interactuar con los usuarios y proporcionarles una interfaz amigable para la recolección de datos de entrada. Por otro lado, se ha creado una API utilizando la tecnología NodeJS en su versión 18 (LTS) para el backend.

Con la utilización de OpenMapGeoData se llevarán a cabo diversas pruebas en datasets seleccionados de algunos de los portales gubernamentales de datos abiertos más relevantes de Argentina. Estos datasets han sido cuidadosamente seleccionados y formarán parte de una recolección de datos definida previamente como muestra.

El objetivo de estas pruebas es asegurar la calidad de los datos abiertos, mediante la utilización de las métricas de calidad.

OpenMapGeoData permitirá llevar a cabo estas pruebas de manera eficiente y efectiva, asegurando la precisión de los resultados obtenidos.

La funcionalidad general del software OpenMapGeoData consiste en seleccionar un archivo dataset del tipo de formato Json/GeoJson y efectuar la validación de las métricas propuestas.

- Las funcionalidades detalladas de la herramienta son:
- Identificación del cumplimiento de las métricas críticas;
- Detección y detalle de los casos que no concuerdan con el formato válido para el tipo de dato decimal;
- Cálculo estimativo de los tipos de datos de los campos del dataset validado;
- Identificación de la cantidad de casos que poseen campos con registros Vacíos (Sin Datos y con espacios en los campos).

- Visualización del detalle de los casos de registros con decimales post coma que n la validación.
- Identificación del formato correspondiente.
- En la Bitácora se proporcionará un análisis detallado de los resultados de la validación de los datos del dataset, incluyendo información sobre los campos que no han aprobado las validaciones. Se ofrecerán filtros de búsqueda para permitir a los usuarios buscar en el dataset según los criterios que elijan, y también se podrán buscar palabras clave dentro de los campos que no han aprobado las validaciones. De esta manera, se proporcionará un nivel adicional de información y análisis a los usuarios para que puedan comprender mejor la calidad de los datos que están utilizando.
- Creación de un Mapa utilizando las tecnologías de OpenStreetMap y la inclusión de marcadores que indican la ubicación de los puntos cardinales seleccionados en el archivo validado. Además, se mostrará el detalle correspondiente a cada uno de estos puntos para que el usuario pueda obtener información adicional. De esta forma, se podrá visualizar de manera clara y precisa la información geográfica contenida en el archivo validado en un mapa interactivo y detallado.
- Implementación de una funcionalidad de generación de mapas que utiliza tecnologías de OpenStreetMap, las cuales permiten la visualización de rutas a partir de la información validada y procesada. De esta manera, se habilita la generación de mapas detallados y precisos que muestran la información geoespacial relacionada con el conjunto de datos validado, lo que permite una mejor comprensión de la información y su posible análisis.
- Implementación de una funcionalidad de generación de mapas que utiliza tecnologías de OpenStreetMap, las cuales permiten la visualización de polígonos a partir de la información validada y procesada.

## Conclusiones

En la actualidad, el concepto de Gobierno Abierto tiene un papel fundamental en la conexión entre organismos gubernamentales y ciudadanos, lo que permite una mayor transparencia y participación. Las Tecnologías de la Información y la Comunicación (TICs) son fundamentales en la gestión de los datos abiertos públicos en geolocalización, y es importante verificar constantemente los conjuntos de datos compartidos en los portales web gubernamentales. Por eso mismo, es esencial contar con guías, métricas y herramientas para mejorar la calidad de los datos de geolocalización y validarlos rápidamente para evitar errores y garantizar su confiabilidad.

La herramienta OpenMapGeoData, desarrollada y propuesta en este estudio, expone diversas métricas para analizar la calidad de los conjuntos de datos en geolocalización y mitigar posibles errores en su tratamiento, lo que es importante para la interoperabilidad del software y la reutilización de datos abiertos. Esta herramienta podría utilizarse en los portales de las diferentes agencias gubernamentales para validar sus datos antes de ser publicados y así mejorar su calidad.

Es esencial tener en cuenta la calidad y confiabilidad de los datos abiertos gubernamentales para que los ciudadanos y las organizaciones puedan monitorear activamente las políticas públicas involucradas. Aunque existen portales de datos abiertos que mantienen actualizados sus conjuntos de datos, esto no asegura que la información obtenida sea completa y correcta. Por lo tanto, es necesario realizar un análisis periódico del estado de los países que se guían por el enfoque de Gobierno Abierto para mejorar la calidad de los datos ofrecidos al público.



Es importante compartir buenas prácticas, guías de aprendizaje, foros y herramientas para permitir un mejor análisis y tratamiento de los datos abiertos. La herramienta OpenMapGeoData permite validar rápidamente el estado de un conjunto de datos geolocalizables y analizar sus métricas para garantizar su calidad y confiabilidad.

## Referencias bibliográficas

- » ISO 3166-1. (2013). ISO 3166-1:2013. Obtenido de Codes for the representation of names of countries and their subdivisions: <https://www.iso.org/standard/63545.html>
- » Araujo. (2012). Desafíos comunes Tomo I: Retrato de la sociedad chilena y sus individuos. LOM Ediciones.
- » Boulos et al. (2011). Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management. *International journal of health geographics*.
- » Chen et al. (2012). Smart city and the applications. *International Conference on Consumer Electronics, Communications and Networks*.
- » Datos Gob AR. (27 de 05 de 2023). Datos Argentina. Obtenido de <https://datosgobar.github.io/paqueteapertura-datos/guia-interoperables/#divisiones-o-unidades-territoriales-internas>
- » De Luca, A. E. (2019). Carta Internacional de Datos Abiertos. Quito, Ecuador: Flacso Ecuador.
- » Goodchild et al. (2007). Public participation in geographic information systems: A literature review and framework. *Association of American Geographers*.
- » Haklay et al. (2008). *Web mapping 2.0: the neogeography of the GeoWeb*. Londres: Geography Compass.
- » Instituto Nacional de Estadística y Censos. (27 de 05 de 2023). INDEC. Obtenido de INDEC: <https://www.indec.gob.ar/>
- » Janssen, M., & Charalabidis, Y. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management. ISM*.
- » Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*.
- » Martínez, M. I. (2021). Métricas de calidad para validar los conjuntos de datos abiertos públicos gubernamentales. Buenos Aires, Argentina: Universidad Nacional de la Plata.
- » Moore, K. &. (2017). Open data, open cities: Open government data and the reconfiguration of urban governance. *GeoJournal*, 757-765.
- » Munte-Kunigami, A., & Serale, F. (2018). *Los datos abiertos en América Latina y el Caribe*. Banco Interamericano de Desarrollo.
- » Naser, A. &. (2011). *El gobierno electrónico en la gestión pública*. Cepal.
- » Nicolás, M. A., & Catachura, C. M. (2020). Gobierno abierto: análisis de websites de datos abiertos gubernamentales en Argentina, Brasil y Paraguay. *Desenvolvimento em Debate*.