

Vinculación de portales abiertos mediante API

Juan Ignacio Torres^{ORCID}, Ariel Pasini^{ORCID}, Silvia Esponda^{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)
{jitorres, apasini, sesponda}@lidi.info.unlp.edu.ar

Abstract. El presente trabajo se desarrolla en el marco de la Práctica Profesional Supervisada de la carrera de ATIC (Analista en Tecnologías de la Información y la Comunicación) de la Facultad de Informática - UNLP. En los últimos años, los gobiernos buscan generar herramientas que involucren a los ciudadanos a participar de forma activa en las decisiones de gobierno. Una forma de incrementar la participación ciudadana es brindar, a la comunidad, acceso a la información y permitir que ellos mismos sean capaces de generar nuevos aportes que asistan a las decisiones del gobierno. Los diferentes niveles de gobierno ponen al alcance de los ciudadanos una gran cantidad de información en formatos abiertos. Sin embargo, pueden existir incompatibilidades a la hora de acceder y analizar esta información. Por lo que se propone desarrollar una herramienta que permita compatibilizar diferentes fuentes de datos disponibles mediante APIs, de forma sencilla para el ciudadano, facilitando la comparación, el procesamiento y análisis de la información de manera generar valor agregado.

Keywords: Datos Abiertos, Gobierno abierto, API, vinculación de datos.

1. Introducción

En los últimos años se popularizó un nuevo paradigma de gestión pública, llamado gobierno abierto, el cual se basa en el acceso a la información por parte del ciudadano permitiendo controlar y/o generar nuevos aportes al gobierno. Un ejemplo común en los diferentes niveles de gobierno es la búsqueda de transparencia en las cuentas públicas, poniendo a disposición de los ciudadanos información tal como balances, contrataciones, licitaciones, etc. para que ellos puedan ver el destino de los fondos públicos. Por otro lado, los ciudadanos podrían tomar dicha información y analizarla desde diferentes puntos de vista generando aportes constructivos al gobierno. Este proceso genera mayor transparencia en la gestión y en consecuencia mayor confianza del ciudadano.

Para lograr mayor transparencia ante sus ciudadanos, los diferentes niveles de gobierno ponen una gran cantidad de datos públicos en formatos abiertos a disposición de los ciudadanos de manera tal que éstos analicen y procesen la información para

generar valor agregado, apoyando el proceso de toma de decisiones de manera inteligente.

Sin embargo, al obtener información de distintos proveedores (ya sean agencias gubernamentales, organizaciones internacionales u ONGs) o incluso de un mismo proveedor mediante APIs, se pueden encontrar problemas de compatibilidad en los formatos de las respuestas, como por ejemplo diferentes unidades.

Es por eso, que, mediante el uso de una herramienta que vincule las APIs de las fuentes de datos, se buscará dar una solución a estos problemas, de manera tal que se pueda generar un nuevo dataset, en el cual los datos sean comparables y analizables, permitiendo a los usuarios participar en el proceso de generación de valor agregado.

Este trabajo se desarrolla en el marco de la Práctica Profesional Supervisada de la carrera de ATIC (Analista en Tecnologías de la Información y la Comunicación).

2. Problemática

A nivel global, los ciudadanos esperan cada vez más de sus gobiernos y gobernantes. Estas demandas, junto con la revolución tecnológica generaron un cambio drástico en la interacción entre un gobierno y sus ciudadanos. El uso inteligente de la información registrada por los gobiernos en formatos abiertos permite mejorar la calidad de vida de los ciudadanos. Estos datos abiertos, pueden ser accedidos mediante la comunicación con una API.

No obstante, puede suceder que cuando se intenta acceder a esta información, procedente de diferentes proveedores o incluso de un mismo proveedor, el usuario no esté conforme con el formato de los datos recibidos.

Además, pueden existir diferencias en los formatos de respuesta o incompatibilidad de unidades entre las distintas fuentes de datos. Un ejemplo de estas diferencias se evidencia en la extensión de las líneas de subte en Argentina y en los Estados Unidos, donde en Argentina la extensión se mide generalmente en kilómetros, mientras que en los Estados Unidos se mide en millas.

En algunos casos, los usuarios podrían requerir operaciones sobre los datos de cada una de las fuentes, como filtrarlos, agruparlos bajo algún criterio, o simplemente realizar operaciones aritméticas.

Aplicando las operaciones requeridas por quien analice los datos y solucionando las diferencias mencionadas, se podría facilitar la comparación y el procesamiento de la información, permitiendo así la generación de valor agregado por y para parte de los ciudadanos.

3. Desarrollo propuesto

Se desarrollará una herramienta que permitirá brindar una solución a la incompatibilidad de formato y/o unidades entre fuentes de datos accesibles mediante APIs, que el usuario desee comparar y analizar.

Esta herramienta, pretende estandarizar la información a comparar mediante una conversión del formato de respuesta (en caso incompatibilidad) o un factor de

corrección (en el caso de incompatibilidad de unidades), generando un nuevo dataset que contendrá información significativa para el análisis del usuario.

Para lograr este resultado, es necesario pasar por un proceso de 4 pasos:

1. Búsqueda de la información: el usuario selecciona las fuentes de datos disponibles mediante comunicación vía API que sean de su interés a través de portales de datos abiertos.
2. Conexión con la API: Una vez elegidas las fuentes de datos, el usuario cargará en la herramienta las URL correspondientes. La comunicación con las APIs se realizará mediante HTTPS, y en caso de ser correcta la conexión, se obtendrán los datos en formato JSON.
3. Análisis previo: El usuario podrá visualizar en paralelo la información elegida en cada dataset para su análisis. En esta etapa, podrá realizar diferentes operaciones con los datos (como por ejemplo filtrado o agrupamiento). Posteriormente, seleccionará las columnas a comparar y el factor de corrección a aplicar (en caso de que haya problemas de compatibilidad en las unidades).
4. Visualización: una vez seleccionadas las columnas a comparar, las operaciones aplicadas y el factor de corrección en caso de ser necesario, se podría generar un nuevo dataset que contenga la información a comparar de manera estandarizada. El usuario podrá ver en pantalla esta información y podrá descargar un archivo .csv para su futuro análisis y procesamiento.

Al finalizar este proceso, el usuario obtendrá información comparable y procesable, solucionando así la incompatibilidad de formato y/o unidades.

Volviendo al ejemplo de la extensión de las líneas de subte, en caso del usuario haber elegido ver la información medida en kilómetros, obtendrá un dataset con la extensión de las líneas de Argentina y Estados Unidos medidas en kilómetros.

4. Líneas de investigación futura

El principal desarrollo a futuro es la implantación de un sitio de gestión para la vinculación de diferentes portales de datos abiertos que provea datos mediante API, que permita cargar y almacenar la información de conexión con las API de los diferentes portales de datos abiertos para que el ciudadano los pueda vincular de manera sencilla.

5. Bibliografía básica

- Attard, J., Orlandi, F., Scerri, S., Auer, S. (2015). A Systematic Review of Open Government Data Initiatives. *Government Information Quarterly*. 32. 10.1016/j.giq.2015.07.006.
- Braunschweig, K., Eberius, J., Thiele, M., & Lehner, W. (2012). The state of open data. *Limits of Current Open Data Platforms*.
- Greco, A. O. (2020). "IndiMaker: Una herramienta para la construcción de indicadores personalizados en tableros de control de SGC". *Tesina de Licenciatura en Sistemas, Facultad de informática, UNLP*. <http://sedici.unlp.edu.ar/handle/10915/118501>
- J. S. Preisegger, A. Greco, A. C. Pasini, M. Boracchia, and P. M. Pesado, Marco de vinculación de datos abiertos aplicado al contexto de datos medioambientales. *Actas del XXVI Congreso*

- Argentino de Ciencias de la Computación (CACIC 2020), ISBN: 978-987-44-1790-9, págs. 756-766, 2020.
- Janssen, M., Charalabidis, Y., Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government, *Information Systems Management*, 29:4, 258-268, DOI: 10.1080/10580530.2012.716740
- Lathrop, D., Ruma, L. (2010). *Open government: Collaboration, transparency, and participation in practice*. United States of America: O'Reilly Media, Inc.
- Naser, A., Ramirez, A., Rosales D: Desde el gobierno abierto al Estado abierto en América Latina y el Caribe. Libros de la CEPAL - Planificación para el Desarrollo No. 144.
- 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>.
- Ubaldi, B. (2013), "Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives", OECD Working Papers on Public Governance, No. 22, OECD Publishing, Paris, <https://doi.org/10.1787/5k46bj4f03s7-en>
- Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, 31(1), 17–29. <https://doi.org/10.1016/j.giq.2013.04.003>