

EDITORIAL

Oportunidades en datos abiertos

El concepto *código abierto* surgió como una forma de describir la idea que “el acceso libre y compartido al código fuente de un programa permite mejorarlo y, por consiguiente, evolucionar”. Esta forma de colaboración ha ido creciendo hasta convertirse, actualmente, en una *cultura* que permea en otros ámbitos de la sociedad como es el *gobierno abierto*. Este último concepto corresponde a la influencia que la búsqueda del bien común tiene sobre las instituciones y sus gobernantes, así como la relación de estas con el desarrollo tecnológico, económico y humano de las sociedades.

Los *datos abiertos* son antecedentes que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona. Se encuentran sujetos, a lo más, al requerimiento de atribución y/o de compartirse de la misma forma en que aparecen originalmente. En esta editorial damos una mirada a cómo la utilización de *datos abiertos*, provenientes desde distintas fuentes, contribuye al análisis de eventos, por ejemplo territoriales, con el fin de construir, entre otros, indicadores que aporten a atender problemas de planificación urbana, segregaciones socioespaciales, pobreza territorial y valoración de la efectividad de las políticas públicas en general.

Los datos provenientes desde ministerios, intendencias, gobernaciones, municipios, universidades y colegios, entre otros, están asociados a una gran diversidad de información “pública” (capital de propiedad compartida) que puede ser de interés a variados usuarios. Estos antecedentes gozan, en su mayoría, de confiabilidad estadística, al disponerse de los documentos metodológicos de su obtención. Esta cualidad se agradece, pues tanto en economía como en otras áreas la escasez e inexistencia de estadísticas en niveles microterritoriales se abordan, principalmente, con técnicas de predicción y/o desagregación.

A nuestro parecer, la importancia de los datos abiertos radica en un conjunto de razones que describe la Open Knowledge Foundation y que ligan a este concepto con el de gobierno abierto:

1. **Transparencia.** En una sociedad democrática, adecuadamente funcional, los ciudadanos necesitan saber lo que su gobierno está haciendo. Para ello deben ser capaces de acceder libremente a los datos e información gubernamentales y de compartir esta información con los demás ciudadanos.
2. **Aporta valor social y comercial.** En la era digital, los datos son un recurso clave para las actividades sociales y comerciales. Por medio de la apertura de los datos, los gobiernos pueden ayudar a impulsar la creación de empresas innovadoras y servicios que aportan valor social y comercial.
3. **Gobierno participativo.** Por medio de la apertura de los datos, los ciudadanos están facultados para estar más directamente informados e involucrados en la toma de decisiones. Esto es más que la transparencia: es acerca de construir una cabal sociedad de “lectura/escritura”, no solo estando consciente de lo que ocurre en el proceso de gobierno, sino que siendo capaz de contribuir a él.

Desde el 2009 Chile cuenta con una ley específica sobre el acceso a la información pública, incorporándose a la Alianza de Gobierno Abierto. Este organismo internacional requiere que los países miembros desarrollen políticas en cuatro áreas prioritarias: la promoción de estándares de integridad profesional, aumento en el acceso a las nuevas tecnologías para la apertura y rendición de cuentas, el acceso a la información sobre las actividades gubernamentales y el apoyo a la participación ciudadana por medio de la transparencia sobre la formulación de políticas y la toma de decisiones. Como ejemplo, el acceso a la

información de actividades gubernamentales puede realizarse en la plataforma *datos.gob.cl*, que dispone de datos publicados por diferentes instituciones públicas bajo un esquema estandarizado que facilita su análisis comparado. Asimismo, la Biblioteca del Congreso Nacional también dispone su información en el sitio web *datos.bcn.cl*.

El desarrollo de soluciones provenientes de la Ingeniería de Datos, la Inteligencia de Negocios y los Sistemas de Información Geográfica (SIG) ha permitido a los usuarios obtener nueva información, diferente a la fijada inicialmente por quienes proveen los datos. Esto se logra mediante el cruce, la clasificación y el aprendizaje o integración de datos desde distintas fuentes. En este ámbito, nuestro equipo ha abordado una línea de trabajo que integra tecnologías y datos abiertos desde las siguientes aplicaciones:

- A partir de la evolución de los datos, contrastando información y analizando, por ejemplo, los efectos que las políticas públicas, con el tiempo, han producido sobre una cierta población.

En este ámbito se ha trabajado en integrar la información de los censos nacionales de personas, viviendas y hogares en Chile a partir de 1992. Para esto nos hemos apoyado en el uso de tecnología de Bases de Datos y Extracción, Transformación y Carga (ETL).

- En segundo lugar, es interesante situar estos datos espacialmente, lo que implica observar, por ejemplo, cómo los cambios en el entorno han afectado a una población, y cómo los cambios en una población han afectado a su entorno.

En este ámbito se ha trabajado en situar espacialmente la información del censo utilizando, esencialmente, tecnología de Bases de Datos Espaciales, SIG y Servidores de Mapas Web.

- Finalmente, sostenemos que la integración de datos provenientes de diferentes fuentes produciría un efecto de sinergia, permitiéndonos ver aspectos que se encontraban ocultos de las miradas individuales.

En este ámbito se ha trabajado en integrar información de los censos en Chile con la Encuesta de Caracterización Social y Económica (CASEN). Para esto se ha utilizado una técnica de extrapolación de datos denominada Matching Espacial, que busca en la encuesta una (o más) observación(es) en el censo que es (son) estadísticamente similar para extrapolar los datos en unidades microterritoriales no disponibles en la encuesta. La solución desarrollada ha incorporado tecnología ETL, almacenes de datos, procesamiento analítico en línea (OLAP, en inglés) y *dashboards*.

A futuro, el uso de tecnologías como la minería de datos, la inteligencia computacional u otras afines se ven como potenciales alternativas de ser consideradas.

Como hemos dicho, las oportunidades están. Analizar datos abiertos de diferentes fuentes, especialmente gubernamentales, es posible. Depende de nosotros obtener un uso eficiente de ellos.

Ricardo Valdivia

Escuela Universitaria de Ingeniería
Industrial, Informática y Sistemas
Universidad de Tarapacá
Arica-Chile
rvaldivi@uta.cl

Mónica Navarrete

Escuela Universitaria de
Administración y Negocios
Centro de Estudios
Regionales (CEUTA)
Universidad de Tarapacá
Arica-Chile
mnavarre@uta.cl

Diego Aracena

Escuela Universitaria de Ingeniería
Industrial, Informática y Sistemas
Universidad de Tarapacá
Arica-Chile
daracena@uta.cl