

## DESARROLLO DE HERRAMIENTAS PARA ACCESO A BASES DE DATOS HETEROGÉNEAS BASADO EN ONTOLOGÍAS EN EL CONTEXTO DE LA ENTREGA DE SERVICIOS PÚBLICOS DIGITALES



Laboratorio de Ingeniería de Software y Sistemas de Información (LISSI) | Instituto de Investigación en Informática (III-LIDI)

Expositores: Dr. GÓMEZ, Sergio A. (LISSI) | Mg. PASINI, Ariel (III-LIDI)

Grupo de trabajo: P. Fillottrani, E. Estévez, P. Pesado, R. Muñoz, P. Thomas

<https://lissi.cs.uns.edu.ar>  
[sag@cs.uns.edu.ar](mailto:sag@cs.uns.edu.ar)

<http://weblidi.info.unlp.edu.ar>  
[apasini@lidi.info.unlp.edu.ar](mailto:apasini@lidi.info.unlp.edu.ar)

### RESUMEN

El peso de las políticas sociales en los gobiernos municipales bonaerenses ha crecido en los últimos años con el objetivo de alcanzar un mayor grado de inclusión y de desarrollo, mejor distribución del ingreso y reducción de la pobreza. Sin embargo, una mayor asignación de recursos no siempre implica una mejor concreción de los objetivos. El proyecto *Herramientas para el desarrollo y la entrega de servicios públicos digitales de acción social para municipios bonaerenses (PIT-AP-BA 2016)* tiene como objetivo mejorar la eficacia de la implementación de las políticas sociales municipales mediante soluciones informáticas que permitan mejorar la entrega de servicios públicos de acción social basados en Internet, aplicaciones móviles, almacenamiento en la nube y nuevas tecnologías informáticas. Como parte de estos esfuerzos, en este trabajo se describe una aplicación de software, a nivel de prototipo, para la exportación de datos en bases de datos relacionales como datos enlazados en la Web Semántica. La aplicación desarrollada permite exportar el contenido de una base de datos relacional como una ontología expresada en el *Ontology Web Language (OWL 2)*, respetando estándares internacionales. Estos esfuerzos son parte de la construcción de un marco que permitirá llevar a cabo parte de los objetivos planteados por el proyecto PIT-AP-BA 2016.

*Palabras clave:* servicios públicos digitales; gobierno digital para municipios; software para servicios públicos; acceso a datos heterogéneos; acceso a datos basado en ontologías.

### INTRODUCCIÓN

La importancia de las políticas sociales ha crecido en los últimos años en todos los niveles de gobierno (ya sean municipales, provinciales, nacionales e internacionales), puesto que representan una de las herramientas principales para combatir las desigualdades económicas que se dan a nivel global [Bertot *et al.*, 2016] y sirven para atender las necesidades de muchos grupos vulnerables. La prestación de servicios públicos de acción social al ciudadano se convierten en una obligación para los gobiernos, así,

dichos servicios son un derecho humano, tales como lo son el acceso a servicios de agua, energía, salud, educación y otros.

A pesar de la relevancia global, la universalidad en la provisión de servicios públicos es un desafío para cada gobierno, debido a la variedad de contextos en que se prestan dichos servicios, incluidas las necesidades de grupos sociales específicos, las capacidades de cada gobierno y las condiciones específicas del contexto (e.g. territorio, aspectos políticos, culturales, económicos, etc.), (Estévez *et al.*, 2011). En particular, en la provincia de Buenos Aires se observan los siguientes desafíos para la prestación de servicios públicos de acción social: 1) los servicios son prestados por varias dependencias del gobierno municipal y no existe información consolidada sobre cómo se están entregando los servicios; 2) actualmente, existen aplicaciones *ad hoc* que soportan el proceso para la entrega de cada servicio, estas aplicaciones funcionan de manera aislada, sin compartir datos; 3) no existe una estrategia para la entrega de estos servicios utilizando múltiples canales; 4) los canales digitales que podrían utilizarse no se están explotando adecuadamente; 5) no existe una infraestructura de software que permita el desarrollo rápido de aplicaciones para la entrega de servicios de acción social (Fillottrani *et al.*, 2017).

Por lo tanto, es necesario encontrar una manera de publicar datos contenidos en aplicaciones legadas y actuales, desarrolladas y utilizadas en diversas instituciones del Estado en una forma que se puedan integrar, acceder, modificar y consultar en un formato que sea uniforme, distribuido y escalable. En este sentido, las tecnologías de la Web Semántica han madurado lo suficiente para poder considerarlas como una solución viable para la publicación e integración de datos institucionales. En particular, usar la *semántica* implica concebir sistemas donde el significado de los datos es explícitamente especificado y es tomado en cuenta para diseñar las funcionalidades de los mismos. Esta idea se ha vuelto crucial para una gran variedad de aplicaciones de procesamiento de información y ha recibido mucha atención en las comunidades de la inteligencia artificial, bases de datos, web y minería de datos.

En los últimos diez años, ha surgido un nuevo paradigma para modelar e interactuar con sistemas de integración de datos llamado Acceso a Datos basado en Ontologías (OBDA). De acuerdo al OBDA, el cliente del sistema de información es liberado de entender como los datos y los procesos están estructurados en los recursos concretos (bases de datos, aplicaciones de software, servicios, etc.) e interactúa con el sistema expresando sus consultas en términos de una representación conceptual del dominio de interés, llamada *ontología*.

En este trabajo presentamos un sistema prototípico que tiene el objetivo de facilitar el desarrollo de ontologías a partir de su extracción automática (proceso conocido como *bootstrapping*) a partir de bases de datos relacionales. El sistema tiene las ventajas de respetar la especificación del mapeo directo propuesto por la W3C (Arenas *et al.*, 2012) y permite elegir el perfil del lenguaje OWL en el que se almacena la ontología de salida, lograda a partir de una base de datos relacional MySQL. Estos esfuerzos son parte de la construcción de un marco que permitirá llevar a cabo parte de los objetivos planteados por el proyecto *Herramientas para el desarrollo y la entrega de servicios públicos digitales de acción social para municipios bonaerenses* PIT-AP-BA 2016 (Fillottrani *et al.*, 2017) y de la línea de investigación original planteada en Gómez & Fillottrani, 2017.

El resto del trabajo se estructura como sigue. En la sección 2, presentamos la aplicación en desarrollo junto con casos de estudio sobre el que fue testada. Finalmente, en la sección 3 concluimos el trabajo.

## APLICACIÓN PROTOTIPO PARA EL BOOTSTRAPPING DE ONTOLOGÍAS

El acercamiento utilizado en este trabajo para el razonamiento ontológico sobre fuentes de datos heterogéneas, consiste en la materialización de una ontología a partir de un conjunto de tablas almacenadas en una base de datos relacional H2. La aplicación desarrollada es un prototipo operacional implementado en el lenguaje de programación Java usando OWL API (Horridge & Bechhofer, 2011) que permite especificar los parámetros para acceder a una base de datos mediante JDBC. Luego, el programa recupera las tablas en dicha base de datos. El usuario puede entonces especificar un filtro SQL para expresar un subconjunto de la base de datos o bien de una tabla particular. Entonces el programa, luego de mostrar al cliente el esquema de la tabla, permite exportar los registros especificados por el filtro SQL como una ontología en formato OWL XML y OWL Turtle (Hitzler *et al.*, 2012) mediante la metodología dada (Arenas *et al.*, 2012).

## CASO DE ESTUDIO

Como caso de estudio del funcionamiento de la aplicación desarrollada, tomemos como ejemplo tres tablas basadas débilmente en el detalle de los beneficiarios de todas las ayudas sociales en el período seleccionado en la Municipalidad de Bahía Blanca (ver <http://datos.bahiablanca.gob.ar/dataviews/74266/ayudas-sociales>), donde tenemos una tabla llamada *Programa* (*id*, *nombre*) con los programas de ayuda social, otra con los beneficiarios llamada *Beneficiario* (*id*, *dni*, *monto*, *fecha recepción*, *programa de ayuda*) y una tercera llamada *Persona* (*dni*, *apellido nombre*) con los datos de las personas inscriptas en los planes. El esquema de la tabla de programas contiene el identificador (el cual es la clave primaria) y nombre del programa. La tabla de beneficiarios de programas de ayuda contiene la identificación (*id*) del beneficio (el cual es la clave primaria de la tabla), el número de documento de la persona beneficiada (el cual es una clave foránea), el monto recibido, la fecha de recepción y el programa de ayuda por el que recibió la ayuda (el cual es una clave foránea). La tabla de personas contiene el número de documento de la persona (el cual es clave primaria) su apellido y su nombre.

El prototipo desarrollado permite ver las tablas de la base datos y a partir de ellas (o de un subconjunto de ellas si fuera necesario) y una URI base para referenciar los datos en Internet (en este caso <http://datos.mbb.gob.ar/>) el sistema produce la ontología OWL la cual contiene una terminología indicando información conceptual sobre el esquema de la base de datos junto con información asercional de los registros de las tablas de la base de datos. La terminología, modelando el esquema de las tablas de la base de datos, es presentada en la Figura 1, con la notación visual para ontologías computada por <http://www.visualdataweb.de/webvowl/#>. Esta ontología puede ser consultada utilizando razonadores estándares sobre lógicas descriptivas.

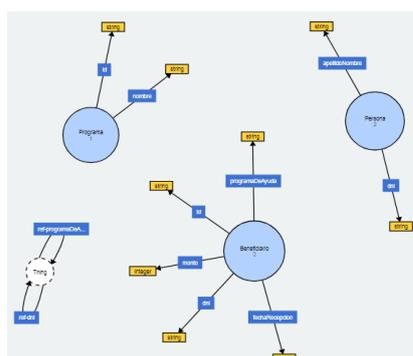


Figura 1.  
Ontología obtenida  
a partir de la base  
de datos

## CONCLUSIONES

Este trabajo presentó una aplicación para el *bootstrapping* de ontologías que es parte de los esfuerzos para llevar a cabo, parte de los objetivos planteados por el proyecto PIT-AP-BA 2016. El prototipo presentado permite especificar un conjunto de tablas y filtros para obtener, en caso de necesidad, subconjuntos de las mismas, recupera los datos contenidos en las tablas y los exporta como una ontología OWL. Esta ontología puede ser consultada por razonadores estándares e integrada con otras ontologías. Estos resultados representan un paso adelante en el objetivo de integrar los datos de gobierno en un formato abierto que sea escalable y permita su consulta en forma uniforme. Así, las ontologías computadas tienen la potencialidad de ser consumidas por agentes de software inteligentes que realicen inferencias sobre sus contenidos, permitiendo el razonamiento automatizado e inteligente sobre los datos publicados.

## BIBLIOGRAFÍA

- (Arenas *et al.*, 2012) MARCELO ARENAS, ALEXANDRE BERTAILS, ERIC PRUD'HOMMEAUX, JUAN SEQUEDA. *A Direct Mapping of Relational Data to RDF - W3C Recommendation 27 September 2012*, (2012), <http://www.w3.org/TR/rdb-direct-mapping/>
- (Bertot *et al.*, 2016) J. BERTOT, E. ESTEVEZ, T. JANOWSKI. *Universal and Contextualized Public Services - Digital Public Service Innovation Framework*, *Government Information Quarterly*, ISSN 0740-624X, vol. 33, pp. 211-222, (2016).
- (Estévez *et al.*, 2011) E. ESTÉVEZ, P. FILLOTTIRANI, T. JANOWSKI, A. OJO. *Government Information Sharing – A Framework for Policy Formulation*, capítulo 2, págs. 23-55 en “*E-Governance and Crossboundary Collaboration: Innovations and Advancing Tools*”, Pin-Yu, Chu y Yu-Che, Chen editores. IGI Global, (2011).
- (Fillottrani *et al.*, 2017) PABLO FILLOTTIRANI, ELSA ESTÉVEZ, KARINA CENCI, PATRICIA PESADO, ARIEL PASINI, PABLO THOMAS. Herramientas para el desarrollo y la entrega de servicios públicos digitales de acción social para municipios bonaerenses, IV Congreso Internacional Científico y Tecnológico-CONCYT 2017, (2017).
- (Gómez & Fillottrani, 2017) SERGIO ALEJANDRO GÓMEZ, PABLO RUBÉN FILLOTTIRANI. Completitud de los métodos de acceso a datos basado en ontologías: enfoques, propiedades y herramientas, XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires), pp. 772-776, (2017).
- (Hitzler *et al.*, 2012) PASCAL HITZLER, MARKUS KRÖTZSCH, BIJAN PARSIA, PETER F. PATEL-SCHNEIDER, AND SEBASTIAN RUDOLPH. *OWL 2 Web Ontology Language Primer (Second Edition)*, *W3C Recommendation 11 December (2012)*.
- (Horridge & Bechhofer, 2011) Matthew Horridge and Sean Bechhofer. *The OWL API: A Java API for OWL Ontologies*. *Semantic Web*, 2(1):11–21, (2011).