

Integración de Bases de Datos Heterogéneas para Servicios Web en Gobierno Electrónico

Lic. Angela Cesetti - Dr. Pablo Fillotrani - Mg. Elsa Estevez

Departamento de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur
Av. Alem 1253 - 8000 Bahía Blanca - ARGENTINA
phone: +54 291 459 5100 - ext 2613, 1031
fax: +54 291 459 5136
e-mail: [abc, prf, ece]@cs.uns.edu.ar

1. Gobierno Electrónico

Gobierno Electrónico (GE) se define como el uso de las Tecnologías de la Comunicación y la Información (TICs) en ambientes de gobierno como una plataforma para el intercambio de información. Para esto, provee servicios y permite realizar transacciones con ciudadanos, negocios y otras ramas de gobierno. El GE ha evolucionado en aplicaciones más transaccionales e integradas y los gobiernos han comenzado a agregar recursos organizativos y tecnológicos.

Los servicios ofrecidos están pensados para tres grupos que interactúan: usuarios, negocios y gobierno. Algunas características que el GE debería proveer, para lograr una comunicación eficiente son:

- mejorar el acceso de los ciudadanos a los servicios de GE y al uso de la información
- ofrecer privacidad y seguridad

Actualmente, es una realidad que la mayoría de los gobiernos ofrecen un portal donde agrupan determinada información: en principio solo aquella de carácter informativo. El siguiente paso, no del todo desarrollado solo en casos muy particulares, sería ofrecer servicios a partir de un portal, que permita realizar diversas transacciones.

Otro aspecto relacionado y que merecerá un buen análisis es como afectará la participación ciudadana en las decisiones de gobierno, a partir de la publicación de la información por parte de las entidades.

A otro nivel, existen preguntas del dominio del GE concernientes al reto que implica lograr interoperabilidad entre las bases de datos heterogéneas existentes, dadas las diferencias semánticas de interpretación; la diversidad de leyes, regulaciones, servicios ciudadanos, procesos administrativos; la educación e incorporación de mas ciudadanos en el uso de estas aplicaciones y como este uso afectará a las aplicaciones web.

Dada la interacción entre los ciudadanos y los servicios de GE, resultará en evolución y beneficios para ambos.

2. Web Semántica

La *Web Semántica* (WS) surge como consecuencia de la necesidad de diferentes grupos (gobiernos, negocios, ciudadanos) de proveer un ambiente que pueda permitir fácil acceso a la información, servicios de transacción e interoperabilidad entre aplicaciones. En un principio, la idea básica detrás de la WS es que las páginas web se identifiquen como recursos informativos, identificando el tipo de información que incluye la página web y clasificarlas utilizando lenguajes de marcado semántico sobre las páginas web. Los beneficios de esta mejora, es que facilitan buscar y encontrar cualquier clase de información.

La WS involucra la publicación de datos en un lenguaje, Resource Description Framework (RDF) y esto permite categorizar los datos según la percepción humana, pero permitiendo ser comprendida por las computadoras. Por ejemplo, la World Wide Web (WWW) se basa principalmente en documentos escritos en Hypertext Markup Language (HTML), un lenguaje de marcado que codifica el cuerpo de un texto intercalado con objetos multimedia.

3. Servicios Web

La W3C define a un *Servicio Web* (SW) como 'un sistema de software diseñado para soportar interacción interoperable entre máquinas sobre la red'. Los SW frecuentemente son solo Interfaces de Programación de Aplicaciones (APIs) web, que suelen accederse desde Internet, y los servicios requeridos se ejecutan sobre un sistema de hosting remoto. Otras alternativas con casi igual funcionalidad son Object Management Group's (OMG), Common Object Request Broker Architecture (CORBA), Distributed Component Object Model (DCOM) de Microsoft o Java/Remote Method Invocation (RMI) de SUN.

Los SW apuntan a la comunicación del tipo cliente/servidor sobre la Web, mediante el uso del protocolo HTTP. Los servicios que se ofrecen caen en uno de dos grandes grupos: el primero utiliza mensajes de Lenguajes de Marcado Extensible (XML), que toma como estándar el Protocolo de Acceso de Objetos Simples (SOAP), donde, en general, existe una descripción machine-readable con las operaciones ofrecidas por el servicio escritas en Lenguaje de Descripción de Servicios Web (WSDL). Este lenguaje no es un requisito para un endpoint de SOAP, pero es un prerequisite para la generación automática de código del lado del cliente en muchos frameworks SOAP, Java y .NET.

4. Integración de Bases de Datos mediante el uso de ontologías

Una de las cuestiones más importantes del GE es como lograr la interoperatividad de las distintas bases de datos. Para esto, analizaremos el uso de las ontologías.

Una *Ontología* es la representación de un conjunto de conceptos dentro de un dominio y la relación que mantienen. Los componentes comunes de las ontologías incluyen: instancias u objetos, clases, atributos y sus relaciones, reglas que describen las aserciones de inferencias lógicas de forma lógica que incluye la ontología de su dominio de aplicación.

Algunos de los componentes para desarrollar aplicaciones exitosas son:

- Lenguajes de marcado: XML, RDF, OWL
- Herramientas de Edición/marcado: para la construcción y uso de las ontologías
- Motores de inferencia: para deducir nuevo conocimiento del conocimiento ya especificado

Las tecnologías de SW y los ambientes de desarrollo relacionados proveen una base para el desarrollo de un sistema de aplicación. Sin embargo, el desarrollo de sistemas basados en semántica aumenta las preguntas concernientes a los aspectos técnicos y organizativos que todavía necesitan resolverse. Existen diferentes respuestas dependiendo del dominio y del tipo de aplicación.

Hay que considerar algunas cuestiones de la SW, como:

- La experiencia frustrante del usuario,
- Los problemas con la interoperabilidad, causado por diferencias en los esquemas conceptuales de dos aplicaciones que intentan cooperar,
- La administración pobre de documentos, y
- Las barreras en la recuperación de información: estandarización de metadatos.

Para una administración exitosa de los proyectos en SW, se debe considerar:

- El costo/beneficio organizativo,
- La participación del usuario,
- La integración técnica, y
- La estrategia de implementación.

Aquí, es importante destacar los avances que se están realizando en el mundo sobre el desarrollo de infraestructuras que soporten lo anteriormente mencionado, por ejemplo:

- *Providing Integrated Public Services to Citizens at the National and Pan-European level with the use of Emerging Semantic Web Technologies. SemanticGov Project* FP6-2004-IST-4-027517 is funded by the European Commission within the INFORMATION SOCIETY TECHNOLOGIES (IST) Programme. **SemanticGov**¹ is a thirty six month EU-funded research and development project that aims at building the infrastructure (software, models, services, etc) necessary for enabling the offering of semantic web services by public administration (PA).
- *Data, Information and Process Integration with Semantic Web Services (Project DIP*²): DIP's objective has been to develop and extend Semantic Web and Web Service technologies in order to produce a new technology infrastructure for Semantic Web Services (SWS) - an environment in which different web services can discover and cooperate with each other automatically. DIP's long term vision is to deliver the enormous potential benefits of Semantic Web Services to e-Work and e-Commerce.
- The *Cabinet Office*³ of the UK government has based its technical guidance in the eGovernment Interoperability Framework (eGIF), which was first issued in 2000, and updated to its version 6.1 in March 2005. eGIF mandates sets of specifications and policies for joined-up and web enabled government. It covers four areas: interconnectivity, data integration, e-services access and content management. The eGIF contains a Technical Standard Catalogue, which is revised and updated every six months.
- The French *ADAE*⁴, formerly known as ATICA, published "Le Cadre Commun d'Interoperabilité" (CCI) in January 2002 and its latest version (2.1) in September 2003. CCI comprises the recommendations for strengthening public electronic systems coherence and for enabling multi-agency electronic service delivery.

¹ <http://www.semantic-gov.org/>

² <http://dip.semanticweb.org/>

³ <http://www.cabinetoffice.gov.uk/e-government/>

⁴ <http://www.adae.gouv.fr/>

5. Portales Web Semántica

Un *portal web* (PW) es un punto de entrada para la presentación de información e intercambio sobre Internet, usado por una comunidad de interés. Es por eso, que es necesario ofrecer un soporte eficiente para la comunicación y que permita compartir la información.

Actualmente, las tecnologías Web presentan diversas limitaciones en la búsqueda, acceso, extracción, interpretación y procesamiento de información. Estas limitaciones atentan contra los procesos de comunicación y el compartir de la información entre los miembros de la comunidad. Para resolver estos inconvenientes, se pueden hacer uso de las tecnologías de SW y esto permitirá que los portales actuales evolucionen en PW mejorados semánticamente. Bajo este contexto, resulta fundamental el uso de las ontologías, ya que permiten modelar la estructura del portal.

Existen algunos temas en los cuales es importante ahondar, por ejemplo:

- los portales semánticos no emplean múltiples ontologías y no están diseñados para interoperar con portales relacionados, con lo cual se hace difícil compartir información. No hay que olvidar que uno de los principales objetivos es lograr la interoperabilidad. Los portales WS deberían evolucionar en portales P2P, para permitir compartir información y mejorar las capacidades de exportación/importación de la información. Es fundamental que estos portales permitan manejar la heterogeneidad en términos de múltiples ontologías y de múltiples portales.
- no existe la capacidad de mantener versiones. En un ambiente tan dinámico como la Web, las ontologías pueden cambiar en el tiempo. Es por eso que resulta necesario lograr que los portales WS utilicen mecanismos de versión para mantener la información consistente, permitiendo que los cambios que se realicen se puedan rastrear.
- actualmente, los portales no presentan suficiente funcionalidad para soportar SW entre portales. Lo que se debería considerar en un portal SW para obtener una funcionalidad completa es:
 - Los portales existentes y las tecnologías de administración de contenidos deberían ser reutilizables,
 - Proveer diferentes niveles de búsqueda: búsqueda basada en palabras claves, basadas en ontologías, agregando diferentes niveles de flexibilidad y haciendo un uso apropiado de las tecnologías de la SW para mejorar dichas búsquedas.
 - Apuntar a la estandarización de los portales SW, para permitir alcanzar un alto grado de comunicación y la reutilización de estructuras de conocimiento.
 - Comenzar a trabajar en la incorporación de Servicios de SW. Esto representará un punto clave para alcanzar interoperatividad entre portales.

Existen herramientas que permitan evaluar los portales, para detectar sus debilidades y luego analizar cuales estrategias se usarán para su mejora. Una aproximación que se puede utilizar es la que se ofrece desde la perspectiva funcional, que muestra un esquema de evaluación de portales de tres capas:

- *Capa de acceso a la información*: representa las características que permiten la interacción usuario-sistema.
- *Capa de procesamiento de la información*: muestra la forma en la cual el portal procesa los ítems de información: creación, publicación, organización, acceso y mantenimiento.
- *Capa base*: abarca todas las tecnologías necesarias para soportar las características de las capas superiores, tanto las tecnologías de Semántica Web, como las tecnologías de sistema.

6. Propuesta de trabajo

En estos momentos, se está trabajando para la incorporación de GE junto con diversos servicios web en la Secretaría de Relaciones Internacionales y Planeamiento, de la Universidad Nacional del Sur. A partir de los resultados que se obtengan, se evaluarán viabilidad y plazos para la implementación paulatina en el resto de los sectores de la Universidad.

A continuación se mostrarán algunos resultados obtenidos.

○ Existe hoy en la Universidad el interés y el poder político que se necesita, para llevar a cabo todas las acciones necesarias para avanzar sobre los objetivos que propone el Gobierno Electrónico. En mi humilde entender, es uno de los aspectos mas importantes para poner las cosas en marcha. Además,

- En estos tiempos resulta de suma importancia concebir nuestra Universidad como un ente competitivo. Esto se puede lograr mediante el uso de las TICs en su amplia definición.
- Existe hoy ya una infraestructura informática instaurada con diversos niveles de seguridad, que permiten un buen funcionamiento de todas las dependencias (desarrollos SIU).
- La Universidad posee la inmejorable oportunidad debido a su área de influencia, de poder estar en contacto con muchas localidades que están involucrados distintos emprendimientos. Además, la ciudad posee un puerto muy importante. En estos momentos, se deben aprovechar las posibilidades económicas y académicas que ofrece esta situación, pudiéndose potenciar.
- Recientemente se ha logrado que el alumnado se inscriba a las materias desde la página web de la Universidad, siendo este un gran logro, puesto que debido al gran aumento del alumnado, el sistema que se utilizaba hasta hace poco se había transformado en un cuello de botella.

○ En la actualidad, es fundamental que la Universidad, que goza de gran prestigio académico, no se quede fuera de los avances no solo tecnológicos sino académicos que puede aportarle la incorporación de GE a su estructura funcional completa. Se apunta a lograr obtener gran beneficio en distintos niveles, como por ejemplo,

- Permitir el acceso a información pública
- Lograr menores tiempos en la búsqueda de información
- Lograr disminuir los tiempos de compras y contrataciones, para así además, lograr mayor transparencia y permitir que mas empresas puedan participar de los procesos.
- Como consecuencia de lo demás, lograr mayor ahorro y eficiencia de los costos operativos (papel), tiempo de los empleados, horarios de acceso puesto que la pagina estaría disponible 24/7. Así, también se favorecería la operatividad de los empleados, pudiéndose reasignar recursos tanto humanos como económicos.
- Al difundir más información concerniente a la Universidad, se espera, una mayor participación del alumnado, docentes e investigadores en cuestiones que nos atañen y afectan.
- La Universidad además de proveer instrucción a la comunidad, podría comenzar a ser rentable: al mostrar todas las investigaciones que se están realizando, lograr patentes conjuntas con empresas que permitan, por ejemplo, obtener fondos para equipamiento, pago de becarios, etc. Y de esta manera lograr financiamiento para otros proyectos.
- Lograr una proyección internacional que permita que la Universidad participe de más redes y proyectos internacionales.
- Lanzar en mayor escala la plataforma de Educación a Distancia, que posee la Universidad con una plataforma de desarrollo propio.

- Es necesario implementar más servicios los necesarios para poder obtener feedback de todas las personas que utilizan el portal de la Universidad, que puedan dejar sus inquietudes, sabiendo que serán atendidos y tomados en cuenta a la brevedad.
- Resulta de gran necesidad, para la incorporación exitosa de SW, lograr la interoperabilidad de todas las bases de datos que se manejan en la Universidad.

Son muchas las expectativas que se tienen respecto a los resultados que se puedan alcanzar con la migración a un sistema que aplique las ventajas que ofrece el GE. Se espera que se termine con al primera etapa, a fin de este año, que consta de la implementación de la página web de la Secretaría. Luego de analizar las herramientas que se utilizarán para el desarrollo de la misma, más la implementación de la mayor cantidad de lineamientos que atañen al GE, se espera en mediano plazo, alrededor de 5 años, la compra de una infraestructura especial y comienzo de la migración de todo lo que se encuentra implementado, a la nueva plataforma.

A largo plazo, alrededor de 10 años, se espera incrementar los servicios que hoy brinda la Universidad a todos los actores (personal administrativo, alumnos, docentes, investigadores, empresas).

Para que este proceso sea un éxito, será necesario implementar diversos cursos de perfeccionamiento en las nuevas herramientas, para todos los niveles.

Se espera que con la incorporación de las TICs en todos los niveles de la estructura universitaria, se logren mejorar el funcionamiento de la estructura de Universidad y alcanzar mayor transparencia en todos los aspectos, internos y externos. Además, que los tiempos de operación se reduzcan, con una consecuente disminución de los gastos operativos.

Así, se podrá poner más énfasis en el crecimiento parejo de todas las áreas y ampliar la participación de la Universidad no solo a nivel académico sino también a nivel económico. Algunos ejemplos podrían ser, lograr la incorporación de nuestros alumnos a emprendimientos de la ciudad y la región, incrementar la movilidad estudiantil a nivel de grado y posgrado, obtener mas patentes que sean rentables, etc.

7. Bibliografía

- Electronic Government from Ispedia. J. Ramón Gil-Garcia
- Digital government. Gary Marchionini, Hanan Samet, Larry Brandt
- Semantic Web for e-government. Ralf Klischewski
- Semantic Web Technologies for Information Management within e-Government Services. Ralf Klischewski, Martti Jeenicke
- An Infrastructure for E-government Based on Semantic Web Services. Liuming Lu, Guojin Zhu, Jiaxun Chen.
- Infrastructure for E-government Web Services. Brahim Medjahed, Abdelmounaan Rezgui, Athman Bouguettaya, Mourad Ouzzani.
- Customized Delivery of E-Government Web Services. Brahim Medjahed, Athman Bouguettaya.
- Directions in E-Government: Processes, Portals, Knowledge. Roland Traunmüller, María Wimmer.
- Top down or Bottom up? How to establish a common ground for semantic interoperability within e-government communities. Ralf Klischewski.

- A Semantic Web Service-based Architecture for the Interoperability of E-government Services. Alessio Gugliotta, Liliana Cabral, Jhon Domingue, Vito Roberto, Mary Rowlatt, Rob Davies.
- Semantic Web Enables E-Government Services. Brahin Medjahed, Athman Bouguettaya, Mourad Ouzzani.
- The Semantic Electronic Government: Knowledge Management for citizen relationship and new assesment scenarios. Miltiadis D. Lytras.
- WSMO-PA: Formal Specification of Public Administration Service Model on Semantic Web Service Ontology. Xia Wang, Tomas Vitvar, Vassilios Peristeras, Adrian Mocan, Sotirios K. Goudos, Konstantinos Tarabanis.
- IRS-III: A Broker for Semantic Web Services based Applications. Liliana Cabral, John Domingue, Stefania Galizia, Alessio Gugliotta.
- E-government administrative and semantic cooperation: the intelligent document role. F. Corradini, A. Polzonetti, R. Pruno.
- Exploring e-Government Evolution. The influence of Systems of Rules on Organizational Action. J. Ramón Gil-Garcia, Ignacio J. Martinez-Moyano.
- E-Government Cross-Agency and Intergovernmental Initiatives Research Project: Web Survey Results. Jane E. Fountain, Robin McKinnon, Eunyun Park.
- Survey of Graph Database Models. R. Angles, C. Gutierrez.
- Queryng RDF Data from a Graph Database Perspective. R. Angles, C. Gutierrez.
- Foundations of Semantic Web Databases. C. Gutierrez, C. Hurtado, A. O. Mendelzon.
- Minimal Deductive Systems for RDF. S. Muñoz, J. Perez, C. Gutierrez.
- The Semantics of Complexity of SPARQL. J. Perez, M. Arenas, C. Gutierrez.
- An Evaluation of Semantic Web Portals. Ruben Lara, Sung-Kook Han, Holger Lausen, Michael Stollberg, Ying Ding, Dieter Fensel.
- Semantic Web Portals - State of the Art Survey. Holger Lausen, Michael Stollberg, Ruben Lara Hernández, Ying Ding, Sung-Kook Han, Dieter Fensel.