

La Política de Datos Abiertos del Ayuntamiento de Zaragoza: datosabiertos.zaragoza.es

María Jesús Fernández-Ruiz¹, Jose Manuel Alonso², Martín Álvarez², Víctor Morlán-Plo¹, María Jose Perez Perez³, F F. Javier Zarazaga-Soria⁴

¹ Ayuntamiento de Zaragoza, Vía Hispanidad 20, 50071 Zaragoza

² W3C España, C/ Ada Byron, 39 - Edificio Centros Tecnológicos, 33203 Gijón (Asturias)

³ GeoSpatiumLab s.l., Carlos Marx 6, 50015 Zaragoza

⁴ Universidad de Zaragoza, María de Luna 1, 50018 Zaragoza

mjferuiz@zaragoza.es, josema@w3.org, mlvarez@w3.org, mjperez@gesolab.com,
vmorlan@zaragoza.es, javy@unizar.es

Resumen. Durante los últimos cinco años, la ciudad de Zaragoza ha venido trabajando en la implantación de tecnología semántica con el objeto de facilitar a la ciudadanía el acceso a los diferentes servicios ofertados por la Sede Electrónica. El último hito que se fijó la Institución fue la implantación de una política de gestión en la Web para poder ofrecer los conjuntos de datos en formatos abiertos. Este trabajo tiene su foco en la presentación de la metodología que se sigue en estos momentos, así como los resultados de aplicación de la misma y que se encuentran accesibles a través del sitio Web datosabiertos.zaragoza.es

Palabras clave: Open Data, Web municipal, Ayuntamiento, Sede Electrónica

1 Introducción

El objeto de la presente comunicación es la descripción de la Política en materia de “Open Data” de la Web Municipal, Sede Electrónica del Ayuntamiento de Zaragoza. Esta Política ha sido desarrollada en cumplimiento de la Ley 37/2007, de reutilización de la información del sector público, por la que se insta a las administraciones a poner sus datos de carácter público a disposición de cualquier parte interesada.

Datos de Zaragoza es un proyecto del Ayuntamiento de Zaragoza que se traduce en una serie de servicios que se publican a través de <http://datosabiertos.zaragoza.es> para fomentar una apertura efectiva de los datos públicos que obran en su poder, para facilitar la reutilización de la información por parte de la ciudadanía, las empresas y otros organismos, lo que ofrece un aumento de la transparencia de la administración, el incremento de la participación ciudadana y la posibilidad de crecimiento económico en distintos sectores.

Datos de Zaragoza nace sobre la base de todo un conjunto de experiencias desarrolladas con anterioridad por el Ayuntamiento en la línea de puesta a disposición pública de los datos de la Entidad. Los antecedentes más destacados son:

- Servicios Web de IDEZar. IDEZar es una iniciativa del Ayuntamiento de Zaragoza para el desarrollo de servicios internos (departamento de policía, cuerpo de bomberos, emergencias, ...) y externos (principalmente la ciudadanía y turismo) alrededor de sus características geográficas. Desde un punto de vista de la tecnología, la base de IDEZar son estándares internacionales, principalmente los propuestos por el Open Geospatial Consortium (OGC, <http://www.opengeospatial.org>). Estos estándares son la puerta ofrecida por el Ayuntamiento para facilitar el acceso a su información geográfica o georreferenciada [9][10][11].
- Servicio Planifica tu Visita (Aplicación desarrollada en 2007) Mediante una ontología se ha definido un modelo de datos que permite integrar de forma coherente la información disponible en las diferentes bases de datos y los servicios web de IDEZar, la infraestructura de datos espaciales del Ayuntamiento de Zaragoza. A través de una serie de adaptadores los datos se transforman al formato RDF y se combinan en un modelo semántico. Es una aplicación elegida como Buena Práctica por el W3C. http://www.zaragoza.es/turruta/Turruta/index_Ruta
- Canales Rss y GeoRSS

El mapa del sitio web “Datos Abiertos Zaragoza” <http://datosabiertos.zaragoza.es> :

- Portada: Servicios Destacados
- Catálogo de Conjuntos de Datos
- Formatos
- SPARQL endpoint
- Terminos de Uso
- Buenas Prácticas Aplicaciones
- Consulta Ciudadana

zaragoza.es / Aprovecha tu Web / Datos Abiertos de Zaragoza

EL AYUNTAMIENTO LA CIUDAD CULTURA PARA LA GENTE TURISMO

Texto Buscar

Presentación
Catálogo
SPARQL
Formatos
Términos de uso
Aplicaciones
• Desarrolladas
• Tus aplicaciones
Consulta Ciudadana

DATOS ABIERTOS DE ZARAGOZA

DESTACADOS

SPARQL
• Acceso
• Consulta de datos en Formato Semántico

Catálogo de Datos
• Estado de tráfico
• Busca tus datos

IDEZar
• Servicios Cartográficos

Consulta Ciudadana
• Resultados
• Consulta

© Ayuntamiento de Zaragoza webmunicipal@zaragoza.es Mapa Web Accesibilidad Miembro W3C XHTML 1.0 CSS 2.0 Aviso Legal Política de privacidad

Fig. 1. Página principal de Datos Abiertos Zaragoza.

Este trabajo tiene su foco en la presentación de la metodología aplicada por el Ayuntamiento de Zaragoza, y en mostrar los resultados de su aplicación a través del sitio Web datosabiertos.zaragoza.es

2 Metodología

La Metodología de trabajo tiene como objetivo la descripción de las acciones necesarias para la exposición de “Conjuntos de Datos” de carácter público en formatos que permitan el procesamiento automático, utilizando para ello tecnologías abiertas y estándar.

Esta metodología pretende establecer las pautas y características de diseño que se deben seguir para el desarrollo y mantenimiento de un esquema de URIs estable, persistente y extensible, de forma que proporcione la base mínima necesaria para habilitar el mecanismo de Reutilización de la Información Pública en el Ayuntamiento de Zaragoza.

Los conjuntos de datos que pueden ser expuestos para su reutilización pueden ser de dos tipos: datos en crudo, o ficheros estáticos que incluyen los datos en formato XML, CSV, etc; y datos “Linked Data”, expresados semánticamente siguiendo los principios de la Web Semántica. Para la publicación de estos datos es necesario un sistema, una metodología que permita la persistencia y gestión de los mismos.

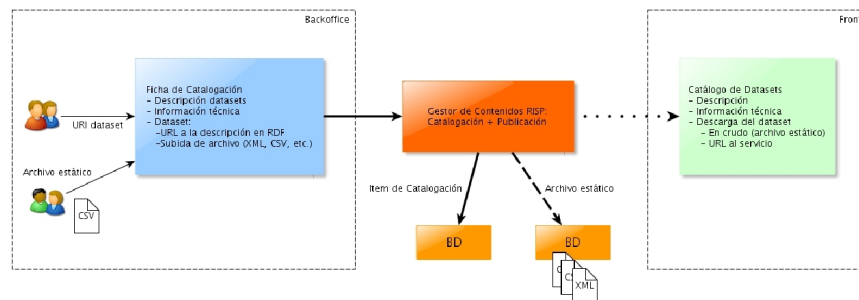


Fig. 2. Esquema de la herramienta de Gestión y Publicación del Catálogo de datasets

Seguidamente se detallan los pasos principales

2.1 Identificación y Selección

El objetivo de esta fase es la identificación de los conjuntos de datos que serán expuestos para su reutilización. Se estudian todas las variables técnicas y administrativas que influyen en la obtención y publicación de los datos. Los dominios de información, su procedencia, formatos, posibles limitaciones técnicas de los sistemas, o licencias de uso, algunos datos son más adecuados para la publicación que otros.

En esta fase es importante analizar la viabilidad de incorporar los conjuntos de datos y los formatos solicitados por la ciudadanía a través “Consulta Ciudadana”.

Variables

A continuación se especifican los parámetros que se tienen en cuenta para la selección de los conjuntos de datos más adecuados para su exposición:

- Relevancia de la información para la ciudadanía. Teniendo en cuenta los distintos perfiles de usuario que existen (si tienen conocimientos técnicos o sólo son usuarios básicos, si pertenecen a distintos colectivos sociales, si tienen discapacidades, etc.).
- Se indica si ha habido peticiones sobre algún dominio de información en concreto.
- Relevancia de la información para las empresas. Se valora el impacto y la posible utilidad de la información para las empresas que harán uso de los datos. Es una aproximación que tiene en cuenta las distintas organizaciones y el uso que harán de los datos (para consumo propio, con fines lucrativos, etc.)
- Relevancia de la información para la administración pública. Se valora el impacto y la posible utilidad de la información para el propio Ayuntamiento de Zaragoza u otro organismo público. Un ejemplo de uso puede ser la federación de contenidos por parte de distintos departamentos y servicios.
- Frecuencia de actualización de los datos. La frecuencia de actualización y la previsión de crecimiento del conjunto de datos es otro factor clave a tener cuenta, ya que desde el punto de vista de la reutilización los conjuntos dinámicos ofrecen normalmente mayor valor y posibilidades de explotación que los conjuntos estáticos.
- Volumen de información contenida en los conjuntos. Es importante valorar el volumen de información de los conjuntos de datos candidatos, ya que, en general, a mayor cantidad de información más valor tendrá el conjunto y mayor valor añadido se podrá obtener con su explotación y reutilización posterior.
- Recomendaciones de la UE para la publicación de los datos. Se tiene también en cuenta si la temática del conjunto de datos se corresponde con las sugerencias realizadas desde la Unión Europea para la publicación de información pública a través del informe MEPSIR1 (Measuring European Public Sector Information Resources). MEPSIR es un estudio que define, prueba y aplica una metodología para medir la Reutilización de Información del Sector Público en los Estados Miembro y Noruega, sirviendo como base para la Directiva 2003/98/EC, sobre la infraestructura de interoperabilidad entre Miembros de la UE.
- Reutilización por parte de los agentes de reutilización en Zaragoza. Se tiene en cuenta también el potencial interés por parte del sector de I@s agentes reutilizadores de información en Zaragoza en cuanto temática y

formato de los conjuntos de datos a publicar, ya que ellos serán una parte importante a la hora de añadir valor a esos datos. A través de la consulta ciudadana se ha recogido el interés y la opinión de las personas y empresas interesadas en la reutilización de los datos públicos.

- Fuente de Información. Localización de los datos. La localización de la información será importante a la hora de valorar la facilidad de acceso a la misma. Puede ser que la información no esté al alcance de la “Oficina de Gestión de la Web” y por lo tanto se alarguen los procesos y por tanto la inmediatez de la publicación del conjunto de datos, todo ello aún existiendo una declaración global a nivel institucional.
- Política de acceso a los datos. Se primarán aquellos casos en los que los datos ya sean considerados previamente de algún modo de acceso público para evitar problemas adicionales que puedan venir derivados de cuestiones relacionadas con los derechos y/o la privacidad, así como para agilizar los procesos de exposición ante la posibilidad de que haya que solicitar permisos de acceso temporal para las personas que trabajen en el proyecto.

Cualquier dato es susceptible de ser expuesto para su reutilización, salvo que se considere:

1. Información excluida por intereses públicos: actuaciones políticas, defensa y seguridad, política monetaria, investigación de delitos, información con materias clasificadas o información registral.
2. Información excluida por intereses de terceros: información que contenga datos de carácter personal, información cuyo acceso pueda afectar a intereses de terceros o a intereses colectivos. Debemos tener en cuenta el Procedimiento PR-14 de protección de datos de carácter personal, de forma que no se haga ningún tipo de tratamiento que vulnere la Ley Orgánica 15/1999 de 13 de diciembre de Protección de Datos de Carácter Personal.

Siempre y cuando se cumpla con las restricciones anteriores y los recursos lo permitan, cualquier dato debería ser publicado. Se ha elaborado una clasificación de los dominios recomendados para ser publicados en función de la valoración de los criterios establecidos. Se establece una cuantificación de la relevancia de cada criterio pudiendo tomar los valores Bajo, Medio y Alto.

Nótese que existen criterios que se tendrán en cuenta para la selección final de los conjuntos de datos y no aparecen especificados en el registro. Estos criterios son: Localización de los datos ; Política de acceso a los datos.

2.2 Análisis

Una vez se ha identificado y seleccionado un conjunto de datos en esta fase se analizará y valorará la forma más adecuada de preparar las salidas de la información. En esta fase se tendrá en cuenta el origen de la información, la frecuencia de

actualización, el formato demandado por l@s agentes reutilizadores y los formatos estándares del world Wide Web Consortium (w3c).

Siempre que los recursos técnicos y humanos de la Oficina de Gestión de la Web lo permitan, se publicará la información en varios formatos.

Los criterios que se tienen en cuenta para elegir un mecanismo o un formato están relacionados con favorecer los intereses de los agentes reutilizadores (ciudadanía, empresas, otras administraciones públicas) y del propio ayuntamiento. Pero también están relacionados con aspectos técnicos:

- **Contenidos Dinámicos:**
 - La información se encuentra en Base de Datos se integrará directamente en el Bus LGD. SPARQL endpoint. Se determinará si utilizamos el Bus LGD, o bien si es necesario preparar salidas ad-hoc para el conjunto de datos.
 - La información georreferenciable se publica en formatos geoRSS y geoJson, utilizando la aplicación georref, que consiste en un API REST que permite obtener los resultados en XML, JSON, etc.
- **Contenidos estáticos:** Si la información no se encuentra en BBDD se comprobará si el formato en el que se encuentra puede ser incorporado directamente en el catálogo de datos abiertos, o bien es necesario hacer alguna transformación.

Formatos Utilizados

Para la representación de los formatos, se pueden utilizar los siguientes formatos:

- CSV (Comma-Separated Values)
- ODF (Open Document Format)
- RSS (Really Simple Syndication)
- GEORss. (RSS + Coordenadas gráficas)
- JSON (JavaScript Object Notation)
- RDF (Resource Description Framework)
- SHP (Shapefile)
- SPARQL JSON. Resultado de la consulta al punto SPARQL en formato JSON
- SPARQL XML Resultado de la consulta al punto SPARQL en formato XML
- RDF-XML. Resultado de la consulta al punto SPARQL en formato RDF
- XML (eXtensible Markup Language)
- XLS (Microsoft Office Excel)
- WMS (Web Map Service de OGC)
- WFS (Web Feature Service de OGC)

2.3 Modelado

El objetivo de esta fase, siempre que sea viable técnicamente, es la generación de una base de conocimiento semántico en RDF¹ ('Triplestore'), basándose en información procedente de bases de datos relacionales, y la creación de un punto de consulta SPARQL sobre esos datos.

Además, para el servicio de la información se realiza la gestión del mapeo de las URIs de los recursos a sus representaciones usando las recomendaciones del W3C. El sistema, en función del agente que hace la petición, permitirá servir dos tipos de documentos a través de redirecciones 303: HTML legible para personas y RDF (RDF/XML) para el procesamiento automático.

Debemos realizar una tarea de modelado, en la que según el conjunto de datos con el que se trabaje se tendrá que elegir el/los vocabularios y las clases con las que se va a describir los datos, siempre basándose en otros vocabularios de referencia. Entre las ontologías base están los Términos de Dublin Core, FOAF, OE-Gov y Geonames.

Debemos tener en cuenta los conjuntos de datos ya modelados para reutilizarlos y ampliarlos, además de enlazar nuestros datos con fuentes externas para que posean una mayor riqueza. Entre las fuentes utilizadas para enlazar a datos externos está la dbPedia, representación semántica de la Wikipedia, y Geonames, base de datos sobre lugares geográficos de gran utilidad para enriquecer la descripción de lugares en Zaragoza. Dependiendo del caso específico, como el de la incorporación de referencias a Geonames, se incluyen referencias (rdfs:seeAlso) a los recursos de Geonames. Se utilizan scripts que hacen llamadas al API de Geonames para obtener el identificador que compone el URI del recurso.

Actualmente existen los siguientes conjuntos de datos modelados:

Turismo

- Agenda de Actividades
- Monumentos
- Alojamientos
- Edificios Históricos
- Zonas Verdes

Conjunto de Datos: Oferta de Empleo

- Plazas ofertadas con sus diferentes estados
- Grupos profesionales
- Plantillas de personal
- Tipos de examen
- Turnos de ingreso
- Escalas profesionales
- Anuncios relacionados

Conjunto de Datos: Centros Municipales

¹ <http://www.w3.org/TR/rdf-concepts/>

- Centro de servicios
- Edificio
- Organismo público
- Áreas organizativas
- Poblaciones destinatarias
- Materias y tipos de centros.

Conjunto de Datos: Trámites y Servicios

- Procedimiento
- Formas de tramitación
- Normativa
- Lugar de presentación
- Preguntas frecuentes
- Materias
- Destinatarios

2.4 Exposición

Tras seleccionar los datos que serán expuestos y el formato en el que se expondrán, un Gestor de Contenidos (agente publicador) pondrá a disposición del público los distintos conjuntos de datos.

Exposición de datos en archivos estáticos

En cuanto se detecte un archivo estático susceptible de ser expuesto, el gestor de contenidos deberá publicar los datos utilizando el gestor de contenidos que ofrece una URL al recurso en sí. Además de la publicación del archivo de los datos, catalogará dicha información para que el Sistema de Gestión de la Web pueda incluir dicho conjunto de datos en su catálogo.

Se recomienda convertir dicho archivo a un formato abierto, si es que no lo es. Y además, formatearlo según algún esquema semántico estándar, si es posible.

Exposición de datos provenientes de canales RSS o Servicios Web

En el caso de que se detecte un conjunto de datos dinámico susceptible de ser expuesto, como puede ser un canal RSS, un Servicio Web, o una consulta a una base de datos, el Gestor de Contenidos debe asegurarse que el acceso a dicho servicio está disponible para cualquier consulta externa. El servicio se identifica y se accede desde una URL, que será la que se incluya en el catálogo de conjunto de datos.

Cualquier nuevo RSS o Servicio Web público deberá ser registrado e incluido en el Sitio Web Datos Abiertos Zaragoza.

Exposición de datos en RDF y SPARQL

En el caso que exista la posibilidad de modelar información en formato RDF, la publicación de la información se deberá hacer a través del Sistema de Gestión de la

web, utilizando lo que se conoce como el BUS LGD (Linked Government Data), una base de datos que permite almacenar, gestionar y publicar información en RDF.

Existen varios métodos de inserción de la información, dependiendo de la fuente de datos: si procede de una base de datos relacional, de un fichero XML, de otra fuente RDF, etc. Para cada uno de los métodos y para cada tipo de información específica se deberán configurar las herramientas que convierten y estructuran de forma semántica la información.

Además de la conversión e introducción de la información semántica en el BUS LGD, el agente publicador deberá registrar cada uno de los conjuntos de datos para su correcta catalogación.

La publicación de los datos siguiendo este sistema se ofrece en distintos formatos: SPARQL-JSON, SPARQL-XML, RDF/XML, RDF/N3, RDF/Turtle y HTML.

La consulta a los puntos SPARQL se puede realizar en la página <http://www.zaragoza.es/datosabiertos/sparql>

Nombre	Grafo	Tripletas
Equipamientos Ciudad	http://www.zaragoza.es/datosabiertos/catalogo/centros	549359
Oferta de Empleo Municipal	http://www.zaragoza.es/datosabiertos/catalogo/ofertaempleo	20485
Tramites y Servicios Municipales	http://www.zaragoza.es/datosabiertos/catalogo/tramites	72218
Turismo	http://www.zaragoza.es/datosabiertos/catalogo/turismo	103025
Catalogo	http://www.zaragoza.es/datosabiertos/catalogo	122879

Tabla 1: Número de tripletas por modelo

Exposición de Datos en RDFa

En determinados casos en los que se quiera fomentar la indexación de la información publicada en el sitio web, la información representada en HTML se puede enriquecer con meta-información semántica para permitir un procesamiento automático más eficiente y conciso. Esto puede hacerse a través de la tecnología estándar RDFa, que incrusta marcado semántico en RDF dentro del código de las páginas HTML.

Esta forma de publicación puede ser contemplada cuando se disponga de información modelada semánticamente y su información equivalente –legible– en la web.

El propio portal de datos abiertos ofrece la representación del catálogo y de sus fichas en formato RDFa, modelada siguiendo un vocabulario internacionalmente reconocido, propuesto como parte del trabajo del Grupo de Interés de eGovernment

del W3C (<http://www.w3.org/egov/>), conocido como dcat –Dataset Catalog– Vocabulary (<http://vocab.deri.ie/dcat>)

2.5 Catalogación

Siempre que se pueda publicar, cualquier Conjunto datos deberá ser descrito y publicado en el catálogo, utilizando los formatos básicos disponibles –en crudo– o bien transformándolos a un formato procesable automáticamente, más sofisticado y que permita una reutilización más sencilla, completa o inmediata.

El catálogo se define usando un vocabulario estándar, conocido como "dcat" (Dataset Catalog), que permite la interoperabilidad con cualquier vocabulario del mundo. Y se generará automáticamente mediante los datos introducidos por los catalogadores a través de una aplicación que esta integrada en el sistema de gestión de contenidos de la web del Ayuntamiento de Zaragoza.

Gracias a esto, cualquier agente reutilizador puede encontrar de forma automática los Conjuntos de Datos publicados por el Ayuntamiento de Zaragoza, así como hacer búsquedas o combinar los resultados con otros catálogos.

Los conjuntos de datos que pueden ser expuestos/publicados para su reutilización pueden ser de dos tipos: datos en crudo, o ficheros estáticos que incluyen los datos en formato XML, CSV, etc; y datos “Linked Data”, expresados semánticamente siguiendo los principios de la Web Semántica. Para la publicación de estos datos es necesario un sistema que permita la persistencia y gestión de los mismos.

Catálogo de Conjunto de Datos: http://www.zaragoza.es/ciudad/risp/buscar_Risp

Cuando se vaya a actualizar el catálogo, el editor podrá editar un formulario Web, con la información de las fichas de catalogación de la entrada del catálogo. El editor tiene la posibilidad de enlazar la dirección donde se encuentra el dataset (p.e, un servicio web o un volcado de un SPARQL) o subir un archivo estático (p.e., fichero XML, CSV o PDF).

El gestor de contenidos del Sitio Web Datos Abiertos Zaragoza se ocupará de la persistencia del archivo, si fuese necesario, y de la catalogación del dataset. Posteriormente, se ocupará de permitir el acceso a ese repositorio de archivos y de generar el catálogo en diferentes vistas para l@s agentes reutilizadores finales.

Fichas descriptivas para catalogación

- Cada dataset estará identificado por una URI única, que sigue el esquema de URIs.
- El dataset puede contener datos en dos formatos: en crudo, archivos estáticos en diferentes formatos (CSV, XLS, XML, etc.); y servicio HTTP, información obtenida tras hacer una llamada a un URL (servicio web, endpoint SPARQL, etc.)
- Un dataset puede tener varios formatos asociados (ficheros en crudo en distintos formatos y/o URLs HTTP).

- Se almacenan los ficheros en crudo que contienen en un sistema de persistencia
- Los ficheros en crudo serán publicados y accesibles para descarga de l@s agentes reutilizadores.
- Este sistema de persistencia funcionará como un repositorio de versiones de los ficheros que contienen los conjuntos de datos.
- Cada dataset puede estar materializado en distintos formatos (p.e., un informe medioambiental sobre la calidad del aire puede estar representado en un fichero CSV y en XML, aunque se trate de la misma información).
- Cualquier versión del dataset será añadida al catálogo.
- Las entradas del catálogo pueden ser modificadas si se trata de errores tipográficos.
- Las entradas del catálogo pueden ser borradas, al igual que los registros y ficheros asociados, aunque no se recomienda salvo necesidad.

2.6 Reutilización

Como fase final de la metodología, una vez publicados y catalogados los Conjuntos de Datos, se llevan a cabo las estrategias de difusión y fomento de su reutilización por parte de l@s agentes reutilizadores. Estos agentes pueden ser cualquiera de las personas u organizaciones que usen la información para uso personal o para beneficio comercial, entre los que se encuentra el propio Ayuntamiento de Zaragoza.

Como prueba de concepto de reutilización se ofrecen aplicaciones sencillas que utilizan los conjuntos de datos expuestos. Estas aplicaciones podrán ser tomadas por cualquier agente reutilizador como base para sus propios desarrollos, o incluso para su uso directo si es de su interés. Entre los servicios piloto desarrollados se encuentran: un buscador de centros públicos de Zaragoza integrado como aplicación de la red social Facebook; un buscador simple de los trámites del Ayuntamiento de Zaragoza; y dos widgets –aplicaciones ligeras– para incorporar en sitios personales, páginas de inicio, móviles, etc.

Además de los ejemplos de reutilización, se está llevando a cabo una labor de difusión de la iniciativa a través de una consulta ciudadana que se mantiene activa y a través de la aparición en medios de comunicación y en foros temáticos relacionados con RISP. Concretamente, durante la celebración del Día del W3C en España 2010 se realizó la presentación oficial del proyecto Datos Abiertos Zaragoza, en la sede Zaragoza Activa y fue organizada conjuntamente por la Oficina W3C en España, CTIC y el Ayuntamiento de Zaragoza.

Indicar también que cada vez que se publica en el portal un nuevo conjunto de datos se comunica en los diferentes canales de las redes sociales del ayuntamiento de Zaragoza.

3 Servicios del Sitio Web Datos Abiertos Zaragoza

Esta sección presenta los principales servicios ofrecidos por datosabiertos.zaragoza.es que fue presentado en noviembre de 2010.

Primera Página: Servicios Destacados

Catálogo de Conjunto de Datos: incluye los conjuntos de datos (publicados en la web), ordenados por temática, etiquetas o "tags" y por formatos de representación. L@s agentes reutilizadores podrán buscar en función de sus preferencias. Se pretende ofrecer cualquier dato público y en el mayor número de formatos posibles, siempre buscando los formatos abiertos para no limitar el uso. Este catálogo incluye también, conjuntos de datos en formato "Linked Data" (Tecnología Web 3.0 o Web Semántica). Estos conjuntos de datos están representados mediante formatos estándar y vocabularios internacionalmente reconocidos, lo que permite una reutilización automática desde cualquier parte del mundo, ya que ofrece una interoperabilidad total con cualquier sistema estándar.

El catálogo se muestra como una rejilla con los distintos conjuntos de datos y con la información más relevante representada: breve descripción, formatos disponibles, procedencia y licencia.

Es importante que desde la página del catálogo se encuentra una referencia a los términos de uso de los datos, para que cualquier agente de reutilización pueda acceder a los mismos. Cada conjunto de datos tiene asociada una ficha con sus detalles específicos: descripción, fecha de creación, fecha de última actualización, tipo de licencia de los datos, formatos que se pueden encontrar, etc.

Dentro de los conjuntos de datos existentes en el portal los que se han utilizado para crear aplicaciones son:

- Paradas de autobús, en formato WFS
- Estaciones de bicicleta, en formato WFS
- Estaciones de tranvía, en formato WFS
- Farmacias, en formato WFS
- Gasolineras, en formato WFS
- Paradas de taxi , en formato WFS
- Puntos Wifi, en formato geoJson y geoRss
- Incidencias en la vía pública (cortes de agua, cortes de tráfico), en formato geoRSS
- Agenda de actividades en formato geoJson y geoRss
- Estado del tráfico en tiempo real en formato Json

Además de los anteriores los conjuntos de datos que más están siendo utilizados son:

- Catálogo de procedimientos
- Información turística
- Oferta de empleo público
- Perfil de Contratante

SPARQL: Como hay información que se modela en formato RDF, se incluye el punto de acceso SPARQL para realizar consultas sobre los datos expuestos usando tecnologías semánticas. Ya que esta tecnología es novedosa para muchos desarrolladores, se indican varios enlaces a tutoriales y guías de referencia rápida con ejemplos de utilización del lenguaje SPARQL.

Formatos: sección dedicada a los Formatos Utilizados (<http://www.zaragoza.es/ciudad/risp/formatos.htm>), en donde se hace una breve descripción de los formatos utilizados en el catálogo de datos.

Términos de Uso: todos los conjuntos de datos que ofrece el Ayuntamiento de Zaragoza, si no se indica lo contrario, se publican bajo los *términos de la licencia ColorFURIS*.

La licencia permite a l@s agentes reutilizadores

- Copiar, distribuir y comunicar de forma pública la obra de forma parcial o total, permanente o provisional, por medios alámbricos o inalámbricos, en cualquier soporte, respetando los derechos morales del Excmo. Ayuntamiento de Zaragoza (paternidad e integridad, entre otros).
- Citar esta obra para análisis, comentario o juicio crítico - para uso privado y sin ánimo de lucro y/o con fines docentes o de investigación - y/o como tema de actualidad.
- Hacer obras derivadas.

La licencia no permite a l@s agentes reutilizadores

- Atribuirse la autoría sobre esta obra ni modificar, contra la voluntad del autor, el modo en que éste se ha atribuido la autoría
- Realizar actos que supongan atentado contra la integridad de la obra.
- Impedir la modificación de la obra original por parte del autor.
- Impedir al autor retirar la obra del comercio.
- Impedir al autor el acceso al ejemplar único o raro de la obra.

Consulta Ciudadana: es un foro para la ciudadanía, que fomenta el diálogo y participación de cualquier parte interesada en los temas relativos. Este foro fue abierto el 8 de julio de 2010, siendo presentado mediante una nota de prensa que tuvo gran impacto en la ciudadanía.

La plataforma utilizada para realizar esta encuesta popular es una herramienta comercial llamada “Uservoice”, donde se configuraron dos foros para poder opinar sobre dos temas principales:

1. Conjuntos de Datos que la ciudadanía desearía reutilizar
2. Visualizaciones o aplicaciones que desearía tener la ciudadanía a su disposición.

Cualquier persona tiene la oportunidad de participar con sus sugerencias o votando las propuestas existentes. 41 ideas y más de 500 votos. Los sectores más demandados son: Negocio y comercio, Cartografía, Medioambiente, Datos sociodemográficos, Transporte, etc.

En el Portal se publica la relación de los conjuntos de datos solicitados, si son de titularidad municipal y el estado de ejecución del servicio que dará respuesta a la petición realizada por la ciudadanía .

Esta sección es de suma relevancia para tener un contacto directo con l@s agentes reutilizadores. (<http://www.zaragoza.es/ciudad/risp/consulta.htm>)

Buenas Prácticas: Ejemplos de aplicaciones que utilizan los conjuntos datos publicados en el portal. Estas aplicaciones pueden ser las propuestas por el Ayuntamiento de Zaragoza o las desarrolladas por los agentes de reutilización. Hay un formulario que facilita a l@s agentes reutilizadores comunicar al ayuntamiento información sobre la aplicación desarrollada.

4 Reconocimientos

Los sistemas de información ofertados a través de la Web del Ayuntamiento de Zaragoza han tenido diferentes reconocimientos bacionales e internacionales. Los más desatacados son:

- Planifica tu Visita (Aplicación desarrollada en 2007) Mediante una ontología se ha definido un modelo de datos que permite integrar de forma coherente la información disponible en las diferentes bases de datos y los servicios web de IDEZar, la infraestructura de datos espaciales del Ayuntamiento de Zaragoza. A través de una serie de adaptadores los datos se transforman al formato RDF y se combinan en un modelo semántico. Es una aplicación elegida como Buena Práctica por el W3C
- DatosAbiertos.zaragoza.es: Este proyecto se considera como uno de los pioneros en todo el mundo en la utilización de las tecnologías Linked Data para representar y exponer los datos públicos en formatos abierto, pudiéndose considerar como uno de los pocos que alcanzan las 5 Estrellas Linked Data, que es el máximo nivel en el baremo propuesto por Sir Tim Berners-Lee -Director del W3C- en la clasificación del estado de estas iniciativas a nivel mundial.

5 Estado del arte

En la actualidad existen más de un centenar de iniciativas de datos abiertos en todo el mundo [1], de los cuales únicamente cinco sirven datos a través de plataformas basadas en tecnologías de la Web Semántica: el Gobierno Federal de EEUU [2], el Catálogo Nacional del Reino Unido [3], el del Gobierno del Principado de Asturias [4]; la Biblioteca del Congreso Nacional de Chile [5] y el del Ayuntamiento de Zaragoza.

El proyecto Government Linked Data pionero y que sirve como referencia mundial es el impulsado desde el Gobierno del Reino Unido. Esta iniciativa ha contado con

Tim Berners-Lee como consejero técnico y ha sido desarrollado basándose en la publicación dual de datos en crudo "raw data" (archivos y volcados de datos directos) y datos en formato semántico modelados en RDF mediante ontologías consensuadas y expuestas al público para su reutilización.

Los conjuntos de datos basados en Linked Data en data.gov.uk se centran en sectores básicos para la ciudadanía y las empresas reutilizadoras, como es la educación, transporte, legislación e información geográfica. Además de datos y vocabularios específicos de estos sectores, desde esta iniciativa se han ofrecido puntos de consulta de los datos SPARQL y construido una herramienta que facilita la utilización y aprendizaje de estas tecnologías de consulta estándar ofreciendo a los desarrolladores un API (REST) tradicional que sirve información tras hacer consultas SPARQL sobre el RDF almacenado.

El Gobierno del Reino Unido también ha publicado especificaciones donde se recomienda a todos sus organismos a seguir un esquema de URI definido para mantener la coherencia, persistencia, legibilidad y facilidad de uso de las mismas.

La catalogación de los conjuntos de datos se lleva a cabo mediante una herramienta llamada CKAN [6]. El catálogo se representa en formato legible para las personas a través del portal basado en Drupal [7], y en RDF procesable automáticamente usando el vocabulario DCAT [8].

6 Conclusiones

Durante los últimos cinco años, la ciudad de Zaragoza ha venido trabajando en la implantación de un sistema de gestión certificado por AENOR que tiene como objetivo el que los datos sean únicos, compartidos, accesibles y reutilizables por terceros, lo que ha facilitado el que la implantación de esta metodología semántica se realizara sin afectar a los procedimientos de actualización de contenidos de la sede electrónica

Como consecuencia de la puesta en marcha del portal datosabiertos.zaragoza.es se ha conseguido: ser más transparentes al permitir el acceso a terceros a los datos para que estos puedan ser reutilizados o consumidos por los interesados; que se hayan creado nuevos servicios y aplicaciones generando riqueza y valor para la ciudadanía sin coste adicional para el Ayuntamiento; además de facilitar el flujo de la información dentro de la propia Administración.

Los siguientes pasos que se van a realizar van a ser: continuar mejorando los servicios, sobre todo ofreciendo más documentación sobre los conjuntos de datos; por otro lado vamos a impulsar la presencia de RDFa en las salidas de BBDD, seguiremos incorporando nuevos conjuntos de datos y procurando que sea en formatos Linked Data.

Se va a hacer un esfuerzo en la difusión de los conjuntos de datos existentes ya que hasta ahora no somos conscientes de que haya aplicaciones desarrolladas que hagan uso de Linked Data, sino que lo normal es utilizar salidas predeterminadas en WFS o geoRSS.

7 Bibliografía

1. <http://datos.fundacionctic.org/sandbox/catalog/faceted/>
- 2.- <http://www.data.gov>
- 3.- <http://data.gov.uk>
- 4.- <http://risp.asturias.es>
5. <http://datos.bcn.cl>
6. <http://ckan.net>
7. <http://drupal.org>
8. <http://vocab.deri.ie/dcat-overview>
9. López Pellicer, F.J.; Álvarez, P.; Muro-Medrano, P.R.: "IDEZar: Procesos, Herramientas y Modelos Urbanos Aplicados a la Integración de Datos Municipales Procedentes de Fuentes Heterogéneas". Avances en las Infraestructuras de Datos Espaciales. Treballs D'informàtica I Tecnologia. Castelló de La Plana: Universidad Jaime I De Castellón, 2006, P. 105-113. ISBN 84-8021-590-9.
10. Fernández, M.J.; Álvarez, P.; López, F.J.; Muro, P.R.: "IDEZar: Un Ejemplo de Implantación de una IDE en la Administración Local" Actas de las IX Jornadas Sobre Tecnologías de la Información para la Modernización de las Administraciones Públicas (Tenimap 2006). Sevilla, España, 30-May - 2-Jun, 2006.
11. Portolés-Rodríguez, D.; Álvarez, P.; Béjar, R.; Muro-Medrano, P.R.: "IDEZar: An Example of User Needs, Technological Aspects and the Institutional Framework of a Local SDI". Proceedings Of The 11th EC-GI&GIS Workshop: ESDI: Setting The Framework. 2005, P. 56-58.