

Los servicios de catálogo, la oveja negra de los servicios OGC

Por qué son clave en la implementación de INSPIRE

ALMANZA, Josu; BUSTO, Natalia; GONZÁLEZ, Nagore; GUINEA, Alejandro; LÓPEZ DE TURISO, Olga; PASCUAL, Estibaliz

Hay muchos servicios geográficos puestos a disposición del público por los proveedores de datos de toda Europa. El protocolo WMS, para visualizar mapas, está tremendamente extendido en sitios web, catálogos, revistas, cursos, congresos, artículos, etc. El protocolo WFS, para acceder a los datos vectoriales, también está bastante extendido. Ambos servicios son vistosos y agradables, se pueden ver y usar los mapas de forma independiente del dispositivo y convencer incluso a los escépticos de las ventajas de usar estándares.

Hay muchas herramientas que son ya compatibles con estos servicios, tanto de software privativo como libre. Todo el mundo puede acceder a mapas a lo largo y ancho de Europa. ¡Bien hecho!

Si buscamos herramientas que usen CSW (CSW + Herramienta) obtenemos el 10 % de los resultados que encontramos si buscamos herramientas que usen WMS. Esto es sólo un ejemplo rápido que muestra cómo la implementación de los servicios CSW no está tan extendida como otros servicios OGC. Sin embargo, los protocolos CSW son clave para construir aplicaciones interoperables reales, que puedan automatizar tareas y obtener valor añadido de la Información Geográfica.

Los servicios CSW actualmente publicados adolecen de una serie de problemas que los hacen difíciles de utilizar. En comparación con los servicios WMS que son fáciles de ver y utilizar, los problemas se van resolviendo poco a poco, algo que es más lento con los servicios CSW.

Se pueden encontrar aproximadamente unos 150 catálogos CSW en Europa, que almacenan miles de conjuntos de datos y servicios. Esta presentación expondrá la necesidad de utilizar el protocolo CSW para utilizar adecuadamente los recursos, y los problemas y buenas prácticas encontrados en los catálogos utilizados.

PALABRAS CLAVE

INSPIRE, OGC, CSW, Catálogos, IDE, implementación.

INTRODUCCIÓN

Los servicios de catálogo OGC® (Open Geospatial Consortium) son un estándar orientado a facilitar la publicación, búsqueda y utilización de datos espaciales, servicios y otros objetos, enfocado tanto a personas como a procesos informáticos. Son esenciales en el intercambio de datos en una comunidad de la información [1].

La versión actual del estándar de implementación (versión 3.0, publicada en junio de 2016) tiene la característica de estar dividida en dos partes: el modelo general y el protocolo para operaciones HTTP o CSW (Catalogue Service for the Web).

En el protocolo CSW las operaciones de peticiones y respuestas entre clientes y servidores son realizadas a través de métodos HTTP GET y/o HTTP POST. Se especifican tres clases de operaciones de servicio: OGC_Service, Discovery y Manager. Las operaciones pertenecientes a OGC_Service

permiten conocer los metadatos del propio servicio, así como de los registros contenidos a través de su identificador, mientras que las de la clase Discovery ofrecen información sobre la ejecución del catálogo, así como la posibilidad de recuperar información sobre los registros, pudiendo definir filtros o subconjuntos en la información a extraer. Las operaciones de Manager permiten editar los registros del catálogo [2].

Las operaciones definidas son:

- GetCapabilities (OGC_Service)
- GetRecordById (OGC_Service)
- GetRecords (Discovery)
- GetDomain (Discovery)
- Transaction (Manager)
- Harvest (Manager)
- UnHarvest (Manager)

Las clases más interesantes para el usuario son las dos primeras: OGC_Service y Discovery, y en concreto, estas operaciones:

- GetCapabilities: Operación que devuelve el documento de capacidades, con los metadatos del servicio.
- GetRecordById: Operación para la obtención de los metadatos de un registro. Es necesario conocer previamente el identificador del recurso.
- GetRecords: Operación que devuelve un listado de registros que cumplen las condiciones establecidas en la petición. Aquí se obtienen los identificadores de los registros.

Por otro lado, la Directiva INSPIRE define los servicios de descubrimiento (o localización) o *Discovery Services* de forma que cumplan los requerimientos establecidos por la directiva y la regulación de los servicios de red o *Network Services*. Estos están basados en el estándar ISO anterior, pero con ciertas extensiones referidas a operaciones del servicio, metadatos y multilingüismo [3]. La diferencia se aprecia en el documento de capacidades del servicio (*Capabilities*).

ESTADO ACTUAL

Con el avance de la implementación de la directiva INSPIRE, cada día más CSW son accesibles para los usuarios de las Infraestructuras de Datos Espaciales (IDE). Su publicación es obligatoria y la mayoría de los grandes proveedores europeos, así como proveedores regionales u otro tipo de organizaciones ya ofrecen estos servicios. Con los servicios de catálogos, el usuario tiene acceso a la colección de datos de un proveedor, es capaz de visualizar la información disponible y la forma en la que se puede consumir. Son esenciales por tanto a la hora de enfrentarse a los datos.

De acuerdo a la directiva, los países tienen que presentar unos indicadores anuales sobre la implementación y uso de las IDE. En el apartado *Monitoring and Reporting* dentro de la página Web de INSPIRE se pueden consultar estos datos, así como en las páginas de las fuentes, que incluyen información sobre la existencia, accesibilidad y conformidad de los metadatos y servicios. La existencia de los servicios de catálogos se puede comprobar aquí.

Alemania es un buen ejemplo en cuanto a la publicación de sus indicadores, en diferentes formatos para que sea más sencilla la lectura, de forma centralizada, teniendo en cuenta que se trata de un

estado federado y la mayor parte de la información proviene de los estados.

La información publicada a través de los catálogos puede ser consumida a través de diverso *software*, permitiendo la consulta y el trabajo con esta. Algunos de los más conocidos, como ArcMap o QGIS permiten la instalación de extensiones para obtener datos de los CSW. Por ejemplo, se puede buscar un texto en el catálogo como “ortoimágenes” u “orthoimagery”, seleccionar un conjunto de datos o *dataset*, leer sus metadatos y cargar un servicio de visualización, todo desde la misma aplicación.

VENTAJAS Y DESVENTAJAS DEL USO DE SERVICIOS DE CATÁLOGOS

Desde el punto de vista de un usuario tal vez es más amigable el acceso a un catálogo Web, donde se pueden realizar búsquedas de datos utilizando distintas herramientas, como búsquedas por título, filtros por temas o palabras clave o búsquedas por localización geográfica. El uso del protocolo CSW no es sencillo para un usuario estándar, ya que las peticiones se realizan a través de código y los resultados se muestran en formato XML, tedioso para la lectura. Hay que conocer las distintas peticiones y el método (GET/POST) que soporta el catálogo.

Entonces, ¿cuál es la ventaja del uso del CSW? Sin duda, la automatización en la lectura de catálogos (y la monitorización).

Empleando soluciones orientadas a este tipo de funcionalidad como pueden ser Geonetwork [4] (*opensource*) o Voyager [5], se puede extraer la información de los catálogos y tratarla de forma automática. En caso de que estas herramientas no sean suficiente y sea necesaria alguna funcionalidad extra, el proceso de extracción puede ser implementado empleando una herramienta de transformación de datos como FME [6]. Realizando un seguimiento de la información extraída, además, se puede llevar un control de la evolución de los catálogos: las altas y bajas en los registros, las modificaciones en los metadatos de estos o las modificaciones en los metadatos del propio catálogo.

Antes de realizar estas operaciones habrá que tener en cuenta las restricciones de acceso al catálogo, ya que se puede dar la posibilidad de que no se permita el tratamiento masivo de la información.

Otra ventaja es que el acceso a los registros se hace a través de su identificador que es el mismo que el identificador de fichero (*fileIdentifier*) y este es invariable. Se da con frecuencia en los catálogos Web el hecho de que tras una actualización de la plataforma o de los metadatos se cambia el identificador (ID) asignado a una serie, conjunto de datos o servicio, perdiéndose el rastro del recurso. Como el identificador asignado en la Web no coincide siempre con el identificador de fichero, hay que realizar una búsqueda, comprobar que el recurso es el mismo y posteriormente comprobar si ha sufrido modificaciones.



The image shows a browser window with the URL `www.ign.es/csw-inspire/srv/spa/xml_iso19139?id=3998`. Below the address bar, a message states: "This XML file does not appear to have any style information associated with it. The document tree is show". The XML document tree is displayed with the following structure:

```
<gmd:MD_Metadata xmlns:gmd="http://www.isotc211.org/2005/gmd" xmlns:gco="http://www.isotc211.org/2003/gco" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.isotc211.org/2003/gmd http://www.isotc211.org/2003/gmd/gmd.xsd">
  <gmd:fileIdentifier>
    <gco:CharacterString>spainCCserie201307190000</gco:CharacterString>
  </gmd:fileIdentifier>
  <gmd:language>
    <gmd:LanguageCode codeList="./resources/codeList.xml#LanguageCode" codeListValue="spa">spa</gmd:LanguageCode>
  </gmd:language>
  <gmd:characterSet>
    <gmd:MD_CharacterSetCode codeList="./resources/codeList.xml#MD_CharacterSetCode" codeListValue="utf8">utf8</gmd:MD_CharacterSetCode>
  </gmd:characterSet>
  <gmd:parentIdentifier gco:nilReason="missing"/>
  <gmd:hierarchyLevel>
    <gmd:MD_ScopeCode codeList="./resources/codeList.xml#MD_ScopeCode" codeListValue="series">series</gmd:MD_ScopeCode>
  </gmd:hierarchyLevel>
</gmd:MD_Metadata>
```

Figura 1: Diferencia entre ID y fileIdentifier.

Los mayores problemas en el tratamiento automático de la información se deben a errores en el esquema de los metadatos incluidos en el catálogo, no a defectos en el propio servicio. A pesar de que hay que seguir unos estándares, la información no está siempre relacionada de la misma manera. Por ejemplo:

- Existen catálogos que no incluyen apenas servicios, ya que estos están incluidos como recursos dentro de los conjuntos de datos o *datasets*.
- En otros catálogos, en cambio, la jerarquía es la inversa. Partiendo de los metadatos de los servicios se llega a los distintos datasets (a través de una relación a través del atributo *operatesOn*).
- En otras ocasiones, los servicios de catálogo son *catálogos de catálogos* y contienen información de distintas fuentes.

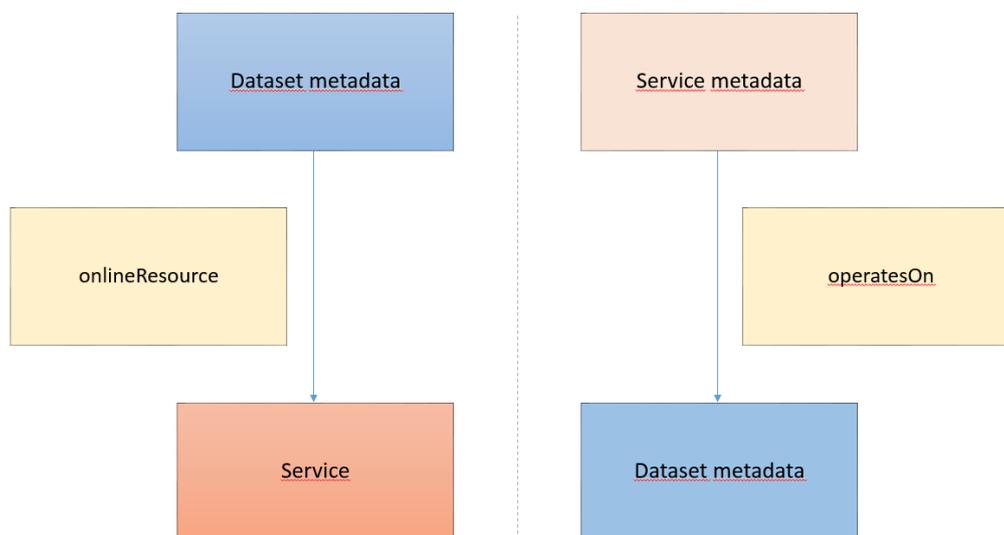


Figura 2: Distintas relaciones entre registros dentro de un catálogo.

CONCLUSIONES

A pesar de la existencia de numerosos servicios de catálogos, no es fácil encontrar herramientas que permitan su explotación de forma sencilla. Además, debido a que aún no se han publicado todas las guías técnicas para asegurar la interoperabilidad, cada proveedor tiene ciertas especificidades de aplicación de la directiva INSPIRE a la hora de clasificar los datos, lo que dificulta el acceso a los contenidos.

La aplicación de la directiva está propiciando el aumento de este tipo de servicios, por lo que probablemente en los próximos años surjan nuevas herramientas que faciliten el acceso y explotación de los servicios de catálogo, aún infrautilizados.

Los servicios de catálogo permiten automatizar la búsqueda e ir más allá accediendo de forma automática a la información geográfica, algo muy difícil de conseguir sin utilizar catálogos.

REFERENCIAS

- [1] OGC® Catalogue Services 3.0 - General Model, <http://docs.openeospatial.org/is/12-168r6/12-168r6.html>

- [2] OGC® Catalogue Services 3.0 Specification - HTTP Protocol Binding, <http://docs.openeospatial.org/is/12-176r7/12-176r7.html>
- [3] Technical Guidance for the implementation of INSPIRE Discovery Services, http://inspire.jrc.ec.europa.eu/documents/Network_Services/TechnicalGuidance_DiscoveryServices_v3.1.pdf
- [4] Geonetwork, <http://geonetwork-opensource.org/>
- [5] Voyager, <https://www.voyagersearch.com/>
- [6] FME, <https://www.safe.com/>

AUTORES

Josu ALMANZA

Josu.almanza@geograma.com
Geograma

Natalia BUSTO

Natalia.busto@geograma.com
Geograma

Nagore GONZÁLEZ

Nagore.gonzalez@geograma.com
Geograma

Alejandro GUINEA

Alejandro.guinea@geograma.com
Geograma

Olga LÓPEZ DE TURISO

olga.lopezdeturiso@geograma.com
Geograma

Estibaliz PASCUAL

Estibaliz.pascual@geograma.com
Geograma