

GeoLake Search

El futuro de la IDE está en mejorar su catálogo

SERGIO MARTIN SEGURA
IAAA, Universidad de Zaragoza
segura@unizar.es
FRANCISCO J. LOPEZ-PELLICER
IAAA, Universidad de Zaragoza
fjlopez@unizar.es
JUAN VALIÑO GARCÍA
IAAA, Universidad de Zaragoza
juanv@unizar.es
F. JAVIER ZARAZAGA-SORIA
IAAA, Universidad de Zaragoza
javy@unizar.es

RESUMEN: Es tiempo de preguntarnos cómo los nacidos en el siglo XXI verán las Infraestructuras de Datos Espaciales (IDE) y, especialmente, los geoportales como medio por el cual los datos y servicios de dichas infraestructuras están disponibles. Los nacidos en el siglo XXI tienen que interactuar en las IDE con un entorno diferente al que están acostumbrados, muy marcado por el peso que tienen los productores de datos en la misma y por su propia idiosincrasia tecnológica. ¿Cuál puede ser su opinión sobre la misma idea de infraestructura? ¿Cómo reaccionarán ante nuestros portales? Podrían preguntarse, por ejemplo, por qué en el Catálogo de Metadatos solo pueden buscar conjuntos de datos mediante sus descripciones (los metadatos) y no aprovechando también su contenido (los datos). Acostumbrados a portales como Google es racional esperar que piensen así.

Actualmente las aplicaciones y servicios de descubrimiento en el dominio de las IDE no aprovechan adecuadamente las tecnologías que actualmente son el estado del arte en la búsqueda de información. En ellas, donde antes sólo era posible procesar los metadatos que describen colecciones de datos ahora podemos procesar todos y cada uno de los datos de esa colección y extraer información tanto de ellos como de los metadatos de sus colecciones para construir mejores índices de búsqueda.

En este trabajo proponemos una aproximación novedosa a los servicios de descubrimiento en el contexto de las IDE. Esta consiste en extender el tradicional modelo basado en la indexación de metadatos ISO 19115 con la indexación de los datos contenidos en las colecciones de datos identificadas por dichos metadatos. Esta extensión permite que la búsqueda de colecciones de datos, dado por ejemplo una temática y lugar, se base en buscar y ordenar colecciones de datos utilizando un criterio de relevancia con respecto a la búsqueda que combina la del conjunto de datos según su metadato junto con la de cada uno de sus datos. La propuesta se completa con la presentación de un prototipo operativo de catálogo denominado GeoLake Search que demuestra la viabilidad de esta aproximación, así como las posibilidades que abre este nuevo enfoque.

Es comprensible que, en su origen hace 20 años, las tecnologías de la época obligaran a

hacer concesiones en la granularidad de los servicios de descubriendo. Ahora no podemos permitirnos que dichas limitaciones nos sigan lastrando. Soluciones como la propuesta pueden ser el inicio de una nueva generación de geoportales donde los datos sean totalmente localizables, accesibles e, incluso, agregables bajo demanda desde los servicios de descubrimiento, donde cualquier usuario pueda entenderlos, pueda explotarlos o, simplemente descubrirlos, donde, simplemente, cumplan su propósito de ser útiles para la sociedad.

PALABRAS CLAVE: geoportal, catálogo, metadato, servicio de descubrimiento, modernización.