

Mejorando la Usabilidad en las Plataformas de Infraestructura de Datos Espaciales

La evolución tecnológica de la IDE Ucuena

Villie Morocho

Departamento de Ciencias de la Computación, Centro IDI
Universidad de Cuenca
Cuenca, Ecuador
villie.morocho@ucuenca.edu.ec

Resumen—La disponibilidad de la información geoespacial a través de herramientas como las Infraestructuras de Datos Espaciales (IDE) permite una relación con el ciudadano y que es especialmente buscada por las autoridades. Además, investigadores, directores de proyectos, administradores y todos quienes trabajan sobre planeación territorial, y áreas relacionadas, necesitan administrar información geoespacial en estructuras adecuadas. Uno de los objetivos primordiales de las IDE es cubrir las necesidades descritas anteriormente, sin embargo, la complejidad de las plataformas desarrolladas o heredadas con *cores* de IDEs extranjeras han alejado la consecución de esta meta. En este documento se hace la presentación de la versión 3.0 “*work in progress*”, de la plataforma IDE UCuenca resultado de un análisis de usabilidad de las IDE. Además, se ha integrando ideas innovadoras para mejorar los indicadores donde se propone búsquedas basadas en relacionamientos semánticos.

Infraestructura de Datos Espaciales; Sistemas de Información Geográfica, Ontologías; Planeación Territorial; Sensor Ciudadano

I. INTRODUCCIÓN

En una analogía con la evolución de la Internet donde se comparte gran cantidad de información de diversos tipos pero que en sus inicios fue textual, “deberíamos entender que el primer rol de compartición de información fue de su propio diseño y operación a través de documentos RFC (*Request for Comments*)” [1] es esencial **estudiar los avances en relación a la accesibilidad de la información geoespacial a través de la World Wide Web**. La cantidad de información geoespacial que está conectada a través de Internet y que podría compartirse, no necesariamente es la cantidad de información que puede ser “alcanzable”. Analizando varias fuentes de estadística sobre usos del Internet, es notorio que el motor predominante de búsqueda a nivel mundial es Google®, con pequeñas excepciones de los países como China donde predomina Baidu® y en la federación Rusa Yandex® [2]. Por lo tanto, hablar de búsqueda de información es hablar de búsqueda por Google con las ventajas y desventajas que ello representa.

En consecuencia la información alcanzable, es la información “googleable” por lo que existe la información invisible o “The Invisible Web” [3]. En un intento de que la

información geoespacial sea alcanzable y por lo tanto compartida por la Web, nacen las Infraestructuras de Datos Espaciales. Según revisiones realizadas data como fecha del nacimiento de las IDE (o SDI por sus siglas en inglés) el año 1994 por W.J. Clinton [4]. Ya para entonces se había propuesto el uso de los metadatos como una parte esencial de los *clerlinghouse* de datos espaciales. Por lo tanto, las búsquedas de la información geoespacial dependen específicamente de la calidad de los metadatos y de la estructura de los contenedores de información geoespacial.

La propuesta de mejorar la calidad de servicios y la disponibilidad de la información geoespacial a través de un IDE contribuye a las iniciativas que buscan “sacar de las sombras” a este tipo de información, no clásica. Adicional a esto, la introducción de relaciones semánticas que apoyen a la búsqueda de información basada, no solamente en metadatos, es una tendencia en avances de software a nivel mundial aunque con enfoques hacia la interoperabilidad[8]. En este documento se aborda dicho enfoque en el contexto del IDE UCuenca mejorando la usabilidad del IDE tanto para que pueda integrar nueva información, como en el momento de búsqueda. Si se logra mejorar la búsqueda de la información geoespacial, con técnicas de anotación semántica, esto permitirá evitar que estudios, proyectos, y toda la información geoespacial generada se mantenga bajo el velo de la Web Invisible. Este artículo hace una presentación de las actualizaciones propuestas en la versión 3.0 de IDE UCuenca y aborda la caracterización hecha en la búsqueda con relaciones semánticas centrada en OntoRisk, una Ontología sobre riesgos en desastres naturales que actualmente se encuentra en desarrollo.

II. LA EVOLUCION DEL IDE UCUENCA

En España las IDE aparecen a mediados del 2002, y se puede considerar como el motor de promoción de las IDE a nivel de Iberoamérica, de hecho, varios de los proyectos IDE a nivel de Sudamérica parten de especialistas formados en Universidades de España y que promueven la IDE en países latinoamericanos. La IDE UCuenca en la Universidad de Cuenca, Ecuador, es uno de los primeros proyectos IDE que consigue levantar servicios e integrar fuentes autónomas y

Este artículo es parte del proyecto “Nuevos métodos y tecnologías para fortalecer la participación ciudadana en caso de alertas tempranas, potenciando IDE RedCEDIA” - IDE REDCEDIA AT, cofinanciado por el CEDIA a través del CEPRA 2013. Nov2013-Dic2014

misma, pueda presentar mejoras para su uso, es decir, que los administradores de la IDE puedan crear visualizadores de manera más dinámica y con facilidades de gestión. En la versión actual el registro de capas .shp son almacenados en la base de datos y son manipulados de forma gráfica. Adicional a ello, se puede registrar capas de un servicio WMS externo de manera intuitiva. La IDE UCuenca v3.0 está basada en la versión 3 de Google Map API, debe apuntarse que la versión 2 de dicho API quedó obsoleta en noviembre del 2013 y que dejó obsoletas varias instalaciones réplicas de la IDE UCuenca o basadas en dicha IDE. La facilidad de gestión de los visores que pueden crearse desde la plataforma actual permite mejorar la usabilidad por parte de los usuarios técnicos administradores. Estas mejoras apuntan directamente a los indicadores 1, 3 y 5 de [10].

B. Incremento en la Democracia

En el proyecto actual “Nuevos métodos y tecnologías para fortalecer la participación ciudadana en caso de alertas tempranas, potenciando IDE RedCEDIA” uno de los principales objetivos es capturar información del ciudadano como “sensor social” o, valga la redundancia, “sensor ciudadano”. Por lo tanto, aunque el indicador 7 de [10], referente a la cantidad de servicios Web disponibles ha ido incrementando paulatinamente en la medida que nuevos proyectos fueron desarrollados basados en la IDE UCuenca, el indicador 9, sobre interacción del ciudadano, resultará efectivamente incrementado por la posibilidad de interacción con la plataforma y además abre nuevas líneas futuras de interacción ciudadana. Entre los proyectos más recientes y relevantes que ha influenciado en algunas de las ideas plasmadas en el nuevo proyecto del que este documento es parte, está el proyecto CEDIA-PNUD[12].

Por lo tanto, la versión 3.0 (ver. Figura 2) en sus variaciones contará con la incorporación más directa de información en línea generada por los mismos ciudadanos y que espera conseguir mejorar la participación ciudadana, pero también el interés por la información que por ellos mismos están incorporando. Si la incorporación de información de interés tanto para el ciudadano como para los GAD son bien dirigidas, el resultado esperado es que el incremento del indicador 7 se multiplique exponencialmente análogo a lo que sucede con las redes sociales. En el proyecto se prevé también utilizar las redes sociales como información de entrada y visible en la IDE, en aspectos de interés del campo de estudio.

C. Incremento en la Efectividad

Aunque en el estudio [10] no se realizó la separación entre los niveles de usuarios, sin embargo, el indicador 10 se ve mejorado tanto desde el punto de vista de usuario administrador, como usuario no técnico. En el usuario administrador la facilidad con que la IDE UCuenca v3.0 le permite la creación de nuevos visores y principalmente para la creación dinámica del árbol de las capas en el visor de mapas, consigue un incremento directo. Adicional a ello, para el usuario final o “no técnico” el cambio en la lógica para la conformación del árbol y la consulta al servidor de mapas, en el momento de ejecución de la IDE en el cliente, da como

resultado un incremento potencial en la velocidad de consulta de información.

Se realizaron pruebas comparativas entre la versión 2.0 de la IDE UCuenca contra la versión actual 3.0, con la misma cantidad de capas consultadas (alrededor de 700). La versión 2.0 necesitaba un tiempo mayor a 5 min para cargar el visor completo. Actualmente la versión 3.0 necesita menos de 20 segundos. (<http://ide.ucuenca.edu.ec>)

Si se considera el índice 12 referentes a descargas, debe apuntarse que en la versión 3.0 se prevé la posibilidad de descarga de la información con usuario y clave directamente desde la IDE. Finalmente sobre el índice 14, la facilidad de creación de visores, sobre la versión 3.0, permitirá acrecentar la posibilidad y el interés principalmente de los GADs y posiblemente de otras empresas privadas, para poder colocar servicios del IDE en su sitio web.

D. Mejora en la Educación

La posibilidad de que la IDE se convierta en una herramienta donde sea factible plasmar investigaciones, o sirva de apoyo a otras, en áreas tan diversas que van desde Ingeniería de Software hasta de Ciencias de la Tierra, permite que el interés de profesores y estudiantes se vea incrementado. En la versión 3.0 del IDE UCuenca se tiene previsto varios aspectos relevantes de investigación que serán plasmados especialmente en las formas de anotación semántica y las búsquedas por coincidencia de las ontologías.

Si se considera que uno de los problemas más relevantes de investigación en el ámbito de las IDEs, que todavía no ha sido tratado a cabalidad, es conseguir que la información geoespacial pueda ser “buscable”, como con Google, para que así deje de ser parte de la “Web invisible”. Por ello, facilitar la operación de búsqueda es una de las tareas primordiales en el proyecto actual.

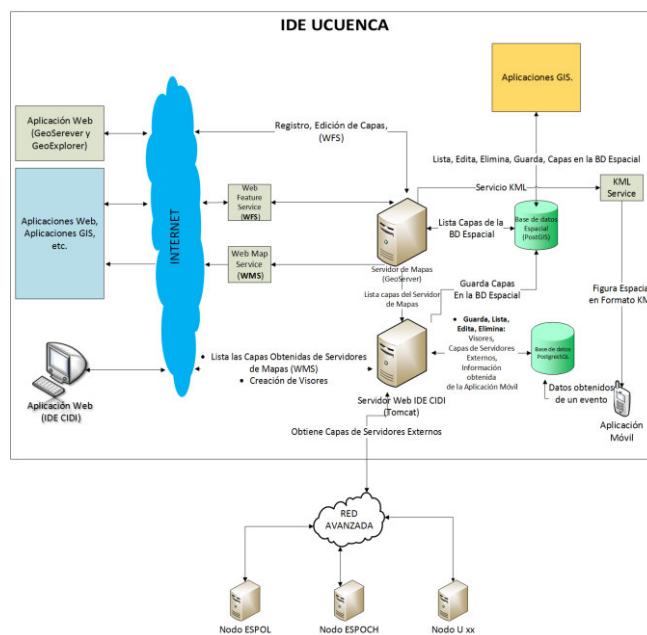


Figura 2 Esquema de funcionamiento del IDE UCuenca

IV. RELACIONES SEMANTICAS PARA LA BUSQUEDA GEOESPACIAL

El proyecto IDE RedCEDIA AT, tiene enfoque hacia alertas tempranas en desastres naturales, y considera tres escenarios. Uno referente a inundaciones provocadas por ríos, estudiada por el grupo de trabajo de la Universidad de Cuenca, un segundo en peligro volcánico, estudiada por el grupo de la ESPOCH, y un tercero, sobre aguaje y tsunamis abordado por el grupo de la ESPOL. Estos tres escenarios consideran a la IDE como un elemento central para compartir información de relevancia como apoyo a la gestión de desastres naturales. Es en este punto donde halla cabida la propuesta de integración de relaciones semánticas para apoyar a la búsqueda de información en los nodos de la IDE RedCEDIA. La IDE UCuenca, replica su implantación en los diferentes nodos de la IDE RedCEDIA, es decir, los adelantos que se están dando en la IDE UCuenca, serán implementados en este proyecto, en los diferentes nodos de la IDE RedCEDIA, inicialmente en los nodos ESPOCH y ESPOL, posteriormente en los nodos pertenecientes a la Red Avanzada (CEDIA).

A. La importancia de contar con mayor cantidad de información para gestión de riesgos

Es clara la necesidad de estudiar de manera conjunta la mayor cantidad de información geoespacial que pueda estar disponible especialmente para casos relacionados a desastres naturales[16]. Aunque, la esencia del presente artículo no se centra de ningún modo en la discusión sobre el significado de “desastre natural”, inclusive considerando enfoques actuales en los que se habla de que el desastre no es natural sino es consecuencia del abuso del hombre respecto a su actividad sobre la tierra, sin embargo, se puede hacer referencia a la definición de riesgo con la formulación en la que varios autores están de acuerdo[16].

$$R = A \times E \times S \times P \quad (1)$$

O haciendo una simplificación donde Vulnerabilidad que es la relación entre Susceptibilidad y Valor

$$R = A \times E \times V \quad (2)$$

Inclusive, se puede medir las consecuencias derivadas de la Exposición y Vulnerabilidad

$$R = Pr \times C \quad (3)$$

Dónde: R = Riesgo, A= Amenaza, E=Exposición, S= Susceptibilidad, P= Peso, V = Vulnerabilidad, Pr = Probabilidad, C = Consecuencia

Estos factores, en cualquiera de las ecuaciones anteriores (1)(2)(3), pueden ser representados por información geoespacial. Es esencial, por lo tanto, conseguir realizar operaciones entre la mayor cantidad de información relevante que pueda ser “alcanzable” y por ende “buscable” desde las diferentes fuentes de información geoespacial. Si sobre una IDE a nivel subnacional como es el caso de la IDE RedCEDIA, se consigue efectuar búsquedas para determinar información geoespacial que contribuya sobre un tema en específico, como es el caso de gestión de desastres naturales, y en un ámbito de interés, (inundaciones, peligro volcánico, aguajes-tsunamis) entonces es totalmente relevante determinar metodologías de

búsqueda con mayor eficacia que solamente los basados en palabras clave sobre metadatos.

B. Incrementando semántica a través de ontologías

OntoFire [13], en el dominio de incendios, hace un enfoque de integración de conocimiento por medio de ontologías, en ese caso integra conocimiento relacionado a los metadatos proponiendo una búsqueda de la información, no solamente basada en técnicas de palabras clave que se pueden hallar en los metadatos, sino que incrementando el conocimiento del metadato al usar relaciones semánticas.

En el presente artículo se plantea incrementar el conocimiento al relacionar una ontología, no solamente con la información de metadatos, sino inclusive, con la base de datos geoespacial y aprovechando WFS (*Web Feature Service*). En la versión 3.0 del IDE UCuenca y sus variaciones se está utilizando GeoServer en lugar de MapServer, especialmente por la posibilidad de servicios WFS y no solamente WMS (*Web Map Service*) (ver. Figura 3).

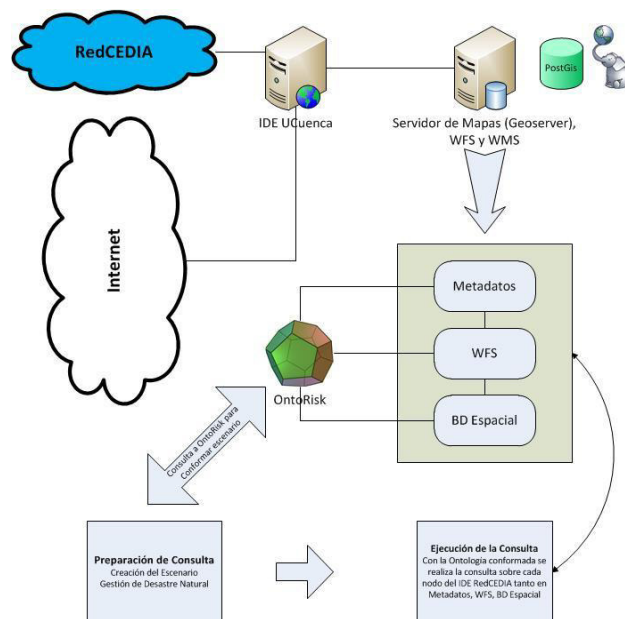


Figura 3 OntoRisk una ontología para desastres naturales

C. Conformación de la Ontología

OntoRisk es la ontología en el campo de desastres naturales que está siendo conformada basada en información extraída de técnicos especializados en los diferentes campos relacionados. Tiene punto de partida en el esquema desarrollado como resultado de la interacción de los integrantes de la comunidad LatinIDE (www.redclara.net), comunidad conformada por investigadores de seis países de Latinoamérica. Ya que el conocimiento que refleja una Ontología se deriva de la experiencia de expertos, en este sentido OntoRisk pretende integrar conocimiento extraído de la comunidad antes mencionada unida al trabajo realizado por la Secretaría de Gestión de Riesgos [17] y que una experiencia piloto de aplicación del esquema fue plasmado en el documento [12]. El uso de ontologías para mejorar los resultados en interoperabilidad de fuentes espaciales no es nuevo [8][14][15],

sin embargo, el uso de ontologías para mejorar las búsquedas de información geoespacial es un campo que todavía está abierto a propuestas [13]. Además, la propuesta de este artículo busca la implementación sobre IDE RedCEDIA considerándolo como un caso de estudio exitoso de implementación de una IDE de carácter subnacional.

V. CONCLUSIONES

Aunque las primeras investigaciones donde se propone utilizar relaciones semánticas, y en ese caso específico ontologías, para interoperabilidad e integración de fuentes de información geoespacial, fueron dados alrededor del 2004 [14], no fue sino hasta la maduración de las IDE cuando se propone la aplicabilidad de dicha teoría. La experiencia ganada en los años de implementación de las IDEs en Ecuador, desde la IDE UCuenca hasta las IDE RedCLARA pasando por IDE RedCEDIA, por más de seis años, permite describir las características esenciales que una IDE en nuestra región debería cumplir para poder ser una herramienta realmente utilizada por los usuarios a los que está dirigido. Adicionalmente, con el planteamiento hecho en los avances de la IDE UCuenca en su nueva versión 3.0 y con las variaciones de la versión que están descritas en este documento, pretende, no solamente llegar a ser una herramienta adecuada para docentes/investigadores, estudiantes, ciudadano, usuario-técnicos de GADs, usuario-no técnico de GADs, y otros similares; sino también propone avances metodológicos que generan oportunidades de aprovechar recursos tecnológicos y que integra tendencias actuales en investigación. Las ideas plasmadas en este documento consiguen, además, que la investigación propuesta termine siendo plasmada en una IDE con carácter Subnacional. Trabajo que ha sido avalado por diferentes entes de gobierno y que conseguiría aportar de forma efectiva a la gestión de desastres naturales. El multi relacionamiento semántico propuesto en este documento aporta a las técnicas de búsqueda de información geoespacial, que permitirá acrecentar la posibilidad de “localizar” este tipo de información, evitando así que continúe siendo parte de la “Web Invisible”. Este es un aporte significativo que podría terminar generando en lo futuro, no solamente una herramienta como ayuda a la gestión de desastres naturales, sino que podría llegar a ser un “Spatial- Google” donde toda la información geoespacial conectada a la Web pueda ser visible. Queda claro que la validación de la información, así como sucede en la Web actual, es a criterio del usuario pudiendo tener fuentes de información unas más confiables que otras. Los resultados alcanzados permiten mejorar drásticamente las velocidades de carga y acceso de la información, lo que devenga en un acertado enfoque a la usabilidad de la plataforma porque promueve el uso de la IDE para que los GADs sean los principales, aunque no los únicos, usuarios y beneficiados de estas funcionalidades. Otra ventaja es que el uso de la plataforma está liberada en la web lo que contribuiría notablemente a incrementar la relación ciudadana, con ahorro económico que podría llegar a ser a nivel nacional, inclusive.

AGRADECIMIENTOS

Proyecto cofinanciado con fondos de la Dirección de Investigación de la Universidad de Cuenca. En el proyecto

colaboran los grupos de trabajo de la ESPOCH, liderado por Fernando Romero y la ESPOL, liderado por Juan Carlos Pindo. Además, participan los ayudantes técnicos del proyecto de la Universidad de Cuenca, Fabián Santander y Sonia Barreto. Especial agradecimiento a la Secretaría de Gestión de Riesgo, INOCAR y ECU911-Cuenca, por su colaboración.

REFERENCIAS

- [1] Barry M. Leiner, Vinton G. Cerf, David D. Clark, Robert E. Kahn, Leonard Kleinrock, Daniel C. Lynch, Jon Postel, Larry G. Roberts, and Stephen Wolff. 2009, “A brief history of the internet”. SIGCOMM Comput. Commun. Rev. 39, 5 (October 2009), 22-31.
- [2] Torres F-, Del Valle M.J., “Ranking de buscadores por países en 2011” <http://www.10puntocero.es/blog/66-seo-sem/184-ranking-buscadores-por-paises-2011.html> consultado feb-2014
- [3] The invisible web: Uncovering Information Sources Search Engines Can't See. Information Today, Inc. Medford, NJ, USA. 2001
- [4] Clinton, W.J. Coordinating geographic data acquisition and access to the National Geospatial Data Infrastructure. Executive Order 12096, Federal Register, 17671-4. Washington: D.C. 1994.
- [5] Kelly M. C., Haupt B. J., Baxter R.E., “The Evolution of SDI Geospatial Data Clearinghouses”, Encyclopedia of Data Warehousing and Mining, John Wang Montclair State University, USA, vol.2, pp.802-809, 2009.
- [6] Guimet J., “Integración de los municipios en la IDE Regional. Primeros resultados y conclusiones”, Avances en las Infraestructura de Datos Espaciales. Publicaciones de la Universidad Jaume I, pp.125-135, 2006.
- [7] Morocho V., Garcia-Almirall P., Queraltó P., Morales A., “Infraestructura de Datos Espaciales de la Universidad de Cuenca: una herramienta para el desarrollo local”, Mapping Interactivo, Revista Internacional de Ciencias de la Tierra; ISSN 1.131-9.100. Ago, 2010
- [8] Yunpeng Y.; Li, J.; Zhengmin H., "Research on Ontology Based Semantic Integration Model in Spatial Data Sharing," Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on , vol., no., pp.2872,2875, July 31 2006-Aug. 4 2006
- [9] Morocho V., Morales A. “El vertiginoso crecimiento de la IDE UCuenca hacia la IDE REDCEDIA: Un estudio de caso exitoso de IDE SUBNACIONAL”, Proceedings of CISCi 2011, 19-22, Orlando, Florida, USA, Julio 2011.
- [10] Morocho V., Morales A., Feijoo A., Sarmiento A., Garcia-Almirall P., Queraltó P., “Estudio del impacto social, económico y académico de la Infraestructura de Datos Espaciales de la Universidad de Cuenca”, ACE: Architecture, City and Environment = Arquitectura, Ciudad y Entorno ISSN 1886-4805. Vol 19. Pág: 265 – 284, España, Junio 2012.
- [11] Álvarez M., Morocho V., Morales A., Rosanigo Z., Jones G., López-Álvarez L., “Contributions to the SDI from Latin American Universities - Some Undertaken Initiatives”, Global Geospatial Conference 2012 Spatially Enabling Government, Industry and Citizens Pag 1- 17. Quebec, Canada , Mayo 2012.
- [12] “Aportes a la lectura de la vulnerabilidad, a partir de los resultados de los análisis realizados a nivel cantonal”, Morocho V., CEDIA. 2012. Consultado en <http://dspace.cedia.org.ec>
- [13] K. Kalabokidis et al.: OntoFire: an ontology-based geo-portal for wildfires, Nat. Hazards Earth Syst. Sci., 11, 3157–3170, 2011. www.nat-hazards-earth-syst-sci.net/11/3157/2011/
- [14] Morocho V., Pérez-Vidal LI. and Saltor F, “Database Schema Detection and Mapping on Mobile Applications: An Ontology-based Approach”, Proceedings of the IASTED International Conference on Databases and Applications. Innsbruck, Austria. Feb, 2004
- [15] Morocho V., Saltor F. and Pérez-Vidal LI., “Schema Integration on Federated Spatial DB across ontologies”, Proceedings of Engineering Federated Information Systems, Coventry United Kingdom. Jul,2003
- [16] Castillo-Rodriguez J.T., et al, “The value of integrating information from multiple hazards”, Nat. Hazards Earth Syst. Sci., 14, 379-400, 2014. www.nat-hazards-earth-sys-sci.net/14/379/2014
- [17] “Guía para Implementar el Análisis de Vulnerabilidades a Nivel Cantonal”, Hallo A., 2012. ISBN 9942-9887-3