

Hacia la Implementación de un Nodo Regional de una Infraestructura de Datos Espaciales: Plataforma Tecnológica

MSc. Oscar Víquez A. MSc. Leonardo Víquez A.

oviquez@itcr.ac.cr

lviquez@itcr.ac.cr

Instituto Tecnológico de Costa Rica, Sede San Carlos
Carrera de Ingeniería en Computación
Santa Clara, San Carlos, Costa Rica
Junio-2014

Resumen

Actualmente la necesidad de información geoespacial es un tópico que está presente en múltiples aspectos de la vida diaria. Factores como falta de actualización de la información, los pocos mecanismos de divulgación de datos geoespaciales y la duplicación de esfuerzos para generar nueva información provocan que no se dé el aprovechamiento máximo de este valioso recurso.

Las Infraestructuras de Datos Espaciales son herramientas para la gestión de este tipo de información que facilitan la erradicación de este tipo de problemas. La implementación de este tipo de soluciones se puede dar de manera relativamente simple y accesible utilizando software libre.

En la Zona Norte de Costa Rica se trabaja realizando la implementación de una solución de infraestructura como la mencionada logrando como primeros objetivos la interacción entre actores locales estratégicos, productores de información geoespacial, y la creación de una primera versión de la plataforma regional.

Palabras clave: Infraestructura de Datos Espaciales, IDE, Nodo, San Carlos.

1. Introducción

La Región Huetar Norte de Costa Rica posee características únicas en ámbitos variados

tales como el turismo, la agroindustria, tecnologías de información y comunicación entre otras, que conforman gran parte de modelo socio-económico de la región y que

por ende son de interés para diversos grupos de actores que a la vez requieren que exista un gestión adecuada de esta información.

Instituciones de gobierno, empresas privadas, gobiernos locales entre otros son los principales proveedores de información que alimentan las bases de datos que se pueden encontrar en esta región. La diversidad de estos datos es muy variada tanto en contenido como en formato, por lo que en algunos casos se requiere utilizar mecanismos especializados para la gestión de los mismos.

En el caso de la información de carácter geoespacial que existe en la región, esta puede encontrarse en diversas fuentes, normalmente de forma centralizada y no relacionada entre sí por lo que limita el potencial que se puede tener al cruzar información tan valiosa de distintas fuentes que en algunos de los casos ni se conocen. El tratamiento de los datos para este fin pasa primeramente por gestionar alianzas entre actores para el manejo conjunto de la información y la implementación de plataformas tecnológicas para la gestión y el establecimiento de una cultura de manejo de los datos a través de las plataformas definidas para dicho tema.

El Gobierno de Costa Rica decreta la creación de un ente rector de políticas para el intercambio de información geográfica en todas las entidades que administran información del territorio, denominado Sistema Nacional de Información Territorial (SNIT - <http://snitcr.org>) [1]; implementando para esto normativas, estándares nacionales y una plataforma tecnológica base dónde consultar datos.

Con el objetivo de contribuir en la mejora de la gestión de datos geográficos de los diferentes actores locales de la región, el Instituto Tecnológico de Costa Rica (ITCR) a través de la Carrera de Computación sede San Carlos lidera, junto con el gobierno local de San Carlos y otras instituciones, la

creación a partir de estas políticas nacionales de una plataforma nodo del SNIT con la información de carácter regional de mayor impacto para la zona norte del país.

Dicha creación implica inicialmente el desarrollo tecnológico de una plataforma, así como el establecimiento de los vínculos iniciales con los proveedores de datos geográficos para ofrecer una serie de servicios en línea. Este artículo describe el camino seguido hacia la creación de dicha plataforma desde un punto de vista informático, sin dejar de lado las definiciones técnicas de conceptos que son necesarios para el entendimiento del producto final.

2. Marco Teórico

Son variados los conceptos que se interrelacionan con la definición final de una Infraestructura de Datos Espaciales, seguidamente se detallarán algunos de ellos con el propósito de inducir a la lectura de estos temas para profundizar en ellos.

2.1 Qué es un Sistema de Información Geográfico

Fazal [2] define un Sistema de Información Geográfico (SIG o GIS por sus siglas en inglés) como un sistema para capturar, almacenar, chequear, integrar, manipular, analizar y desplegar información que está espacialmente referenciada a la Tierra y para lo cual normalmente se involucra una base de datos georeferenciada y aplicaciones de software especializadas para este fin.

La información que se gestiona utilizando un SIG es tan variada que toca múltiples aspectos en prácticamente todos los ámbitos de acción de las sociedades actuales. Es así como vemos mapas digitalizados en temas tan importantes como el manejo de los recursos naturales, sociales, económicos e inclusive hasta información del ámbito político. Hoy en día la proliferación de esta información está tan arraigada en nuestro quehacer diario, que hasta nos resulta

impensable no disponer de un mapa en nuestros dispositivos móviles cuando queremos hacer consultas triviales como a qué lugar ir a comer o la ruta para llegar al destino deseado.

Los SIG se convierten de este modo en herramientas para la toma de decisiones en todos los ámbitos mencionados y su adecuada gestión es clave para asegurar la calidad de los datos que se manejan. Por ejemplo, es común encontrar información geográfica de una región o ciudad, relacionada con características propias de dicha área geográfica. Variables como la población, los medios de transporte, los puntos de interés, las zonas protegidas, entre otros; son representados en un SIG por capas que contienen líneas, puntos y polígonos. Cada una de estas capas deben ser cuidadosamente manipuladas, de manera que cada uno de sus componentes calce adecuadamente al superponerse con el resto para que la información que el consolidado arroje sea lo más fidedigna posible. Gráficamente puede observarse este fenómeno en la Imagen 1.

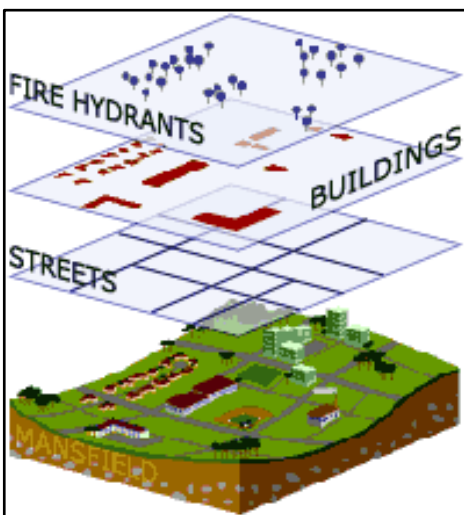


Imagen 1- Interpolación de componentes geográficos. (Fuente: <http://www.mansfield-tx.gov/departments/gis/>)

Son diversas las herramientas, tanto propietarias como de software libre, que hoy

en día se encuentran disponibles para el manejo de los SIG, entre las cuales se pueden resaltar ArcGIS, GRASS, gvSIG, MapInfo y QGIS. Esta última, es una herramienta con gran potencial, que cuenta con muchas librerías con funciones para realizar todo tipo de procesamientos a partir de información local y remota; con conectividad a diversos tipos de bases de datos y con un presente de uso significativo a nivel mundial y un futuro muy promisorio.

Bases de Datos Geográficas

Hoy en día las capas geográficas pueden ser almacenadas en bases de datos, estas bases de datos cuentan con librerías que permiten el manejo de datos geográficos tales como atributos localizaciones y topologías [6].

Actualmente, uno de los manejadores de bases de datos más utilizados para este tema es PostgreSQL (www.postgresql.org), específicamente por ser una base de datos de código abierto.

Esta herramienta es de uso general en la gestión de bases de datos, y cuenta con un módulo denominado Postgis (<http://postgis.net/>) que como ya se mencionó es específico para el manejo de objetos geográficos de tipo vectorial y convierte a PostgreSQL en un potente gestor de información geográfica. Algunas de las principales características de PostGIS son: funciones básicas de topología, transformación de coordenadas, importación de formatos shape, entre otras.

En la actualidad no resulta suficiente para las organizaciones y usuarios de este tipo de información contar un SIG y una base de datos geográfica pues en la dinámica actual debe existir una constante interrelación entre los actores para el intercambio de información con este tipo de características y es ahí donde surge y toma fuerza el concepto de Infraestructura de Datos Espaciales (IDE).

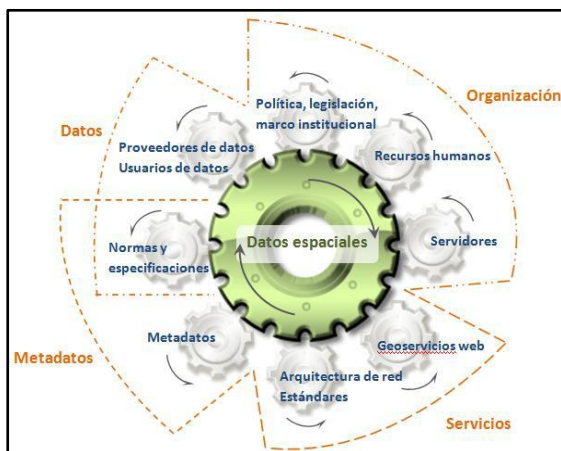
2.2 Infraestructura de Datos Espacial

Algunas de las dificultades que tradicionalmente han enfrentado los SIG son la localización y el acceso de una manera sencilla y eficaz a la información geográfica actualizada, existente sobre un territorio.

Con el objetivo de cubrir estas dificultades surgieron las Infraestructuras de Datos Espaciales (IDE) que intentan resolver de manera estructurada no solamente lo relacionado con la necesidad de publicar y compartir información geográfica, evitando duplicaciones de esfuerzos y reduciendo costos al momento de creación de nueva información, sino que también la definición de políticas y mecanismos para la gestión de esta información.

Kuhn[3] define la IDE como:

“...una serie de acuerdos acerca de estándares tecnológicos, acuerdos institucionales y políticas que permiten el descubrimiento y uso de información geoespacial por parte de usuarios y con propósitos diferentes de aquellos para los que fue creada esa información.”



Una IDE está conformada por tres componentes fundamentales como se muestra en la Imagen 2: los datos como representación o característica de una entidad, los metadatos que es la información que describe los conjuntos y servicios de datos espaciales y servicios que son las funcionalidades accesibles mediante los recursos de Internet

A nivel técnico está compuesta por elementos de hardware tales servidores con alto poder de cómputo, almacenamiento de la información y conectividad, así como y elementos de software necesarios para el acceso a la información. Aplicaciones como visualizadores de mapas, servicios de transferencia de información, gestores de datos geográficos, búsquedas a partir de catálogos de metadatos, entre otros, son ejemplos de herramientas que deben estar presentes en una IDE. Para ello, las opciones de software libre son una seria alternativa para el establecimiento rápido y accesible de estas aplicaciones [4][5].

Respecto del área geográfica que abarcan, las IDE forman una estructura escalonada en diferentes niveles en función de la extensión y de la información geográfica que gestiona, en la que los niveles superiores integran a los inferiores, de tal forma que podemos encontrar IDEs en el nivel continental, nacional, regional y local como se muestra en la siguiente figura:

Imagen 2 - Componentes fundamentales de una IDE. (Fuente: www.egeomapping.com/)

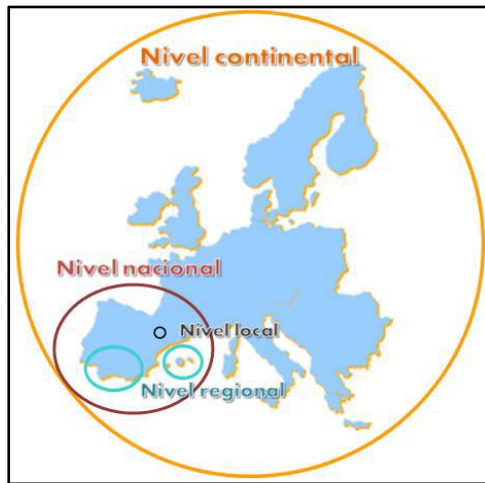


Imagen 3 - Niveles organizativos de una IDE. (Fuente: www.egeomapping.com/)

A partir de estos conceptos, se puede concebir a la IDE no sólo como la iniciativa o el instrumento facilitador para el uso de los SIG de forma colaborativa en el ámbito de las instituciones públicas de un territorio, sino también puede ser vista como el disparador de la cultura del uso de estos recursos en la población en general. De darse esto, los entes generadores de información geográfica se verán en la obligación de generar productos, actualizar información y brindar servicios geoespaciales para cumplir con la creciente demanda de los usuarios y las exigencias de la IDE.

2.3 Plataforma de Geoservicios

Una plataforma de geoservicios es justamente el componente software que se establece para brindar las capacidades que debe contemplar una IDE enfatizando en aquellas que permiten la gestión de información geográfica en tiempo real.

Existen diversos tipos de plataformas de geoservicios disponibles para su implementación. Específicamente las opciones de software libre son de las más robustas y mayormente utilizadas para dar solución a ejemplos tan complejos como lo es la IDE del gobierno Español (www.idee.es).

Los servicios de información geográfica que brindan estas plataformas son puestos a disposición de los usuarios en Internet a través de lo que se denomina un geoportal.

Geonode es una solución que conjuga una serie de productos de software libre integrados entre sí para la implementación del portal de geoservicios necesarios en el establecimiento de una IDE. Entre las principales herramientas que integra se encuentran PostGIS para el manejo de los elementos en la base de datos (o en su defecto las soluciones de carácter privativo), GeoServer para el manejo de las capas y mapas temáticos así como de recursos disponibles en línea como los brindados por los servicios de Google Maps o Microsoft Bing entre otros [8], GeoNetwork para el manejo de los metadatos de la información, django como administrador de contenidos relacionados entre las capas y sus metadatos y OpenLayers como la herramienta de interacción con el usuario que permite consumir recursos provenientes de diversas fuentes (incluida las bases locales y recursos externos) para la visualización de los datos. Su integración puede observarse en la siguiente figura.

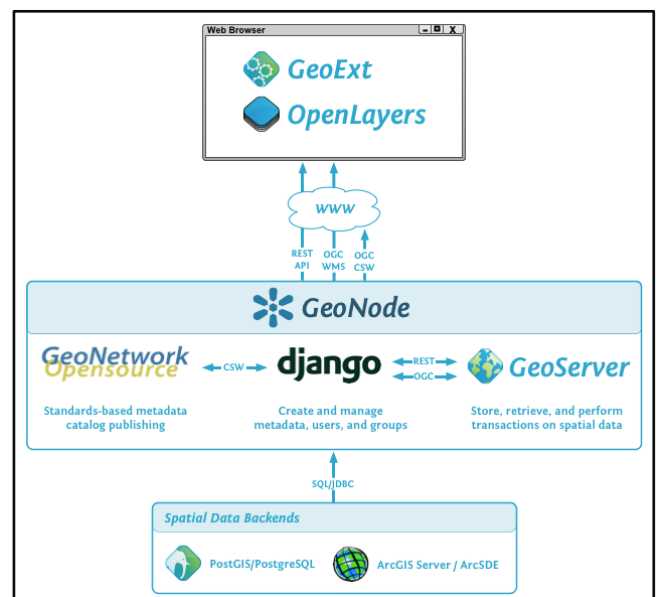


Imagen 4 – Integración de herramientas OpenSource para geoportal Geonode (Fuente: <http://docs.geonode.org/en/1.2/developers/architecture.html>)

Entre los geoservicios que contempla una herramienta como Geonode y que deben estar en todo Geoportal que dé vida a una IDE, podemos destacar los servicios definidos por el Open Geospatial Consortium (<http://www.opengeospatial.org/>) para el intercambio de información entre plataformas generadoras de datos geográficos. Los tres estándares de geoservicios de mayor difusión y uso son:

- WFS (Web Feature Service) que proporciona acceso a capas vectoriales para su manejo y permite además la modificación de dichas capas a partir de cambios hechos por los usuarios autorizados.
- WCS (Web Coverage Service) que proporciona para brindar servicios de acceso interoperable a coberturas terrestres principalmente en formato raster.
- WMS (Web Map Service) que es el más utilizado de todos y que proporciona mapas en formato de imágenes para ser consumidos por las aplicaciones que lo demanden.

Todos estos servicios tienen la característica que pueden ser consumidos tanto por clientes web normalmente definidos en la propia IDE o en otras IDE externas, así como en clientes locales tales como los gestores de SIG mencionados en apartados anteriores.

El medio más común de acceso a información geográfica para usuarios de las IDE son los visores Web. OpenLayers (<http://openlayers.org/>) es un visor que permite colocar un mapa dinámico en cualquier sitio web. Tiene la capacidad de desplegar variados tipos de información geográfica así como la integración de servicios geoespaciales mencionados anteriormente [7].

3. Creación de la Infraestructura

El producto principal del trabajo realizado consiste en la implementación de la IDE Regional para el cantón de San Carlos.

El primer aspecto en esta implementación consistió en el desarrollo de los elementos tecnológicos necesarios para la implementación de la IDE para lo cual se capacitó al equipo de desarrollo y participantes externos del proyecto en los temas relacionados con la implementación de elementos de software y hardware para una infraestructura de datos espaciales. Así mismo se han realizado investigaciones sobre el uso de herramientas tecnológicas para el desarrollo de aplicaciones que utilicen la información que la plataforma provea.

El diseño y desarrollo de la IDE ha requerido la definición de las características mínimas del servidor, la arquitectura a utilizar, los servicios que serán ofrecidos, la publicación de capas de datos geográficas y la personalización del portal que alberga la información administrada por la IDE.

El primer proceso para lograr determinar el tipo de infraestructura a desarrollar es el análisis de las necesidades propias de la región a impactar y los agentes involucrados en el proceso.

Se ha determinado un área de afectación regional para lo cual el vínculo con el gobierno local fue clave para la generación de servicios de información geográfica de la IDE, debido al nivel organizativo y de producción de información geoespacial con que el municipio cuenta.

3.1 Determinación de servicios a ofrecer

Según las necesidades identificadas en la región, el servicio que ofrece la IDE implementada es el de publicación de capas a través de WMS para la comunicación entre las organizaciones participantes y la integración de servicios.

El siguiente gráfico muestra la estructura básica de organización de la infraestructura de datos espacial implementada.

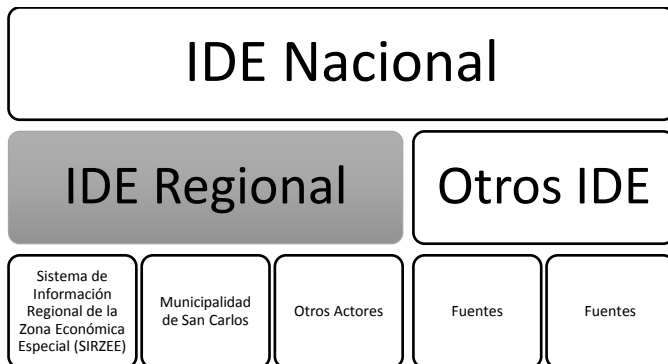


Imagen 5 - Organización de actores en el IDE Regional

La arquitectura utilizada implementa múltiples geoservicios, sin embargo, inicialmente se utiliza WMS para compartir información de capas para evitar su manipulación por agentes externos. La IDE regional se alimenta de las fuentes geográficas administradas por el Sistema de Información Regional de la Zona Económica Especial (www.sirzee.itcr.ac.cr) y la municipalidad de San Carlos, y se mantiene la infraestructura necesaria para la incorporación de otros geoservicios y otros actores responsables de la generación de información geográfica para la región de interés.

3.2 Instalación y configuración de componentes para la IDE

Se ha utilizado GeoNode como plataforma para la implementación del IDE por mostrar una adecuada administración de la información geoespacial he integración de componentes que complementas los servicios a ofrecer.

El proceso de instalación requiere como mínimo 6GM de Memoria Principal, un procesador de al menos 2.2 GHz y 1GB en disco. Para el desarrollo del IDE se dispone de un servidor en producción con dos procesadores Intel(R) Xeon(R) de 2500 GHz, y un total de 8GB RAM y sobre plataforma Ubuntu Server 12.04 para brindar una respuesta adecuada a los servicios que se ofrecen.

La instalación sobre plataforma Ubuntu requiere agregar el repositorio de paquetes para Geonode, automáticamente el proceso de instalación busca las librerías necesarias para la instalación de los componentes requeridos entre ellos: PostgreSQL, OpenLayer y Geoserver.

```
$ sudo add-apt-repository
ppa:geonode/release

$ sudo apt-get update

$ sudo apt-get install geonode
```

Imagen 6 - Comandos para la instalación de Geonode

Posterior a la instalación es necesario actualizar el IP del Servidor a utilizar y crear el primer usuario administrador de la plataforma. De aún no contar con un dominio puede accederse a Geonode de modo local indicando como IP 127.0.0.1.

```
$ sudo geonode-updateip 127.0.0.1
$ geonode createsuperuser
```

Imagen 7 - Comando para la configuración básica del Geonode

La siguiente imagen detalla la implementación del sitio Web del IDE Regional desde una ruta local y la publicación temporal de capas de interés.

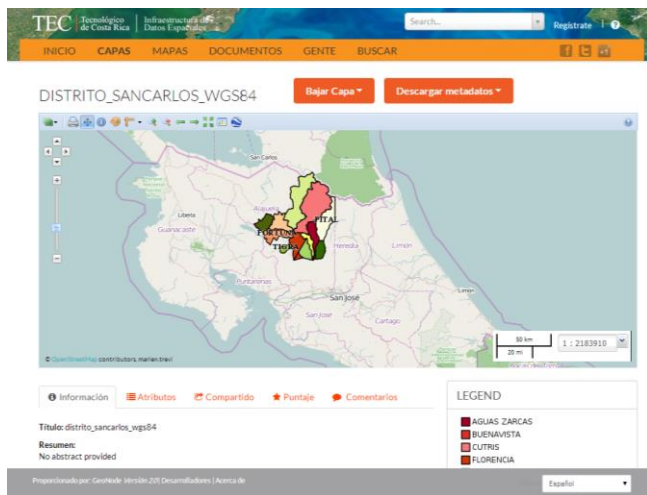


Imagen 8 - Plataforma de IDE regional instalada

3.3 Seguridad de la información

Cada nueva capa geográfica que se ingrese a la plataforma será registrada en Postgis como una nueva tabla geográfica y accedida por los usuarios mediante Geoserver como gestor de geoservicios.

Las capas ingresadas a Geoserver pueden compartirse con los distintos usuarios por medios de los servicios WCF, WFS y WMS según el propósito con que fueron ingresadas. Para ello cada capa puede asociarse a un rol de usuario determinado con los límites de acceso definidos por el usuario administrador.

3.4 Consumiendo recursos

La arquitectura de IDE es transparente para los usuarios finales que acceden a capas geográficas o mapas publicados desde un visor web. La IDE como servidor de datos centralizado cuenta con un visor OpenLayer provisto por Geonode, sin embargo cada entidad puede brindar sus geoservicios externos al IDE.

4. Resultados

Se ha logrado la vinculación con distintos actores estratégicos tales como la Municipalidad de San Carlos, el ITCR y el Ministerio Nacional de Ambiente y Energía

(MINAET) como participantes activos en el establecimiento de la IDE Regional en el rol de proveedores de información base que se distribuiría como geoservicios. Así mismo se logró el vínculo con el IDE Nacional administrado por el SNIT.

La arquitectura en hardware y software para soportar el IDE se encuentra desarrollada y ha sido probada con capas de datos geográficos suministrados por el SIRZEE y por la Municipalidad de San Carlos.

Producto del trabajo realizado en la implementación de la IDE, se brindó capacitación a funcionarios de la Municipalidad de San Carlos y el MINAET en el desarrollo de IDEs locales con la finalidad de propiciar la creación de este tipo de plataformas.

5. Conclusiones

La implementación de una IDE provee los mecanismos necesarios para integrar a diferentes actores en el panorama regional propiciando una cultura de trabajo colaborativo entre las organizaciones que velan por el desarrollo local utilizando información geoespacial para la toma de decisiones.

Ante el reto que significa la creación de una IDE, existen soluciones adaptables para un conjunto de necesidades propias. Las soluciones de Software Libre constituyen una opción interesante para su implementación al ser estas altamente configurables y evidentemente accesibles en términos económicos.

Opengeo ha probado ser una herramienta robusta porque permitió gestionar capas de información de esta región, algunas de ellas con diferentes proyecciones geográficas, sin mayores problemas.

6. Trabajos Futuros

Los productos obtenidos a la fecha permiten establecer una plataforma robusta pero que requiere de trabajo adicional para su establecimiento definitivo.

Se deberá establecer un esquema de seguridad para las distintas relaciones que se generen entre los gestores de información geográfica que permita formalizar el uso de la información y los servicios que se ofrezcan y posibles gestiones colaborativas de los datos geográficos.

Estructurar la organización administrativa de las capas geográficas y los servicios brindados por el IDE para la estandarización entre las organizaciones participantes del IDE.

Se espera lograr una ampliación de geoservicios en función del crecimiento espontáneo que se llegue a tener y de las relaciones existentes entre las fuentes de información geográfica.

Referencias

[1] Gobierno de Costa Rica. Decreto 37773 JP-H-MINAE-MICIT. Gaceta N° 134. Julio 2013.

[2] Fazal, S. (2008). *GIS basics*. Daryaganj, Delhi, IND: New Age International.

[3] Kuhn, W. (2005) *Introduction to Spatial Data Infrastructures*. Institute for Geoinformatics, University of Münster. Alemania.

[4] Stefan Steiniger, Andrew J. S. Hunter (2012). *Free and Open Source GIS Software for Building a Spatial Data Infrastructure*. Geospatial Free and Open Source Software in the 21st Century pp 247-261.

[5] Bunzel, K., Ager, A., & Schrader-Patton, C. (2010). *Up in the air: Adventures in serving*

geospatial data using open source software and the cloud. Washington, D.C.: ACM.

[6] Gutiérrez, M (2006). *El Rol de las Bases de Datos Espaciales en una Infraestructura de Datos*. Global Spatial Data Infrastructure Conference. Santiago, Chile.

[7] Hazzard, E. (2011). *OpenLayers 2.10 beginner's guide*. Olton, Birmingham, GBR: Packt Publishing.

[8] Youngblood, B., & Iacovello, S. (2013). *GeoServer Beginner's Guide*. Olton, Birmingham, GBR: Packt Publishing.