

GeoINTA: compartiendo datos por medio de la Infraestructura de Datos Espaciales del INTA.

Santiago Banchemo¹ and Bellini Saibene, Yanina²

¹ Investigador, INTA, Instituto de Clima y Agua, Buenos Aires, Argentina.

² Investigador, INTA, Estación Experimental Anguil, La Pampa, Argentina.
{banchemo.santiago, bellini.yanina}@inta.gob.ar

Resumen GeoINTA es la Infraestructura de Datos Espaciales del INTA. Se trata de un proyecto federal que se maneja en red, fortaleciendo las vinculaciones hacia adentro y fuera del INTA. Desde adentro se ha implementado una red de geoservidores (Red GeoINTA) conectados a través de servicios de interoperabilidad, con esta visión distribuida se ha podido formar recursos humanos a través de capacitaciones en las temáticas ligadas a tecnologías propias de las IDE. También permite que desde todas las dependencias del INTA que producen información georreferenciada puedan contribuir con productos de calidad a la Red GeoINTA y alimentar tanto a las herramientas de visualización propias de la institución como también al resto de los interesados en este tipo de información. Este proyecto constituye un caso de éxito y fortalecimiento institucional, debido a que ha logrado integrar en una única plataforma gran parte de la información geográfica desarrollada por una institución que tiene presencia en todo el país.

Keywords: Bases de datos geográficas, IDE, Interoperabilidad, Open-Data

1. Introducción

Existe una enorme demanda de datos relacionada a los Recursos Naturales, Ambientales y Agropecuarios a través de Internet por parte de distintos usuarios, por lo que el Instituto Nacional de Tecnología Agropecuaria (INTA) desarrolló en los últimos 10 años su “Infraestructura de Datos Espaciales (IDE): GeoINTA”, cuyo objetivo principal es desarrollar y organizar un sistema de geoinformación en red que incluyó: un inventario de los datos, cartografía, bases de datos georreferenciadas y datos satelitales y aerofotográficos. Esta información, que está en constante actualización, permite a usuarios internos y externos contar con datos de base para la toma de decisiones, evaluación, seguimiento e inventariado de Recursos Naturales, Ambientales y Agropecuarios, disponibles en Internet, permitiendo su uso según sus necesidades. Tiene un alcance horizontal amplio, al generar una plataforma y procesos para multiplicidad de temáticas. La IDE ofrece gran variedad de datos georreferenciados, posibilitando una investigación más profunda que la simple consulta del dato, permitiendo asociaciones, análisis

espaciales y consumo en línea desde otras aplicaciones. GeoINTA contribuye a generar nuevas formas de colaboración e intercambio, explotando y maximizando la producción e interacción entre distintos grupos técnicos distribuidos a lo largo de Argentina, ofreciendo una red descentralizada para volcar información geográfica relacionada con el sector agropecuario y agroindustrial.

2. Situación-Problema u Oportunidad.

El INTA es un organismo estatal, dependiente del Ministerio de Agroindustria de la Nación. Desarrolla acciones de investigación y extensión para mejorar la competitividad y el desarrollo rural sustentable del país y genera información relacionada al sector Agroalimentario y Agroindustrial utilizada para la toma de decisiones, tanto en el ámbito agrario, ambiental, ordenamiento territorial y mitigación de riesgo. La institución tiene presencia en todo el territorio argentino, a través de una estructura que comprende: una sede central, 15 centros regionales, 6 centros de investigación, 21 Institutos de investigación, 53 estaciones experimentales y más de 300 Unidades de Extensión [1].

Una gran mayoría de los proyectos del INTA utilizan herramientas de teledetección o Sistemas de Información Geográfica (SIG) [3] para la investigación, así como para la difusión y publicación de sus productos o indicadores, las temáticas que más uso hacen de estas herramientas son Recursos Naturales (RRNN) y Ordenamiento Territorial.

El contexto tecnológico actual con la masificación de las comunicaciones y la adopción consolidada de tecnologías de la información en organismos como el INTA, han propiciado la producción (automática o manual), el relevamiento y la distribución de datos e información georreferenciada en gran parte de sus dependencias. Así mismo la incorporación de instrumentos de formalización como las tecnologías para el desarrollo de IDEs, sobretodo en el contexto del software libre, han propiciado el surgimiento de soluciones que facilitan la manipulación de datos espaciales, más la liberación a nivel internacional de datos, ha ido multiplicando los potenciales usuarios de información espacial [8].

Debido a la característica territorial de INTA y a partir de una pluralidad de proyectos y temáticas, la información generada en diferentes unidades de la institución estaba dispersa y aunque existían algunos emprendimientos para la publicación de productos y resultados espaciales en Internet, el mayor porcentaje de la información se encontraba almacenado en las computadoras de los técnicos que la generaban. Esta situación dificultaba la gestión de búsqueda y localización, y obligaba muchas veces a realizarla de manera personalizada en cada lugar con la pérdida de tiempo y duplicidad de esfuerzos. Asimismo la falta de estandarización adecuada hacía que muchos datos no pudieran ser utilizados por otros usuarios, tanto internos como externos y que se contara con diversas versiones de la misma capa de información (figura 1), dificultando la generación de información a niveles regionales y nacionales que implican el trabajo en red de varias Estaciones Experimentales Agropecuarias (EEA).

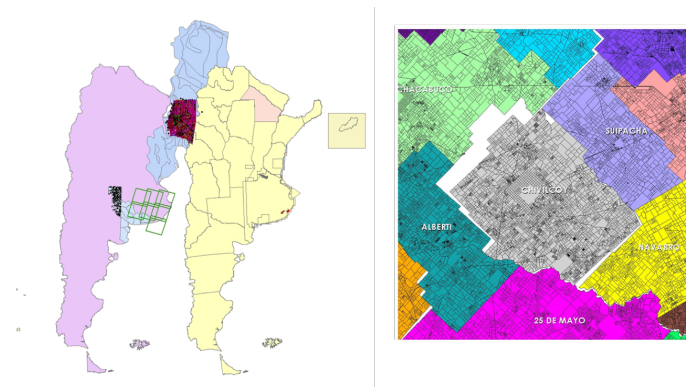


Figura 1. Ejemplos de capas de información que deberían coincidir realizadas por diferentes unidades y técnicos.

Tampoco se disponía de un catálogo de los datos espaciales producidos por INTA que estuviera actualizado y público a nivel nacional [10] para conocer la información y los datos existentes, como tampoco una plataforma que permitiera la publicación en línea de estos productos más allá de un mapa estático por medio de imágenes. Esta situación, sumado al interés en la utilización de procedimientos y protocolos de trabajo acorde a normas estándares de calidad, el explosivo desarrollo de herramientas de software libre y código abierto que permitieron cubrir adecuadamente las funcionalidades de los programas de base comercial, la demanda de los usuarios tanto internos como externos de acceso a información actualizada y de calidad, llevó al desarrollo de una IDE [8].

3. Solución

Las IDE son el conjunto de políticas, estándares, datos, metadatos y tecnologías relacionadas de forma sistémica que sustentan el descubrimiento, el acceso y el uso compartido de información interoperable en ambiente Web [11]. El proyecto se inicia en 2006 debido a la necesidad de compartir datos e información georreferenciada acerca de los RRNN, necesarios para el planteo productivo, los estudios de sostenibilidad, potencial productivo y riesgos para el sector agropecuario y a la demanda pública de acceder a esos datos e información vía Web. Uno de los objetivos de GeoINTA fue el diseño de una IDE dinámica e interactiva con niveles de acceso y actualización desde Internet con la finalidad de contribuir a minimizar la dispersión de información geoespacial actualmente existente y a la vez facilitar la investigación científica, la toma de decisiones, así como la generación de nueva información geoespacial de calidad [10].

El proyecto se basó en cuatro pilares: personal especializado, tecnología, disponibilidad de los datos y soporte institucional. Se necesitó trabajar en cambiar una cultura instalada que no promovía el intercambio, uso compartido, disponibilidad de los datos y consulta pública [8]. GeoINTA ofrece incentivos para

quien genera la información geoespacial potenciando los resultados y la difusión de datos e información, al tener como premisa que subir los datos en la IDE no implicaría la pérdida de la autoría del dato original, ni la posesión física por parte de la unidad productora del mismo, sino que incrementa las capacidades de cada grupo para la publicación de información y datos geoespaciales en línea y una plataforma informática para el intercambio y la coordinación de desarrollos. Se desarrollaron cuatro módulos de trabajo: Catálogo de metadatos, estándares y protocolos, Sistemas de consulta y acceso en-línea y Sistema de recuperación de datos históricos.

Catálogo de metadatos: primero se generó un catálogo de metadatos que permitió inventariar los mapas (papel y digital), bases de datos, aerofotografías e imágenes satelitales disponibles y producidas por INTA. Para ello hubo que consensuar estándares de metadatos, creándose el perfil de metadatos INTA compuesto por aproximadamente treinta elementos de acuerdo a normas ISO 19115 [12]. En cuanto a la cartografía digital, se definieron las referencias espaciales comunes, promoviendo los estándares de interoperatividad para intercambiar información y se proporcionaron las herramientas y conocimientos para que se ponga al alcance del público la información espacial que se ha ido generando. El diseño de la red informática propuesta y la funcionalidad del trabajo nodal requirieron de una adecuada conectividad a Internet en cada uno de los nodos participantes [8].

Estándares y protocolos: Se avanzó con la estandarización del software de la IDE y las características básicas de las capas de información que se publican en la misma (por ejemplo: todas deben estar en coordenadas geográficas (latitud, longitud), Datum WGS84 que no es más que un sistema de coordenadas geográficas a nivel mundial, que permite localizar cualquier punto de la Tierra [5]; y se deben completar una serie de metadatos). También se realizaron guías para la instalación, el uso y administración de los nodos y la información que gestiona y los servicios web que brinda.

Sistemas de consulta y acceso en-línea: el acceso a la información de GeoINTA se hace por medio de un sitio web (<http://geointa.inta.gob.ar>) que presenta los links a cada uno de los servicios y desarrollos de la IDE. Su principal exponente es el Visor GeoINTA, el cual ha simplificado sobremanera la divulgación de información por medio de unidades organizativas llamadas proyectos, las cuales presentan de forma ordenada y unificada una serie de capas de información para su consulta. Al inicio presentaba una administración centralizada que se convirtió en un cuello de botella para la publicación de nuevos proyectos; con el desarrollo del Gestor GeoINTA, una herramienta que integra todos los servicios de interoperabilidad que están en la IDE y genera visualizaciones de manera dinámica, se federalizó la publicación de proyectos en el visor (figura 2).

Finalmente, la creciente adopción de dispositivos móviles y el fácil acceso a redes de datos inalámbricas, han propiciado la aparición de herramientas para que las personas puedan usufructuar la información geográfica de manera más simple, en este sentido se desarrolló una interfaz Web responsive para móviles (<http://:movil.geointa.inta.gob.ar>), que facilita el acceso y la consulta de la in-

formación existente en GeoINTA. La herramienta es multiplataforma y accesible desde cualquier navegador Web. Permite consultar varios servicios de la red de geoservidores y recibir las respuestas acordes a la posición consultada, separando los contenidos en locales y nacionales para agilizar el proceso de consultas [7].

Árbol de capas

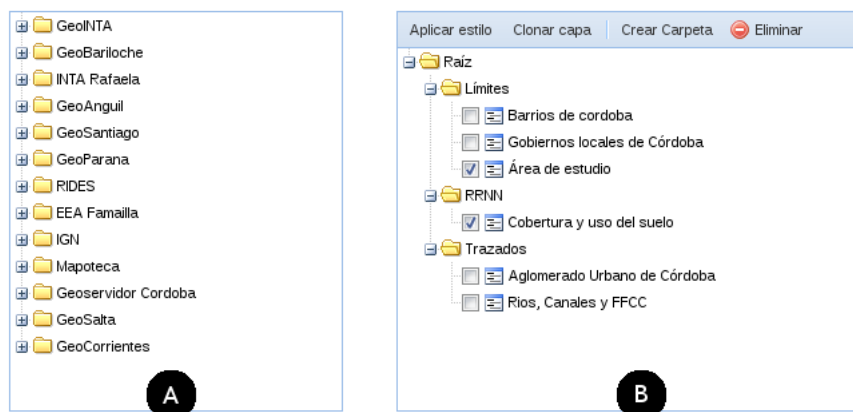


Figura 2. Gestión GeoINTA permite vincular una serie de repositorios IDEs A) que pueden ser parte de la Red GeoINTA o externos - como el Instituto Geográfico Nacional (IGN) - y facilita la creación de árboles de capas B) donde es posible organizar y definir diferentes parámetros de configuración, como por ejemplo si la capa será provista por un cache, o con qué estilo (SLD - *Styled Layer Descriptor* [4]) se quiere desplegar esa capa, qué atributos aparecen en la consulta, etc.

Sistema de recuperación de datos históricos: los datos históricos aerofotográficos y cartográficos se digitalizaron y georreferenciaron. Son de gran valor para muchas aplicaciones del INTA y uno de los contenidos más consultados.

Para realizar la implementación del proyecto se realizaron reuniones con los productores de datos espaciales, referentes de cada EEA o laboratorio de recursos naturales con el fin de transmitir los beneficios del proyecto, generar el perfil de metadatos del INTA y discutir la ubicación de los nodos, los cuales se fueron incorporando gradualmente. A estas actividades, se suman la organización de capacitaciones para los participantes de la red y los productores de datos, ser miembros de IDERA (IDE de Argentina) y principalmente realizar concientización, enfocados a vencer la idiosincrasia de los productores de datos, que hace difícil el compartir la información desinteresadamente.

4. Innovación e Inédito

A diferencia de otras iniciativas de construcción de infraestructuras de datos espaciales GeoINTA ha sido implementada como una IDE formada por nodos IDEs distribuidos por todo el país. Esto se debe a la naturaleza federal de INTA que permite vincular a diferentes grupos de trabajo formados por técnicos de diferentes disciplinas que producen información georreferenciada. El método de replicación de nodos se centra principalmente en la implementación de un ecosistema de tecnologías de software libre y en la recuperación de equipos servidores subutilizados existentes en algunas dependencias de INTA. Esta solución ha permitido incrementar sobremanera el número de nodos publicadores distribuidos a lo largo y ancho del país. La ventaja estratégica de contar con nodos IDE es que además de contribuir a la red GeoINTA son los articuladores directos con las IDEs locales que requieren la información geográfica de INTA.

La integración de todos los nodos de la Red GeoINTA es a través de un sistema web de geoinformación para consulta y procesamiento de multiplicidad de datos georreferenciados del país. Allí se incluye cartografía, bases de datos geográficas, datos satelitales y aerofotográficos, disponibles para evaluación, seguimiento e inventariado de recursos naturales, ambientales y agropecuarios [6]. Este sistema permite un rápido acceso a la información existente, en tiempo real, desde cualquier sitio, en cualquier momento y con la mejor calidad disponible que se publica desde origen, poniendo a disposición del público y de los tomadores de decisión datos espaciales y mapas interactivos que facilitan el conocimiento de nuestros recursos naturales, ambientales y agropecuarios, que a su vez sirven de medio para actividades de planificación, permitiendo acercar al productor de datos e información geoespacial con el usuario, de manera directa.

5. Beneficiarios

Los primeros beneficiarios son los propios Investigadores del INTA generadores de información geoespacial y otras instituciones de ciencia y técnica al poder intercambiar y aumentar la información disponible. La posibilidad de realizar la edición compartida de una capa de trabajo simplificó las fases operativas de generación y actualización de un contenido desarrollado en red entre diferentes EEA, como también la publicación en línea de los datos.

El desarrollo de sistemas informáticos que necesitan información territorial o la capacidad de consultar y analizar la información desde una perspectiva espacial (ej: <http://rian.inta.gob.ar/prets>) tienen el acceso facilitado por medio de servicios web de GeoINTA, asegurando su acceso y actualización. Ya no es necesario conocer al productor de datos para poder acceder a los mismos y tampoco es necesario revisar cual es la última versión de la capa de interés.

También las cadenas productivas, productores, empresas, consultores, inversores, inmobiliarias, contratistas, desarrolladores de *DataProduct*, al contar con la última información espacial del INTA disponible desde un único lugar y sin necesidad de intervención humana. Toda la comunidad educativa de escuelas

técnicas agropecuarias, escuelas en niveles primarios y secundarios (contenidos de Recursos Naturales, Ambiente y Agropecuario-Agroindustrial) y comunidad universitaria. Los medios de comunicación también son usuarios de la plataforma. Finalmente, las diferentes instancias de gobierno, al contar con esta información para el ordenamiento territorial y la toma de decisiones.

6. Relevancia para el Interés Público

Reconociendo la responsabilidad de desarrollar y mantener una IDE para la gestión de la información geoespacial del INTA y dadas las implicancias de la organización federal del INTA, es requerimiento indispensable una adecuada articulación institucional para llegar a ser integralmente eficaces. Para el desarrollo de este sistema, las tecnologías, estándares y normas que se emplean deben ser replicados en cada Unidad de INTA generadora de datos espaciales. Ordenar y coordinar las responsabilidades al respecto conlleva un trabajo de cooperación e intercambio entre los profesionales responsables del procesamiento de dichos datos en cada Unidad, evitando la superposición de esfuerzos, generando la reducción de costos, favoreciendo el uso compartido, ordenando el cúmulo de información, y materializando un sitio de filosofía “clearinghouse” [6], [8]. GeoINTA presenta un portal de datos espaciales abiertos institucional, cumplimentando el derecho de acceso a la información pública, generada por INTA, multiplicando su uso y por ende haciendo un uso más eficiente de los recursos del estado; pero no se queda allí, facilita una plataforma tecnológica que contribuye a generar nuevas formas de colaboración e intercambio, explotando y maximizando la producción e interacción entre distintos grupos humanos y técnicos distribuidos a lo largo de la República Argentina, ofreciendo una red descentralizada para volcar información geográfica.

7. Efectividad

Internamente, en el transcurso de estos 10 años, se han visto muchos avances en el compartir la información, se puede ejemplificar en el hecho que al inicio había que ir a buscar a los productores de datos y hoy muchos productores de datos vienen a ofrecer sus resultados a GeoINTA. Se inició con tres nodos, actualmente se cuenta con once en funcionamiento (Salta, Famaillá, Santiago del Estero, Saenz Peña, Manfredi, Rafaela, Paraná, Anguil, Bariloche, Corrientes e Instituto de Clima y Agua de Castelar) y cuatro (San Luis, Balcarce, Alto Valle y Mendoza) en proceso de instalación, aprovechando hardware existente en muchos casos y sin necesidad de adquirir licencias por el uso de software libre. Desde el portal GeoINTA se puede acceder a:

- Catálogo de metadatos geoespaciales, que cumplen con estándares de la ISO 19115 e ISO 19139 [9] para consulta de información disponible y Clearinghouse de metadatos implementado con Geonetwork Opensource.
- Normas y estándares de calidad para la publicación de datos.

- Bases de datos geográficas o geodatabases con Mapas de Suelos. Mapoteca Digital, Visor integrado, Mapas climáticos, Mapas de cobertura de la tierra, etc. Actualmente se cuenta con cerca de 500 capas de información disponibles en todos los nodos de la red.
- Recursos OGC (Open Geospatial Consortium) [2]. La plataforma brinda acceso a través de WMS, WFS, CSW, WCS, entre otros servicios de interoperabilidad y también permite acceder a información en formatos estandarizados y abiertos como Shapefiles, GeoJSON y CSV entre otros. Se pueden acceder desde la sección de servicios del portal GeoINTA a todos los WebServers de cada uno de los nodos.

El sitio web tiene un promedio de 6.000 visitas mensuales, el visor de GeoINTA tiene 500 visualizaciones diarias y los servicios de la IDE responden unos 10 millones de peticiones de mapas a la semana.

8. Facilidad de Reproducción

El hecho de utilizar software libre y estándares abiertos le da a GeoINTA una capacidad de réplica y reproducción muy alta. Cualquier institución que desee desarrollar una IDE puede acceder a la experiencia de GeoINTA, o bien desde uno de sus nodos o desde la coordinación. Se han publicado las diversas experiencias en congresos y jornadas de IDEs, como los manuales y guías de trabajo.

9. Ambiente de Hardware y Software

Todo el diseño de GeoINTA es acorde a los estándares técnicos de presentación de datos espaciales, reuniendo las normas ISO, *Federal Geographic Data Committee* (FGDC) y nacionales del IGN. Asimismo se asienta en la experiencia y sugerencias de desarrollo de OGC. La figura 3 presenta el ecosistema GeoINTA: ambiente de hardware y software. El almacenamiento de los objetos geográficos está implementado utilizando un gestor de datos espaciales PostgreSQL/Postgis que simplifica el manejo de información georreferenciada. La principal ventaja de esta tecnología es que permite un almacenamiento basado en estándares abiertos e interoperables, que permiten su uso y publicación.

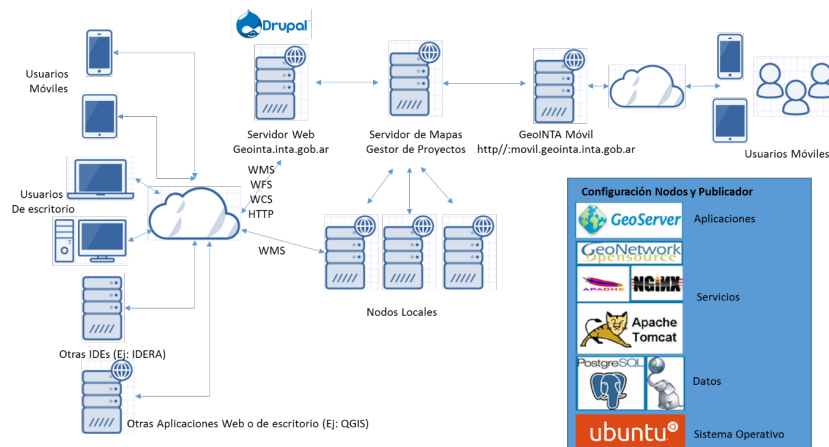


Figura 3. Ecosistema GeoINTA: detalle del mapa tecnológico utilizado

La implementación de los servicios de interoperabilidad están basados en Geoserver utilizando Tomcat como servidor de aplicaciones y Nginx como proxy reverso. A su vez se utiliza GeoWebCache, hoy en día contenido en Geoserver, para mejorar los tiempos de respuesta y optimizar los recursos de cómputo.

En el nodo primario de GeoINTA la configuración de Geoserver se ha realizado a través de un cluster horizontal donde se encuentran corriendo tres instancias del servidor de mapas en diferentes equipos y delante de ellos un módulo Load-balancer de Nginx realiza el balanceo de carga. Esto permite afrontar ciertos picos de consumo en la plataforma que atentan contra su estabilidad. Sobre todo cuando se consulta información crítica con una alta cantidad objetos espaciales, donde el renderizado o la generación de estos objetos en determinados formatos es extremadamente costoso computacionalmente.

Otra aplicación existente en el ecosistema tecnológico es el gestor de metadatos Geonetwork. Esta herramienta permite cargar y gestionar metadatos y perfiles de metadatos como también facilita el proceso de cosecha (o harvesting) entre diferentes servicios de catálogo (CSW – Catalog Service Web) u otros servicios de interoperabilidad.

10. Comentarios Finales

Este proyecto constituye un caso de éxito y fortalecimiento institucional, debido a que ha logrado integrar en una única plataforma gran parte de la información geográfica desarrollada por una institución que tiene presencia en todo el país. Donde a través de la Red GeoINTA se ha dotado a cada nodo con capacidades para publicar desde origen sus contenidos georreferenciados y se ha formado recursos humanos para el operar y mantener esos nodos. El método de

replicación de los nodos utilizando tecnologías de software libre ha facilitado que la red escale en cantidad de nodos y garantiza la sustentabilidad en el tiempo.

Referencias

1. ¿Qué es el INTA? | INTA :: Instituto Nacional de Tecnología Agropecuaria. <http://inta.gob.ar/queeselinta>, accedido el 2016-07-08
2. About OGC | OGC (2016), <http://www.opengeospatial.org/ogc>, accedido 2016-07-11
3. Sistema de información geográfica - Wikipedia, la enciclopedia libre (2016), https://es.wikipedia.org/wiki/Sistema_de_informaci%C3%B3n_geogr%C3%A1fica, accedido el 2016-07-11
4. Styled Layer Descriptor | OGC (2016), <http://www.opengeospatial.org/standards/sld>, accedido el 2016-07-11
5. WGS84 (May 2016), <https://es.wikipedia.org/w/index.php?title=WGS84&oldid=91237388>, 00000 Page Version ID: 91237388
6. Banhero, S., Elena, H.J., López, J., Pizarro, M.J., Umaña, F.J.: La red geointa una solución para la gestión de datos espaciales del inta. Jornadas de La Infraestructura de Datos Espaciales. 8. 2013 11 06-08, 6, 7 Y 8 de Noviembre de 2013. San Carlos de Bariloche, Río Negro (2013)
7. Banhero, S., Marrone, A., Elena, H., Lopez, J., Umaña, F.: Experiencia en aplicaciones para dispositivos móviles utilizando servicios de interoperabilidad. Jornadas de La Infraestructura de Datos Espaciales. 10. 14 y 15 de mayo de 2015. Mendoza (2015)
8. Cruzate, G., Mercuri, P., B.S.P.M.: The GEOINTA spatial data system. presented at the sharing enviromental data information: issues of open enviromental data in latin america. Porto Alegre, Brasil. (2012)
9. Maganto, A.S., Iso, J.N., Ballari, D.: Normas sobre metadatos (iso19115, iso19115-2, iso19139, iso 15836). en Mapping (123), 48–57 (2008)
10. Mercuri, P., Cruzate, G., Pizarro, M.J., Banhero, S., Walkman, C., Gómez, L.: Sistema de información geoespacial del inta–geo-inta. Seminario de Tecnología de Información en Geociencias. Reunión Latinoamericana de Especialistas en Geoinformación de Servicios Geológicos. 1. 2009 06-07 30-02, 30 de junio al 2 de julio de 2009. Buenos Aires. AR. (2009)
11. Olaya, V.: Sistemas de información geográfica. Cuadernos internacionales de tecnología para el desarrollo humano (8), 15 (2009)
12. Ramos, L., Mercuri, P., Bellini Saibene, Y., Pizarro, M.J.: Catálogo de datos geoespaciales del instituto nacional de tecnología agropecuaria. Jornadas Argentinas de Informática. 37. Jornadas de Informática Industrial. 2008 09 8-12, 8 al 12 de septiembre de 2008. Santa Fe. AR. (2008)