

LA CONSTRUCCIÓN DESCENTRALIZADA DE DATOS ESPACIALES: RIESGOS PARA LA CALIDAD DE LA INFORMACIÓN GEOGRÁFICA

Francisco Javier ARIZA LÓPEZ

Departamento Ingeniería Cartográfica, Geodésica y Fotogrametría. Universidad de Jaén

Recibido: 09/03/2015

Aceptado: 17/07/2015

RESUMEN: En este trabajo se define en qué consiste un marco de «construcción descentralizada» de datos espaciales, el cuál presenta matices diferenciadores respecto a lo que se ha venido denominando «producción descentralizada». Para entender lo que pueden suponer los riesgos de la mala calidad, primeramente, se refuerza la idea de la importancia de los datos espaciales en la toma de decisiones. Posteriormente se identifican y presentan los que entendemos que, al día de hoy, son los principales retos: producción (comprometerse en mucho), acceso (permitir accesos de calidad), interoperabilidad (posibilitar la utilización y uso legal), metadatos (describir bien los productos), trazabilidad y metacalidad (mantener el linaje e información sobre la calidad de la calidad) y la preservación (asegurar del uso futuro). El trabajo acaba con unas breves conclusiones en las que se identifican las principales líneas de trabajo (normalización, formación, liderazgo, etc.) para evitar estos riesgos.

PALABRAS CLAVE: datos espaciales, calidad, riesgos.

DECENTRALIZED SPATIAL DATA CONSTRUCTION: RISKS FOR QUALITY OF GEOGRAPHIC INFORMATION

ABSTRACT: This paper defines what constitutes a «decentralized construction» framework of spatial data. This framework presents some differences with respect to what has been called «decentralized production» framework. First, in order to understand the possible risks due to poor quality, we reinforced the importance of spatial data in decision-making. Subsequently we identify and present some challenges that we understand are the main challenges today: production (commit much), access (accessing with quality), interoperability (enabling the use and legal use), metadata (describing well the products), traceability and metacalidad (maintaining the lineage and information about quality) and preservation (ensuring the future use). The work ends with a brief conclusion in which are identified the main lines of work (standardization, training, leadership, etc.) in order to avoid these risks.

KEY WORDS: spatial data, quality, risk.

I. INTRODUCCIÓN

El título de este trabajo hace referencia a los siguientes conceptos: datos espaciales, construcción descentralizada, calidad, riesgos; los cuáles conviene

aproximar antes de entrar en mayores profundidades.

Los datos espaciales son cualquier tipo de dato que se encuentre referido, directa o indirectamente, a un espacio. De esta definición se deduce que cualquier dato (p.ej. un dato estadístico, un número de móvil, etc.) es un dato espacial, lo que va en la línea del pensamiento actual relativo a la ubicuidad de los recursos TIC (Tecnologías de la Información y de las Comunicaciones). No obstante, tradicionalmente se ha considerado como datos espaciales a aquellos expresados en forma de puntos, líneas, áreas o superficies. En nuestro caso preferimos hablar de «Información Geográfica» (IG) en lugar de «datos espaciales» pues consideramos que la forma en que se almacenan (bases de datos espaciales) y la mayoría de los productos en que se ofrecen son algo más que simples datos, hay modelización de la realidad y un importante conocimiento incluido.

La construcción descentralizada es uno de los aspectos claves del encargo que pretende cubrir este trabajo. La construcción descentralizada viene a ser sinónimo de producción descentralizada, pero con el matiz de que una construcción supone la interacción final de los elementos y por ello deber ser, finalmente, funcional. La construcción descentralizada está unida a la idea del ensamblaje de un componente mayor, de mayor valor añadido y capacidades, a partir de muchos y pequeños elementos que se deben ensamblar de manera adecuada. Por ejemplo, un vehículo es un ejemplo de construcción descentralizada en cuanto al origen de sus partes. Este aspecto no queda del todo claro si se habla de producción descentralizada. Por ejemplo, la producción descentralizada puede ser de materias primas (p.ej. áridos, cereales, etc.), las cuáles se pueden producir de una manera descentralizada. El aspecto descentralizador de la producción no es novedoso y ha existido siempre en el campo de la IG, así, por ejemplo en el caso de España, a lo largo de las últimas décadas hemos vivido una situación de producción descentralizada de IG, tanto en las temáticas (p.ej. el Instituto Geográfico Nacional en cartografía básica, el Instituto Geológico y Minero en cartografía geológica, etc.), como en lo territorial (p.ej. el Instituto Cartográfico de Cataluña en Cataluña, el Instituto de Estadística y Cartografía de Andalucía en Andalucía, etc.). Bajo mi punto de vista, lo que ocurre en la actualidad es que, debido a la Web 2.0 y al marco que se establece con las Infraestructuras de Datos Espaciales (IDE), el aspecto de «descentralización» es aún más radical. Así, aparecen aquí dos maneras distintas de crear datos espaciales:

- ✓ La VGI (*Volunteered Geographic Information*, Goodchild 2007) o colaborativa. Aquí el término «voluntarios» debemos entenderlo en un sentido amplio y no sólo ligado a personas individuales, también se pueden considerar organizaciones. Claves para entender y definir esta realidad

- son la democratización de la cartografía, la apertura en los datos, el desarrollo de aplicaciones colaborativas Web 2.0 y de herramientas FOSS (*Free Open Source Software*) (STEINIGER Y HUNTER, 2012) y las posibilidades de captura «*low cost*» de datos, así como el auge de movimientos VGI (p.ej. OSM).
- ✓ La interadministrativa. Esta forma de producción de datos espaciales siempre ha existido, pero con motivo de Inspire, tanto en Europa como en España, se está estableciendo una nueva forma de producir datos. La idea de productor de todos los datos asociados a un mapa (p.ej. Mapa Topográfico Nacional) se diluye dando paso a los productores de capas temáticas (p.ej. red viaria, red hidrográfica, etc.) por parte de las organizaciones que realmente tienen competencia en esas materias y que, además, en muchos casos son afectadas por otras directivas (p.ej. Marco de Aguas).

En este trabajo se considerará que en un producto creado de manera descentralizada cada instancia de dato espacial puede tener un origen distinto. Bajo la perspectiva de construcción de productos de datos espaciales el objetivo es que los datos procedentes de distintas fuentes se ensamblen adecuadamente para ofrecer un producto de mayor valor y adecuación a un uso concreto. Lo que realmente se plantea aquí es un problema de interoperabilidad bajo distintas perspectivas. Esta será la principal temática a desarrollar aquí.

Por otra parte, el concepto de calidad ya es algo cercano a todos, se utiliza en el lenguaje coloquial y es entendido y aceptado por todos. En general, podemos decir que una obra bien hecha tiene calidad o es de calidad. El Diccionario de la Lengua Española la define como *Propiedad o conjunto de propiedades inherentes a una cosa, que permiten apreciarla como igual, mejor o peor que las restantes de su especie*. Esta definición aclara pues que la calidad no tiene por qué limitarse a una única propiedad del objeto que se considera, sino que pueden ser varios factores los que entren en juego para definir la calidad. Por otra parte, lo inherente es lo propio o inseparable de las cosas, y aquí cabe matizar que dentro de lo inherente hay factores que son más evidentes, o explícitos, que otros que poseen un carácter más implícito. Otro aspecto de interés de esta definición es el que hace referencia a que la calidad debe ser apreciada y comparada. Esto significa que cada individuo podrá tener una apreciación distinta, y que se utiliza la comparación para valorarla.

Desde la perspectiva de los procesos de producción industriales, que es donde mayor experiencia hay en los temas de calidad, el concepto de calidad ha ido evolucionando durante los últimos siglos, desde definiciones muy limitadas hasta ideas que se constituyen como verdaderas filosofías. En el principio de la industrialización, la calidad se consideraba simplemente como la

adecuación del producto a las especificaciones establecidas. En la actualidad una de las definiciones más aceptadas y completas de calidad se refiere a la totalidad de las características de un producto o servicio tal que le confieren su aptitud para satisfacer unas necesidades explícitas e implícitas. Subyace aquí la idea de uso que se utiliza también para definir la calidad de un producto o servicio como su aptitud o idoneidad para el uso (*fitness for use*).

Un riesgo lo podemos definir coloquialmente como la contingencia o posibilidad de que ocurra algo que resulta perjudicial. Dado que este artículo se centra en la IG de manera general, este riesgo lo entendemos aquí como la contingencia de que las posibilidades de uso de la IG, o sus propios rasgos intrínsecos de calidad, puedan verse afectados por el hecho de una construcción descentralizada.

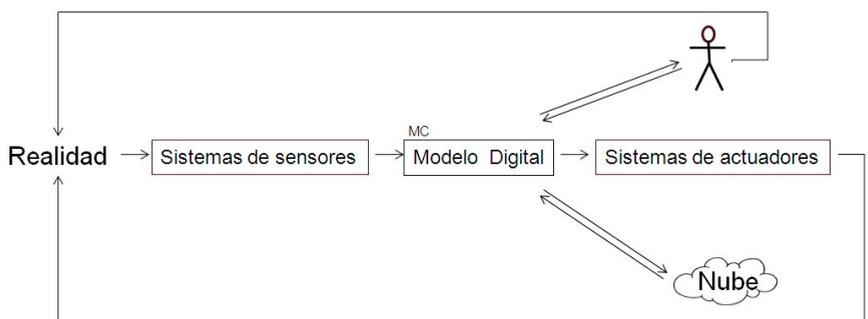
Ya definidos, aunque brevemente, los conceptos fundamentales que titulan este trabajo, también opino que es importante en esta introducción plantear unos esbozos relativos al futuro inmediato de lo que concierne y rodea a la IG. Así, considero que el futuro que condiciona los aspectos de la calidad de la IG va más allá del marco establecido por las IDE y que se conforma como un nuevo paradigma, que denomino Geomática 3.0, que es un sistema dinámico de interacción entre el mundo real y el mundo virtual (FIGURA 1) sin pérdida de continuidad por la necesidad de la intervención humana. En este sistema los sensores realizan el seguimiento del mundo real en tiempo real y la IG modela la realidad de manera detallada, enriquecida y contextualizada. Se trata de una «realidad virtual», que funciona como copia digital del mundo real y que permite la interacción con las personas, con sistemas de Inteligencia Artificial y con la Nube. Además, mediante actuadores el sistema puede influir en el mundo real, modificar la realidad y generar la necesidad de realimentar los procesos. De tal forma que la IG ya no es algo para uso e interpretación exclusiva de los operarios humanos, la IG también es la base de sistemas inteligentes y automatizados para intervenir en el mundo real, a distintas escalas y niveles de detalle (p.ej. casa, edificio, ciudad, territorio, etc.).

En relación a este nuevo paradigma y a la IG, de forma resumida, y bajo mi punto de vista personal, los siguientes rasgos son los más definitorios:

- ✓ Captura de datos cada vez más masiva, barata y descentralizada.
- ✓ Mayor participación de los ciudadanos y organizaciones no cartográficas en la creación de IG.
- ✓ Desarrollo de redes de sensores de todo tipo apoyadas en la Internet de las cosas (Sensor Observation Services (OGC, 2012), etc.), etiquetas de radiofrecuencia, huellas digitales, VGI (GOODCHILD, 2007), etc.

- ✓ Datos 4D nativos. Inclusión de los ciclos de vida de los fenómenos 3D en los modelos.
- ✓ Integración definitiva de la IG en todo tipo de modelos, particularmente en aplicaciones de la Inteligencia Artificial.
- ✓ Desarrollo del Territorio Inteligente (*smart cities* (Wikipedia, 2013)).
- ✓ Producción de datos oficiales alineada con la filosofía Inspire de confederación de productores y basada en «datos únicos».
- ✓ Adopción de identificadores globales únicos (IIGUU), de la trazabilidad en la IG y de los procesos de generalización automatizados.
- ✓ Mayor responsabilidad legal de los productores de datos.
- ✓ Mayor necesidad de normalización de los aspectos de la calidad.
- ✓ Integración de la IG en la web semántica (BERNERS-LEE, 2001) (datos espaciales enlazados).
- ✓ Incremento del desarrollo de servicios de procesamiento.
- ✓ Mayor cantidad de productos de IG, mayor desarrollo de productos de valor añadido.
- ✓ Mayor necesidad de técnicos especializados en calidad de la IG.

Figura 1. Evolución del modelado cartográfico y su uso



G3.0 = interoperabilidad + integración + almacenamiento + inteligencia + actuación en tiempo real

!!!Este ya es un sistema continuo realidad <-- --> virtualidad!!!



FUENTE: Elaboración del autor

Además de esta introducción, este trabajo incluye un apartado dedicado a explicar la importancia de la cartografía y de su calidad y, posteriormente, otro apartado dedicado a presentar los que, al día de hoy, considero son los principales riesgos de la construcción descentralizada. Aquí se tratará sobre: la producción, el acceso, los metadatos; en el fondo, sobre la interoperabilidad. En este apartado también se presentará la necesidad de incorporar los aspectos de trazabilidad, metacalidad como elementos que permitan gestionar la calidad

en un paradigma descentralizado. El trabajo acaba con unas conclusiones finales e indicando unas líneas de actuación para minimizar los riesgos presentados.

II. IMPORTANCIA DE LA CALIDAD DE LA IG

Considero que la actividad productiva de IG ha de entenderse en el campo de las tecnológicas y que, como tal, debe tener una clara orientación a la producción de utilidades, ya sean bienes o servicios, y, más concretamente, a la producción de modelos virtuales sobre cualquier realidad que nos rodea con el objetivo de ayudar en la toma de decisiones, por ejemplo:

- ✓ Diseño de anteproyectos, proyectos de ejecución de infraestructuras e intervenciones sobre el territorio.
- ✓ Explotación del territorio y de infraestructuras.
- ✓ Inversiones inmobiliarias y asentamiento de empresas.
- ✓ Planificación del trabajo (rutas, logística...).
- ✓ Planificación del ocio (turismo...).
- ✓ Defensa nacional y protección civil (emergencias, riesgos).
- ✓ Defensa de la naturaleza (protección de especies).
- ✓ Política agraria y forestal (política agraria común...).
- ✓ Etc.

La FIGURA 2 muestra el paradigma en el que tradicionalmente se ha desarrollado la Cartografía como hacer al servicio de la sociedad, en concreto cuando es utilizada por parte de las administraciones públicas. Pensando ahora en IG, *mutatis mutandis*, este marco de toma de decisiones sigue siendo válido, tanto a nivel de las administraciones como de las personas individuales. Básicamente el proceso mostrado es el siguiente:

- ✓ Del mundo real, mediante un proceso de abstracción, modelización, captura y depuración se obtienen bases de datos espaciales.
- ✓ Las bases de datos espaciales se someten a análisis e incorporación de nuevas informaciones y conocimientos.
- ✓ A partir de las bases de datos espaciales se desarrollan políticas, planes y proyectos y se modifica la realidad que ha de volver a ser capturada en forma de nuevos datos geográficos.
- ✓ En resumen, la creación de datos espaciales consiste en hacer modelos de la realidad. Estos modelos se plasman en IG que se presenta en forma de productos (p.ej. mapas o conjuntos de datos espaciales) y servicios (p.ej. rúters en la Web) que se utilizan como base o información cuyo análisis permite el desarrollo de proyectos, planes y políticas, es decir, actuaciones y programación de inversiones. Por tanto, se puede afirmar que la

cartografía es un hacer importante, que tiene consecuencias reales sobre el medio y la sociedad, y que afecta a individuos concretos en su vida, bienestar, derechos y obligaciones y por ello las administraciones encargadas de crear y gestionar IG deben procurar que ésta alcance los niveles de calidad adecuados.

Figura 2. La cartografía como herramienta para la actuación



FUENTE: Elaboración del autor

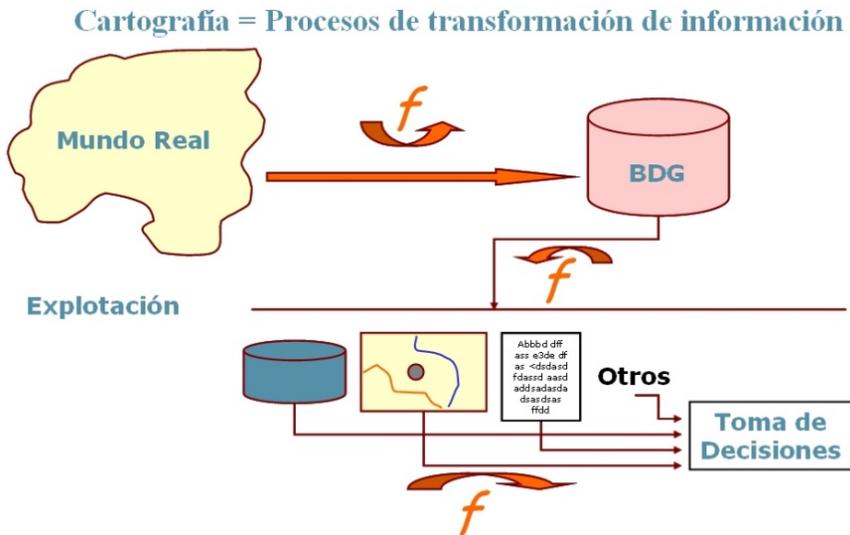
De una forma abstracta, y muy general, podemos decir que el hacer cartográfico (FIGURA 3) es un proceso de transformación de datos e informaciones del mundo real en un producto o productos cuya explotación sirve para la toma de decisiones, muchas veces con importantes consecuencias económicas, sociales, ambientales, etc.

Es justamente por este papel de elemento base para la toma de decisiones, que la calidad es importante para la IG, y cuanto más importantes sean las consecuencias del uso de la IG, más importante debe ser la atención prestada a sus aspectos de calidad.

La IG impregna todo el hacer cotidiano de las personas y de los gobiernos. Así, el informe anual del *Federal Geographic Data Committee* correspondiente a 2006 indicaba que entre el 80%-90% de la información gestionada por los gobiernos tiene una componente geográfica. Igualmente, la *Geospatial Information and Technology Association* informaba que entre el 70%-80% de la información gestionada en los negocios está vinculada a posiciones espa-

ciales. La importancia económica de la IG ha quedado evidenciada en numerosos informes, entre los que destacan el estudio realizado por la Comisión Europea (PIRA Y OTROS, 2000) que estima el valor económico de la PSI en Europa en 60 – 70 mil millones de euros al año, de los que más de la mitad corresponden a IG. El estudio de OXERA (1999) indicaba que el valor añadido de las actividades soportadas por la IG alcanzaba en GB aproximadamente 100 000 millones libras esterlinas. En el informe de OXERA (2013) preparado para Google planteado a escala mundial se establece un rango entre 150 y 270 billones de dólares, llegando a establecer incluso estimaciones de su impacto en el ahorro de tiempo, de combustibles, respuesta frente a emergencias, educación y competitividad.

Figura 3. Cartografía como proceso de creación de modelos y transformaciones de la información para el apoyo a la toma de decisiones



Ejemplos más sencillos relativos a la importancia de la calidad de la IG son los que se derivan directamente de las consecuencias de su mala calidad. Un campo de mucho interés al respecto es el de la cartografía náutica pues su realización tiene unas obligaciones y responsabilidad bien delimitadas. Valgan los siguientes ejemplos:

- ✓ Caso Urquiola: En mayo de 1976 el petrolero monocasco Urquiola quedó embarrancado en la Coruña. Una mala señalización de una aguja rocosa en la cartografía náutica hizo que este gran buque sufriera un siniestro. Según Sentencia del Tribunal Supremo (18 de julio de 1983,

de la Sala 4^a), se reconoce un anormal funcionamiento del servicio público de Cartografía Marina y de información sobre el mar y litoral motivado por la no incorporación de la presencia de las «agujas» en el fondo marino. Con motivo de lo anterior, se condenaba a la Administración General del Estado Español a indemnizar por la situación causada.

- ✓ **Caso Rockness:** En enero de 2004 el M/V Rocknes golpeó con un banco en cerca de Bergen (Noruega) tras sólo unos minutos después de salir de puerto, muriendo 18 miembros y vertiendo gran cantidad de fuel. Durante muchos años, y en numerosas cortes de justicia, se estuvo determinando si el Estado era el responsable, pues en la cartografía no había constancia de ese banco (http://www.hydro-international.com/issues/articles/id1053-The_Rocknes_Drama.html).

Figura 4. Identificación de un mismo terreno, incluido erróneamente en dos términos municipales diferentes y con identificación y número de plantas de olivar no coincidentes



FUENTE: Elaboración del autor

Los casos anteriores pueden parecernos alejados de nuestra vida cotidiana; sin embargo, la IG también nos afecta en nuestra vida diaria, valgan como ejemplo los dos casos siguientes:

- ✓ **Seguros.** Para el caso de las inundaciones, las empresas aseguradoras tienen en cuenta la situación de los inmuebles. Puede ocurrir que la mala calidad de una cartografía afecte o beneficiando/perjudicando bien a la empresa aseguradora o al asegurado. Si la cartografía no refleja bien el nivel de riesgo será la empresa aseguradora la que salga perdiendo con los seguros y si la cartografía sobreestima el nivel será el asegurado el que deberá pagar más o se quedará sin posibilidad de suscribir una

póliza. En Australia ha llegado a ocurrir que las aseguradoras no aceptan clientes en las zonas con mala calidad de la cartografía (<http://www.news.com.au/finance/business/bad-maps-reason-for-lack-of-flood-insurance/story-e6frfkur-1226037037453>).

Subvenciones al campo. Gran parte del presupuesto de la Unión Europea se dedica al sostenimiento de la Política Agraria Común. Las ayudas al campo se gestionan en gran medida por medio de las parcelas catastrales, las cuáles están asignadas a un polígono catastral y a un término municipal. En casos como el olivar o de los cítricos cada una de las plantas está contabilizada. Así, casos como el de la FIGURA 4 generan problemas a la hora del cobro de estos subsidios (gestiones que debe realizar el propietario, informes para subsanar los errores, retrasos en el cobro de las ayudas, etc.).

III. PRINCIPALES RIESGOS

La situación descrita en los apartados anteriores nos permite considerar los siguientes aspectos como aquellos que, para el caso de la construcción descentralizada de IG, concentran los mayores riesgos inmediatos:

- ✓ Producción: no comprometerse con todo lo que significa.
- ✓ Acceso: no permitir accesos de calidad.
- ✓ Interoperabilidad: no posibilitar la utilización y uso legal.
- ✓ Metadatos: no describir bien los productos.
- ✓ Trazabilidad y metacalidad: no mantener el linaje e información sobre la calidad de la calidad.
- ✓ Preservación: no asegurar el uso futuro.

A continuación se presentan en mayor detalle cada uno de estos aspectos.

III.1. Producción: no comprometerse con lo que significa

Podemos pensar que actualmente la producción de IG es sencilla pues está muy facilitada por las numerosas y cada vez más baratas tecnologías de captura. Pero esto no es del todo cierto. La producción de datos espaciales en la actualidad requiere más compromisos que los que eran asumidos anteriormente. Los modelos de datos que se están proponiendo y utilizando son cada vez más sofisticados, por ejemplo, hoy en día es usual hablar de: ciclo de vida de los fenómenos del mundo real, ciclos de vida de los datos, ciclos de vida de los productos, dato único, identificadores únicos, sellos temporales, datos enlazados, sindicación de datos espaciales, microdatos, reglas de implementación etc.

La captura es mucho más sencilla que antes, pero ahora el valor añadido es también mucho mayor que antes. Se trata de una producción mucho más tecnificada, compleja y dependiente del conocimiento, que ha de ser escalable y también capaz de evolucionar más ágilmente que nunca para estar al nivel de exigencia de las tecnologías TIC. Realmente no se producen datos, se producen modelos digitales que ofrecen distintas perspectivas virtuales de la realidad.

La producción supone, ahora más que nunca, un compromiso mantenido a lo largo del tiempo, y una visión sistémica de todo el proceso de producción. Desde esta perspectiva de la producción descentralizada de los datos, los principales riesgos que vislumbramos se relacionan con el fraccionamiento y debilidad de los productores nóveles. Entre otros los problemas principales pueden ser:

- ✓ Relativos a la producción continuada:
 - Falta de perspectiva temporal. Básicamente ocurre en agentes nóveles, que consideran la creación de datos desde una perspectiva episódica y aislada. La creación de datos espaciales debe suponer un compromiso a largo, pero ¿esto es asumible/asumido por todos los nuevos agentes?
 - Falta de viabilidad institucional. Hoy se producen datos espaciales por unos intereses concretos (p.ej. personales), porque es barato y está de moda, pero la institución/organización no asume ese compromiso de cara al futuro por lo que, al menor cambio, se deja de producir. Esta situación ya se ha vivido en nuestro país.
- ✓ Relativos a los aspectos tecnológicos
 - Exigencias tecnológicas crecientes. La creación/captura puede ser cada vez más sencilla y económica pero el mantenimiento, preservación, inclusión de ciclos de vida, creación de unos buenos metadatos, evolución tecnológica sostenible de los datos, etc., suponen cada vez más exigencias añadidas a la producción. Sólo unos pocos con financiación y personal altamente cualificado pueden acometerlo con éxito.
 - Acceso por servicios. En el paradigma de producción de la cartografía tradicional, e incluso de datos digitales sobre soportes físicos (p.ej. DVD, CDROM, etc.), el producto se materializaba en un recurso físico accesible aunque el productor hubiera dejado de existir. En el paradigma actual los datos se ofrecen por medio de servicios, si es el productor quien los ofrece y ese productor deja de existir la cuestión es ¿se seguirán ofreciendo esos servicios?
- ✓ Relativos a la fragmentación de la producción.

- Idiosincrasia. Si bien las organizaciones tradicionales productoras de datos espaciales son bien distintas y presentan un comportamiento distinto debido a sus experiencias vitales, personal, know-how, etc., con la llegada de nuevas formas de producción, como es la IG de voluntarios, aparecen patrones muy distintos que hay que saber entender para utilizar bien sus productos. Así, si bien la información VGI se ha puesto de moda como forma de crear datos espaciales (p.ej. OSM), es de destacar que los datos creados por estos movimientos tienen notables limitaciones (p.ej. falta de cubrimiento homogéneo de todo el espacio geográfico, grandes niveles de omisión, no existe una clasificación única de los elementos, etc.).
- Interoperabilidad. La fragmentación de la producción y aparición de nuevos productores supone graves riesgos para la interoperabilidad de los datos si no se refuerzan las acciones de coordinación. En un marco de producción descentralizada se requiere un mayor compromiso de todas las partes para que los datos sean interoperables y así conseguir la sinergia que posibilita el uso conjunto. El aspecto de interoperabilidad se tratará más adelante.

III.2. Acceso: no permitir accesos de calidad

En la actualidad los productos de datos espaciales, la IG, se ofrecen en forma de servicios vía web, por lo que también se ha de pensar en éstos como elementos críticos. En un paradigma de construcción descentralizada también se debe considerar un paradigma descentralizado en la oferta de servicios de acceso a la IG.

En este sentido, dado que en otro apartado se tratará sobre la interoperabilidad, y que en el epígrafe anterior se ha esbozado el mantenimiento de los servicios a lo largo del tiempo, en este epígrafe nos queremos centrar en la calidad del servicio. Muchos productores de IG, muchos de ellos con escasos recursos (p.ej. pequeñas municipalidades, ONG, privados, etc.), podrán llegar a generar IG de calidad e interés para la comunidad; pero la capacidad de ofrecer servicios con accesos de calidad (p.ej. disponibilidad, tiempo de respuesta, latencia, seguridad, fiabilidad, etc.), puede quedar fuera de su alcance. De esta manera, el riesgo consiste en no tener una disponibilidad efectiva del recurso de IG por limitaciones en la prestación del servicio debido a que no todos los productores de datos espaciales tengan capacidad económica ni tecnológica para ofrecer servicios de calidad.

III.3. Metadatos: no describir bien los productos

Hasta hace pocos años la cartografía era producida en forma de monopolio u oligopolio (instituciones cartográficas nacionales) y sólo era utilizada por técnicos, científicos y militares, los cuales tenían una formación que les capacitaba para su uso, y además disponían de información relevante sobre los productos. Sin embargo, junto a la revolución TIC ha existido una clara democratización en el uso y producción de datos espaciales. Esto ha generado que nuevos grupos sociales accedan al uso de estos productos sin la formación e información adecuada, en un entorno de mayor abundancia y volatilidad de la oferta de productos de datos. En esta situación los metadatos son un elemento clave.

En un paradigma de producción descentralizada de IG, para que el sistema sea realmente eficaz en el uso de los recursos existentes, se requiere que toda la IG generada disponga de metadatos. Los metadatos han de permitir conocer la existencia de un conjunto de datos espaciales y evaluar, a priori, la posibilidad de usarlos (p.ej. extensión, calidad, licencias, etc.). Podemos afirmar que lo que no esté metadatado no existe, pues no se podrá encontrar adecuadamente en la red, y que lo que no esté bien metadatado es un riesgo, pues puede inducir usos inadecuados. Bajo nuestro punto de vista se presentan dos riesgos fundamentales relativos a los metadatos:

- ✓ No existencia. Se genera un problema de localización del recurso. Esto puede suponer incurrir en nuevos gastos para generar ese mismo recurso, o bien, no poder realizar ciertos análisis basados en él, o realizar esos análisis de manera incompleta por la falta de esos datos.
- ✓ Mala calidad. Los metadatos son unos elementos incómodos de crear, gestionar y mantener. La norma de metadatos para la IG (ISO 19115-1, -2) es una especificación mostrenca y compleja. Estos metadatos que resultan presentan notables problemas: no son intuitivos, no tienen en cuenta las comunidades de interés, son difíciles de mantener, mucha información que podría ser relevante es opcional, usualmente se adopta una perspectiva minimalista para rellenarlos, están desfasados, etc.

III.4. Interoperabilidad: no posibilitar la utilización y uso legal

Bajo nuestro punto de vista, los mayores riesgos de la construcción descentralizada son los relativos a la interoperabilidad.

La interoperabilidad es la base de las TIC y por ello es la piedra angular de las IDE y de la Geomática 3.0 tal como se ha definido anteriormente. Según RODRÍGUEZ Y SEVILLA (2013), la interoperabilidad es la idea clave hacia la que se orientan todas las actividades de normalización y estandarización en el

campo de la IG. ISO la define formalmente (ISO, 2008) como: «capacidad para comunicar información, ejecutar programas o transferir datos entre unidades funcionales de manera que el usuario solo necesite tener un conocimiento escaso o nulo sobre las características particulares de tales unidades [ISO/IEC 2382-1:1993]»

En el campo de las IDE una definición de interés es la que proporciona la Directiva Inspire (PARLAMENTO Y CONSEJO EUROPEOS, 2007): «posibilidad de combinación de los conjuntos de datos espaciales y de la interacción de servicios, sin intervención manual repetitiva, de forma que el resultado sea coherente y se aumente el valor añadido de los conjuntos y servicios de datos».

Para conseguir la interoperabilidad se requieren acuerdos entre las partes. Estos acuerdos se materializan en forma de normas internacionales, especificaciones, estándares, reglamentos, etc. Por lo que la interoperabilidad está ligada íntimamente a las acciones de coordinación y liderazgo que resulten en documentos como los indicados anteriormente.

La interoperabilidad abarca distintos campos que van desde lo más tecnológico o duro «*hard*» (p.ej. formatos, sistemas de referencia, semánticas, etc.) a lo más suave «*soft*» (p.ej. aspectos institucionales, aspectos legales de licencias de uso, etc.) y no por ello menos costoso. La falta de interoperabilidad en los datos es posiblemente el riesgo más evidente y que más hemos sufrido en el trabajo con IG a lo largo de las últimas décadas. Ejemplo de esto son los problemas de intercambio de formatos, los problemas de integración de modelos de datos y semánticas distintas, los problemas de interoperabilidad debidos a discrepancias posicionales, etc. Siguiendo con lo estrictamente relativo a los datos, las técnicas de confluencia o también denominadas de integración o fusión de datos, son las encargadas de solventar estos problemas (p.ej. confluencia vertical, confluencia horizontal, (RUIZ-LENDÍNEZ y col. 2011)). Pero también hay otros aspectos no menos relevantes, como las limitaciones legales, que pueden surgir del uso conjunto de datos con modelos de licencia que no resulten compatibles. En el caso de la producción descentralizada llevada cabo por las administraciones, también se puede hablar de interoperabilidad institucional, la cual puede abarcar muchas facetas que impidan el uso conjunto efectivo de datos.

En principio, la descentralización se puede considerar un movimiento de incremento de la entropía y, como tal, que va en contra de alcanzar acuerdos técnicos, que son la base de la interoperabilidad. Un mayor número de productores de datos espaciales, muchos de ellos de carácter sectorial, otros amateurs, otros esporádicos, etc., tendrán tendencias naturales a crear «desorden». Por ello, la construcción descentralizada de datos espaciales supone notables

riesgos de interoperabilidad si no existe una labor eficaz de liderazgo para alcanzar consensos que la favorezcan. Sin embargo, en contraposición a todo lo indicado previamente, ejemplos como Inspire nos demuestran que un liderazgo fuerte permite aprovechar la variedad y diversidad de los sistemas descentralizados para aprovechar lo mejor de cada uno creando un proyecto conjunto. Una de las claves del éxito en el liderazgo de un sistema descentralizado es facilitar especificaciones, y herramientas que resulten cómodas para que nadie se salga de la corriente principal.

III.5. Trazabilidad y metacalidad: no mantener el linaje e información sobre la calidad de la calidad

Podemos definir la trazabilidad como la capacidad de conocer el historial, ubicación y trayectoria de un producto o lote de productos a lo largo de la cadena de suministros en un momento dado, a través de unas herramientas determinadas. La trazabilidad de los productos es algo que está asumido en muchos campos de la industria (p.ej. automovilística, espacial, alimentaria, electrónica, etc.), y que también ha de llegar al campo de la IG. La trazabilidad de datos ya se ha introducido en otros campos como el *master data management*. Por ello cabe esperar que sea una exigencia más de cara al futuro en el caso de la IG. En este caso, la creación descentralizada supone un reto tecnológico importante en la gestión de la trazabilidad.

La falta de trazabilidad genera numerosos riesgos en los productos. El no conocer los suministradores, procesos, operarios, métodos, etc., deja al productor sin capacidad de pedir responsabilidades y sin capacidad de mejorar de cara al futuro. En la industria (p.ej. de componentes electrónicos) la falta de trazabilidad está estrechamente ligada al denominado mercado informal o paralelo (*grey market*), donde es usual que los productos y procesos no cumplan realmente con las especificaciones del producto de marca. Por ello, las grandes multinacionales no dejan de controlar el aspecto de la trazabilidad de sus productos y suministros, por lo que, por analogía, una producción de IG de calidad internacional debe cuidar este aspecto.

La metacalidad es un nuevo elemento que aparece en la normativa internacional (ISO 19157). La calidad de la IG es una preocupación de los productores de datos espaciales, pero una vez que la calidad relativa a los productos se va asentando surge la necesidad de conocer y ofrecer mayor información, y más transparente, sobre la calidad de los datos relativos a la calidad. Esta necesidad ya fue apuntada de manera pionera por la prenorma Europea 12656 (CEN, 1997), y más recientemente se ha incorporado en la norma ISO 19157 donde se incluye en forma de: confianza, representatividad y homogeneidad. Sin embargo, la metacalidad es algo que existe fuera del ámbito geomático

desde hace mucho tiempo. Uno de los ámbitos donde la idea de metacalidad está presente es, a modo de ejemplo, el campo metrológico, donde existe una jerarquía de laboratorios que van asegurando encadenadamente la calidad de los instrumentos de medición. Incluso en ISO 9001, las exigencias sobre los equipos de medición, inspección y ensayo suponen asegurar la metacalidad instrumental, y de las medidas realizadas con ellos.

La evaluación de la metacalidad conlleva mayores exigencias que las de la calidad, y por ello hacen falta avances conceptuales, metodológicos y herramientas. Una producción muy descentralizada obliga a un mayor esfuerzo en la caracterización de los productores debido a su mayor número y diversidad y, como consecuencia, un muchísimo mayor esfuerzo en lo relativo a la metacalidad.

Por otra parte, la falta de información relativa a la metacalidad es un riesgo en sí misma. La caracterización de unos parámetros de calidad de un CDE sin información sobre su metacalidad puede ocasionar dos situaciones nefastas. Por un lado, si el producto es realmente bueno pero existe una subestimación de la calidad del CDE, al no disponer de información sobre la calidad de las estimaciones de la calidad lo lógico será no utilizar ese recurso, con las consecuencias de despreciar una IG de valor en una toma de decisiones. Por otro lado, si el producto es realmente malo y existe una sobreestimación de la calidad del CDE, al no disponer de información sobre la calidad de las estimaciones de la calidad lo lógico será utilizar ese recurso, con las consecuencias de incorporarlo en una toma de decisiones.

III.6. Preservación: no asegurar el uso futuro

La debilidad del almacenamiento digital ha llevado a acuñar el término *digital dark age* como aquella posible situación futura en la que sería imposible o muy difícil acceder a documentos del pasado por estar almacenados en soportes o formatos digitales obsoletos. A nivel mundial existe verdadera preocupación por este asunto y se están tomando iniciativas vinculadas al desarrollo de las denominadas e-infraestructuras. En esta línea existen iniciativas generales lideradas por la *Library of the Congress* (EEUU), *National Archives and Records Administration* (EEUU), *UK Data Archive*, etc.

La IG es un tipo de contenido digital de gran valor económico y que interesa conservar por motivos legales, administrativos, científicos, culturales, sociales y de toda índole. Sin embargo, desde la perspectiva de la IG, las iniciativas son menos numerosas, sin que exista todavía ninguna norma nacional o internacional al respecto. No obstante, el nivel científico y técnico alcanzado en diversos proyectos directamente relacionados con los datos espaciales (p.ej.

National Digital Information Infrastructure And Preservation Program - EEUU-, *National Geospatial Digital Archive* -EEUU-, *Long Term Data Preservation* -*European Space Agency*-, Suiza -BOS y Col., 2010-) permite afrontar esta tarea con seguridad y confianza debido a la experiencia acumulada. La situación actual queda descrita en detalle en ARIZA Y COL. (2012).

La calidad está relacionada con el uso adecuado, satisfactorio y fructífero. Sin embargo, un factor limitativo del uso y beneficio futuro del uso de la IG es su preservación o curación. Para ello se requiere un conjunto de metadatos especiales (distintos de los propuestos en ISO 19115 o en *Dublin Core*) que son los que deben permitir la caracterización de los contenidos digitales y, a partir de ellos, permitir la gestión de los procesos de preservación. Junto a los metadatos para la preservación también se requiere una organización y gestión adecuada del sistema que, en parte, puede ser parecido al paradigma IDE. Este marco debe cubrir adecuadamente:

- ✓ La necesidad de incluir la perspectiva de preservación y explotación futura de los datos desde el diseño de los productos de datos espaciales mediante el diseño de un ciclo de vida de producto.
- ✓ Considerar la necesidad de establecer distintos alcances temporales en la preservación, como por ejemplo el corto (5 años), medio (20 años) y largo plazo (100), según la importancia del dato.
- ✓ Establecer un conjunto de metadatos especiales pensados para la preservación e integrables con las normas ya existentes (p.ej. ISO 19115, en todas sus partes).
- ✓ Establecer las bases para la prueba de los sistemas de preservación que puedan implantarse.
- ✓ Establecer las bases para una infraestructura federada de preservación a nivel nacional.
- ✓ Evaluar la calidad (eficacia y eficiencia) de los sistemas de preservación.

La producción descentralizada ocasiona productores y producciones más dinámicas y efímeras y heterogeneidad en la preservación. Estos dos aspectos dificultan la preservación futura de la IG y suponen un claro riesgo de ocultación o desaparición de conjuntos de datos.

IV. CONCLUSIONES

En cierta medida, los riesgos para la calidad son los mismos con los que nos hemos venido manejando hasta ahora, pero con diferencias, antes, la producción estaba en manos de organismos oficiales con personal técnico altamente

formado, pero ahora, la democratización de las técnicas de captura y de procesamiento ponen en mano de numerosos inexpertos la posibilidad de crear nuevos conjuntos de datos espaciales. La construcción descentralizada es un nuevo paradigma donde queremos que los datos procedentes de muchas fuentes diversas, y sin importar su origen, se ensamblen de manera sencilla, eficiente y eficaz para proporcionar nuevos productos o servicios de datos espaciales. De esta manera surgen riesgos en las fases de producción, acceso, descripción, evaluación y mantenimiento a largo plazo de los conjuntos de datos espaciales. Consideramos que para limitar estos riesgos se debe:

- ✓ Incrementar las actividades de normalización de la IG y su difusión a todos los usuarios de IG.
- ✓ Incrementar la formación de aquellos que desean participar en la producción.
- ✓ Concienciar de las des-economías que se generan por una producción no interoperable.
- ✓ Concienciar y divulgar guías de buenas prácticas y ejemplos.
- ✓ Informar sobre los costes de la mala calidad que puede generar una IG poco adecuada.
- ✓ Potenciar la presencia de líderes que coordinen de manera efectiva la producción.
- ✓ Potenciar la aparición de certificaciones de productores y productos.

AGRADECIMIENTOS

Agradecemos a la Junta de Andalucía el apoyo económico que por medio del PAIDI ha venido prestando al Grupo de Investigación en Ingeniería Cartográfica (*TEP-164*) desde su creación.

BIBLIOGRAFÍA

- ARIZA-LÓPEZ, F.J.; ARIZA-LÓPEZ, R.M.; UREÑA CÁMARA, M.A.; JOAQUÍN CORTÉS, J. y UREÑA LÓPEZ, L.A. (2012): «Preservación de la Información Geográfica: Perspectivas y situación en España», *GeoFocus (Artículos)*, nº 12, 171-200.
- BERNERS-LEE, T.; HENDLER, J. y LASSILA, O. (2001): «The Semantic Web». *Scientific American*. Mayo; 29-37
- BOS, M.; GOLLIN, H.; GERBER, U.; LEUTHOLD, J. y MEYER, U. (2010): *Preliminary study on the archiving of geodata*. Berna: Federal Department of Defence, Civil Protection and Sport, Federal Office of Topography Swisstopo, Federal Department of Home Affairs, Swiss Federal Archives.
- CEN (1997): ENV 12656: *European Prestandard (ENV) 12656. Geographic information - Data description - Quality del Comité Européen de Normalisation (CEN)*.
- COMBER, A.J., FISHER y P.F., WADSWORTH, R.A. (2010): «User-focused metadata for spatial data, geographical information and data quality assessments». *10th AGILE Conference*

- GOODCHILD, M.F. (2007): «Citizens as sensors: the world of volunteered geography». *GeoJournal* 69 (4): 211-221. DOI: 10.1007/s10708-007-9111-y.
- ISO (2008): ISO/TS *Geographic Information–Terminology*. Ginebra.
- ISO (2013): ISO 19157:2013. *Geographic information. Data quality*.
- OGC (2012): *OGC Sensor Observation Service (SOS) Standard Version 2.0 Adopted*
- OXERA (1999): *The economic contribution of Ordnance Survey GB*. Oxford Economic Research Associates Ltd. Oxford.
- OXERA (2013): *What is the economic impact of Geoservices?* Prepared for Google.
- PIRA (2000): *Commercial Exploitation of Europe's Public Sector Information*. Pira International, Ltd., University of East Anglia, and Knowledge. Luxembourg.
- RODRÍGUEZ, A.F.; SEVILLA, C.; ABAD, P.; LÓPEZ, E. y SÁNCHEZ A. (2013): *Panorama de la estandarización en el campo de la información geográfica: la importancia de los estándares abiertos*. CIGMA 2013, México.
http://cigma2013.mx/es/agenda/doc/Paper_Antonio_Federico_Rodriguez_Pascual.pdf
- RUIZ, J.J.; ARIZA, F.J.; UREÑA, M.A. y BLÁZQUEZ, E.B. (2011): «Digital map conflation: a review of the process and a proposal for classification», *International Journal of Geographical Information Science*, 25:9, 1439-1466, DOI:10.1080/13658816.2010.519707
- STEINIGER, S. y HUNTER, A.J.S. (2012): «The 2012 free and open source GIS software map – A guide to facilitate research, development, and adoption», *Computers, Environment and Urban Systems*, Volume 39, May 2013, Pages 136-150.
<http://dx.doi.org/10.1016/j.compenvurbsys.2012.10.003>.
- Wikipedia (2013): *Smart City*.
http://en.wikipedia.org/wiki/Smart_city