

[RCG] Revista Catalana de Geografia

Revista digital de geografia, cartografia i ciències de la Terra

[Inici](#) > [Articles](#) > Preservación de la i...

[Pàgina principal](#)

[Imprimir](#)

[Enviar per e-mail](#)

[Descàrrega en PDF](#)

[Comentaris \(0\)](#)

Autor/s: Ariza, F. J.
Universidad de Jaén
Títol: Preservación de la información geográfica
Temàtica: Cartografia històrica
Publicat a: Revista Catalana de Geografia
IV època / volum XVII / núm. 46 / octubre 2012
Font: V IBERCARTO. Santander, 4-5 octubre 2012
URL: <http://www.rcg.cat/articles.php?id=246>

PRESERVACIÓN DE LA INFORMACIÓN GEOGRÁFICA

Francisco Javier Ariza López
Universidad de Jaén

1. Introducció

La debilidad del almacenamiento digital ha llevado a acuñar el término *digital dark age*, situación futura en la que sería imposible o muy difícil acceder a documentos del pasado por estar almacenados en soportes o formatos digitales obsoletos. En esta situación tienen responsabilidad tanto las administraciones públicas (en todos sus niveles), como el sector privado. Se están creando numerosísimos contenidos digitales (CCDD) y, en concreto, datos geocientíficos (p.e. de observación de la Tierra), e información geográfica (IG), en un volumen sin precedentes. A nivel mundial existe preocupación por este asunto y se están tomando iniciativas como las denominadas e-infraestructuras.

La importancia de la IG y de la Geomática en la Sociedad de la Información ha hecho que se alcance valores de negocio elevados en todos los países avanzados. La IG es información con unas características especiales (voluminosa, dependiente de escala, borrosa, dinámica, encadena numerosos procesos y algoritmos) (APCIG, 2008) que la distinguen. Además, es información costosa de generar y de mantener actualizada y, a la vez, muy necesaria para la gestión que realizan las administraciones públicas (Pira, 2000).

Acceder a la IG actual, más reciente, es el mayor deseo de la mayoría de los usuarios, pero las decisiones económicas, medio ambientales y sociales muchas veces se deben basar en estudios temporales que den soporte a decisiones técnicas robustas. Según la ESA (2009), el interés y necesidad de acceder a datos históricos se ha incrementado fuertemente en los últimos 10 años, y se espera que esta tendencia se incremente en el futuro.

Aspectos que caracterizan la IG actual, a su entorno y forma de uso, y que suponen riesgos de cara al futuro son (Morris, 2005):

- Los productores de IG, y la mayoría de los usuarios, se centran sobre los datos actuales
- Los avances tecnológicos hacen que el soporte futuro de ciertos formatos quede cuestionado.
- La posibilidad de fallo (p.e. de los soportes de la información) es un riesgo.
- La preservación requiere de metadatos específicos, aspecto que no suele ser tratado por los productores.
- Existe una evolución de productos, desde el estilo clásico, hacia servicios o suministros de datos en flujo, lo que diluye aún más las posibilidades de control y reconocimiento de los conjuntos de datos geográficos (CDG) originales.

El problema ha sido apuntado de manera general, y se han desarrollado esfuerzos para preservar CCDD de carácter general, pero la IG todavía no ha recibido suficiente atención. Esto ha llevado a algunos países a plantearse la necesidad de programas y acciones para preservar CCDD. Ejemplo es el *National Digital Information Infrastructure And Preservation Program* (NDI&PP) de EEUU (Friedlander, 2002) y, de una forma más específica, los programas dedicados a la IG, como es el caso del *National Geospatial Digital Archive* (NGDA) (Sweetkind-Singer y col., 2006).

El principal objetivo de este trabajo es presentar y explicar el problema de la preservación de la IG y conocer su estado actual en España por medio de una encuesta a bibliotecas y archivos.

2. La preservación

Hasta los años 80 del siglo XX los archivos y bibliotecas habían gestionado mayoritariamente documentos analógicos. Con la aparición de los CCDD fueron sumando a sus fondos soportes digitales (p.e. disquetes, CDROMs, etc.). Sin embargo la gestión de esta información, a largo plazo, lo que denominamos preservación, es mucho más compleja que para el caso analógico. La diversidad de medios de almacenamiento, de formatos y versiones, el tamaño de los ficheros, los aspectos de propiedad, la vinculación de los CCDD con software específico, etc., hacen que la preservación sea especialmente costosa y complicada. La situación, descrita de forma general, es aún más crítica en el caso de la IG. Aunque se puede afirmar de manera general que los archivos y bibliotecas están ya acostumbrados a la gestión de CCDD, también se puede indicar que la preservación a largo plazo es un aspecto poco considerado hasta la fecha, y aún menos en el caso de la IG digital. Así el uso y beneficio futuro de estos CCDD queda comprometido a largo plazo, a menos que se creen e implementen políticas y procedimientos para asegurar su disponibilidad (Brown y col. 2005). Además, la preservación puede y debe ser una actividad ligada a la demanda y creación de valor a lo largo del tiempo. Así, hasta ahora se ha hablado de preservación, pero un paso más es el de curación. La curación de los CCDD es la actividad de mantener y añadir valor a los CCDD para su uso actual y futuro (Janeé, 2009). La curación cubre el ciclo de vida de la información de manera completa (creación, uso, archivo, reutilización).

Un concepto fundamental para entender la preservación digital es el largo plazo. En la línea de ISO/TR 18492 el largo plazo se entiende cuando el periodo de conservación excede a expectativa de vida de la tecnología (hardware y software) utilizada para crear y mantener la información: un tiempo suficientemente amplio para preocuparse por el impacto de los cambios tecnológicos, media, formatos o cambios en la comunidad de usuarios (ISO/PDTR 26102). Se pretende la accesibilidad en el largo plazo, entendida ésta como la inteligibilidad de la información, la confidencialidad en el acceso (si procede), la autenticidad de los contenidos y la integridad de la información.

La información digital es difícil de preservar a largo plazo por: deterioro de los medios de almacenamiento, obsolescencia de los medios de almacenamiento, obsolescencia de los soportes lógicos (software), obsolescencia de los soportes físicos necesarios para ejecutar el software, fallos en la documentación de los formatos (p.e. fallos en metadatos), problemas en la gestión a largo plazo de los datos (Brown y col. 2005).

Las soluciones a la obsolescencia pueden orientarse según varias líneas generales (Arms, 2000), las cuáles admiten cierto grado de mezcla. La Tabla 1 presenta algunas opciones y sus ventajas/desventajas. Se denomina arqueología digital a las acciones de recuperación de contenidos cuando no se han realizado acciones sistemáticas de preservación.

Técnica	Ventajas	Desventajas	Periodo de preservación
Preservación	La información se almacena en su formato nativo	Mantener la tecnología original es costoso y difícil. La información sólo es accesible por medio del hardware y software original.	Solución a corto plazo e imposible de implementar a largo plazo.
Emulación	Se preserva la funcionalidad original de la aplicación y su apariencia.	Requiere el acceso al código fuente original y se replica la funcionalidad en un entorno tecnológico nuevo.	El recrear la funcionalidad es costoso y difícil. Solución a corto y medio plazo.
Migración	No se necesita mantener el software original.	El riesgo de pérdida de información es pequeño. Existe riesgo de alterar los registros en la migración así como probabilidad de pérdida de integridad. La metodología es un	Solución a corto y medio plazo.

		proceso continuo en el tiempo y por ello costoso. Se requieren metadatos administrativos.	
Encapsulación	Se preservan tanto los registros como la información sobre los registros (metadatos). No se modifica el formato original.	Se requiere invertir un tiempo considerable para la encapsulación y desarrollo del sistema.	Solución a largo plazo.

Tabla 1. Posibles soluciones para gestionar la obsolescencia.

Fuente: Brown y col. (2005)

Las opciones indicadas no son del todo adecuadas si se adopta una perspectiva a largo plazo. Si se desea una preservación duradera (p.e. al menos cien años), es necesario plantear modelos de sistema, y que estos sistemas sean independientes de cada momento (tecnologías), lo que obliga a un alto grado de interoperabilidad entre una secuencia o cadena de sistemas de preservación. Con esta idea Janeé y Frew (2008) establece los siguientes criterios a adoptar en los sistemas de preservación:

- Principio de transmisión: El sistema de preservación debe soportar el traspaso de su contenido al próximo sistema de la cadena, es decir soportar su propia migración (no sólo la de los contenidos).
- Principio de vuelta atrás: El sistema de preservación debe soportar alguna forma de paso hacia atrás de su contenido en situaciones en la que el propio sistema no está funcional.
- Principio de resurrección: Supone revivir la información. Este principio se opone a la inmortalidad la cual consiste en mantener los contenidos vivos en todo momento (Janeé y Frew, 2008).

Autores como Waugh y col. (2000) añaden otras exigencias:

- La autosuficiencia: Minimización de las dependencias del sistema, datos o documentos respecto a otros elementos.
- La autodocumentación/encapsulado: Inclusión como parte del documento de la información para que pueda ser interpretado/comprendido por un usuario (aplicación o persona).

2.2. Preservación de la IG: motivos y problemas

En primer lugar hay que preguntarse si la IG debe o merece ser preservada. De esta forma se puede establecer un conjunto de razones específicas para preservar la IG:

- Mandato legal: Caso de aquellos organismos, productores o usuarios, que así lo tengan dispuesto en sus regulaciones.
- Maximizar beneficios: Aprovechando toda la capacidad de generación de riqueza de los datos a lo largo de su ciclo de vida.
- Apoyar la investigación: Básicamente el análisis de cambios y modelización como forma de aprender del pasado para controlar mejor el presente y futuro de nuestras sociedades y del entorno.
- Proporcionar evidencias históricas: Una perspectiva algo distinta a la indicada en el punto 3º y que permite el cartografiado de la historia y conocimiento del pasado.
- Desastres y cooperación: Alto valor para la planificación y coordinación de acciones frente a desastres y en proyectos de cooperación.
- Costes de reproducción: Coste de producción elevado.

Sin embargo, el propio carácter de la IG la convierte en una información difícil de preservar (Morris, 2009). Janeé y Frew (2008):

- Modelos de datos no uniformes: Se puede hablar de modelos ráster y vectoriales, y dentro de este último caso de múltiples variantes.
- Formatos propietarios: Muchos formatos SIG (p.e. Shp, dxf, dwg, dgn, E00, etc.) son propietarios. El que sean controlados por firmas comerciales y no por normas internacionales o especificaciones abiertas los somete a las tensiones de un mercado muy competitivo.
- Tamaño de grano variable: Los datos espaciales se gestionan con una granularidad muy variable, desde objetos individuales a capas temáticas que cubren zonas como hojas de series

de mapas, divisiones administrativas, etc., los datos pueden agregarse, desagregarse con facilidad.

- Sistemas relacionales: Cada vez más el almacenamiento SIG se realiza sobre bases de datos relacionales, heredando sus desventajas en cuanto a la preservación (copias sobre sistemas en ejecución, formatos de copia propietarios, formatos de copia complejos, ficheros grandes y monolíticos).
- Gran tamaño de los ficheros: El tamaño de ciertos conjuntos de datos espaciales puede ser gigantesco (gigabytes y terabytes).
- Programas de larga duración: Algunos datos proceden de proyectos con una larga vida (p.e. Landsat, SPOT), lo que genera grandes volúmenes de almacenamiento, la existencia de cambios en el proyecto, la necesidad de planificación de la preservación desde el principio, etc.
- Contexto extensivo: Los datos espaciales requieren más información de contexto (metadatos) que otras tipologías de datos.
- Contexto implícito: En muchos casos el contexto de los datos está implícito y embebido en comunidades científico-técnicas aisladas y pequeñas.
- Datos dinámicos: Algunos datos requieren que con cierta periodicidad sean reprocesados para incorporar las correcciones y mejoras en los modelos que los originaron (p.e. comportamiento de los sensores). Ello obliga a incluir información relativa a esta circunstancia.

A las que, a partir de otros trabajos (Arms y Fleischhauer (2005), Morris (2009) y Morris y col. (2009)), entre otros, consideramos se pueden añadir:

- Metadatos: Muchas veces no existen o son insuficientes y, en todo caso, cuando existen no están pensados para facilitar la preservación. Es crítica la falta de perspectiva de ciclo de vida del producto en los metadatos (p.e. según ISO 19115). Otro problema es la falta de sincronía entre los metadatos y los datos.
- Versionado: La IG vectorial con requisitos temporales (p.e. catastro, construcciones) hacen que el problema de la preservación adquiera una dimensión mayor.
- Identificadores únicos: Los sistemas de identificadores únicos no están totalmente extendidos, ni aún menos los de versiones y sellos temporales.
- Complejidad de los formatos: Muchos formatos de IG se conforman como un conjunto de ficheros interrelacionados, con formatos, estructuras y relaciones diversas (p.e. SHP = SHP + DBF + SHX + PRJ).
- Paquetes contenedores: No existen paquetes contenedores de IG difundidos y asumidos de manera general, tales que permitan manejar de manera conjunta conjuntos complejos de datos de IG.

Para finalizar este subapartado se pueden incluir las conclusiones del primer seminario organizado por la asociación EUROS DR sobre esta temática (EUROS DR, 2012):

- La preservación comienza durante la creación del dato, y afecta durante todo su ciclo de vida, no es sólo un asunto a tener en cuenta al final.
- Es mejor ser muy selectivo en el conjunto de datos a preservar.
- Se deben especificar tres tipos de duraciones, de 1 año, con carácter operacional, de 10 años para acceso público y de 100 años para el mantenimiento de la memoria.
- El modelo OAIS es un punto de partida, aunque resulta demasiado general y debe ser especificado.

3. Los metadatos y la preservación

La preservación no es simplemente mantener una copia de seguridad, supone disponer de unos metadatos y capacidades adecuadas para poder acceder a la información en el futuro, para preservarla frente a cambios de formatos, al daño de sus soportes, accesos indeseados, pérdidas fortuitas, etc., asegurando sus propiedades. Todo esto cuesta muchos recursos y por ello debe ser selectiva. De esta forma se puede afirmar que sin metadatos la preservación no es posible. Los metadatos de preservación deben ser suficientes para permitir cambios en los registros a través de generaciones de hardware y software, y permitir la trazabilidad y auditoría a lo largo del ciclo de vida de un registro. Los metadatos de preservación son críticos para conseguir la preservación a largo plazo.

Según McGarva y col. (2009) aspectos importantes que deben cubrir los metadatos de preservación son, entre otros:

- El archivo de información relativa a los derechos.
- Los metadatos administrativos relativos a la forma de adquisición.
- Metadatos técnicos relativos a los métodos de transferencia, incluyendo la garantía de la integridad de los datos.
- Metadatos relativos a cualquier transformación de los datos realizada en los procesos de archivo.
- Los resultados de cualquier evaluación de la validez de los datos o de cualquier evaluación del riesgo asociada a los datos.

Si bien existen algunas iniciativas pioneras de metadatos de preservación en el ámbito de los registros electrónicos de datos espaciales (CIESIN, 2005), en el ámbito más general de los CCDD también existen iniciativas más consolidadas. El proyecto PREMIS (*Preservation Metadata: Implementation Strategies*) (BNE, 2010) ha sido una de las principales iniciativas en preservación digital. PREMIS define un modelo de datos relativamente sencillo basado en cinco entidades: objetos, acontecimientos, aspectos intelectuales, derechos y agentes (ver Caplan 2009 para una visión general del sistema).

En el ámbito de la IG existe una gran sensibilidad, reciente, hacia los metadatos. Este reconocimiento viene potenciado por el desarrollo de las IDE (Infraestructuras de Datos Espaciales) y la difusión de las normas ISO 19100, en concreto de ISO 19115. Por ello, dada la amplia difusión de los metadatos en el campo de la IG, en primer lugar se hace necesaria una revisión crítica de las normas de metadatos desde la perspectiva de la preservación. En esta línea ya se han realizado algunos trabajos.

Hoebelheinrich y Banning (2008) analizan la norma CSDGM (*Content Standard for Digital Geospatial Metadata*) del FGDC y vienen a apuntar dos conclusiones las cuales consideramos también son válidas para la norma ISO 19115. En primer lugar, la riqueza del CSDGM puede ser considerada como una debilidad dado que el elevado número de elementos que incluye pueden resultar abrumador y confuso. La complejidad del estándar y la resistencia general a la creación de metadatos se combinan para resultar en una tendencia de que los registros de metadatos contengan sólo el mínimo de elementos rellenos. Esto puede ser suficiente para la descripción y localización, pero no para la preservación a largo plazo. Desde la perspectiva de la preservación, la segunda debilidad del CSDGM es el énfasis en registrar la situación del recurso en el momento de su creación con muy pocas oportunidades de describir eventos importantes en el ciclo de vida del recurso cuando se gestiona para su preservación a lo largo del tiempo. Por su parte, en GeoMAPP (2009a,b) se presenta un análisis del CSDGM frente al Dublin Core y cuya conclusión final va en la línea de proponer el uso combinado de los dos estándares para aprovechar las capacidades y soslayar las deficiencias. Igualmente, se llegan a identificar (GeoMAPP, 2011) los campos de mayor interés de la norma CSDGM respecto a la preservación.

Si se consideran los metadatos como informadores acerca de la preservación de elementos de IG ni el CSDGM, ni ISO 19115, ni Dublin Core disponen de elementos adecuados para la preservación de la IG (Shaon y Woolf, 2011). De hecho, ninguno de los perfiles desarrollados (p.e. el Núcleo Español de Metadatos, NEM) trata este aspecto. Para que la preservación de la IG sea efectiva se necesitan metadatos, y para que éstos puedan ser consultados, actualizados y usados de una manera interoperable es necesaria la inclusión de la preservación en las normas internacionales, o en los perfiles derivados. Shaon y Woolf (2011) proponen la creación de un perfil de metadatos que incluya la preservación, indicando los siguientes elementos:

- Información de representación: Incluiría una descripción de la forma de representación de las entidades cartografiadas y la descripción del contexto, con la idea de acceder a los objetos preservados en un contexto de significado pleno.
- Anotaciones adicionales: Se utilizan para ello textos libres, etiquetas, atributos cuantitativos, etc., que actúan como una especie de metadatos para dar mayor comprensión al elemento de información al que se refieren.
- Información del versionado de datos: Todos los datos poseen un ciclo de vida durante el que pueden sufrir cambios. Los datos deben poseer un identificador único y una referencia a su versión para estar perfectamente identificados.
- Procedencia: Como un aspecto adicional del linaje donde se indican los diferentes agentes desde un punto de vista amplio, como pueden ser: proveedor inicial o productor, propietarios, posteriores organismos de preservación, etc.

- Información de eventos: Se refiere a los eventos que genera cada proceso de preservación (p.e cambio de sistema de preservación, de plataforma informática, de tecnología de almacenamiento, etc.).
- Autenticación: Debe incluir dos elementos, uno de firma y otro de anclaje. Ambos tratan de asegurar la procedencia de la información.

Aunque esta propuesta está bastante elaborada es insuficiente y no trata todas las posibles soluciones. Se centran en el desarrollo de nuevas entidades sin incluir nuevos atributos para las ya existentes. En definitiva, se puede indicar que las normas de metadatos (CSDGM, ISO 19115, Dublin Core, NEM) carecen de la perspectiva de la preservación, y que ello conlleva que no se almacenen datos de preservación. Así, cabe pensar que en un futuro será obligada una revisión de todas las normas dando lugar a una nueva generación que incluya los aspectos de la preservación. Es más, al igual que ya se ha asumido que los metadatos del producto (p.e. tipo ISO 19115) deben ser diseñados y creados a la par que el producto, y no como algo posterior, la preservación y sus metadatos tienen que ser tenidos en cuenta en el momento de diseño y creación de la IG. Sin embargo, esta situación no se da ni en España ni en otros muchos países.

4. Iniciativas y experiencias de preservación de IG en Europa

Las iniciativas y experiencias en preservación de la IG son relativamente recientes y abarcan como mucho una década. La mayor parte de ellas han tenido lugar en EEUU y, como suele ser común, son las más documentadas y conocidas a nivel internacional debido a la política de transparencia y divulgación al uso en los EEUU. Gran parte de estas iniciativas se han ido presentando en los apartados anteriores. Sin embargo, en Europa también existen proyectos dignos de mención, entre ellos, por sus perspectivas bien distintas, el proyecto suizo desarrollado sobre su sistema federal de IG y el proyecto LTDP (*Long Term Data Preservation*), auspiciado por la ESA, para datos relativos a la observación de la Tierra.

Centrado específicamente en la IG general se destaca el proyecto desarrollado en Suiza que pretende una solución integral para la totalidad de la administración federal; y que se encuentra muy bien documentado en Bos y Col. (2010). Lo interesante de este proyecto no son sólo las conclusiones del mismo, sino la metodología seguida para su desarrollo. Se parte de no haber encontrado una solución asentada al problema, y por ello de la necesidad de desarrollar una solución específica para la realidad helvética. El proyecto está basado en la colaboración entre especialistas en IG y archivos, sobre un amplio análisis interno y externo, así como en la adopción del marco OAIS y en la realización de numerosas pruebas para ir ajustando y depurando la solución. Aún siendo un proyecto general, se llega a un gran nivel de detalle como es la determinación de formatos aceptables y rechazados para la preservación (p.e. formatos rechazados: JPEG2000, MrSID, USGS DEM, NITF, etc.) y aceptados provisionalmente; GeoTIFF, ML, SHP, INTERLIS 1 y 2).

Con una visión distinta la ESA ha venido liderando en los últimos años una iniciativa orientada a desarrollar una política y unas guías para la preservación a largo plazo que fuera válida para los productores, proveedores y archivadores de este tipo de datos. Se trata también de una iniciativa muy bien documentada (ESA, 2009, 2010). Una parte importante de la metodología fue la definición de las necesidades de los usuarios, que se recogieron por medio de una encuesta y entrevistas. Lo más importante de este proyecto es que llega a concretar un modelo que debe considerarse la base para el diseño, rediseño e implantación de iniciativas de preservación. Las directrices desarrolladas sobre los aspectos más importantes de la preservación se agrupan en las siguientes temáticas: composición de los conjuntos de datos a preservar, organización y operación de los archivos, seguridad de los archivos, ingestión de los datos, mantenimiento de los archivos, acceso a los datos e interoperabilidad, explotación y reprocesado de datos, evaluación de datos y prevención de purgas.

Un aspecto común a ambos proyectos, y que toda iniciativa futura debiera tener en cuenta, es la importancia de las actividades de cooperación necesarias (políticas, tecnología, metodologías y desarrollos, normalización, soluciones operacionales y explotación de datos). Así, podemos afirmar que estos proyectos deben tratarse como proyectos de e-infraestructuras, tal y como ocurre con las IDE.

5. Preservación de la IG en España

España ha sido pionera en el desarrollo de iniciativas de digitalización de fuentes cartográficas como, por ejemplo, la desarrollada por el Instituto de Cartografía de Andalucía durante los años 90

(Cortés-José, 1997), que ha permitido digitalizar de más de 132.000 mapas, planos y vistas (más de 165.00 imágenes), y también en la apertura del acceso de estos catálogos vía Internet (p.e. siguiendo con el ejemplo anterior, el Catálogo de Digital de Cartografía Histórica de Andalucía). Por otra parte, el proyecto HISPANA (<http://hispana.mcu.es/es/inicio/inicio.cmd>) liderado por el Ministerio de Educación, Cultura y Deporte ha sido la versión española del macroproyecto EUROPEANA de la Unión Europea, que promueve los archivos abiertos digitales de carácter general. Con una perspectiva centrada en la IG se pueden destacar otras iniciativas. Así, de manera paralela al desarrollo de las IDE se ha participado el proyecto europeo DIGMAP, y se ha desarrollado el proyecto español CartoVIRTUAL. Ambos son relativos a cartotecas virtuales y están centrados tanto en la incorporación de CCDD de IG, como en el desarrollo de herramientas informáticas (captura, georreferenciación, cambios de proyección, metadatos, etc.) compatibles con las IDE (Fernández y Bernabé, 2011). Sin embargo, ninguno de estos proyectos es específico de preservación; concepto y preocupación que no aparecen explícitamente en sus objetivos.

Dada la falta de referencias específicas sobre proyectos de preservación en España, y con el objetivo de conocer el estado de la cuestión en nuestro país, y con carácter exploratorio, se ha desarrollado una encuesta de la que se presentan sus planteamientos y resultados en este apartado. Frente a otras posibles opciones (p.e. entrevistas), se consideró que una encuesta breve podría ser la mejor manera de realizar un acercamiento exploratorio a la situación. Para ello, en primer lugar se planteó cual podría ser la población objetivo (aquí las denominaremos organizaciones). En relación a este aspecto se consideró que se debía cubrir: el archivo nacional, bibliotecas y cartotecas regionales y universitarias, y las pertenecientes a centros de producción cartográfica. Con este criterio y con la idea de cubrir, en la medida de lo posible, el territorio nacional, se compiló una selección o muestra de 26 organizaciones a encuestar: 13 bibliotecas universitarias, 4 bibliotecas regionales, 2 cartotecas, 5 cartotecas de centros de producción de IG y 2 archivos. Se debe advertir que debido al carácter exploratorio ya indicado, no se ha realizado un diseño estadístico en cuanto al número ni manera de selección, por lo que los resultados deben ser entendidos con estas limitaciones.

Dado el carácter de aproximación y de iniciación, la encuesta debía ser breve pero también abrir la posibilidad a aportaciones relevantes. Por ello se decidió incluir sólo 5 preguntas que se contestaban con un SI/NO, y dejar libre la opción de explicar en cada una las actividades de los centros consultados, así como de incluir cualquier comentario, documentos o enlaces relacionados. La primera pregunta (P1) era relativa a la existencia o no de cartoteca en la organización. Se pretendía averiguar el grado de especificidad en cuanto a la gestión de la IG. Las dos siguientes preguntas (P2 y P3) se referían a contenidos digitales convencionales, inquiriendo sobre la existencia de actividades de preservación y curación. Las dos últimas (P4 y P5) eran paralelas a las dos anteriores pero relativas a la IG.

Cada organización fue contactada telefónicamente en la persona de su director con el objetivo de presentar el proyecto y la encuesta, así como para solicitarle su participación. En la mayoría de los casos han sido los propios directores quienes han contestado las encuestas, en otros casos han derivado al personal que ellos han considerado más apropiado. En muchos casos ha sido necesaria cierta insistencia para obtener las contestaciones, obteniendo un resultado final de 20 encuestas completadas (77%). El resumen es el que se presenta en el Tabla 3.

Pregunta	Contestación	
	Sí	No
P1: Disponen de cartoteca	61.5	38.5
P2: Preservación de contenidos digitales generales	42.3	57.7
P3: Curación de contenidos digitales generales	7.7	92.3
P4: Preservación de información geográfica digital	19.2	80.8
P5: Curación de información geográfica digital	7.7	92.3

Tabla 3. Resultados de la encuesta relativa a las acciones de preservación y curación de contenidos digitales generales y relativos a la Información Geográfica

Nota: Encuesta realizada a 26 organizaciones, por contacto y explicación telefónica y contestación vía correo electrónico de Junio a Noviembre de 2011.

Fuente: Elaboración propia

En cuanto a los resultados de P1, en un 61.5% las organizaciones consultadas disponen de cartoteca, con independencia de cómo se organice ésta (servicio, sección, de manera independiente, etc.). En muchos de los casos relativos a las universidades, donde se ha indicado que no existe cartoteca, se matiza que son los departamentos (p.e Geografía), los que gestionan las fuentes cartográficas, ya sean analógicas o digitales, y que las bibliotecas centrales carecen de información y conocimiento relativo a las prácticas de preservación y curación que éstos pudieran realizar.

La P2 ha sido contestada afirmativamente en un 42.3%. El propósito de esta pregunta era obtener una referencia para comparar los niveles de preservación de la IG. A pesar de que los problemas de preservación de CCDD son bien conocidos por parte de los cuerpos de archiveros y bibliotecarios, sólo algo más de un tercio de las organizaciones afirman realizar actividades de este tipo. Es cierto que en varios casos se ha indicado que se estaba en vías de implantar sistemas para gestionar estas actividades, por lo que en un futuro cercano se ampliará el número

La P3 ha presentado algún problema en su entendimiento. En varios casos los entrevistados no conocían exactamente el sentido del término. Detectada esta situación se consideró conveniente introducir en las encuestas enlaces a documentos con la definición de ambos términos (preservación y curación). Esta situación nos introduce cierta duda sobre las primeras contestaciones recibidas para esta pregunta. En cualquier caso, el porcentaje de contestaciones afirmativas es muy reducido (7.7%) (dos casos) y, por supuesto, se da en una organización que ya desarrolla actividades de preservación. Este porcentaje tan bajo indica que, respecto a CCDD generales, las organizaciones realmente no han dado ese paso que significa la curación, manteniendo una actitud más pasiva que proactiva.

Las dos preguntas siguientes (P4 y P4) se formularon de manera paralela a las relativas a los CCDD de carácter general tratados en P2 y P3. Un 19.2% de los casos indican que se realiza preservación de IG (P3). Es decir, alrededor de la mitad de los casos que indican preservación de CCDD generales. Además, salvo un único caso (relativo a una cartoteca), el resto de organizaciones que indican preservación de IG realizan también preservación de CCDD de carácter general. El perfil del resto de organizaciones que contestan afirmativamente a la pregunta es consistente con la afirmación (cartoteca y centros regionales de primera línea). Respecto a la P4 sólo existen dos contestaciones afirmativas (7.7%) y se dan en una cartoteca, que destaca por la realización de actividades formativas y divulgativas sobre el uso de recursos de IG, y en una biblioteca regional.

Fuera de los resultados estrictos de la encuesta, la conversación mantenida en su presentación y los comentarios aportados por las contestaciones nos permiten indicar:

- Que hay centros con gran experiencia en la preservación de contenidos analógicos mediante sucesivas digitalizaciones (distintas resoluciones, en B&N y color, etc.) conforme han ido avanzando las posibilidades técnicas.
- Que en algunos casos la preservación se entiende de manera muy simplista como el mantenimiento de copias de seguridad periódicas que cumplen ciertas condiciones en su almacenamiento.
- Que en algunos casos los objetivos de la encuesta debieran ser directamente los departamentos universitarios generadores de la IG, los centros regionales de producción de IG (empresas públicas o instituciones cartográficas) o los centros de cálculo e informáticos de esas instituciones.
- Que en el ámbito universitario existe una desconexión total entre los departamentos que generan IG y las bibliotecas y archivos de estas instituciones.
- Que en muchos casos las contestaciones indican que no se adquiere IG digital.
- Que los formatos mencionados son imagen (TIFF y JPEG) y PDF.
- Que se indica que se siguen estándares internacionales para los repositorios, pero no se indican cuáles.
- Que en el ámbito de la curación se menciona la creación de indexaciones y el uso de la herramienta ABSYS.

6. Líneas de trabajo futuro en preservación de la IG

Si bien en los últimos años se ha avanzado mucho en preservación de CCDD y en preservación de la IG, son todavía muchos los frentes abiertos en los que se puede trabajar de manera aplicada e investigar en este campo y, de una manera especial, en nuestro país por la poca o nula experiencia detectada. Morris (2009) indica las siguientes líneas:

- En colaboración con los cuerpos normativos (ISO, OGC, etc.), desarrollo de normas y estándares en temas como empaquetado de contenidos y archivo de perfiles.
- Desarrollo de mecanismos de transferencia y gestión distribuida en colaboración con las comunidades de computación distribuida.
- Mejorar de los métodos para incorporar datos temporales e históricos en los entornos de localización y acceso.
- Explorar todas las posibles formas de manejar y retener contenido temporal y de contexto dentro de los entornos de soporte a la decisión basados en servicios web.
- Investigación sobre GeoPDF para la preservación a largo plazo, especialmente para documentos complejos y representaciones.
- Exploración de los aspectos de preservación relacionados con los recursos de información espacial emergentes y basados en la localización.
- Exploración de las tecnologías bibliotecarias establecidas y emergentes como METS, PREMIS, OAI-ORE en conexión con la preservación de datos espaciales.

Otra cuestión importante es la preservación de la IG en los servicios en la red (Lazorchak y Sewash, 2011). Morris (2006) indica que el desplazamiento hacia servicios web de distribución de la IG puede ser una amenaza para la conservación a largo plazo de los contenidos. Sin embargo, esos mismos servicios también podrán ser utilizados para desarrollar los propios archivos.

7. Conclusiones

La conservación de contenidos analógicos es una tarea rutinaria en archivos y bibliotecas desde hace siglos. Sin embargo, la preservación de CCDD es una necesidad imperiosa de nuestra sociedad y todavía presenta problemas. Los principios básicos de la preservación digital ya existen y son aplicados en numerosos archivos y bibliotecas de carácter nacional (Librería del Congreso (EEUU), Biblioteca Nacional (ES), etc.). Los metadatos son fundamentales para la preservación a largo plazo, los metadatos deben estar pensados para esta actividad.

La IG es una información valiosa económicamente y de interés para el futuro, por lo que también es una necesidad preservarla. Sin embargo, la IG tiene una problemática especial por lo que su preservación se hace aún más complicada, costosa y menos extendida.

En el ámbito de la IG el paradigma es productivista, el foco se centra en la producción y en la actualidad de los datos. Muy pocos profesionales han dedicado tiempo a pensar sobre el archivo de la IG a largo plazo. Un aspecto crítico es que la producción de IG muchas veces está desligada de un verdadero diseño del ciclo de vida del producto de IG. Sería muy conveniente que todos los productores de IG hicieran una fuerte reflexión sobre el ciclo de vida de sus productos y que esta perspectiva se incluyera en los grados universitarios relacionados con la producción cartográfica.

En EEUU y Europa existen varios proyectos de preservación de IG a escala estatal y nacional, algunos de ellos se han presentado en este trabajo. Las conclusiones de los mismos indican que: los metadatos específicos para la IG no están pensados para acciones de preservación. Dada la falta de referencias sobre proyectos de preservación de la IG en España, se ha desarrollado una encuesta exploratoria. Se ha considerado una muestra de 26 organizaciones (bibliotecas, cartotecas y archivos) de la que se han recibido 20 contestaciones (77%). Las actividades de preservación y curación de CCDD generales indican que más del 40% de los casos realizan preservación y que sólo el 7.7% curación. En el caso de contenidos de IG sólo el 19% realizan preservación, alcanzando la curación un 7.7% de los casos. Es decir, las actividades relacionadas con la preservación de la IG están en el orden de la mitad de los casos de las correspondientes a los CCDD generales.

Existe un amplio campo para la investigación en preservación de la IG que, además, debería desarrollarse conjuntamente con las IDE. La experiencia de las IDE, donde los aspectos de coordinación entre organizaciones y de interoperabilidad son muy importantes, puede ser muy valiosa para los proyectos de preservación de IG a nivel nacional, donde la planificación y estrategia con vistas al futuro son aspectos clave.

Finalmente, no quisiéramos acabar sin hacer una llamada a la necesidad de que en el Sistema Cartográfico Nacional se tome en consideración la preservación de la IG. La simple revisión de la

legislación cartográfica más reciente (Ley 14/2010, de 5 de julio, RD 1545/2007 de 23 de noviembre, RD 1637/2009 de 30 de octubre, RD 663/2007 de 25 de mayo), y de las legislaciones y planes cartográficos autonómicos (p.e. para Andalucía el Decreto 141/2006, Plan Cartográfico de Andalucía 2009-2012) permite afirmar que se trata de un aspecto totalmente olvidado hasta la fecha. Por todo ello, en la actualidad, son fundamentales las acciones de divulgación, sensibilización y apoyo relativas a la preservación de la IG,

8. Agradecimientos

Este trabajo ha sido cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER), por el proyecto P08-TIC-4199 de Excelencia de la Junta de Andalucía, y por las ayudas que recibe el Grupo de Investigación en Ingeniería Cartográfica (PAI-TEP-164). Además, deseamos expresar nuestro agradecimiento a todas las personas e instituciones que amablemente nos han contestado a la encuesta.

9. Referencias

- APCIG (2008): Sistemas y Tecnologías de la Información. Madrid. Asociación Profesional del Cuerpo de Ingenieros Geógrafos, Instituto Geográfico Nacional.
- Arms, W. (2000): Digital Libraries. Cambridge, Massachusetts, USA: MIT Press.
- Arms, C.R.; Fleischauer, C. (2005). Digital Formats: Factors for Sustainability, Functionality, and Quality. Washington, DC, USA. Office of Strategic Initiatives, Library of Congress.
- BNE (2010): Diccionario de Datos PREMIS de Metadatos de Presección Versión 2.0. Madrid, España: Biblioteca Nacional de España, 2010.
- Bos, M., Gollin, H., Gerber, U., Leuthold, J., 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.
- Brown, D., Welch, G., Cullingworth, C. (2005): Archiving, Management and Preservation of Geospatial Data. Summary Report and Recommendations. GeoConnections, Working Group on Archiving and Preserving Geospatial Data. [En Línea]. . [21 de diciembre de 2011].
- Caplan, P. (2009): Entender PREMIS. Washington, DC, USA: Library of the Congress.
- CIESIN. (2005): Data Model for Managing and Preserving Geospatial Electronic Records Version 1.00. Palisades, NY, USA: Center for International Earth Science Information Network, Columbia University.
- Cortés-José, J. (1997): "Cartografía Histórica de Andalucía. Asociación de Imágenes a un gestor Documental. Catálogos de Cartografía Histórica Multimedia". Vitoria Gasteiz. Ed. Congreso Internacional sobre Sistemas de Información Histórica. Juntas Generales de Álava.
- ESA. (2009): European Framework for the Long Term Preservation of Earth Observation Space Data. Noordwijk, The Netherlands: European Space Agency.
- ESA. (2010): Long Term Preservation of Earth Observation Space Data. European LTDP Common Guidelines. Noordwijk, The Netherlands: European Space Agency.
- EUROSDR (2012): GI+100: Long term preservation of digital Geographic Information - 16 fundamental principles. Euro Spatial Data Research. [En Línea]. . [21 de diciembre de 2011]
- Fernández-Wytenbach, A., Bernabé-Poveda, M. (2011): "La difusión del patrimonio: los servicios de mapas". Revista Patrimonio Histórico. Instituto Andaluz del Patrimonio Histórico, nº 77, pp. 132-138.
- Friedlander, A. (2002): The National Digital Infrastructure Preservation Program. Expectations, Realities, Choices and Progress to Date. D-Lib Magazine. [En línea], vol. 8, nº 4. . [21 de diciembre de 2011]. ISSN 1082-9873.
- GeoMAPP. (2009a): FGDC and Dublin Core Metadata Comparison. Geospatial Multistate Archive and Preservation Partnership. Raleigh, NC, USA: NC Center for Geographic Information and Analysis.
- GeoMAPP (2009b): Metadata Comparison. Raleigh, NC, USA: NC Center for Geographic Information and Analysis.
- GeoMAPP. (2011): Utilizing Geospatial Metadata to Support Data Preservation Practices. Geospatial Multistate Archive and Preservation Partnership. Raleigh, NC, USA: NC Center for Geographic Information and Analysis.
- Hoebelheinrich, N.; Banning, J. (2008): An investigation into metadata for long-lived geospatial data formats. Washington, DC. Library of Congress.
- Janeé, G. (2009): Digital Curation. In Liu, M. y Özsu, M.T. (ed). Encyclopedia of Database Systems. Springer, p. 816-817.

- Janeé, G.; Frew, J. (2008): Relay-supporting Archives: Requirements and Progress. International Journal of Digital Curation, vol. 4, nº 1, pp. 57-70.
- Lazorchak, B.; Sewash, J. (2011): Dust Free Data: Developing Dynamic Geospatial Data Archives. NSGIC Midyear Meeting. Annapolis, MD, USA: National States Geographic Information Council.
- LC (2009): Strategy for Framing a National Preservation and Access Strategy for Geospatial Data. Meeting Minutes. Washington, DC. Library of Congress.
- McGarva G., Morris S, Janeé G. (2009): Preserving Geospatial Data. Technology Watch Series report 09-01. Heslington, York, USA. Digital Preservation Coalition.
- Morris, S. (2005). Collection and Preservation of At-Risk Digital Geospatial Data NCSU Libraries and NCCGIA. [En línea]. . [21 de diciembre de 2011].
- Morris, S. (2006): Geospatial Web Services and Geoarchiving: New Opportunities and Challenges in Geographic Information Services. Library Trends. [En línea]. vol. 55, nº 2. < <http://muse.jhu.edu/journals/lib/summary/v055/55.2morris.html>>. [21 de diciembre de 2011]. ISSN 0024-2594.
- Morris, S. (2009): Preserving Geospatial Data: Challenges and Opportunities. Proceedings Indo-US Workshop on International Trends in Digital Preservation. Human Centered Design and computing Group. Pune, India. CDAC, pp. 11-19.
- Morris, S.; Tuttle, J.; Essic, J. (2009): "A Partnership Framework for Geospatial Data Preservation in North Carolina". Library Trends, vol. 57, nº 3, Winter 2009.
- PIRA (2000). Commercial Exploitation of Europe's Public Sector Information. Luxembourg. Pira International, Ltd., University of East Anglia, and Knowledge.
- Shaon, A.; Woolf, A. (2011): Long-term Preservation for Spatial Data Infrastructures: a Metadata Framework and Geo-portal implementation. D-Lib Magazine. [En línea]. Vol. 17, nº 9/10. . [21 de diciembre de 2011]. ISSN: 1082-9873.
- Sweetkind-Singer, J., Larsgaard, M.L., Erwin, T. (2006): Digital preservation of geospatial data. Library Trends. vol. 55, nº 2, pp. 304-314.
- Waugh, A., Wilkingson, R., Hill, B., Dell'Oro, J. (2000): Preserving Digital Information Forever. Fifth ACM Conference on Digital Libraries). San Antonio, Texas, USA. ACM Press. pp.175-184.