

PRESERVACIÓN DE LA INFORMACIÓN GEOGRÁFICA: PERSPECTIVAS Y SITUACIÓN EN ESPAÑA

FRANCISCO JAVIER ARIZA LÓPEZ¹, ROCÍO DEL MAR ARIZA LÓPEZ¹, MANUEL ANTONIO UREÑA CÁMARA¹, JOAQUÍN CORTÉS JOSÉ², LUIS ALFONSO UREÑA LÓPEZ³

¹Dpto. de Ingeniería Cartográfica, Geodésica y Fotogrametría, Universidad de Jaén

fjariza@ujaen.es, rdmarariza@gmail.com, maurena@ujaen.es

²Instituto de Estadística y Cartografía de Andalucía

Joaquín.cortes.jose@juntadeandalucia.es

³Dpto. de Informática, Universidad de Jaén

laurena@ujaen.es

RESUMEN

La preservación de contenidos digitales es una cuestión sin una solución única ni definitiva. Esta situación es más crítica en el caso de los datos espaciales. Este trabajo presenta definiciones, marcos, métodos, y ejemplos específicos de aspectos y proyectos de preservación de datos espaciales. Se vincula la preservación de los datos espaciales a los metadatos y por ello se analizan las normas de metadatos bajo este punto de vista. Se proponen tres alternativas: nueva norma ISO, desarrollo de perfiles, y extensión de ISO 19115. También se ofrecen algunas directrices para incluir aspectos de preservación de los datos espaciales en sus metadatos. Por medio de una encuesta se determina la situación actual de la preservación de los datos espaciales en España. Esta encuesta ha sido contestada por 20 organizaciones e indica que las actividades de preservación de los datos espaciales están en el orden del 19%, la mitad del nivel de contestación (40%) para las actividades de preservación generales realizadas por las bibliotecas y archivos que han contestado. Finalmente se indican posibles líneas de investigación y trabajo, así como la necesidad de la regulación y legislación a este respecto.

Palabras clave: datos espaciales, información geográfica, preservación, contenidos digitales.

PRESERVATION OF THE GEOGRAPHIC INFORMATION: PERSPECTIVES AND THE SPANISH SITUATION

ABSTRACT

The preservation of digital contents is an issue with neither unique nor definitive solution. This situation is more critical in the case of the spatial data. This paper presents definitions, frameworks, methods, and specific examples of spatial data preservation issues and projects. The preservation of the spatial data is linked to metadata and for that reason the metadata standards are

analyzed under this point of view. Three alternatives are considered: new ISO standard, the development of profiles, and the extension of the ISO 19115. Also some guidelines are offered to include preservation issues of the spatial data in their metadata. By means of a survey the situation of the preservation of the spatial data in Spain is determined. This survey has been answered by 20 organizations and indicates that the activities of preservation of spatial data are in the order of 19%, half of the level of answer (40%) for the general activities of preservation made by the libraries and archives that have answered. Finally possible lines of research and work are pointed out, as well as the necessity of a legal regulation.

Keywords: spatial data, geographic information, preservation, digital contents.

1. Introducción

"Felipe II claramente percibe que la administración de un imperio debe descansar en el control de la escritura, único medio receptor de informaciones y emisor de órdenes. Para cumplir este objetivo construye un edificio y promulga un reglamento. En 1572 manda a Juan de Herrera que haga las trazas de lo que se convertiría en el primer edificio construido para archivo de la época Moderna, y en 1588 firma una instrucción considerada el primer reglamento de archivos del mundo." (Historia del Archivo General de Simancas. www.mcu.es/archivos/MC/AGS/Presentacion/Historia.html).

Los "imperios" actuales continúan descansando en la escritura como medio receptor de información y emisor de ordenes si bien, en la actual sociedad de la información y del conocimiento (SIC), las técnicas de escritura (digital: ofimática) y el soporte han cambiado sustancialmente (magnéticos y ópticos), por lo que su archivo, conservación y custodia, en definitiva su preservación, se enfrenta a nuevos retos y a la toma urgentes de decisiones, tanto en su reglamentación como en el diseño de instrumentos para tratar los contenidos digitales y sus continentes.

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. De esta forma, la necesidad de la preservación de forma masiva de los contenido digitales (CCDD) es algo que ya ha llegado a los medios de comunicación de masas (p.e. Santini, 2011) y que también requiere su atención en espacios más técnicos y científicos de nuestro país pues se trata de un fenómeno que puede tener una gran repercusión social, cultural y económica. Es algo que puede afectar grandemente a la memoria histórica y condicionar en gran manera el futuro de nuestra civilización. La tarea de preservación es un reto tremendo, basta sólo imaginar la necesidad de verificar sucesivamente volúmenes de datos del orden de yottabytes ($2^{80} = 1$ septillón de bytes o un millón de trillones de megabytes) (UC, 2005).

En esta situación tienen responsabilidades tanto las administraciones públicas, en todos sus niveles (supranacionales, nacionales, regionales, locales), y el sector privado. Todos ellos están

creando numerosísimos 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 verdadera preocupación por este asunto y se están tomando iniciativas vinculadas al desarrollo de las denominadas e-infraestructuras. También hay otras acciones lideradas por las agencias espaciales estadounidense (NASA) y europea (ESA) que se llevan enfrentando décadas a los problemas de preservación y acceso a largas secuencias de datos temporales relativos a la observación de la Tierra.

La importancia de la IG y de la Geomática en la SIC ha hecho que pase de ser un sector bastante reducido a alcanzar valores de negocio nada despreciables en todos los países avanzados. La IG es una tipología de información con unas características especiales (voluminosa, dependiente de escala, borrosa, dinámica, encadena numerosos procesos y algoritmos) (APCIG, 2008) que la distinguen del resto. Además, se trata de una información tradicionalmente costosa de generar, costosa de mantener actualizada y, a la vez, muy necesaria para la gestión que realizan las administraciones públicas (Pira, 2000).

Acceder a la información actual, más reciente, es el mayor deseo de la gran mayoría de los usuarios de la IG, pero las decisiones económicas, medio ambientales y sociales se deben basar muchas veces en estudios temporales que descubran tendencias y den soporte a decisiones técnicas lo más robustas posibles. Según las ESA (2009), el interés y necesidad de acceder a datos históricos de observación de la Tierra se ha incrementado fuertemente en los últimos 10 años y se espera que esta tendencia se incremente aún más en el futuro. Por todo ello se hace necesario preservar estos datos sin limitaciones temporales y hacerlos accesibles y explotables. Sin embargo, en este mundo digital se comienza a tener problemas de acceso a IG del pasado por problemas con sus soportes, formatos, o falta de los programas adecuados para que funcionen, todo lo cual parece un contrasentido en el seno de la SIC, donde la rapidez de los avances pueden llevar a perder parte de la información relevante del pasado.

Aspectos que caracterizan la IG actual, su entorno, 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.

Este problema ha sido apuntado de manera general, y se han desarrollado esfuerzos para preservar CCDD de carácter general, pero la IG en formato digital es un campo que todavía no ha recibido suficiente atención y que, a nuestro parecer, presenta una complejidad mayor que otros.

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* (NDIIPP) 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). Si bien la situación ha cambiado bastante, el trabajo de Beagrie (2003) puede ser útil por cuanto presenta diversas iniciativas en todo el mundo (Australia, Francia, Holanda, Reino Unido, etc.).

El convencimiento de que "*the preservation and re-use of digital data and information forms both the cornerstone of future economic growth and development, and the foundation for the future memory*" (Ross, 2000), nos mueve a presentar este trabajo en la línea de la evangelización sobre este nuevo asunto dentro del sector de la IG en España. Con los siguientes objetivos:

- Explicar el problema de la preservación de los contenidos digitales.
- Explicar las peculiaridades de la preservación de la IG, presentando ejemplos en este ámbito.
- Conocer el estado actual de la preservación de la información geográfica en España por medio de una encuesta a bibliotecas y archivos.
- Apuntar algunas posibles líneas de investigación y acciones que se podrían tomar en nuestro país.

En cuanto a la estructura, además de esta introducción, el documento cuenta con siete apartados principales. El primero realiza un acercamiento al concepto de preservación y sus implicaciones. Dada la importancia de los metadatos en la preservación, se dedica un apartado a este aspecto haciendo un análisis específico de los problemas que presentan algunos estándares de metadatos respecto a este asunto. El apartado siguiente presenta algunas iniciativas de otros países que interesa conocer. Tras tener conocimiento de lo que ocurre en el extranjero, en el siguiente apartado se pretende obtener una idea de qué ocurre en España y para ello se presentan los resultados de un primer estudio *ad hoc* basado en una encuesta. A partir de aquí, en los dos apartados finales se formulan posibles líneas de investigación y las principales conclusiones del trabajo.

2. La preservación

Hasta mediados de los años 80 del siglo XX los archivos y bibliotecas habían gestionado mayoritariamente documentos analógicos. Con la aparición de los CCDD también han ido sumando a sus fondos soportes digitales (disquete, CDROM, DVD, memoria, etc.). Sin embargo la gestión de esta información de cara al futuro, a largo plazo, a lo que denominamos preservación, es mucho más compleja que para el caso analógico donde bastaba con tener unas condiciones de almacenamiento adecuadas. Los CCDD no se preservan de cara al futuro como se hace con los documentos analógicos (sobre papel en su mayoría). 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 en comparación con los contenidos analógicos. Esta situación, descrita brevemente de forma general,

es aún más crítica en el caso de la IG dada su especificidad. Por tanto, 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 de éstos es un aspecto poco considerado hasta la fecha, y aún menos en el caso de la IG digital. Por tanto, el uso y beneficio futuro de este tipo de contenidos queda comprometido a largo plazo, a menos que se creen e implementen las políticas y procedimientos adecuados para asegurar la disponibilidad de dicha información (Brown y col., 2005). Según lo indicado, la preservación también puede y debe ser una actividad ligada a la demanda y creación de valor a lo largo del tiempo.

2.1 De la conservación analógica a la preservación digital

Los archivos de documentos tienen dos misiones fundamentales: la custodia y la conservación de la documentación; y en torno a ellas giran sus reglamentos, instrumentos de descripción, instalaciones, sistemas de seguridad, etc. Pero "la rentabilidad del documento de archivo no se mide por la amplitud de su difusión, como puede ocurrir con un libro, sino por tenerlo en un momento dado asequible para instrumentalizar una actividad práctica" (Romero-Tallafigo, 1994). La irrupción en los archivos de la era digital ha venido, en gran medida, a facilitar su trabajo al aplicar las herramientas informáticas a distintas tareas de su quehacer diario, pero a la vez les ha planteado nuevos retos en cuanto al tratamiento que deben recibir los documentos de CCDD.

Hasta ese momento de cambio de era y del objeto de la preservación, no fue necesario distinguir en los archivos el documento analógico del digital, durante mucho tiempo sólo el primero fue el principal objeto de su custodia (guardia y vigilancia), y de su conservación (manteniéndolo de tal manera que no se perdiera ni se deteriorara). Pero al documento de CCDD se ha convenido en aplicarle la acción de preservar, que si bien se utiliza como sinónimo de conservar, el Diccionario de la Lengua Española de la RAE la define de un modo más extenso: "Proteger, resguardar, librar anticipadamente a una persona, animal o cosa, de algún daño o peligro" porque los CCDD y la IG, como se verá, requieren que se libre anticipadamente del daño o peligro de la fragilidad de su soporte, de sus elementos sustentadores de la información y de su irreversible pérdida. Los CCDD, a diferencia de la documentación analógica, no sólo requieren unas instalaciones y condiciones ambientales específicas para su conservación, sino que se debe disponer además del hardware y del software para su lectura y para la recuperación de los datos que contiene; sólo el poder difundirlo y facilitar su uso (instrumentalizar una actividad práctica) justifica su presencia en el archivo.

Los materiales y formatos analógicos (papel, piel, cerámica, metales, etc.) son capaces de durar cientos y, en algunos casos, miles de años. Se trata de materiales que envejecen paulatinamente, por lo que se degradan paulatinamente. Las causas de alteración del material analógico (tipo libro o documento en papel) abarcan tanto aspectos intrínsecos como el soporte (papel, cartón, pergamino), y los elementos sustentados (tintas, barnices, emulsiones, etc.), como otros de carácter externo (luz, humedad, temperatura, contaminación, etc.).

La cartografía como documento analógico de archivo es un tipo de documentación gráfica o figurativa que emplea un lenguaje de "signos naturales de reproducción de formas del objeto de la

información" (Romero-Tallafigo, 1994). La IG que contiene además de emplear ese lenguaje gráfico natural (signos convencionales como ríos, caminos, vegetación...), se expresada también mediante un lenguaje matemático codificado (geodesia, proyección, escala) y un lenguaje alfanumérico (toponimia, cotas altimétricas, distancias...). La información está sobre materiales (papel, pergamino, textiles...) y graficada mediante tintas, barnices, emulsiones, etc. La información se lee directamente sobre el documento analógico y sólo para algunos datos o aspectos puede requerir unos conocimientos expertos o ser necesario utilizar algún instrumento para su interpretación. Su conservación, a excepción de complicaciones debidas a su tamaño, no presenta grandes diferencias con el resto de la documentación escrita. Si bien no son abundantes, sí existen algunas referencias específicas a esta casuística, como por ejemplo los trabajos de Larsgaard (1998), Nichols (1982) y en castellano Colomar-Albájar (2001). Los CCDD y la IG se diferencian sustancialmente por sus soportes (discos, discos flexibles, cintas en bobinas o en casetes...) y por los elementos sustentadores: partículas magnéticas (soporte magnético) o perforaciones de la superficie o perforaciones reflectantes (soporte óptico) codificados en sistema binario, y que siempre requiere de un hardware y de un software para su lectura.

Llegados a este punto aparece la siguiente cuestión: ¿la preservación digital es la preservación de los originales por medios digitales o consiste, por el contrario, en preservar los propios materiales digitales?. Según Kenney (1996) la preservación digital puede ser entendida tanto como la digitalización de los objetos físicos para su preservación, pero también como la preservación de los propios CCDD. La preservación en sí es un fin secundario, puesto que la versión digital de un original, ya analógico ya digital, no se considera como un recurso primario de preservación. El enfoque actual de la preservación de CCDD es que el documento digital pasa a ser el objeto de la preservación, y no el medio. Un paso más es el que aporta el concepto de curación¹. La curación de los CCDD es la actividad de mantener y añadir valor a un conjunto de CCDD para su uso actual y futuro (Janeé, 2009). La curación debe cubrir el ciclo de vida de la información de manera completa (creación, uso, archivo, reutilización). Cada tipo de información tiene necesidades específicas y los datos científicos, dentro de los cuales podemos enmarcar la IG, requieren muchas veces un reprocesado importante así como transformaciones.

Otra característica importante de los CCDD es la siguiente: el acceso a la información y la preservación son, en los materiales tradicionales, dos actividades que se pueden realizar separadamente, pero esto no sucede así en el mundo de los CCDD que requieren, ineludiblemente, una implementación de los mecanismos descriptivos y de acceso que aseguren la posterior recuperación de dichos contenidos. A pesar de los avances, como apuntaba Levy en 1998 (Levy, 1998), la preservación de la información digital es un problema difícil y de difícil comprensión. A diferencia de lo que ocurría con los materiales analógicos, la información digital (a nivel de bit) no se pierde o degrada de manera paulatina, simplemente sus bits son accesibles (existen), o no son accesibles (como si no existieran). La información digital es, por naturaleza, frágil y poco permanente (Brown y col., 2005), especialmente aquella que no tiene contraparte analógica. Esta situación ocasiona, con toda seguridad, que de manera cotidiana miles de recursos digitales importantes a nivel nacional estén en la actualidad en riesgo de perderse o corromperse.

Por tanto, un concepto fundamental para entender la preservación digital es el largo plazo, que ya se ha mencionado con anterioridad en este documento. 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. O en otras palabras, 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). En el fondo lo que pretende la preservación es 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). Como se puede observar en la enumeración anterior, muchos de los aspectos están ligados con la obsolescencia. El deterioro físico de los medios de almacenamiento está ligado a la obsolescencia funcional del soporte. Este deterioro puede causar la pérdida total o parcial de los contenidos si no se refrescan los contenidos (*refreshing bits*). En función de los medios (p.e. cinta de carrete, disquetes, etc.) hay diversas técnicas (p.e. rebobinados periódicos) y condiciones ambientales (temperatura, humedad, campos eléctricos, etc.) que pueden tenerse en cuenta para evitar o, al menos, minimizar este problema. Las obsolescencias de tipo tecnológico (motivadas por la evolución tecnológica), legal (motivadas por cambios legislativos), económico (motivadas por métodos o tecnologías más eficientes) o de estilo (motivadas por modas o estilos) también afectan en gran medida a los CCDD. Muchas veces el soporte físico en el que se almacena el CCDD (p.e. el disquete o el CDROM) puede ser más duradero que el dispositivo que se usa para leerlo, lo que genera problemas de lectura de cara al futuro y ello lleva a implementar procesos de refresco y migración de los datos de un soporte a otro (Brown y col., 2005).

La renovación de soportes, rejuvenecimiento o refresco (UKDA, 2010) es el proceso de transferencia de datos de un tipo de almacenamiento a otro más moderno para asegurar el acceso continuado a la información, pero sin la alteración del formato de los datos. La migración (CLIR, 2000) es la transferencia periódica de contenidos digitales de una configuración de hardware/software a otra, o desde una generación de tecnologías de computadoras a otra generación subsecuente. La migración debe mantener la autenticidad, integridad, fiabilidad y usabilidad de los CCDD (UKDA, 2010). La migración es propensa a errores (CLIR, 2000).

Las soluciones a la obsolescencia pueden orientarse según varias líneas según fueron apuntadas por Arms (2000), las cuáles admiten cierto grado de mezcla. En la [tabla 1](#) se presentan las opciones, y las ventajas y desventajas, de cada una de ellas. Por supuesto, también se pueden considerar enfoques híbridos entre estas opciones. Se denomina arqueología digital a las acciones de recuperación de contenidos cuando no se han realizado acciones sistemáticas de preservación.

Pero las opciones indicadas no son del todo adecuadas si se adopta una perspectiva a largo plazo. Si se desea que la preservación sea duradera (p.e. al menos cien años), es necesario plantear

modelos de sistema, y que éstos 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 en mente, Janée y Frew (2008) establece los siguientes criterios como base de la arquitectura 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: este principio indica que el sistema de preservación debe permitir que por motivos determinados (p.e. económicos) alguna información deje de estar disponible durante un lapso de tiempo, pero debe almacenar y preservar suficientes metadatos e información contextual para permitir la resurrección futura de acceso y uso de esa información. Este principio se opone a la inmortalidad la cual consiste en mantener los contenidos vivos en todo momento (Janée 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 necesaria para que éste pueda ser interpretado/comprendido por un usuario (aplicación o persona física). Es decir, la capacidad de entender y de decodificar la información preservada sin referencia a ninguna otra documentación externa.

En este sentido, mientras la conservación de los materiales tradicionales está orientada a la estabilización del soporte, en el caso de los recursos digitales la obsolescencia de los formatos, de los programas (software) e, incluso, del equipamiento de conversión (hardware) puede devenir con mayor rapidez que la del mismo soporte. Por este hecho, en todo proceso de preservación digital, se hace obligado un flujo o transferencia continua de la información, de lo contrario nos encaminamos a paso seguro, a la destrucción y pérdida de la misma (Brown y col., 2005). Así, se puede afirmar que la preservación de CCDD no depende del soporte en sí, sino de la capacidad que muestre dicho contenido de ser transferido de un soporte/formato a otro soporte/formato dentro de un mismo sistema o entre sistemas. Otro aspecto de especial relevancia para los CCDD en este flujo y cambio de soportes/formatos y/o sistemas son los relativos a su autenticidad e integridad. La autenticidad debe ser preservada, ello supone una cadena de custodia y mantener un enlace permanente entre los metadatos relativos al origen y los propios registros digitales. La integridad se refiere tanto a los CRC (*Cyclic Redundancy Check*, controles de redundancia cíclica) de los datos originales, y de sus réplicas, como a mantener un sistema de nombres y controles persistentes en el tiempo.

En la línea de todo lo anterior, según Janée y Frew (2008) la migración de archivos entre sistemas es una tarea intensiva en personal y recursos que se caracteriza por la necesidad de tomar

decisiones *ad hoc*, tanto de contenido como de contexto. La preservación digital utiliza sistemas específicos como BagIt. Éste es un desarrollo creado a partir de los trabajos de la Biblioteca del Congreso y de la Biblioteca Digital de California dentro de la iniciativa NDII&PP de los EEUU. BagIt es un formato jerárquico de empaquetado de ficheros diseñado para sistemas de almacenamiento basados en discos y transmisión por red. El sistema consiste en una carga y unas etiquetas que actúan como metadatos. En NGDA (2009) se comenta algo este procedimiento y cómo esta utilidad fue mejorándose a lo largo del proceso gracias a las pruebas realizadas de carga de datos.

Otra iniciativa es la denominada *Open Archival Information System* (OAIS) que establece un modelo de referencia de los componentes técnicos y funcionales de los sistemas de gestión y preservación de CCDD. El modelo OAIS se corresponde con la Norma Internacional ISO 14721:2003. NARA (*National Archives and Records Administration*), que es el ente Federal que coordina los archivos de todas las agencias federales de los EEUU. NARA (NARA, 2011) ha adoptado como base para la preservación el modelo OAIS, y para facilitar su adopción ha desarrollado diversos documentos que facilitan su entendimiento y aplicación. El modelo OAIS se basa un conjunto de entidades funcionales y de procesos de negocio (carga, almacenamiento, gestión, administración, plan de preservación y acceso) con los que interactúan tres agentes (productores, clientes y gestores). En Shaon y Woolf (2011) se puede encontrar una buena descripción resumida del modelo.

2.2. Preservación de la IG: motivos y problemas

En primer lugar hay que preguntarse si la IG debe o merece ser preservada. Aun siendo una pregunta obvia no por ello deja de ser necesaria. En la introducción se ha indicado la especificidad y el valor económico y social de la IG, y también se ha indicado que el modo de uso actual supone ciertos riesgos. Puesto que frente a los riesgos los bienes de valor deben ser protegidos, se hace evidente la necesidad de preservación de la IG. 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) indican las siguientes causas:

- 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 (ISO) o especificaciones abiertas (p.e. *Open Geospatial Consortium*, OGC) 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. Además, los datos pueden agregarse, desagregarse y operarse con facilidad.
- Sistemas relacionales: cada vez más el almacenamiento SIG se realiza sobre bases de datos relacionales. Esto supone que se heredan las desventajas de estos sistemas 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 los metadatos no existen o son insuficientes y, en todo caso, cuando existen no están pensados para facilitar la preservación de la IG. Es esta línea es crítica la falta de perspectiva de ciclo de vida del producto de los metadatos al uso en IG (p.e. según ISO 19115). Otro problema que dificulta la preservación es la falta de sincronía entre los metadatos y los datos a los que se refieren. Esta asincronía se refiere a que pueden existir cambios en los datos o su contexto (p.e. una nueva versión, un nuevo responsable o persona de contacto, un nuevo teléfono de contacto) tales que no son actualizados adecuadamente en los metadatos.
- Versionado: la IG vectorial con requisitos temporales (p.e. catastro, construcciones) hacen que el problema de la preservación adquiera una dimensión mayor.

Ariza López, F.J., Ariza López, R.M., Ureña Cámara, M.A., Cortés José, J., Ureña López, L.A. (2012): "Preservación de la Información Geográfica: Perspectivas y situación en España", *GeoFocus (Artículos)*, nº 12, p. 171-200. ISSN: 1578-5157

- 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 EUROSDR sobre esta temática (EUROSDR, 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

Un concepto con el que se trabaja en preservación es el de documento. Es decir una unidad de información auténtica, exacta, fiable e íntegra que se mantiene a lo largo del tiempo. Ello supone mantener también información sobre el contexto y la estructura relativa a la unidad de información preservada. Por tanto, la preservación no es simplemente mantener una copia de seguridad, supone disponer de unos metadatos y capacidades adecuadas para poder acceder a dicha información en el futuro, para preservarla frente a cambios de formatos, al daño de sus soportes, accesos indeseados, pérdidas fortuitas, etc., asegurando las propiedades indicadas anteriormente. Todo esto cuesta muchos recursos y por ello debe ser selectiva. De esta forma se puede afirmar rotundamente que sin metadatos la preservación no es posible. Los metadatos de preservación deben ser suficientes para permitir la realización de cambios en los registros a través de varias generaciones de hardware y software, y permitir la trazabilidad y auditoría a lo largo del ciclo de vida de un registro de IG. 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.

Ariza López, F.J., Ariza López, R.M., Ureña Cámara, M.A., Cortés José, J., Ureña López, L.A. (2012): "Preservación de la Información Geográfica: Perspectivas y situación en España", *GeoFocus* (Artículos), nº 12, p. 171-200. ISSN: 1578-5157

- 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 al respecto, y más consolidadas. El proyecto PREMIS (*Preservation Metadata: Implementation Strategies*) 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). Se trata de un proyecto de eminente orientación práctica en el que trabajaron numerosos expertos internacionales con el objetivo de desarrollar un conjunto fundamental de metadatos de preservación, basados en directrices y recomendaciones de creación, gestión y uso, que pudiesen ponerse en práctica y aplicarse de manera general en una amplia variedad de contextos de preservación digital. Resultado de este trabajo fue la publicación en 2005 del documento *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group* (Diccionario de Datos para Metadatos de Preservación: Informe final del grupo de trabajo PREMIS). Además del diccionario se publicó un esquema XML para facilitar su adopción de manera directa. Tras un proceso de revisión en 2007 se publicó la versión 2 del diccionario. Este material está disponible en español gracias a la traducción realizada por personal de la Biblioteca Nacional de España (BNE, 2010).

En el ámbito de la IG existe en la actualidad una gran sensibilidad hacia los metadatos. Sin embargo, el reconocimiento de su importancia es relativamente reciente. Este reconocimiento viene potenciado por el desarrollo de las IDE (Infraestructuras de Datos Espaciales) y la difusión de las normas ISO de la familia 19100, en concreto de la norma internacional ISO 19115. Como forma de ir avanzado en este tema, y dada la amplia difusión actual de los metadatos en el campo de la IG, en primer lugar se hace necesaria una revisión crítica de los metadatos existentes para la IG pero desde la perspectiva de la preservación. En esta línea ya se han realizado algunos trabajos que se comentan a continuación.

Hoebelheinrich y Banning (2008) analizan el estándar de metadatos del FGDC (*Content Standard for Digital Geospatial Metadata*, CSDGM) y vienen a apuntar dos conclusiones de gran importancia relativas a la preservación, las cuales consideramos también son válidas para la norma ISO 19115 y su nivel de aplicación general. En primer lugar, la riqueza del CSDGM puede ser considerada como una debilidad dado que el elevado número de elementos que incluye puede resultar abrumador y confuso. La complejidad del estándar y la general resistencia 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

(2009 a,b) se presenta un análisis del CSDGM frente al Dublin Core que materializa una tabla de elementos cruzados 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 desde la perspectiva de la preservación. Igualmente, se llegan a identificar (GeoMAPP, 2011) los campos de mayor interés de la norma CSDGM respecto a la preservación.

Shaon y Woolf (2011) proponen tres aspectos fundamentales para la integración del modelo OAIS con las particularidades existentes en la IG y gestionar, de esta forma, la preservación de los metadatos con este tipo de características tan particulares:

- Preservación a largo plazo de los repositorios de datos: conjunto de procesos para mantener en funcionamiento, a pesar de los cambios en la tecnología, los repositorios aunque dejen de utilizarse.
- Modelo basado en la preservación de metadatos: desarrollo del modelo propuesto en ISO 19115 para solventar las dificultades que aparecen al tratar de mantener CCDD a largo plazo.
- Curación de Catálogos de Metadatos en el largo plazo: Idea acuñada por Shaon y Woolf (2008) referida a la necesidad de mantener los catálogos operativos, ofreciendo servicios de calidad a lo largo del tiempo.

Desde un punto de vista más práctico y centrado en los contenidos almacenados en los metadatos de IG de diversos productores nacionales e internacionales y tras revisar más de 50 metadatos de diferentes agencias, podemos afirmar con rotundidad que el aspecto de la preservación no está considerado. Baste como ejemplo los recogidos en la [tabla 2](#) para el caso de las series del Mapa Topográfico Nacional 1/25.000, Mapa Topográfico de Andalucía 1/10.000 y otros productos elaborados por diferentes agencias gubernamentales extranjeras. Ninguno de estos ejemplos se preocupa de la información de preservación. Sin embargo, y no en todos los casos, sí informan de manera profusa sobre su formato de almacenamiento (*MD_Format*), o de los roles de distribuidor (*MD_Distribution*). A pesar de todo ello, muchos de estos productos sufren continuas actualizaciones sin que en ningún momento se haya llegado a considerar que cada actualización es otro objeto de preservación en sí mismo, por lo que directamente se elimina la anterior versión, se publica la nueva y no se informa ni de los cambios realizados ni de los procesos que han intervenido, existiendo en este caso una preservación nula tanto desde el punto de vista de los CCDD como de los metadatos, donde se deberían indicar todos estos cambios. De hecho, algunas agencias ni tan siquiera indican el período de actualizaciones estimado. Es más, de hecho, agencias tan importantes como el *Ordnance Survey* (UK) no disponen de metadatos públicos, y éstos sólo son accesibles tras la descarga de los archivos de pago. Y otras como *Bundesamt für Kartographie und Geodäsie* o *Swisstopo*, pese a tener unos metadatos bastante extensos (cumpliendo ISO 19115), no tienen información del formato de archivos, ni información del mantenimiento, ni información adicional. Con todo, otras entidades públicas como el Instituto de Estadística y Cartografía de Andalucía (IECA), cuando existen cambios sustanciales en las series (por ejemplo en el Mapa Topográfico de Andalucía), sí que permite acceder a cada una de estas versiones de la serie de forma individualizada, existiendo, en cierta medida un nivel de preservación de la IG, al menos desde un punto de vista de obtención de capturas de instantes determinados.

A pesar de todo lo anterior existen iniciativas como la de Swisstopo que pese a no haber implementado con carácter público la preservación dentro de sus IG, hasta la actualidad, ha realizado el esfuerzo de elaborar un estudio preliminar de archivado de IG (Bos y col., 2010) y a principios del año 2011 ha iniciado el proyecto Eclipse para tratar de crear un marco técnico y organizativo para la aplicación de práctica de la preservación a la IG.

Analizando la preservación desde la perspectiva de su inclusión en los metadatos, caben dos opciones, bien como información digital y, por lo tanto, se puede asimilar a otro contenido, con los mismos problemas y dificultades, o desde el punto de vista de un elemento informador sobre el elemento de relevancia a preservar (la perspectiva aquí de interés). Si se consideran los metadatos como informadores acerca de la preservación de elementos de IG, hasta ahora, ninguno de los estándares desarrollados CSDGM, ISO 19115 ni Dublin Core disponen de elementos adecuados para la inclusión de la preservación como un aspecto de la IG (Shaon y Woolf, 2011). De hecho, ninguno de los perfiles posteriormente 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 contenidos propios 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 ([figura 1](#)):

- Información de representación (como añadido al Esquema de aplicación): 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. Por ejemplo, en este sentido y para un elemento correspondiente a una carretera, se indicarían el grosor, patrón, color, o un enlace web para poder observar dicha representación.
- Anotaciones adicionales: este mecanismo permite almacenar información adicional a la propia IG. 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. Para que las anotaciones tengan significado pleno se debe mantener el contexto en el que esta información fue creada y para qué fue creada. Por ejemplo, en el contexto de un nomenclátor, para un nombre geográfico que se ha verificado en campo se podrían incluir comentarios relativos a este proceso.
- 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. En Inspire la forma recomendada (DTDS, 2010) es mediante un sello temporal basado en ISO 8601. En este caso el registro del sello debe tener una longitud de 25 caracteres alfanuméricos para permitir una especificación del tipo "mayor a menor" de forma expandida con: la fecha, la hora y la diferencia horaria.
- 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.). Además, se proponen eventos relativos a la calidad de los procesos: procesos de certificación, de mantenimiento y de control de calidad del proceso de preservación.

- Autenticación de la fuente de datos de cada elemento: debe incluir dos elementos, uno de firma y otro de anclaje. Ambos tratan de asegurar la procedencia de la información, el primero para que cada entidad productora asegure su autoría, y el segundo como un sistema de verificación de la autenticidad de la información recibida (p.e. métodos de comprobación de redundancia cíclica (CRC ó MD5), a modo de comprobadores de transmisión correcta).

Sin embargo, aunque la propuesta de Shaon y Woolf (2011) está bastante elaborada, consideramos que es insuficiente y no ha tratado todas las posibles soluciones. De hecho, se centran en el desarrollo de nuevas entidades sin incluir nuevos atributos a los ya existentes. Por ello, pensamos en al menos tres posibles vías de inclusión de la preservación en los metadatos de IG como alternativas a las ya indicadas:

1. Mediante una nueva norma ISO específica. Dado que el problema de preservación es general para los CCDD, la mejor propuesta sería el desarrollo de una norma ISO que permita la inclusión de toda la información necesaria para la preservación. Sería conveniente la interoperabilidad entre normas para que ésta pudiera ser aplicada de manera independiente sobre Dublin Core, ISO 19115 o cualquier otra norma ISO dedicada a CCDD (p.e. PDF (ISO 32000-1)).
2. Desarrollando un perfil al estilo de lo propuesto por Shaon y Woolf (2011) para la extensión de ISO 19115 con campos específicos. Consiste en aprovechar la estructura actual de ISO 19115 incluyendo nuevos campos dentro de estructura de los paquetes existentes y relacionados con la preservación. Por ejemplo, podrían incluirse aspectos de la preservación dentro de la descripción de la IG (*MD_ContentInformation*), o de la información de mantenimiento (*MD_MaintenanceInformation*). Otros elementos como las restricciones en el acceso (*MD_Constraints*), las restricciones legales (*MD_LegalConstraints*) son exclusivamente referidas a la forma de acceso a los datos desde el punto de vista del usuario y nunca de la entidad legal que se encarga a mantenerlos (salvo para contactar con dicha autoridad para su adquisición dentro del elemento *MD_Distributor*). Además, dentro de los procesos de linaje de cada elemento de IG catalogado se podrían incluir los aspectos de preservación actuales y no sólo los elementos de desarrollo del producto IG publicado. Una propuesta básica que añadiría atributos y/o elementos en diversos paquetes (más allá de lo descrito en *MD_Format* dentro de la propia identificación del documento –que es el aspecto más interesante dentro de la preservación):
 - a. Encargado de la preservación del documento: responsable de la preservación de la IG frente a los responsables de distribución, producción y otros roles diferentes.
 - b. Preservación del documento: linaje de procedimientos de preservación empleado con relación de las entidades encargadas. Estado actual de preservación y forma de acceso.
 - c. Formato de almacenamiento (descripción completa del formato o referencia a un documento accesible indicando el formato de almacenamiento), última fecha de acceso al contenido, si la descripción del formato es externa última fecha de acceso a dicha

descripción, descripción y acceso a los datos, programas empleados para su acceso, creación y/o modificación.

- d. Modificaciones realizadas sobre el documento: catálogo de modificaciones y/o documentos anteriores a los cuales hace referencia. Métodos de preservación empleados.

La inclusión de estos aspectos en ISO 19115, y en perfiles derivados (p.e. NEM) se podría realizar a través de la adición de nuevos campos en el elemento de mantenimiento, añadiendo el rol de organismo y los mecanismos de mantenimiento en estos elementos. Consideramos que estos campos deberían ser obligatorios frente a sus actuales versiones optativas (salvo para la frecuencia de actualización).

3. Mediante la extensión de ISO 19115 con un paquete específico. Es una opción tal vez más inmediata que las anteriores. Respecto a la 1ª opción resuelve el problema dentro del campo de la IG, sin necesidad de esperar a propuestas más generales que seguro luego habría que especificar para este tipo de contenidos tan distintos. Respecto a la 2ª opción también es más inmediata pues no afecta a la parte ya consolidada de ISO 19115, por lo que basta con realizar un añadido. Se basa en las posibilidades de ampliación de ISO 19115 (p.e. en los perfiles). Consistiría en incluir algún paquete específico de preservación donde se almacenasen los aspectos de transformación, métodos de transferencia (con la garantía de integridad de los datos) y resultados de evaluación de la validez (según lo indicado por McGarva y col., 2009) y con la información propuesta en la opción 2. En cualquier caso, las propuestas de McGarva y col. (2009) son poco específicas y con una información demasiado reducida para una verdadera gestión de la preservación, ya que se limita a aspectos muy generales ya tenidos en cuenta en ISO 19115.

En cualquiera de los casos, los contenidos deberían ser, al menos, aquellos indicados por McGarva y col. (2009) desarrollados parcialmente por Shaon y Woolf (2011), y presentados anteriormente. Según lo analizado, la forma de adquisición vienen incluidos en prácticamente todos los metadatos ya que son los que desarrollan el aspecto comercial, sin embargo, el resto de aspectos como los metadatos de adquisición, los métodos de transferencia, garantía de integridad, transformaciones y validez son todos opcionales y en ninguno de los ejemplos se incluyen. El problema de cualquier propuesta en esta línea es que la inclusión de más elementos dificulta la ya compleja elaboración de los metadatos de la IG, tarea costosa y muchas veces relegada a un segundo plano. Por todo ello el diseño debería estar bien balanceado para que los trabajos de metadatos relativos a la preservación no supongan una carga excesiva. En este sentido consideramos más adecuada una filosofía más cercana a la norma Dublin Core que a ISO 19115.

En definitiva, se puede indicar que los modelos actuales de metadatos (CSDGM, ISO 19115, Dublin Core, NEM) carecen de la perspectiva de la preservación, y que esta falta de perspectiva conlleva que no se almacenen datos de preservación. Por ello, siendo la preservación de la IG tan necesaria 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.

Con independencia de la vía que se adopte, es indudable que las actividades de preservación supondrán un trabajo y un coste extra a los productores de IG por esta nueva exigencia de metadatos. Sin embargo, éste es un aspecto no tratado en ninguna de las referencias. Tal vez sea que, al igual que en la actualidad nadie se cuestiona la producción de metadatos para la IG, en un futuro cercano tampoco se podrá cuestionar la creación de metadatos de preservación. Por otra parte, los metadatos de preservación afectan tanto al productor como a los sucesivos agentes involucrados en el uso, almacenamiento y custodia de la IG, por lo que es una situación de esfuerzos compartidos que debe ser interoperable. Consideramos que la incorporación de estos metadatos debe iniciarse por parte de los productores, pues es donde se inicia el ciclo de vida del producto de IG, pero que posteriormente se convierte en una especie de testigo de carrera de relevos que se deberá ir pasando de agente en agente, según vayan interviniendo en la vida de un conjunto de datos concreto. Así, el productor se encargaría de iniciar el proceso, aportando información sobre los contenidos iniciales (procesos, fuentes, aspectos legales, etc.), añadiendo identificadores únicos y de versionado, etc., y los sucesivos organismos propietarios, receptores o depositarios se deberían ir añadiendo registros relativos a los procesos de modificación de los datos y de preservación (p.e. cambio de formatos).

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 o topográfica, se destaca el proyecto desarrollado en los últimos años en Suiza que pretende una solución integral para la totalidad de la administración federal basada en la coordinación sobre IG; y que se encuentra muy bien documentado en Bos y Col. (2010). Lo interesante de este proyecto no son sólo las conclusiones del mismo, que en gran medida están condicionadas por la realidad suiza, sino la metodología seguida para su desarrollo. Se parte de no haber encontrado una solución asentada al problema de la preservación en otros países, 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 (sistemas implantados en la administración helvética, volumen de datos, datos de referencia, datos responsabilidad de los cantones, etc.) y externo (otras iniciativas, estándares, etc.), 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 por estar centrado en datos relativos a los datos espaciales de observación de la Tierra, 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, y que fueran válidas para los productores, proveedores y archivadores de este tipo de datos. Se trata también de una iniciativa muy bien documentada (ESA, 2008, 2009, 2010a, 2010b, 2011). 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. Se apuntan un total de 13 consideraciones (necesidad de preservar, tendencia a preservar todo, necesidad de una visión centrada en los usuarios, la preservación y reuso de la información como camino para el futuro crecimiento del conocimiento, necesidad de datos libres y abiertos, necesidad de armonización, fiabilidad de los datos, servicios de datos simples y sin coste, seguridad, series temporales y necesidad de acceso a datos de más de 15 años, necesidad de reprocesado periódico, etc. También se indicó la necesidad de evaluar el riesgo de pérdida de información y sus impactos asociados. Los métodos de desarrollo de la política y de la guía de preservación también se ha basado en un proceso participativo interno y externo, mediante talleres, encuestas, entrevistas, etc., buscando el consenso entre los actores, tanto en la definición de la política como en el desarrollo de la guía. 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.

4. 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 (<http://www.digmap.eu/doku.php>), y se ha desarrollado el proyecto español CartoVIRTUAL (<http://www.cartovirtual.es/>). 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, a la vez, abrir la posibilidad a descubrir aportaciones de interés. Por ello se decidió incluir tan sólo 5 preguntas que se contestaban con un SI o NO, pero a la vez se dejaba libre la opción de explicar en cada una de ellas las actividades de los centros consultados, así como de incluir cualquier otro comentario e incluso documentos o enlaces relacionados con su trabajo en esta materia. La primera pregunta (P1) era relativa a la existencia o no de cartoteca en la organización, con ello 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, como dos niveles o grados posibles de realizar estas actividades. Las dos últimas preguntas (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, si bien en otros casos nos han derivado al personal que ellos han considerado más apropiado al caso. En muchos casos ha sido necesaria cierta insistencia para obtener las contestaciones, obteniendo un resultado final de 20 encuestas completadas (77%). El resumen de contestaciones para cada una de las preguntas formuladas es el que se presenta en el [tabla 3](#).

En cuanto a los resultados de la P1, en un 61,5% de los casos 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 relativa a la preservación de CCDD generales ha sido contestada afirmativamente en un 42,3% de los casos. El propósito de esta pregunta era servir como nivel de referencia frente al que comparar los niveles de preservación de la IG dados por la encuesta. 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, relativa a la curación de CCDD general, 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 siguientes preguntas (P4 y P4) se formularon de manera paralela a las relativas a los CCDD de carácter general tratados en las preguntas 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.

Ariza López, F.J., Ariza López, R.M., Ureña Cámara, M.A., Cortés José, J., Ureña López, L.A. (2012): "Preservación de la Información Geográfica: Perspectivas y situación en España", *GeoFocus (Artículos)*, nº 12, p. 171-200. ISSN: 1578-5157

- 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.

Finalmente, respecto a la aportación de materiales adicionales, salvo dos casos, las organizaciones encuestadas no han remitido ningún tipo de información complementaria. En estos dos casos la documentación ha sido divulgativa, generalista y de valor limitado (folletos de propaganda de la organización) para nuestros objetivos.

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.

De manera general, la creación de arquitecturas capaces de preservar a largo plazo CCDD es urgente y difícil. Urgente pues cada día aparecen cientos de contenidos nuevos, y difícil no sólo por la cantidad sino por la variedad, por estar distribuida en la nube, y por los múltiples agentes y aspectos que se han de conjugar para alcanzar una interoperabilidad adecuada a largo plazo. Centrados ya en la IG en el interesante trabajo de Morris (2009), se indica las siguientes líneas de trabajo en el ámbito de la preservación:

- En colaboración con los cuerpos normativos (ISO, OGC, etc.), desarrollo de normas y estándares en temas como empaquetado de contenidos y archivo de perfiles (p.e. el grupo de trabajo de OGC en preservación).
- Desarrollo de mecanismos de transferencia y gestión distribuida en colaboración con las comunidades de computación distribuida (*grid computing*).
- 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 de las implicaciones del emergente 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.

Para el autor (Morris, 2009), es fundamental el comprometer en estas acciones tanto a la industria como a las infraestructuras de datos espaciales, y el disponer de una perspectiva basada en infraestructuras (Morris, 2006). 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. Esos mismos servicios también podrán ser utilizados para desarrollar los propios archivos. Es decir, los servicios web podrían utilizarse para crear de manera automática inventarios, extrayendo y transmitiendo sus contenidos. De cualquier forma, esto no deja de ser un deseo pues las especificaciones de los servicios OGC no tienen esta orientación, ni en el caso de los WMS, ni en el caso de los WFS. El mismo autor indica también la necesidad de una capa en la infraestructura de datos espaciales que actúe como el programa *Rsync* y que permita replicaciones completas eficientes, procesos de actualización parcial, selectivos, etc.

Otros autores, como Lazorchak y Sewash (2011) también analizan los retos actuales y coinciden en gran medida con los indicados anteriormente e incluyen otros aspectos relevantes como los relativos a los datos dinámicos o al archivado de proyectos (conjuntos de datos espaciales de distintos modelos y formatos), frente al basado en ficheros.

Desde una perspectiva más aplicada, también queda por explorar las posibilidades de implementación de todas estas cuestiones al nivel de los programas que utilizan los usuarios de IG. En esta línea, ya existen SIG comerciales que incluyen algunas utilidades para la gestión de versiones históricas como es la extensión *History Viewer* de ArcGIS (McGarva y col., 2009).

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. En la preservación de CCDD son importantes los aspectos técnicos, organizativos, los relativos a los derechos de propiedad y al desarrollo de alianzas. Desde el punto de vista de la normativa, se puede afirmar que hay un conjunto de normas suficiente para articular sistemas y desarrollar iniciativas. Los costes de

almacenamiento son constantes y es la gestión es lo que realmente cuesta a lo largo del tiempo (LC, 2009).

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. Por ello un aspecto crucial es que cada responsable realice una reflexión profunda sobre qué se debe preservar, aspecto fundamental del problema.

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, es importante disponer de metadatos pensados para la preservación, los registros de formatos son una línea de acción fundamental, las acciones de preservación son escasas pero se incrementan, el estado de las herramientas es deficitario para su aplicación a la IG.

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, 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.

Como se ha indicado en el apartado dedicado a las líneas de trabajo en preservación de la IG existe un amplio campo para la investigación en preservación de la IG. Además, estas investigaciones deberían tener en cuenta el desarrollo de las IDE, tanto como fuentes para la consulta y acceso a los datos como elementos que podrían conformar parte un sistema de preservación basado en la Red. 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 este trabajo sin hacer una llamada a la necesidad de que dentro del Sistema Cartográfico Nacional se tome en consideración la preservación de la IG. La simple revisión de la legislación cartográfica nacional más reciente (Ley 14/2010, de 5 de julio,

Ariza López, F.J., Ariza López, R.M., Ureña Cámara, M.A., Cortés José, J., Ureña López, L.A. (2012): "Preservación de la Información Geográfica: Perspectivas y situación en España", *GeoFocus* (Artículos), nº 12, p. 171-200. ISSN: 1578-5157

Real Decreto 1545/2007 de 23 de noviembre, Real Decreto 1637/2009 de 30 de octubre, Real Decreto 663/2007 de 25 de mayo), y de las legislaciones y planes cartográficos autonómicos (p.e. en el caso de Andalucía el Decreto 141/2006, Plan Cartográfico de Andalucía 2009-2012) nos 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,

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, desde 1997, 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 realizada y a los revisores que, con sus detalladas sugerencias, han permitido mejorar este documento.

Referencias

APCIG (2008): *Sistemas y Tecnologías de la Información*. Madrid. Asociación Profesional del Cuerpo de Ingenieros Geógrafo, Instituto Geográfico Nacional.

Arms, W. (2000): *Digital Libraries*. Cambridge, Massachussetts, USA: MIT Press.

Arms, C.R. y Fleischhauer, C. (2005): *Digital Formats: Factors for Sustainability, Functionality, and Quality*. Washington, DC, USA. Office of Strategic Initiatives, Library of Congress.

Beagrie, N. (2003): *National Digital Preservation Initiatives: An Overview of Developments in Australia, France, the Netherlands, and the United Kingdom and of Related International Activity*. Washington, DC, USA: Copublished by the Council on Library and Information Resources and the Library of Congress.

BNE (2010): *Diccionario de Datos PREMIS de Metadatos de Presevación Versión 2.0*. Madrid, España: Biblioteca Nacional de España, 2010.

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.

Brown, D., Welch, G., y 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]. <http://www.geoconnections.org/publications/policyDocs/keyDocs/geospatial_data_mgt_summary_report_20050208_E.pdf>. [21 de diciembre de 2011].

Ariza López, F.J., Ariza López, R.M., Ureña Cámara, M.A., Cortés José, J., Ureña López, L.A. (2012): "Preservación de la Información Geográfica: Perspectivas y situación en España", *GeoFocus (Artículos)*, nº 12, p. 171-200. ISSN: 1578-5157

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.

CLIR (2000): Risk Management of Digital Information: A File Format Investigation. *RLD-DigiNews*. [En línea], The Council on Library and Information Resources, vol. 4 (3). <<http://webdoc.gwdg.de/edoc/aw/rlgdn/preserv/diginews4-3.html>>. [21 de diciembre de 2011]. ISSN 1093-5371.

Colomar-Albájar, M.A. (2001): "Conservación del material cartográfico. Preservación y Restauración", en Jiménez-Pelayo, J. y Monteagudo-López-Menchero, J. (Ed.): *La Documentación Cartográfica, Tratamiento, Gestión y Uso*. Huelva, Servicio de Publicaciones, Universidad de Huelva, pp. 421-461.

Cortés-José, J. (1997): "Cartografía Histórica de Andalucía. Asociación de Imágenes a un gestor Documental. Catálogos de Cartografía Histórica Multimedia", Vitoria-Gasteiz, *Congreso Internacional sobre Sistemas de Información Histórica*, Juntas Generales de Álava.

ESA (2008): *1ST Workshop on Long Term Preservation of Earth Observation space data. ESA/ESRIN 27TH - 28TH May 2008. Workshop Report*. Noordwijk, The Netherlands, European Space Agency.

ESA (2009): *European Framework for the Long Term Preservation of Earth Observation Space Data*. Noordwijk, The Netherlands, European Space Agency.

ESA (2010a): *Long Term Data Preservation User Needs Workshop. (LTDP/FIRST) ESA/ESRIN 21ST October 2010. Workshop Report*. Noordwijk, The Netherlands, European Space Agency.

ESA (2010b): *Long Term Preservation of Earth Observation Space Data. European LTDP Common Guidelines*. Noordwijk, The Netherlands, European Space Agency.

ESA (2011): *Project: FIRST Requirements, Assessment & Proposals*. European Space Agency. Noordwijk, The Netherlands.

EUROSDR (2012): *GI+100: Long term preservation of digital Geographic Information - 16 fundamental principles*. Euro Spatial Data Research. [En Línea]. <http://www.eurocdr.net/archiving/Paper_EuroSDR%20Archiving%20Principles.pdf>. [21 de diciembre de 2011]

Fernández-Wytenbach, A. y Bernabé-Poveda, M. (2011): "La difusión del patrimonio: los servicios de mapas", *Revista 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. <<http://www.dlib.org/dlib/april02/friedlander/04friedlander.html>>. [21 de diciembre de 2011]. ISSN 1082-9873.

Ariza López, F.J., Ariza López, R.M., Ureña Cámara, M.A., Cortés José, J., Ureña López, L.A. (2012): "Preservación de la Información Geográfica: Perspectivas y situación en España", *GeoFocus (Artículos)*, nº 12, p. 171-200. ISSN: 1578-5157

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. y Banning, J. (2008): *An investigation into metadata for long-lived geospatial data formats*. Washington, DC. Library of Congress.

Janeé, G. (2009): "Digital Curation", en Liu, M. y Özsü, M.T. (Ed.): *Encyclopedia of Database Systems*, Springer, pp. 816-817.

Janeé, G. y Frew, J. (2008): "Relay-supporting Archives: Requirements and Progress", *International Journal of Digital Curation*, vol. 4, nº 1, pp. 57-70.

Kenney, A. R. (1996): *Digital to microfilm conversion: a demonstration project 1994-1996*. Cornell University Library. Ithaca, NY, USA, Cornell University.

Larsgaard, M.L. (1998): *Map Librarianship: an introduction*. Littleton, CO, USA, Libraries Unlimited, Inc.

Lazorchak, B. y 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.

Levy, D.M. (1998): Heroic measures: reflections on the possibility and purpose of digital preservation. *Third ACM conference on Digital libraries*. Pittsburgh, Pennsylvania, USA. ACM Press, pp. 152-161.

McGarva G., Morris S, y Janeé G. (2009): *Preserving Geospatial Data. Technology Watch Series report 09-01*. Heslington, NewYork, USA. Digital Preservation Coalition.

Morris, S. (2005): *Collection and Preservation of At-Risk Digital Geospatial Data NCSU Libraries and NCCGIA*. [En línea]. <www.lib.ncsu.edu/ncgdap/.../ndiipp_GICC.ppt>. [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.

Ariza López, F.J., Ariza López, R.M., Ureña Cámara, M.A., Cortés José, J., Ureña López, L.A. (2012): "Preservación de la Información Geográfica: Perspectivas y situación en España", *GeoFocus (Artículos)*, nº 12, p. 171-200. ISSN: 1578-5157

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. y Essic, J. (2009): "A Partnership Framework for Geospatial Data Preservation in North Carolina", *Library Trends*, vol. 57, nº 3.

NARA (2011): *Establishing Trustworthy Digital Repositories: A Discussion Guide Based on the ISO Open Archival Information System (OAIS) Standard Reference Model*. Washington DC, USA, Interagency Science Working Group, National Archives and Records Administration.

NGDA (2009): *National Geospatial Digital Archive Project Phase II Data Transfer Report*. Santa Barbara, CA, USA, University of California.

Nichols, H. (1982): *Map Librarianship*. London, UK, Clive Bingley.

PIRA (2000): *Commercial Exploitation of Europe's Public Sector Information*. Luxembourg. Pira International, Ltd., University of East Anglia, and Knowledge.

Romero-Tallafigo, M. (1994): *Archivística y archivos: soportes, edificio y organización*. Carmona, S&C ediciones, Asociación de Archiveros de Andalucía.

Ross, S. (2000): *Changing trains at the Wigan: Digital Preservation and the Future of Scholarship*. London, UK, National Preservation Office.

Santini, S. (2011): "Memoria y obsolescencia de los nuevos soportes tecnológicos", *Diagonal Web Periódico quincenal de actualidad crítica*. [En línea]. nº 146. <<http://www.diagonalperiodico.net/La-obsolescencia-precoz-pone-en.html>>. [21 de diciembre de 2011].

Shaon, A. y Woolf, A. (2008): "An OAIS Based Approach to Effective Long-term Digital Metadata Curation", *Computer and Information Science*, vol. 1, nº 2, pp. 2-12.

Shaon, A. y 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. <<http://www.dlib.org/dlib/september11/shaon/09shaon.html>>. [21 de diciembre de 2011]. ISSN: 1082-9873.

Sweetkind-Singer, J., Larsgaard, M.L. y Erwin, T. (2006): "Digital preservation of geospatial data", *Library Trends*, vol. 55, nº 2, pp. 304-314.

UC (2005): *Preserving the Nation's Endangered Digital Information. Profiles in Research. The 2005 Annual Report of the Office of Research*. Santa Barbara, CA, USA, University of California.

UKDA (2010): *UK Data Archive Preservation Policy*. Essex, UK Data Archive.

Waugh, A., Wilkinson, R., Hill, B. y Dell'Oro, J. (2000): "Preserving Digital Information Forever", *Fifth ACM Conference on Digital Libraries*, San Antonio, Texas, USA. ACM Press, pp.175-184.

Ariza López, F.J., Ariza López, R.M., Ureña Cámara, M.A., Cortés José, J., Ureña López, L.A. (2012): "Preservación de la Información Geográfica: Perspectivas y situación en España", *GeoFocus (Artículos)*, nº 12, p. 171-200. ISSN: 1578-5157

TABLAS

Tabla 1. Posibles soluciones para gestionar la obsolescencia

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 proceso continuo en el tiempo y por ello costoso. Se requieren metadatos administrativos.	Solución a corto y medio plazo.
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.

Fuente: Brown y col. (2005)

Ariza López, F.J., Ariza López, R.M., Ureña Cámara, M.A., Cortés José, J., Ureña López, L.A. (2012): "Preservación de la Información Geográfica: Perspectivas y situación en España", *GeoFocus (Artículos)*, nº 12, p. 171-200. ISSN: 1578-5157

Tabla 2. Ejemplo de metadatos de diversos conjuntos de datos geográficos según ISO 19115 (NEM), FGDC 1998

Recurso de IG	Metadatos relativos a preservación	Valores
Mapa Provincial de España 1/200.000 IGN. (ISO19115NEM) Fuente: http://www.idee.es	Mantenimiento del recurso(s). Frecuencia de mantenimiento y actualización:	Irregular
	Tipo de Representación espacial:	Vector
	Formato de distribución. Nombre:	DGN - Microstation format (Intergraph Corporation)
	Formato de distribución. Versión:	V.5
Mapa Topográfico Nacional 1:25.000 (digital-vector). IGN. (ISO19115NEM) Fuente: http://www.idee.es	Mantenimiento del recurso(s). Frecuencia de mantenimiento y actualización:	Continuamente
	Tipo de Representación espacial:	Vector
	Formato de distribución. Nombre:	DGN - Microstation format (Intergraph Corporation)
Mapa Topográfico de Andalucía 1:10.000 raster. IECA. (ISO19115) http://www.ideandalucia.es/catalogo/info.php?file=%7B086DEA60-3BF1-4DC6-8ACA-8D54BD20FACE%7D_100008_full2052_es	Mantenimiento del recurso: Frecuencia de mantenimiento y actualización:	Irregularmente
	Mantenimiento del recurso: Nota de mantenimiento:	La actualización de la serie se realiza normalmente cada cinco años. Para más detalles, consultar el estado en la página web o en el Departamento de Distribución del Instituto de Estadística y Cartografía de Andalucía.
	Descripción de imágenes: Tipo de contenido:	Imagen
	Formato de distribución: Nombre:	TIFF, MrSID
	Formato de distribución: Versión:	TIFF FAXCITTGrupoIV
DRG - 1:24,000 Topographic Map USGS. (FGDC 1998) http://www.sjrwmd.com/gisdevelopment/docs/metadata/drg2_4.htm	Status: Progress:	Complete
	Status:	None planned
	Maintenance_and_Update_Frequency:	
	Digital_Transfer_Information: Format_Name:	ARCE
	Digital_Transfer_Information: Format_Version_Number:	v8.2
	Digital_Transfer_Information: Format_Specification:	ARC/INFO Export Format - UNIX Compression
	Digital_Transfer_Information: File_Decompression_Technique:	Winzip

Fuente: Elaboración propia

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

Pregunta	Contestación	
	Si	No
P1: Disponen de cartoteca	61,5	38,5
P2: Realizan actividades de preservación de contenidos digitales generales	42,3	57,7
P3: Realizan actividades de curación de contenidos digitales generales	7,7	92,3
P4: Realizan actividades de preservación de información geográfica digital	19,2	80,8
P5: Realizan actividades de curación de información geográfica digital	7,7	92,3
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

FIGURAS

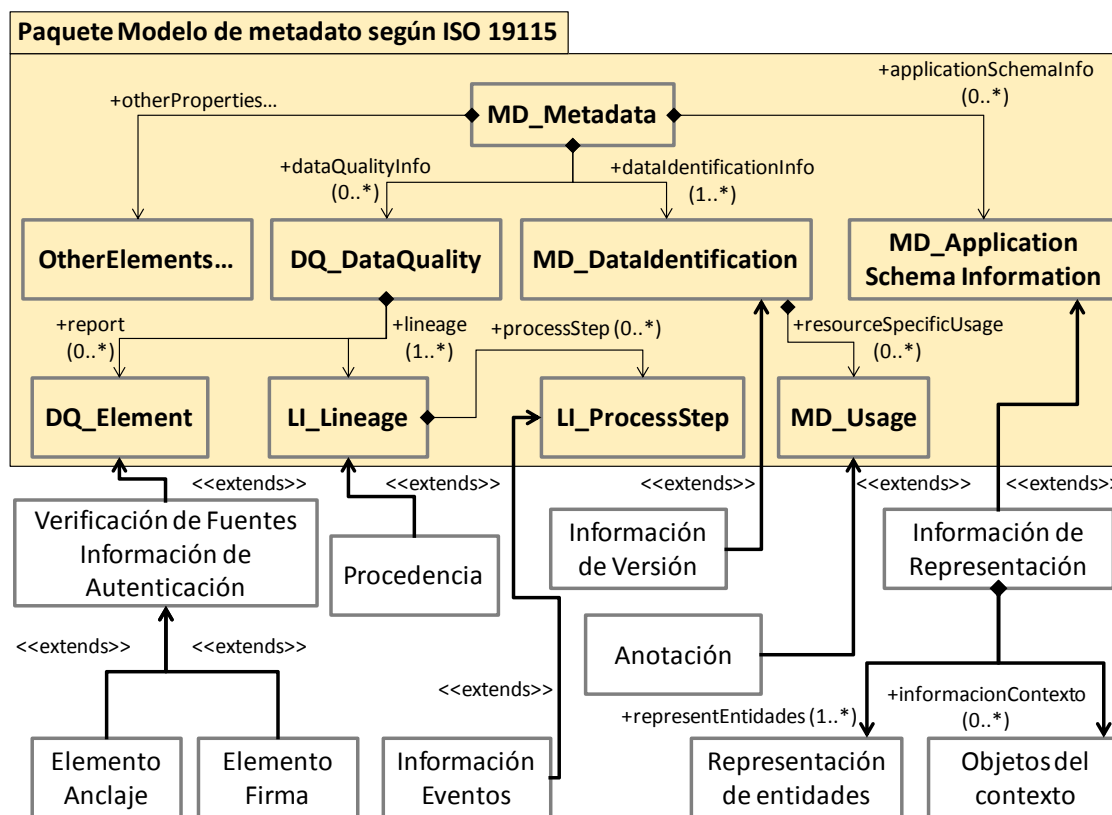


Figura 1. Perfil de preservación para ISO 19115 propuesto por Shaon y Woolf (2011).

¹ Cuanto antes conviene advertir al lector que los términos conservación, preservación y curación pueden generar confusión. Por ello, a continuación, se presenta una aproximación a sus definiciones y una justificación de la decisión que se ha tomado en este trabajo con respecto al uso de ellos. Se considera conservación el hecho de cuidar la permanencia de algo. La preservación supone proteger anticipadamente como consecuencia de una política de prevención, aspecto que puede no estar presente en el caso de la conservación. Según el diccionario de la Real Academia de la Lengua, entre otras acepciones, la curación (acción y efector de curar o curarse) hace referencia a la preparación de materias (p.e. pieles, carnes, pescados, etc.) para la conserva o su conservación. Según el *Digital Curation Center* (UK): "Digital curation involves maintaining, preserving and adding value to digital research data throughout its lifecycle". Es decir, la curación de un paso más sobre la preservación y se considera la existencia del ciclo de vida de los contenidos digitales, incluyendo, de cara al futuro, objetivos de uso y creación de valor añadido por ese contenido. A nuestro modo de ver, este aspecto no se evidencian en la preservación. Por tanto, la curación supone añadir a las actividades de preservación otras relacionadas con el acceso y generación de valor. Este cambio de perspectiva se está adoptando en algunas organizaciones prestigiosas, como por ejemplo la *California Digital Library* (*University of California*) que también ha creado un centro específico (*UC Curation Center*).