

Preservación Digital en España: Situación actual y perspectivas de futuro

*“La preservación del patrimonio digital: conceptos básicos y principales iniciativas”
Madrid, 16 de Marzo de 2006
Ministerio de Cultura*



Internet y los contenidos digitales

¿La nueva biblioteca de Babel?

No hay, en la vasta Biblioteca, dos libros idénticos.

... la versión de cada libro en todas las lenguas, las interpolaciones de cada libro en todos los libros, ...

La Biblioteca de Babel

Jorge Luis Borges

La carta de la Unesco

Preámbulo

La Conferencia General,

Considerando que **la desaparición de cualquier forma de patrimonio empobrece el acervo de todas las naciones,**

Recordando que la Constitución de la UNESCO establece que la Organización debe ayudar a la conservación, al progreso y a la difusión del saber, velando por la conservación y la protección del patrimonio universal de libros, obras de arte y monumentos de interés histórico o científico, que su Programa Información para Todos ofrece una plataforma para el debate y la acción sobre políticas de información y sobre la salvaguardia de los conocimientos conservados en forma documental, y que su programa "Memoria del Mundo" tiene por objeto garantizar la **preservación del patrimonio documental del mundo y un acceso universal al mismo,**

Reconociendo que esos recursos de información y expresión creativa se elaboran, distribuyen, utilizan y conservan **cada vez más en forma electrónica,** y que ello da lugar a un nuevo tipo de legado: **el patrimonio digital,**

Consciente de que el **acceso a dicho patrimonio** brindará mayores oportunidades de creación, comunicación e intercambio de conocimientos entre todos los pueblos,

Entendiendo que este patrimonio digital **se encuentra en peligro de desaparición,** y que su preservación en beneficio de las generaciones actuales y futuras es una preocupación urgente en el mundo entero,

Carta sobre la preservación del patrimonio digital
Unesco, 15 de octubre de 2003

Insta a las organizaciones internacionales, gubernamentales y no gubernamentales, y el sector privado, adopte todas las medidas necesarias para aplicar esta Carta.



La carta de la Unesco

http://portal.unesco.org/es/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html

Artículo 2 - Acceso al patrimonio digital

El objetivo de la conservación del patrimonio digital es que éste sea accesible para el público. Por consiguiente, el acceso a los elementos del patrimonio digital, especialmente los de dominio público, no debería estar sujeto a requisitos poco razonables. Al mismo tiempo, debería garantizarse la protección de la información delicada o de carácter privado contra cualquier forma de intrusión.

Los Estados Miembros tal vez deseen trabajar en colaboración con las organizaciones e instituciones pertinentes para propiciar un contexto jurídico y práctico que maximice la accesibilidad del patrimonio digital. Convendría reafirmar y promover un justo equilibrio entre los derechos legítimos de los creadores y otros derechohabientes y el interés del público por tener acceso a los elementos del patrimonio digital, de conformidad con las normas y los acuerdos internacionales.

Artículo 3 - El peligro de pérdida

El patrimonio digital del mundo corre el peligro de perderse para la posteridad. Contribuyen a ello, entre otros factores, la **rápida obsolescencia de los equipos y programas informáticos** que le dan vida, las incertidumbres existentes en torno a los recursos, la responsabilidad y los **métodos para su mantenimiento y conservación** y la falta de **legislación** que ampare estos procesos.

Los cambios en las conductas han ido a la zaga del progreso tecnológico. La evolución de la tecnología digital ha sido tan rápida y onerosa que los **gobiernos e instituciones no han podido elaborar estrategias de conservación oportunas y bien fundamentadas.** No se ha comprendido en toda su magnitud la amenaza que pesa sobre el potencial económico, social, intelectual y cultural que encierra el patrimonio, sobre el cual se edifica el porvenir.



Artículo 4 - Necesidad de pasar a la acción

A menos que se haga frente a los peligros actuales, **el patrimonio digital desaparecerá rápida e ineluctablemente**. El hecho de estimular la adopción de medidas jurídicas, económicas y técnicas para salvaguardar ese patrimonio redundará en beneficio de los propios Estados Miembros. Urge emprender actividades de divulgación y promoción, alertar a los responsables de formular políticas y sensibilizar al gran público tanto sobre el potencial de los productos digitales como sobre los problemas prácticos que plantea su preservación.

Artículo 6 - Elaborar estrategias y políticas

Es preciso elaborar **estrategias y políticas** encaminadas a preservar el patrimonio digital, **que tengan en cuenta el grado de urgencia, las circunstancias locales, los medios disponibles y las previsiones de futuro**. La colaboración de los titulares de derechos de autor y derechos conexos y otras partes interesadas a la hora de definir formatos y compatibilidades comunes, así como el aprovechamiento compartido de recursos, pueden facilitar esa labor.

Una buena colección digital ...

- ❑ debe **mantenerse en el tiempo**. En particular las colecciones digitales deben disponer de un plan que les permita durar y mantenerse en el tiempo, más allá del proyecto que las originó

Un objeto digital ...

- ❑ es **persistente**. Esto es, que será accesible a lo largo del tiempo a pesar de que haya cambios tecnológicos
- ❑ se digitaliza en un **formato que soporta el uso actual y futuro** o que soporta el proceso de crear copias que faciliten el uso. Un buen objeto se puede intercambiar entre plataformas, es ampliamente accesible y se digitaliza de acuerdo a **las mejores prácticas disponibles**

Los metadatos ...

- ❑ deberían ser los adecuados para **describir** los materiales de la colección, usuarios de la colección y los usos actuales y futuros del objeto digital.
- ❑ Los buenos metadatos **soportan la gestión a largo plazo** de los objetos en las colecciones

Estándares utilizados

CATALOGACIÓN GEOGRÁFICA	CATALOGACIÓN TÉCNICA	CATALOGACIÓN DESCRIPTIVA	CATALOGACIÓN DERECHOS
<ul style="list-style-type: none"> Selección de campos de GeoXML 	<ul style="list-style-type: none"> Activo Físico: Selección de campos de Standard NISO/METS (IMD,AMD...) Activo Digital: Selección de campos de Standard NISO/METS (IMD,AMD...) Proceso: Activo Digital: Selección de campos de Standard NISO/METS (IMD,AMD...) 	<ul style="list-style-type: none"> Base Dublin Core y Dublin Core Qualified Estructura flexible en función del tipo de objeto (digital o agrupación) 	<ul style="list-style-type: none"> Acuerdos de Cesión: Propietario Acuerdos de Distribución: Propietario Expresión XML: A definir(XRML o ODRL)
PROGRAMACIÓN	ALMACENAJE ACTIVOS DIGITALES	ESTRUCTURA PLATAFORMA	WEB/SINDICACIÓN
<ul style="list-style-type: none"> J2EE Webservices Open Source 	<ul style="list-style-type: none"> Imagen 2D: TIFF Audio: PCM Texto: ASCII Video: MPEG-2 Infografía: STL Mapas: Por decidir 360º: QTVR, IPIX Imagen Vectorial: EPS 	<ul style="list-style-type: none"> OAIS Fedora Minerva 	<ul style="list-style-type: none"> Web WAI-2 Minerva Sindicación Propietario Salida XML configurable



Cumplimiento del modelo OAIS

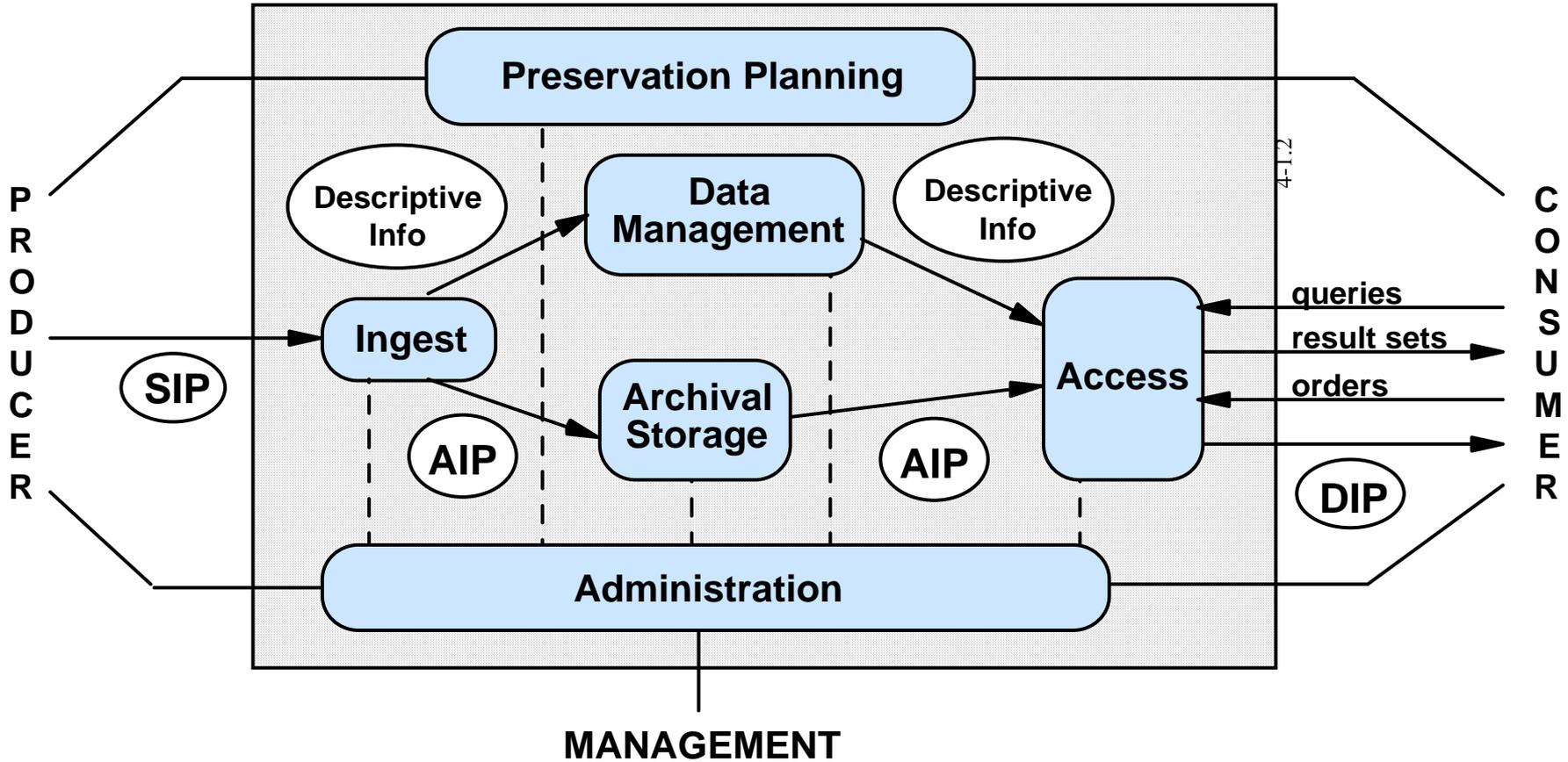
<http://www.rlg.org/longterm/oais.html>

Acorde al estándar METS

- Abierto
 - El modelo de referencia estándar OAIS se desarrolla mediante un proceso público y se consigue de manera libre
 - Cualquier tipo de conocimiento puede ser intercambiado
- Independiente del formato (p.e. físico o digital) que se utilice para representar la información
 - Los datos son las formas de representar la información
- Sistema de archivado de la información
 - Involucra al hardware, software, y la gente que es responsable de la adquisición, preservación y disseminación de la información



OAIS/Mets

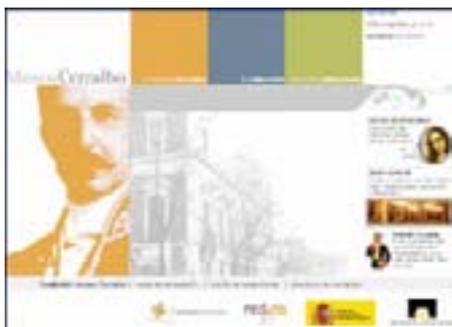


SIP = Submission Information Package
 AIP = Archival Information Package
 DIP = Dissemination Information Package



Proyectos Piloto de Digitalización

Museo Cerralbo



- Digitalización y documentación de selección de salas
- Puesta en valor de contenidos a través de sitio web interactivo del museo

Filmoteca Española



- Digitalización de selección de películas anteriores a 1953 y materiales asociados.
- Masters digitales y derivados de materiales filmicos y no filmicos

Biblioteca Nacional



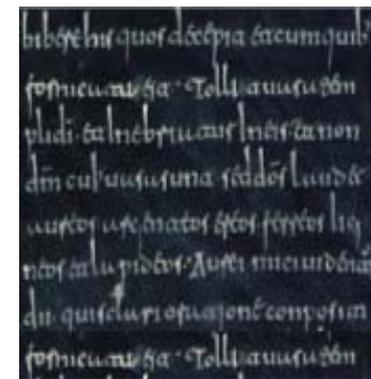
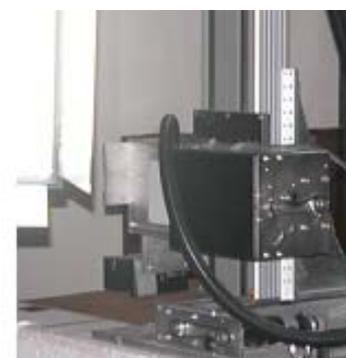
- Replica completa facsímil digital de alta calidad del Beato Emilianense
- Pruebas de impresión de alta calidad en diferentes soportes

Primeras actuaciones

- **Tesoros de Biblioteca Nacional**

- Una **réplica facsímil** digital de alta calidad completa del Beato Emilianense perteneciente a la Biblioteca Nacional

- Una **digitalización de alta calidad** y pruebas de **impresión de alta calidad** con el propósito de poder obtener un facsímil cuasi-original usando **tecnologías de última generación**.



Primeras actuaciones

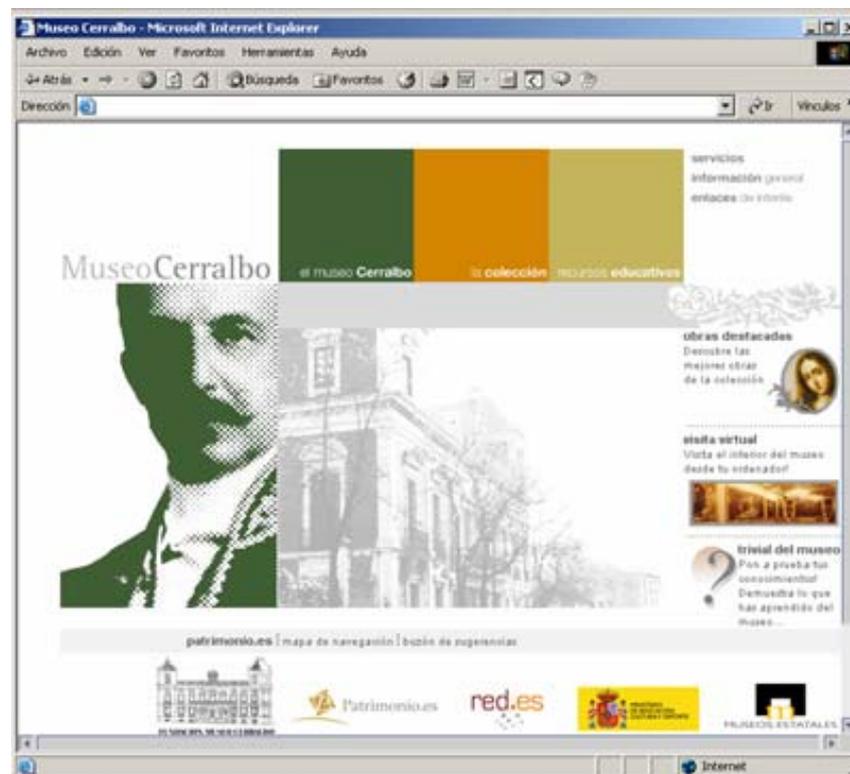
- **Historia del cine español con Fimoteca Española**

El proyecto de digitalización y creación de “masters” de alta calidad de una selección de materiales fílmicos anteriores a 1953



Primeras actuaciones

- **Descubrir Museos: Museo Cerralbo**
- La **digitalización y documentación** de una selección de piezas y salas del Museo, en la que se obtendrán activos digitales de **alta calidad**.
- Transformación y optimización de los masters a formatos de difusión (Internet) para ser integrados con **contenidos multimedia** (audio, texto, etc.) y vinculados de modo **interactivo** a las **visitas virtuales** que se realicen de las salas del Museo.



Prioridades

- Digitalización
- Difusión
- Colaboración con instituciones

Estado Actual

- Convocatoria de proyectos

¿Por qué preservar?

- ↘ Pérdida de información valiosa
 - ↘ Obsolescencia tecnológica
 - ↘ Gestión de grandes conjuntos de datos
- ↘ Preservación del patrimonio digital
- ↘ Reutilización de la Información

¿Quién archiva? / Pocos sistemas de preservación

↘ Las propias instituciones / empresas

- ↘ Control de contenidos (gestores de contenidos)

La mayoría no contemplan preservación

↘ Administraciones nacionales y locales

- ↘ Documentos administrativos

La preservación es un punto importante

- ↘ Patrimonio cultural, lingüístico

Casuística variada

↘ Iniciativas globales

- ↘ Buscadores (google)

- ↘ Archivado de Web (The WayBack Machine)

Preservación, pero problemas legales

Algunos números

Table 1.2: Worldwide production of original information, if stored digitally, in terabytes circa 2002. Upper estimates assume information is digitally scanned, lower estimates assume digital content has been compressed.

Storage Medium	2002 Terabytes Upper Estimate	2002 Terabytes Lower Estimate	1999- 2000 Upper Estimate	1999- 2000 Lower Estimate	% Change Upper Estimates
Paper	1,634	327	1,200	240	36%
Film	420,254	76,69	431,690	58,209	-3%
Magnetic	4,999,230	3,416,230	2,779,760	2,073,760	80%
Optical	103	51	81	29	28%
TOTAL:	5,421,221	3,416,281	3,212,731	2,132,238	69%

Source: *How much information 2003*

2002: 5 exabytes de nueva información

Lyman, Peter and Hal R. Varian, "How Much Information", 2003.

eley.edu/research/projects/how-much-info-2003/

Gigabyte (GB) <i>10⁹ bytes</i>	1 Gigabyte: Una furgoneta llena de libros 20 Gigabytes: Una buena colección de las obras de Beethoven. 100 Gigabytes: Una planta de una hemeroteca científico-académica
Terabyte (TB) <i>10¹² bytes</i>	1 Terabyte: 50000 árboles convertidos en papel e impresos. 2 Terabytes: Una biblioteca de investigación académica. 10 Terabytes: La colección impresa de la <i>U.S. Library of Congress</i> . 400 Terabytes: Base de Datos del <i>National Climactic Data Center (NOAA)</i>
Petabyte (PB) <i>10¹⁵ bytes</i>	1 Petabyte: 3 años de datos del <i>EOS (Earth Observing System) (2001)</i> . 2 Petabytes: Todas las bibliotecas de EEUU de investigación académica. 20 Petabytes: La producción completa de discos duros en 1995. 200 Petabytes: Todo el material impreso.
Exabyte (EB) <i>10¹⁸ bytes</i>	2 Exabytes: Volument total de informacion generada en 1999. 5 Exabytes: Todas las palabras alguna vez dichas por los seres humanos.



El Tamaño de la Web (bytes)

Lyman, Peter and Hal R. Varian,
"How Much Information", 2003.

↘ *Web estática*

2000	20 – 50 terabytes
2003	167 terabytes

↘ *Web dinámica*

Estimada en 400-450 veces la web estática (BrightPlanet)

70.000 – 90.000 terabytes

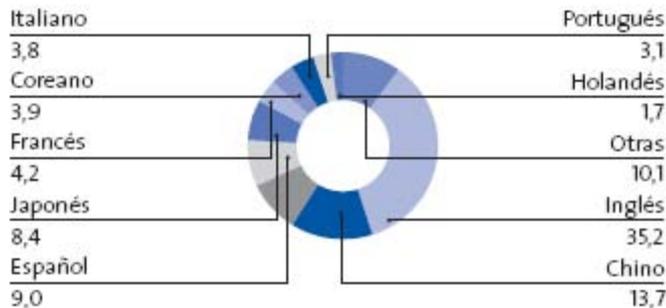
Crecimiento de la Web: 30 % anual

El Tamaño de la Web (páginas)

- ↘ 4.000 millones de páginas
- ↘ 7 millones de páginas nuevas por día
- ↘ 220 idiomas
- ↘ 95 % públicamente accesible
- ↘ Vida media de una página es de 44 días

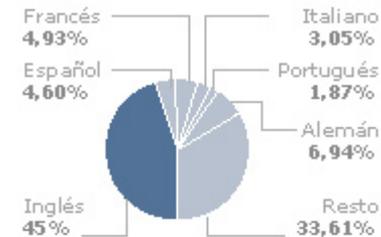
Algunos números

Figura 2-1: DISTRIBUCIÓN DE INTERNAUTAS POR LENGUA MATERNA
(% de internautas)



Fuente: Global Reach, Global Internet Statistics. Septiembre de 2004.

Contenidos en la red
por idiomas



Fuente: Funredes

- 70 millones de internautas hispanohablantes
(9% total, tercer puesto, tras el inglés y el chino)
- Contenidos en español: 4,93%
por detrás del inglés (45%), Alemán (6,94%) y Francés (4,93%)
- El crecimiento de contenidos en español en Internet presenta un “estancamiento” respecto a otros idiomas europeos, como el francés o el alemán.
- Los hispanohablantes: 400 millones de personas.

- ↘ La prioridad en materia de SI es reducir la brecha digital
- ↘ La prioridad del patrimonio cultural digital es digitalizar los fondos en otros soportes
- ↘ Los recursos económicos
- ↘ Los problemas legales
- ↘ El problema es urgente, las posibles soluciones aparentemente no

Conclusiones

- ↘ La información es ya mayoritariamente digital
- ↘ Cualquier previsión, aún la más conservadora, indica que seguirá creciendo a ritmos mayores que los actuales, especialmente los contenidos en lenguas españolas
- ↘ Los nuevos modos de transmisión de información obligan a nuevos modos de gestión.
- ↘ El problema no afecta sólo a bibliotecas y archivos
- ↘ Actualmente, este asunto no se encuentra en las listas de prioridades en materia de SI
- ↘ La colaboración entre sectores y organismos es fundamental
- ↘ El uso y fomento de estándares es vital en este proceso
- ↘ Hay que tener en cuenta los “nuevos” problemas legales que aparecen en este ámbito.

Gracias por su Atención

Ana Nistal
ana.nistal @ red.es