

# nº 47

2007

## Con Firma

Ciro Lluca Fonollosa  
*Archivando la web catalana, el proyecto PADICAT*

## Nombres Propios

Nombramientos  
 Nuevos socios

## Personajes

Entrevista:

Margarita Taladriz Mas  
 Directora del Servicio de Biblioteca. Universidad Carlos III de Madrid

## SEDIC Abierto

II Jornada sobre revistas científicas electrónicas españolas en acceso abierto

.....  
 1ª Jornada Técnica de Bibliotecas de Ciencias de la Salud en el siglo XXI

.....  
 Situación laboral y desarrollo profesional de los socios de SEDIC

.....  
 Asociaciones iberoamericanas en información y documentación

.....  
 Formación continua: resultados 2006

.....  
 El blog de SEDIC cumple un año de vida

Con firma →

## Archivando la web catalana, el proyecto PADICAT

**Ciro Lluca Fonollosa**

Coordinador del proyecto PADICAT  
 (Patrimonio Digital de Cataluña)  
 Biblioteca de Catalunya  
<http://www.padicat.cat>



Como expone la UNESCO en sus [Directrices para la preservación del patrimonio digital](#), los recursos que son fruto del conocimiento o la expresión de los seres humanos, ya sean de carácter cultural, educativo, científico o administrativo, o comprendan información técnica, jurídica, médica o de otro tipo, se generan cada vez más a menudo directamente en formato digital, o se convierten a este formato a partir de material analógico ya existente.

Ello porque las tecnologías de la información y la comunicación han propiciado un crecimiento exponencial de la información de todo tipo y temática publicada en formato digital, en Internet. Pero por la propia naturaleza dinámica de la Red, gran parte de esa información muta a diario: es sustituida o simplemente desaparece. Es decir, la información producida en formato digital, contenida en las páginas web, corre el riesgo de colársenos por las manos como arena fina, si no se instrumentalizan proyectos de preservación y conservación. El reto y también el objetivo es construir repositorios que permitan acceder de manera permanente a los contenidos publicados en Internet.

A tal efecto y desde 1996, administraciones de diversos países han iniciado esos proyectos, llamados mayoritariamente "archivos web", con la meta de garantizar el libre acceso a lo que se considera el patrimonio digital nacional. Suecia ([proyecto Kulturarw3](#)), Australia ([proyecto Pandora](#)), y el ambicioso [proyecto Internet Archive](#), de alcance internacional, fueron los pioneros, y siguieron con el tiempo otros países como Gran Bretaña, Dinamarca, o Japón. En España, la [Biblioteca de Catalunya](#) creó el proyecto PADICAT ([Patrimonio Digital de Cataluña](#)), que ha afirmado su presencia internacional gracias a su reciente inclusión en el [International Internet Preservation Consortium \(IIPC\)](#).

La Biblioteca de Catalunya, cuya misión es recopilar, conservar y difundir la producción bibliográfica catalana, considera patrimonio digital nacional toda aquella publicación en formato digital orientada, en el sentido más amplio, al público de Cataluña. La estrategia se concreta en las páginas web publicadas en lengua catalana u otras lenguas, bajo el dominio .CAT y otros dominios geográficos o temáticos, que estén relacionadas temáticamente con Cataluña.

**SEDIC Joven**

Entrevista: Eva Luque  
 Consultora y directora de  
 operaciones en DOC6,  
 oficina Madrid

**Centros en Marcha**

Biblioteca de Albasanz del  
 CSIC

.....

Centro de Documentación  
 ICO

**Créditos**

ISSN: 1137-0904

sedic@sedic.es

Así nació en junio del año 2005 el proyecto PADICAT, con un presupuesto aproximado de un millón de euros y el objetivo de crear la bibliografía digital catalana, el archivo web de Cataluña, siendo una iniciativa pionera en las comunidades de habla hispana, y con vocación de impulso en la futura e ineludible realización del patrimonio digital del resto de bibliotecas españolas. Siguiendo la pauta generalizada en la mayoría de bibliotecas nacionales que se han sumado a este tipo de proyectos, se optó por la vía híbrida de captura de recursos web. Eso es, en primer lugar, una captura exhaustiva y automatizada de recursos digitales publicados en Internet realizada por el programa informático Heritrix. En segundo lugar, los acuerdos con una amplia selección de los sitios web representativos del entramado que conforma la sociedad civil catalana, tales como empresas, asociaciones profesionales, culturales o deportivas, partidos políticos y sindicatos, universidades, ayuntamientos, etc. En tercer lugar, la promoción de determinadas líneas de investigación futura mediante la integración focalizada de los recursos digitales de determinados acontecimientos, como ha sido la última campaña electoral en Internet para las elecciones al Parlamento de Cataluña, o las elecciones municipales que se celebran en España a finales de mayo de 2007.

Centrándonos en el caso de la campaña electoral de las elecciones al Parlamento de Cataluña para ilustrar la utilidad del proyecto, la BC seleccionó una serie de recursos con presencia en la campaña (todos los partidos que se presentaban a alguna de las circunscripciones catalanas, independientemente de si tenían ya o no representación parlamentaria, los blogs de los principales candidatos, diversas páginas web de fundaciones políticas, web del Parlamento catalán y otras administraciones implicadas en el proceso electoral, la web del President Maragall, y una voluminosa selección de blogs personales, de variopinta temática y espectro político. Tras 40 días dedicados a capturar hasta 652 versiones de 83 recursos digitales, se procedió a publicar los resultados en el portal del proyecto, y darlo a conocer a la comunidad docente universitaria (unos 200 docentes de ciencia política, sociología y comunicación) para asegurar el uso de esos materiales en la actividad docente con interés en análisis de campañas políticas. La acogida, a falta de un proceso de evaluación sobre ese uso, ha sido satisfactoria.

Tras dos años de PADICAT, las anteriores acciones han producido un resultado que permite a la BC ofrecer en abierto, en línea, un total de unos 2.000 recursos digitales capturados en base a la captura sistemática y por los acuerdos con prácticamente 200 instituciones, así como la campaña electoral en Internet en las pasadas elecciones al Parlamento de Cataluña, y lo más importante, la necesaria preparación para encarar con garantías el acceso permanente a esa parte de la Web catalana. En el lado mejorable de la iniciativa, la necesaria gestión personalizada de los acuerdos de cooperación con las instituciones ralentiza camino hacia el objetivo 2009, que pretende abarcar el millar de organizaciones, y el desarrollo del programario dedicado a la captura y gestión de los recursos digitales publicados en Internet necesita aún de mucha dedicación técnica para optimizar sus resultados.

En todo caso, los beneficios de un proyecto como el PADICAT llegan a todos los sectores de la sociedad: para la ciudadanía representa el acceso abierto y permanente a los recursos que son fruto del conocimiento y expresión de los creadores del siglo XXI en un territorio determinado. Para las instituciones, empresas, administraciones y particulares que producen páginas web en ese territorio, la preservación de la propia producción y garantía de acceso, con los condicionantes que

en cada caso la ley regula, a los contenidos y diseños que, de otro modo, desaparecerían. Y para el sistema bibliotecario, posibilidades infinitas de cooperación con el resto de bibliotecas, archivos y museos.



---

#### **REFERENCIAS**

Bibliografía sobre el Patrimonio Digital de Cataluña y otros proyectos de Archivo de la web en el área profesional del PADICAT:

<http://www.padicat.cat/es/areaprof.php>

