

E. REDES SOCIALES Y WEB 2.0

Informe de situación

Evolución de la Web social en 2010

Por **Fernando Juárez-Urquijo**

Juárez-Urquijo, Fernando. "Evolución de la Web social en 2010".
Anuario ThinkEPI, 2011, v. 5, pp. 146-154.



Resumen: *Repaso al estado de la Web en 2010, analizando la situación de las tecnologías 2.0 y su evolución para adaptarse a las nuevas condiciones de generación, acceso y difusión de la información. Teniendo en cuenta la realidad social española se realiza una descripción de los cambios más importantes sucedidos durante el año 2010 y se hacen algunas reflexiones sobre cómo pueden influir en el desarrollo de los servicios bibliotecarios. Finalmente se aventura una breve prospección de lo que puede suceder en 2011.*

Palabras clave: *Redes sociales, Web 2.0, Web social, Bibliotecas, Biblioteca 2.0.*

Title: *Evolution of the Social Web in 2010*

Abstract: *Short summary of the Web situation in 2010, analysing the situation of 2.0 technologies and their evolution to adapt to the new conditions of creation of, access to and dissemination of information. Taking into account the Spanish society, the most important changes in 2010 and how they can influence the services that libraries offer are described. Finally, some predictions about 2011 are envisaged.*

Keywords: *Social networks, Web 2.0, Social Web, Libraries, Library 2.0.*

La web en 2010: una visión rápida

UN VISTAZO RÁPIDO A LAS CIFRAS DE 2010 es suficiente para constatar que el uso de internet ha eclosionado de la mano de los "sitios de redes sociales"; estas plataformas han creado un ecosistema en el que el usuario ordinario encuentra casi toda la información que necesita.

El éxito de la propuesta viene refrendado por los 600 millones de usuarios que alimentan el universo *Facebook*. Muchos usuarios identifican internet con "su facebook" y priorizan la opinión de otras personas para tomar decisiones desplazando a *Google* como punto de acceso en sus búsquedas.

En España continúa la tendencia de años anteriores: crece el número de internautas (el 62,4 % españoles es internauta con una penetración del 94% entre los menores de 24 años) y el uso de las redes sociales. Según el informe *La sociedad de la información en España*, en 2008 el número de internautas españoles que declaraban usar internet para conectarse a una red social era del

9,5%¹. En el informe de 2009² la cantidad es del 28,7% y en 2010³ es ya un 61% (78% entre los menores de 24 años).

"Los usuarios prefieren el acceso al almacenamiento, cosa que potenciará más la nube (cloud computing)"

Sigue el domino de *Facebook* y *Tuenti*. En España *Facebook* pasó de los 7 millones de usuarios en 2009 a los 12 en 2010⁴; parte del éxito se debe a su integración con los principales webs españolas: blogs, medios de comunicación, páginas de servicios, etc. Los plug-ins sociales y, en especial la integración del botón "me gusta", han hecho de esta red social el escaparate perfecto para promocionar empresas y organizaciones. Las bibliotecas siguen abriendo sus perfiles.

Para 2011 uno de los acontecimientos que pueden ser muy relevantes es el lanzamiento en

abierto de *Diaspora*⁵ [sin acento en el original inglés] una red social descentralizada donde nuestros datos los manejamos nosotros mismos, una red social *peer to peer* que tal vez pueda ofrecer un modelo alternativo a *Facebook*.



Figura 1. Maxwell Salzberg, explica como funcionará *Diaspora* (vídeo)⁶

Pero más allá del crecimiento global de las redes sociales, la web en 2010 se ha caracterizado por primar empresas y servicios que han comprendido y asimilado los conceptos que han hecho avanzar el universo internet; la actual aplicación 2.0 es social, colaborativa, gratuita, móvil, multiplataforma, y global. Herramientas “significativas” del momento 2.0 que no han sabido evolucionar adaptándose a los principios de simplicidad, instantaneidad y ubicuidad han desaparecido durante este ejercicio o están perdiendo el favor del público.

Los usuarios prefieren el acceso al almacenamiento. Los cambios de hábito de consumo de información y ocio (desplazando ya a la tv) potenciarán más la nube (cloud computing); los servicios móviles facilitados por el avance de la banda ancha y la aparición de nuevos dispositivos serán clave en este proceso.

“En España un 44% confiesa necesitar internet para su vida cotidiana; un 86% manifiesta que internet mejora el aprendizaje y un 83% de niños la utilizan para su formación”

La red condiciona comportamientos y modos de vida: compra, salud, administración..., el ocio está en la Red y es multimedia. En España un 44% confiesa necesitar internet para su vida cotidiana; un 86% de los españoles manifiesta que internet mejora el aprendizaje y un 83% de los niños la utilizan para su formación.

Importancia del usuario: una web con nuevos patrones

Tras la debacle “punto com” de principio de siglo surgió el concepto “2.0” que no respondía a una nueva tecnología sino a comportamientos de uso. Aprendimos que la eficacia de una herramienta se basa en el uso que hagan de ella los usuarios, no en la tecnología.

En el último tercio de 2010 se sucedieron las noticias sobre la desaparición de servicios y herramientas emblemáticas de la web 2.0: *Bloglines* cerró, *Movable Type* fue vendido, *Delicious* amenazado, *Flickr* perdió terreno frente al todopoderoso *Facebook*, *Ning* dejó de ser gratuito..., y eso a pesar de que los datos de uso⁷ indican que se suben más fotos que nunca, que el número de blogs se incrementa y que los rss siguen facilitando la remezcla.

¿Cómo explicar la desaparición de empresas que lideraron tales servicios? Sencillamente, por el mismo motivo que quebraron muchas de las puntocom: la falta de adaptación. La web actual exige ser instantáneo, omnipresente y, a poder ser, simple. La ubicuidad se consigue reduciendo el tamaño de los dispositivos y adaptando el modo en el que interactuamos con la tecnología. Se priman los dispositivos portables, las interfaces sencillas y el poco tiempo para poder comunicar.

Quien ha demostrado su buena capacidad de adaptación es *Wikipedia* que cerró 2010 con una exitosa campaña de captación de fondos y celebró con muy buena salud su décimo cumpleaños. Esta enciclopedia colaborativa, que para muchos internautas representa los ideales de la cultura libre y abierta, es un ejemplo por su capacidad de respuesta en tiempo real ante las noticias de alcance⁸.

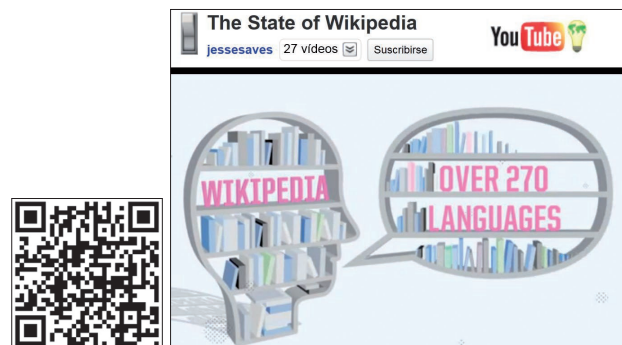


Figura 2. *The state of Wikipedia* (vídeo)⁹

Del ecosistema blog al del microblog: la web instantánea, ubicua e impaciente

En 2004-2005 quien buscaba generar, difundir y recuperar información necesitaba una plata-

forma de blogs, un buscador para encontrarlos y un buen lector de rss para poder consumir el flujo informativo. *Movable Type*, *Bloglines* y *Technorati* eran sitios de uso obligado. En 2010 se ha cerrado el ciclo: *Bloglines* ha desaparecido, *Movable Type* ha sido vendida y *Technorati* es una web irrelevante desde hace años para el posicionamiento web. El paradigma blog, un ecosistema distribuido muy enfocado a crear publicaciones y a consumir información y contenidos¹⁰ cede definitivamente a nuevos patrones de lectura y de consumo de los medios; no se consume todo lo publicado por tal medio (directamente o mediante lector de rss) sino que vamos a un lugar donde nuestros “amigos” nos dicen y recomiendan qué leer. Y queremos que nos lo digan “ya”.

En 2010 se incrementó el número de blogs hasta los 152 millones (*Blogger* y *Word Press* son los servicios de creación de blogs más utilizados, puestos 6 y 20 en *Alexa*) pero su consumo se hace a través de las redes sociales (frente a los 152 millones de blogs había 175 millones de usuarios de *Twitter* y 600 en *Facebook*).

El microblogging se adapta mejor a las nuevas necesidades: es un sistema muy sencillo (lo cual reduce al máximo las barreras de entrada), permite contar cosas en el menor tiempo posible, obtener respuesta inmediata y funciona bien en los nuevos dispositivos móviles. Destacan *Tumblr*, *Posterous* y, sobre todo, *Twitter*, que con 175 millones de usuarios (unos 370.000 nuevos cada día) y más de 1.000 tweets por segundo es el servicio de microblogging de mayor éxito. Ha conseguido ser identificado con el concepto de “web en tiempo real”.

“El microblogging es sencillo, permite contar cosas en el menor tiempo posible, obtener respuesta inmediata y se adapta muy bien a los dispositivos móviles”

Autosuficiencia, sencillez y especialización

La mayoría de las buenas noticias en 2010 las protagonizaron las grandes compañías y sus cifras: aumento del número de usuarios, compra de startups de éxito¹¹ e implementación en sus ecosistemas de servicios novedosos. Son movimientos que denotan el intento de crear una red propia, una intranet, en la que los usuarios encuentren todo lo que necesitan: correo electrónico, intercambio de fotos, geoposicionamiento, almacenamiento,

acceso ubicuo... Los movimientos en España son más modestos y han estado marcados por las compras realizadas por *Amazon* (*BuyVIP*, un “club de compras”) y *Telefónica* (*Tuenti*)¹²

El desmesurado crecimiento de las redes sociales que fagocitan todos los servicios de éxito obliga a muchas empresas a buscar su nicho de mercado en la especialización. Podemos hablar de especialización de supervivencia (*Flickr*), de nuevos servicios (*Foursquare*, *Quora*), por grupos de interés (*Academia.edu*), o de “reposicionamiento” como el de *MySpace* como especialista en música...

Sin embargo algunas aplicaciones y servicios especializados han tenido problemas y desaparecerán a corto plazo. La herramienta de etiquetado social *Delicious* es un ejemplo de aplicación que no ha triunfado porque los usuarios no la utilizan masivamente; a pesar de ocupar un envidiable puesto 247 en *Alexa*, sólo goza de aceptación entre usuarios avanzados y no entre la mayoría de internautas porque es complicada de usar; la noticia de que *Yahoo* (su propietaria) planeaba venderla encendió la alarma y puso de nuevo sobre el tapete el peligro de la fugacidad de algunos negocios en internet.

“La implantación de la ‘biblioteca 2.0’ tuvo que superar barreras de cultura y hábito, tanto en la sociedad a la que servía como entre muchos de sus profesionales”

La masiva entrada de nuevos usuarios en la Red de la mano de *Facebook* y la proliferación de móviles con cámara incorporada afecta a una de las históricas, *Flickr*¹³ (34 en *Alexa*), que pierde “peso” en el negocio de las imágenes. Frente al “todo en uno” de la red social, que seduce a la mayoría que quiere mostrar y compartir fotos entre las amistades, y lo cómodo que resulta utilizar una aplicación desde el móvil (cómo *Instagram*) el futuro de *Flickr* pasa por fidelizar al amante de la fotografía que está dispuesto a invertir más tiempo en una aplicación específica. En este caso la falta de “instantaneidad” no conlleva desaparición pero sí reconversión.

Demostrado el éxito de los sistemas de redes sociales generalistas se constata la aparición de nuevas plataformas que huyen de la masificación porque necesitan la relevancia y la pertinencia de una comunidad de intereses determinada. En grupos de edad superiores van ganando peso las relaciones profesionales: junto a *LinkedIn* (la red

social profesional mayoritaria) aparecen otras como *Academia.edu*, *ResearchGate*, *Nature Network* y *Epernicus* dedicadas a educación, biología, medicina...

<http://www.linkedin.com>

<http://network.nature.com>

<http://www.epernicus.com>

En 2011, dentro de esa tendencia a la especialización, parece que pueden eclosionar servicios que ofrezcan un acceso de más calidad a los contenidos que demandan los usuarios avanzados. La intermediación de contenidos es la apuesta de *Quora*, una red que reúne intereses profesionales (al igual que las más generalistas *LinkedIn* y *Twitter*) y pretende atraer usuarios que necesiten encontrar información generada y filtrada por personas cuyas ideas, conocimientos y opinión valoran. En España su difusión se ha encauzado sobre todo a través de *Twitter* (donde el 75% de sus usuarios lo son por motivos profesionales¹⁴).

Geoposicionamiento

El modelo de servicio emergente en 2010 entre los internautas avanzados fue el geoposicionamiento, gracias al despliegue de redes 3G, banda ancha y la proliferación de dispositivos móviles. A pesar del deseo de los grandes (sobre todo *Facebook Places*) por copar el incipiente mercado de la búsqueda por localización la herramienta estrella ha sido *Foursquare*¹⁵ (6.000.000 usuarios, 862 en *Alexa*).

Foursquare no es la primera compañía en intentar que el usuario comparta su localización, aportando información sobre la misma (contar a los contactos si se está en un restaurante o una biblioteca, dando opinión sobre ellos) pero sí la primera que ha sabido aunar varios de los elementos clave de la web actual: integración con *Twitter* y *Facebook*, aplicación para móviles *Android iPhone* y un juego para incentivar la participación. Su apuesta le sitúa en un buen lugar para liderar el mercado de la publicidad por proximidad.

La nube y sus consecuencias: almacenamiento distribuido y nuevos modelos de negocio

La disparidad de dispositivos al alcance de los usuarios, la movilidad y el cloud computing hacen necesario el almacenamiento y la sincronización de la información de una manera instantánea y ubicua. Tanto los usuarios como las organizaciones que utilizan este tipo de servicios deben tener una estrategia de almacenamiento y recuperación de la información. Una de las webs más valorados

en 2010 por internautas españoles (superando a los de música o lectura) ha sido *Dropbox*¹⁶, un servicio de alojamiento de archivos multiplataforma en la nube. *Dropbox* permite a los usuarios almacenar y sincronizar archivos en línea y entre dispositivos, y compartir archivos y carpetas con otros. Existen versiones gratuitas y de pago, cada una de las cuales con opciones variadas. En *Wikipedia* hay una tabla comparativa:

http://es.wikipedia.org/wiki/Servicio_de_alojamiento_de_archivos

En bibliotecas y centros de documentación una necesidad a corto y medio plazo es aprender a gestionar nuestras propias nubes, y servicios como *Dropbox* pueden ser un buen punto para empezar a familiarizarnos.

“La Web que se intuye moldeará las nuevas estructuras bibliotecarias: hará replantear el diseño arquitectónico... y condicionará la aparición de profesionales constantemente conectados”

Spotify es un servicio de música online por streaming [escucha mientras se está conectado] con ciertos componentes sociales, que es un gran ejemplo de modelo de negocio alternativo gracias al cloud computing. La industria musical fue una de las primeras en evolucionar al ritmo que lo hacía la Red. Tras la pérdida de importancia de los soportes y la consiguiente reconversión de la distribución, desaparición de puntos de venta..., se enfrenta ahora a la transición desde modelos de pago por descarga hacia un modelo de escucha en streaming –gratis con publicidad o libre de ella mediante pago–.



Figura 3. Daniel Ek y Martin Lorentzon, cofundadores de Spotify

Fuente: <http://www.wired.com/epicenter/2009/08/spotify-50-million-closer-to-us-launch/>

El mundo de la edición en 2011 estará marcado por un nuevo modelo de distribución de contenidos. El camino abierto en la distribución de música que propone *Spotify* puede ayudar a comprender el funcionamiento de las nuevas plataformas de distribución y publicación de contenidos para editores que proponen *Apple* y *Google*. El usuario se suscribe a periódicos y revistas a través de estas plataformas¹⁷ para ser leídos desde dispositivos móviles. El papel del editor se diluye frente a las plataformas distribuidoras y las compañías de telecomunicaciones. Las bibliotecas tendrán que replantear el acceso que ofrecen a revistas y periódicos.



Figura 4. Google one pass (video)¹⁸

Nuevo ecosistema de la web ubicua: apps, tablets y smartphones

La eclosión de los dispositivos portables y su evolución condiciona una nueva forma de comunicación. El gran negocio se desplaza a la movilidad: las empresas de software, hardware y contenidos centran sus esfuerzos en copar el negocio de la nube en el bolsillo. Por primera vez el mercado de smartphones supera al de PCs¹⁹ en número de unidades vendidas²⁰ y se habla de un estadio post-PC.

Si éste va a ser el parque de dispositivos mayoritario los usuarios tendremos que adaptarnos a las dimensiones de la pantalla y aprender a teclear, a la conectividad permanente (geoposicionamiento, datos en tiempo real) y a una nueva manera de navegar. Si *Twitter* gestiona el tiempo real es porque su filosofía de mensajes de 140 caracteres encaja perfectamente en este nuevo paradigma.

En 2010 un nuevo binomio ha conquistado el mercado tecnológico y nos alumbrará sobre el camino que va a tomar el desarrollo inmediato: una pantalla táctil que nos lleva al servicio deseado con un simple toque (o lo que es lo mismo, servicios a la carta desde cualquier sitio). Si pusiésemos nombre al binomio ese sería *iPad-App*.

El 23 de octubre de 2001 salió al mercado el primer *iPod*; el 3 de abril de 2010 el *iPad*; entre

ambos eventos casi una década en la que la manera de generar, almacenar, distribuir y consumir información y ocio se ha ido transformando. Si la evolución del binomio *iPod-iTunes* cambió el mercado de la distribución de contenidos musicales, la irrupción del ecosistema de microaplicaciones creado por *Apple* para ser consumidas desde sus dispositivos móviles está modificando el modo de acceso a la Red. Un dato significativo: en 9 meses se descargaron 1.000 millones de aplicaciones en la *App Store*.

Con el modelo *iPad* el usuario obtiene sencillez de uso a cambio de cesión de control. Muchas personas son incapaces de mantener un ordenador personal, y sólo desean un modo sencillo de estar en contacto mediante correo, entrar en *Facebook*, leer el periódico o ver algún vídeo. El administrador del *iPad* está "en la nube" y actúa a través de *iTunes*, único punto de control para descargar e instalar aplicaciones y pasar contenidos. Quien filtra el software, controla actualizaciones, aplica reglas de seguridad y gestiona las versiones del producto es *Apple*, verdadero administrador de la máquina. Aunque de momento son los usuarios avanzados los que han adquirido el dispositivo, el verdadero mercado se encuentra en el gran público. Es probable que en 2011 sean adquiridos masivamente y contribuyan a disminuir determinadas brechas digitales.



Figura 5. Android Market y App Store son las grandes plataformas de aplicaciones móviles

Dos son las grandes plataformas que ofrecen este modelo de acceso: *IOS* y *Android*. *Apple/IOS* defiende un sistema cerrado en el que dispositivo y aplicación son inseparables. En cambio *Google/Android* propone un modelo más abierto a todo tipo de fabricantes. Aunque de momento *Apple* marca la iniciativa, muy probablemente el mercado será de *Android*. Quien no disponga de una aplicación en estas plataformas no podrá afrontar el futuro en condiciones. Las bibliotecas necesitan estar en ambas, bien generando sus propias aplicaciones, bien por compra de las mismas a otros proveedores de datos.

Nuevas condiciones de navegación: estándares abiertos

Nos encontramos ante un nuevo escenario de navegación web: cambian hábitos, objetos de consumo e intermediarios. En 2010 se produce una nueva correlación en la implantación de navegadores, aumenta el acceso en movilidad y se acentúa el predominio de consumo de vídeos.

Por primera vez desde la desaparición de *Netscape*, *Internet Explorer* deja de ser el navegador dominante en Europa; la pugna entre *Google*, *Microsoft* y *Apple* favorece la implementación de estándares abiertos para evitar monopolios. El mayor tráfico de datos en la Red es para consumir vídeos; los movimientos más interesantes se han dado en el respaldo que las diferentes empresas daban a los códecs de vídeo²¹. El anuncio de que *Chrome* (*Google*) dejaba de soportar el códec H264 marcó la pauta; detrás de estas decisiones se encuentra la necesidad de que *YouTube* (25% del tráfico mundial) pueda ser accesible desde todas las plataformas sin problemas.

Tiempo real y los segmentos de edad: ¿Dónde encaja lo “no instantáneo” en las bibliotecas?

Aunque el servicio web más utilizado en 2010 fue el correo electrónico²², un estudio de *ComScore*²³ sobre el tiempo invertido en comunicación interpersonal señala que su uso está en baja para todas las edades excepto en los mayores de 55 años²⁴.

El estudio refleja que las audiencias están cambiando no sólo en cuanto a consumo de contenidos o a generación de los mismos, sino en algo más profundo como es el tipo de comunicación preferido: se quiere estar comunicado siempre.

Esta percepción del uso del correo electrónico dependiendo de la edad nos permite reflexionar sobre el valor de lo instantáneo y la ubicuidad. Mientras aquellos que han crecido siempre con acceso a internet (nativos digitales) valoran la rapidez del flujo pregunta/respuesta en la comu-

nicación y, por tanto, consideran como algo del pasado una herramienta que no ofrece una respuesta inmediata, para el segmento de población que se incorporó a la Red ya con cierta edad (mayores de 55 años) el acceso instantáneo y omnipresente es algo nuevo, muchas veces desconocido, de momento no prioritario y hasta les puede resultar angustioso.

Este dato puede llegar a tener una derivada curiosa en el consumo de lectura. En España, en 2010, los mayores de 55 son los grandes consumidores de ereaders (más que los de 25-34 años). Hacen un consumo de la información con la nueva tecnología pero a la vieja usanza: un dispositivo para cada cosa y cada cosa a su tiempo. Es pronto para interpretar si eso es debido a la edad o si es por una falta de práctica en el uso de las herramientas web.

“En España, en 2010, los mayores de 55 son los grandes consumidores de ereaders”

Una reflexión para las bibliotecas y profesionales es qué modelo de gestión de contenidos y de lectura vamos a poder implantar en nuestros centros: aquel de viejos hábitos en nuevos dispositivos (cada cosa a su tiempo, con un dispositivo sólo para leer, sin posibilidad de intercambio como pretende la industria editorial, obligando al usuario a que se acerque a nuestros centros) o ese otro, más emergente, que marca las pautas de consumo de ocio a nivel mundial, que se sustenta en la nube y que tiene a su favor las nuevas generaciones.

Estadio web, usuarios y profesionales de la información en 2010

Un elemento fundamental para entender el momento en el que se encuentran las bibliotecas con respecto a las demandas de la sociedad es la evolución de la implantación del uso de internet en España. En 2006, cuando se empieza a popularizar el concepto biblioteca 2.0, la mayoría social era ajena al fenómeno internet. El informe *España en la sociedad de la información 2006*²⁵ señalaba que la principal razón que impedía el crecimiento del número de usuarios de internet era la falta de percep-

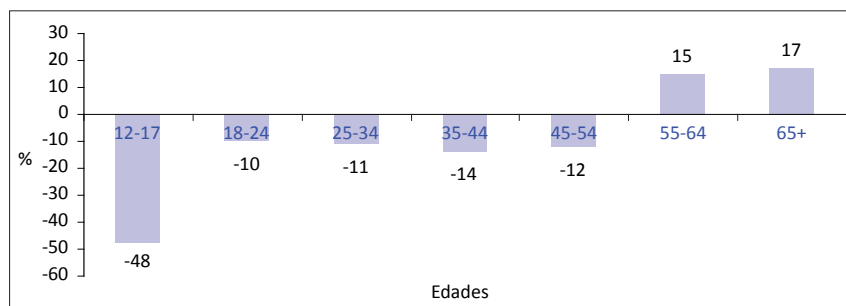


Figura 6. Variación del % de uso del correo electrónico por grupos de edad entre noviembre de 2009 y noviembre de 2010. Fuente: ComScore.



Figura 7. Página principal de la Biblioteca Nacional de España. Anuncia su nuevo espacio de exposición en Flickr. Abajo a la izquierda enlaces a su Facebook, YouTube y blog.

ción de su utilidad por parte de los no usuarios, más del 71% de los cuales aseguraba que no era necesaria.

La implantación de la “biblioteca 2.0” tuvo que superar barreras de cultura y hábito, tanto en la sociedad a la que servía como entre muchos de sus profesionales, al fin y al cabo integrantes de esa misma sociedad; había que demostrar la validez de las nuevas herramientas para poderlas implementar. En el “ecosistema web social” el mercado marca las pautas y los internautas demandan un consumo; ya no hay que convencer, hay que intentar ofrecer eso que “su facebook” o su “twitter” les dan: instantaneidad y ubicuidad en el acceso a la información.

Bibliotecas y centros de documentación continúan adaptando los procesos tradicionales de tratamiento de la información. Las herramientas 2.0 (Margaix²⁶) se implantan progresivamente (ya no es necesario demostrar las virtudes de blogs, rss, redes sociales en las instituciones) y su uso por profesionales facilita la naturalidad en el cambio de actitud hacia las nuevas tecnologías. Pero constatamos que la biblioteca como puerta de acceso a la información sigue perdiendo protagonismo²⁷. Nos encontramos ante una nueva ecología de la información que requiere de unas características, habilidades, competencias informacionales muy diferentes tanto para usuarios como para profesionales²⁸. Si observamos qué ofrecen los servicios que han podido hacer frente al tsunami de las redes sociales tal vez encontremos las claves de lo que necesitamos para revertir la tendencia: especialización, calidad y relevancia en el filtrado de la información.

La Web que se intuye moldeará las nuevas estructuras bibliotecarias: hará replantear

el diseño arquitectónico (pérdida de presencia del fondo documental, liberación de espacios²⁹) para adaptarlo a los nuevos hábitos de acceso a la información (necesidad de banda ancha, pantallas pequeñas, abandono progresivo de la navegación frente a la aplicación específica), y condicionará la aparición de profesionales constantemente conectados a la Red (la dicotomía ámbito profesional-ámbito personal se diluirá mucho más).

La exigencia de conversión bidireccional del entorno web hace necesaria

unas habilidades comunicativas para poder interactuar con los usuarios; será necesario estar, escuchar, participar, compartir, medir, pensar y aprender en las redes sociales; aunque no todas las bibliotecas podrán dedicar recursos humanos específicos a esta tarea (instauración del *community manager*³⁰ para dinamizar el marketing bibliotecario en la web) sí es necesario que ese “espíritu” impregne el día a día de todos los profesionales.

“Las bibliotecas se encuentran en un escenario complicado en el que no pueden ofrecer los contenidos que ya están a disposición de los usuarios por otros medios”

Si hacemos una lectura bibliotecaria de la que se presume la aplicación estrella de 2011 (*Quora*³¹) podemos intuir que la sobreabundancia de información necesita filtros humanos, intermediadores, *curators* que aporten sentido crítico y filtro experto. Si *Diaspora* consigue eclosionar podremos gestionar una red social global con mayores garantías de privacidad y al amparo de los mercados.

En España un elemento preocupante es la incertidumbre ante los retos de la gestión de los contenidos electrónicos. Las estrategias de ámbito nacional no parecen seguir el ritmo del mercado global (los tradicionales proveedores de contenidos no han encontrado el modelo de negocio, y la legislación no favorece el préstamo entre los usuarios). Las bibliotecas se encuentran

en un escenario complicado en el que no pueden ofrecer los contenidos que ya están a disposición de los usuarios por otros medios.

Conclusiones

En España 2010 continúa en la tendencia de años anteriores: crece el número de internautas y el uso de las redes sociales. Se repite el dominio de *Facebook* y *Tuenti*. La mayoría considera internet necesaria. Los mayores de 55 se incorporan a la red y son los grandes consumidores de ereaders (por encima de los de 25-34 años).

Los usuarios prefieren el acceso al almacenamiento. Los cambios de hábito de consumo de información y ocio potenciarán más la nube (cloud computing); los servicios móviles y la aparición de nuevos dispositivos serán clave en este proceso. Los dispositivos móviles con pantalla táctil y aplicaciones específicas ofrecen al usuario sencillez de uso a cambio de cesión de control. Es probable que en 2011 este tipo de dispositivos sean adoptados masivamente y contribuyan a disminuir la brecha digital. Para estar preparadas las bibliotecas deberán ofrecer sus propias aplicaciones en los dos sistemas predominantes: *AppStore* y *Android Market*.

La disparidad de dispositivos, la movilidad y el cloud computing hacen necesario el almacenamiento, la sincronización de la información y tener una estrategia de recuperación de datos. Bibliotecas y centros de documentación deben aprender a gestionar sus propias nubes.

En 2011 primarán servicios que ofrezcan un acceso de más calidad a los contenidos que demandan los usuarios avanzados, entre ellos la intermediación de contenidos. Por primera vez *Google* y *Apple* ofrecerán su plataforma de suscripción de publicaciones.

La web que se intuye moldeará las nuevas estructuras bibliotecarias: hará replantear el diseño arquitectónico para adaptarlo a los nuevos hábitos informacionales y condicionará la aparición de profesionales constantemente conectados a la red con competencias informacionales diferentes. La sobreabundancia de información necesita filtros humanos, intermediadores, *curators* que aporten sentido crítico, y filtro experto; la calidad en la intermediación será la vía para que la biblioteca recupere protagonismo como puerta de acceso a la información.

En España un elemento preocupante es la incertidumbre ante los retos que representa la gestión de los contenidos electrónicos. Las bibliotecas se encuentran en un escenario complicado en el que no pueden ofrecer los contenidos que ya están a disposición de los usuarios por otros medios.

Notas y referencias

1. La sociedad de la información en España: 2008. Madrid: Ariel; Fundación Telefónica, 2009. <http://e-libros.fundacion.telefonica.com/sie08>
2. La sociedad de la información en España: 2009. Madrid: Ariel; Fundación Telefónica, 2009. <http://e-libros.fundacion.telefonica.com/sie09>
3. La sociedad de la información en España 2010. Madrid: Ariel; Fundación Telefónica, 2011. http://e-libros.fundacion.telefonica.com/sie10/aplicacion_sie/ParteA/datos.html
4. <http://www.facebookmarketing.es/infografica-usuarios-facebook-en-espana>
5. **Rubia, David**. "Diaspora, red social open source, ¿plantará cara a Facebook?" *Alt1040*, 13 mayo 2010. <http://alt1040.com/2010/05/diaspora-red-social-open-source-plantara-cara-a-facebook>
6. <http://www.youtube.com/watch?v=9tuuiPHyc8I>
7. Internet 2010 in numbers. <http://royal.pingdom.com/2011/01/12/internet-2010-in-numbers>
8. **Orihuela, José-Luis**. "Estandarte de la cultura libre". *ABC. Medios y redes*. 10 años de Wikipedia. <http://www.abc.es/especiales/wikipedia/jose-luis-orihuela.htm>
9. <http://www.youtube.com/watch?v=gXD1TRGafQ0>
10. **Ortiz, Antonio**. "Ya no quedan 'super consumidores' de información: apuntes ante la muerte de *Bloglines*". *Error 500*, 13 sept. 2010. <http://www.error500.net/articúesarélo/super-consumidores-informacion-bloglines>
11. *Google* es la empresa que más adquisiciones realizó en 2010, seguida de *Facebook*. http://en.wikipedia.org/wiki/List_of_acquisitions_by_Google
12. **Martín, Javier**. "Inversiones y adquisiciones de empresas de la internet hispana en 2010". *Loogic*, 7 enero 2011. <http://loogic.com/inversiones-y-adquisiciones-de-empresas-de-la-internet-hispana-en-2010>
13. Aparecía en la lista de los servicios 2.0 que empleó **Dale Dougherty** cuando acuñó el término. http://es.wikipedia.org/wiki/Web_2.0
14. La sociedad de la información en España 2010. http://e-libros.fundacion.telefonica.com/sie10/aplicacion_sie/ParteA/datos.html
15. <http://foursquare.com>
16. **Manuel, F**. "Dropbox, servicio web más prometededor de 2010". *Genbeta*, 6 enero 2011. <http://www.genbeta.com/genbeta/dropbox-servicio-web-mas-prometededor-de-2010>
17. **Guillermo, Julián**. "Apple y Google anuncian sus servicios de suscripción y publicación para editores". *Genbeta*, 16 febrero 2011. <http://www.genbeta.com/actualidad/apple-y-google->

anuncian-sus-servicios-de-suscripcion-y-publicacion-para-editores

18. http://www.youtube.com/watch?v=qiz_2c_QpOQ

19. Cisco® visual networking index. Global mobile data. Traffic forecast update.

http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.pdf

20. Según los datos de IDC (*Fortune.com*) del último trimestre de 2010. Hablamos de 101 millones de smartphones vendidos en el mundo frente a 92 millones de ordenadores personales, lo que supone un incremento del 87% anual en los primeros y apenas de un 3% de los segundos.

21. **Guillermo, Julián.** "Google liberará el código del códec de vídeo VP8, se abre la 'guerra de códecos'". *Genbeta*, 13 abril 2010.

<http://www.genbeta.com/multimedia/google-liberara-el-codigo-del-codec-de-video-vp8-se-abre-la-guerra-de-codecs>

22. Se enviaron 107 billones de correos en 2010, 294.000 millones al día y aumentaron en 480 los millones de cuentas.

<http://royal.pingdom.com/2011/01/12/internet-2010-in-numbers>

23. **Richtel, Matt.** "E-Mail gets an instant makeover". *The New York times*, 20 Dec 2010.

http://www.nytimes.com/2010/12/21/technology/21email.html?_r=2

24. **Amartino, Mariano.** "Los jóvenes no quieren el email".

<http://www.uberbin.net/archivos/comunidades-online/los-jovenes-no-quieren-el-email.php#more-13034>

25. **Alierta, César.** "La sociedad de la información en España 2006".

http://sociedadinformacion.fundacion.telefonica.com/DYCISHII/Articulos_A_Fondo_-_Sociedad_de_la_Informacion_Espana_2006/seccion=1188&idioma=es_ES&id=2009100116310002&activo=4.do

26. **Margaix-Arnal, Dídac.** "Redes sociales y web 2.0 en 2009". *Anuario ThinkEPI*, 2010, v. 4, pp. 211-216.

27. Perceptions of libraries, 2010: Context and community.

<http://www.oclc.org/reports/2010perceptions.htm>

28. **Rainie, Lee.** "Networked creators how users of social media have changed the ecology of information". *VALA2010 conf.*

http://www.vala.org.au/vala2010/papers2010/VALA2010_Keynote_Rainie_Final.pdf

29. **González-Fernández-Villavicencio, Nieves.** "En bibliotecas universitarias... estanterías ¡fuera!". *Bibliotecarios 2.0*, 30 enero 2011

<http://www.nievesglez.com/2011/01/en-bibliotecas-universitariasestanteria.html>

30. Community manager ¿nueva profesión en la web 2.0?

<http://comunidad20.sedic.es/?p=229>

31. <http://www.quora.com>

E.1. Ciencia 2.0: la aportación de las bibliotecas universitarias

Por **José-Antonio Merlo-Vega**

28 diciembre 2010

Merlo-Vega, José-Antonio. "Ciencia 2.0: la aportación de las bibliotecas universitarias".
Anuario ThinkEPI, 2011, v. 5, pp. 155-159.



Resumen: Se describe el informe de la Red de Bibliotecas Universitarias (Rebiun) "Ciencia 2.0: aplicación de la web social a la investigación" y se destacan las principales webs que permiten a los científicos cooperar para compartir la investigación, los recursos y los resultados. Las bibliotecas universitarias actúan como capacitadoras y difusoras de los entornos 2.0 aplicados a la investigación.

Palabras clave: Investigación, Comunicación científica, Web 2.0, Redes sociales, Bibliotecas universitarias, Academia, Ciencia 2.0.

Title: *Science 2.0: contribution of university libraries*

Abstract: The report of the Spanish Network of University Libraries (Rebiun) "Science 2.0: implementation of the social web to research" is discussed, and the main websites that allow scientists to cooperate or share in the three facets of research, resources and results are listed. University libraries should play an active role in the 2.0 environment to serve scientific research.

Keywords: Research, Scientific communication, Web 2.0, Social networking, University libraries, Academia, Science 2.0.

EL II PLAN ESTRATÉGICO de Rebiun, iniciado en 2007 y aún en vigencia, incluye una línea de trabajo para realizar estudios y proyectos relacionados con los servicios que las bibliotecas universitarias prestan a la investigación.

El grupo de trabajo de Rebiun "Ámbito de la investigación" aprobó como objetivo operacional en 2010 la elaboración de un estudio sobre la utilidad de las tecnologías de la web social en los procesos de investigación científica.

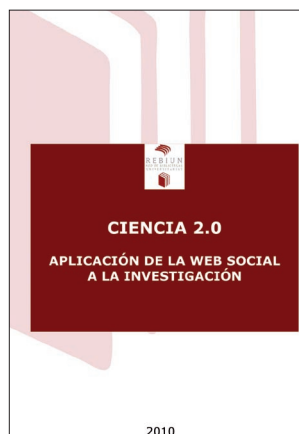
El resultado ha sido el informe *Ciencia 2.0: aplicación de la web social a la investigación*¹, que fue redactado por **Antonia Angosto (UM)**, **Tránsito Ferreras (USAL)**, **José-Pablo Gallo (UMH)**, **José-Ángel Maestro (UPV)**, **Inmaculada Ribes (UPV)** y **José-Antonio Merlo-Vega (USAL)**, quien actuó de coordinador del estudio.

Este documento parte de la necesidad de identificar las distintas manifestaciones y utilidades de lo que se puede considerar ciencia 2.0 o investi-

gación participativa. Para ello se organizan y describen los variados servicios en los que los investigadores pueden emplear las tecnologías de la web social, incorporando en todos los casos ejemplos de aquellos recursos, productos y servicios dignos de destacar por su valor, utilidad y reconocimiento internacional.

Además el informe añade las posibles aportaciones de las bibliotecas académicas, informando en la exposición de los recursos seleccionados sobre cómo pueden aplicarse los diferentes servicios sociales al trabajo bibliotecario, qué pueden hacer las bibliotecas para facilitar el uso de la ciencia 2.0 en cada caso y cómo la biblioteca universitaria puede convertirse en agente promotor de los diferentes recursos 2.0 de interés para la investigación.

Aunque esta nota se basa en el informe, no es un extracto del mismo, por lo que se recomienda su consulta ya que en él se ofrece una completa sistematización de las aplicaciones de la web social a la investigación, junto a un centenar de ejemplos comentados de servicios y una bibliografía selectiva, en la que es conveniente destacar el artículo de **Cabezas-Clavijo** et al sobre herramientas 2.0 para la investigación publicado en *El profesional de la información*², así como el reciente informe de *Research Information Network*³ sobre cómo usan los investigadores la web social.



La ciencia compartida

La web social se caracteriza por la posibilidad de compartir datos e informaciones. Los servicios 2.0 se aplican a los distintos campos de las relaciones sociales, ofreciendo múltiples posibilidades para los métodos utilizados en la investigación científica, ya sea para el fomento de relaciones directas y abiertas entre equipos de investigadores, para la ejecución de proyectos de forma compartida o para la difusión de resultados de manera abierta y participativa. Se puede afirmar que la aplicación de la web social a la investigación se lleva a cabo en tres grandes facetas: compartir la investigación, compartir los recursos y compartir los resultados. En cada una de ellas hay aplicaciones concretas y servicios de indudable calidad y utilidad, que son expuestos en los siguientes párrafos.

“La ciencia 2.0 está encontrando un espacio privilegiado en las plataformas científicas”

Compartir la investigación significa integrar las personas interesadas en el proceso de un proyecto científico concreto, así como la relación activa entre colegios de investigadores que trabajan los mismos temas.

La web social ofrece variados servicios para compartir la experimentación como son las redes sociales, las bases de datos de científicas y sobre todo las plataformas científicas. También hay que incluir en este grupo de aplicaciones los servicios instrumentales participativos. Las redes sociales en la Web son una fiel correspondencia de las relaciones sociales tradicionales, ya que en estos servicios 2.0 se establece una comunicación ágil y un intercambio de recursos e informaciones con actualización constante. Por tanto, no ha sido extraño que se hayan creado redes específicas para las relaciones profesionales entre investigadores.

De igual forma están siendo útiles las bases de datos de científicos, en las que se pueden crear perfiles, aportar publicaciones y seguir a personas que trabajan en un campo de interés. Pero la ciencia 2.0 está encontrando un espacio privilegiado en las plataformas científicas, entendidas como portales que integran sistemas de relación, con medios para compartir recursos, participar en hipótesis y para difundir resultados parciales o finales de experimentos.

Además la web social ofrece diferentes servicios instrumentales participativos, que se aplican en determinadas fases de la investigación y que sirven para compartir archivos, para la realización de encuestas o de investigaciones sociales o bien para la creación y administración de mapas conceptuales.

“La aplicación de la web social a la investigación se lleva a cabo compartiendo la investigación, los recursos y los resultados”

Algunos ejemplos de ciencia 2.0 para compartir la investigación son los siguientes:

Redes sociales científicas

Academia
<http://www.academia.edu>

The screenshot shows the Academia.edu profile for Stephen Hawking. At the top, there is a search bar containing 'Stephen Hawking'. Below the search bar, the profile header includes 'Stephen Hawking' and 'Work being Followed by (1,324)'. The profile is categorized under 'University of Cambridge' with the title 'Faculty Member, Applied Mathematics and Theoretical Physics' and 'Lucasian Professor of Mathematics'. It also lists 'Gonville and Caius College'. A section titled 'Department Colleagues (49)' features four small profile pictures with names: James Grime, TC Hamant, Emma Boland, and Mohammad Esmail. The 'About' section contains a paragraph: 'Stephen Hawking has worked on the basic laws which govern the universe. With Roger Penrose he showed that Einstein's General Theory of Relativity implied space and time would have a beginning in the Big Bang and an end in black holes. These results indicated it was necessary to unify General Relativity with Quantum Theory, the other great Scientific development of the first half of the 20th Century. One consequence of such a unification that he discovered was that black holes should not be completely black, but should emit radiation and eventually evaporate and disappear. Another conjecture is that the universe has no edge or boundary in imaginary time. This would imply that the way the universe began was completely determined by the laws of science.'

<http://cambridge.academia.edu/StephenHawking>

ResearchGate
<http://www.researchgate.net>

Bases de datos de científicos

Researcher ID
<http://www.researcherid.com>

Emerald research connections
<http://info.emeraldinsight.com/research/connections/index.htm>

Plataformas científicas

HUBzero
<http://hubzero.org>

MyExperiment
<http://www.myexperiment.org>

FeelSynapsis
<http://www.feelsynapsis.com>

Servicios instrumentales participativos

Google Docs
<http://docs.google.com>

Prezi
<http://prezi.com>

Survey Monkey
<http://www.surveymonkey.com>

Survey Gizmo
<http://www.surveygizmo.com>

FreeMind
<http://freemind.sourceforge.net>

Mindomo
<http://www.mindomo.com>

Compartir los recursos en la investigación supone tanto ofrecer a la comunidad interesada los enlaces o referencias más válidas para un científico, como permitir que las aportaciones propias se enriquezcan con nuevos recursos, recomendaciones y valoraciones de los datos que se han compartido.

La ciencia 2.0 como sistema para compartir recursos se identifica con tres servicios participativos: la gestión de referencias bibliográficas, el mantenimiento de favoritos sociales y los índices de citas.

Los gestores bibliográficos son esenciales como bases de datos de referencias de documentos. La investigación necesita partir de estudios previos para avalar sus resultados y a menudo, sobre todo en trabajos de investigación exhaustivos como las tesis doctorales, la abundancia de bibliografía obliga al empleo de sistemas de gestión bibliográfica ágiles. Los gestores bibliográficos están incorporando servicios participativos ya sea permitiendo la publicación abierta de las referencias –para que se complementen e integren con otras–, o bien sistemas para incluir recomendaciones, valoraciones y comentarios.

“La publicación de contenidos en blogs y wikis es la forma tradicional de la edición participativa”

Los servicios de favoritos sociales de recursos y referencias bibliográficas ofrecen entornos para compartir enlaces a páginas o documentos en la Web, con opciones para que un colectivo mantenga de forma conjunta un grupo de enlaces, para que los recursos recomendados de un investigador puedan ser seguidos por personas interesadas en el tema o para se establezcan jerarquías en la calidad de los recursos compartidos a partir de las valoraciones del colectivo o del número de personas que recomiendan un recurso.

Por último, los índices de citas 2.0 se basan tanto en los gestores bibliográficos como en los servicios de favoritos sociales, permitiendo obte-

surveygizmo.com

Need help picking a plan? (800) 609-6480 login Free Trial

home pricing features full service help / support about

everything you could want in our advanced feature set

search features or call us

We're easy to reach at (800) 609-6480

- It's Easy**
No experience needed. Jump right in and create an online survey, poll or website form in minutes.
- Over 20 Question Types**
Combine the perfect question types for research, insight, tracking and marketing automation.
- Sophisticated Reporting**
Drill down into your data and create detailed reports you can share with your peers and clients.
- Logic, Branching, Jumping...**
Every advanced feature imaginable – like branching, page jumping, and piping. At a fraction of the cost!
- Enterprise Features**
Manage multiple users, teams, permissions and roles. Experience integration with Salesforce and other SaaS products.
- Absolute Branding Control**
Control the url of your surveys, the look, color and images. Completely obscure the engine and take control!

<http://www.surveygizmo.com>

ner información sobre cuántas citas recibe un documento en línea y cuál es la procedencia de la citación. Algunos ejemplos de ciencia 2.0 para compartir los recursos son los siguientes:

Gestores de referencias bibliográficas

Zotero

<http://www.zotero.org>

Mendeley

<http://www.mendeley.com>

RefWorks

<http://www.refworks.com>

Favoritos sociales

2collab

<http://www.2collab.com>

CiteUlike

<http://www.citeulike.org>

Connotea

<http://www.connotea.org>

Índices de citas

Google Scholar

<http://scholar.google.es>

CiteSeerx

<http://citeseerx.ist.psu.edu>

GetCited

<http://www.getcited.org>

Compartir los resultados de la investigación es también una actividad 2.0, ya que se emplean técnicas y servicios de redifusión de datos, publicación participativa o archivo en acceso abierto. La ciencia 2.0 para compartir resultados se identifica con tres servicios: blogs y wikis científicos, servicios de noticias científicas y servidores de producción científica en acceso abierto.

La publicación de contenidos en blogs y wikis es la forma tradicional de la edición participativa. Las posibilidades de la publicación 2.0 son empleadas también por la investigación para promocionar resultados, informar de proyectos en marcha y recoger comentarios del colectivo interesado.

Es habitual que las entidades y empresas dedicadas a la investigación dispongan de plataformas para crear blogs informativos o wikis colaborativos. En muchos casos la tecnología blog es también la usada por proveedores de noticias científicas, que pueden incluirse dentro de los servicios 2.0, ya que facilitan herramientas para la redifusión de contenidos, para la aportación de comentarios, la recomendación de noticias o para la valoración de las informaciones publicadas.

Es esencial incluir a los repositorios científicos como aportaciones a la ciencia compartida, ya que se trata de los espacios en los que se depositan los artículos, informes, monografías, comu-



<http://www.citeulike.org>

“La ciencia 2.0 presenta a las bibliotecas la oportunidad de cumplir con su compromiso de servicio a la investigación”

nificaciones o cualquier otra forma de transmisión de los resultados de una investigación. De igual forma que los servicios anteriores, los repositorios científicos están empleando mecanismos de sindicación, valoración y comentarios, así como herramientas de valor añadido para compartir datos en plataformas de favoritos sociales, redes científicas y sistemas de gestión bibliográfica.

Algunos ejemplos de ciencia 2.0 para compartir los resultados son los siguientes:

Blogs y wikis

Science blogs

<http://scienceblogs.com>

PLoS blog

<http://www.plos.org/cms/blog>

OpenWetWare

<http://openwetware.org/wiki>

Nature

<http://blogs.nature.com>

Servicios de noticias científicas

Servicio de Información y Noticias Científicas (SINC)

<http://www.agenciasinc.es>

SciTopics

<http://www.scitopics.com>

ScienceDaily

<http://www.sciencedaily.com>

Science News

<http://www.sciencemag.org>

Acceso abierto (recolectores de repositorios)

Scientific Commons

<http://www.scientificcommons.org>

OAister WorldCat

<http://oaister.worldcat.org>

Recolecta

<http://www.recolecta.net>

Hispana

<http://hispana.mcu.es>

Actuación de la biblioteca universitaria

La biblioteca universitaria tiene entre sus principales funciones el servir de centro de recursos y servicios para la investigación. Está por consiguiente al servicio de la investigación, lo cual se expresa tanto ofreciendo recursos de información como ayudando en las distintas fases del proceso de la investigación científica.

Las bibliotecas universitarias y las científicas, gracias a los avances de la sociedad tecnológica, han superado su limitación de ser meras transmisoras de información bibliográfica y de recursos documentales. Sin abandonar su misión de gestoras y provisoras de documentos, las bibliotecas en general y las académicas en particular son agentes promotores de recursos, función que desempeñan diseñando objetos de aprendizaje, produciendo contenidos informativos y realizando acciones formativas.

La biblioteca universitaria es productora de recursos de información participativos para la investigación, ya que se convierte en facilitadora de las aplicaciones, servicios, productos y recursos de utilidad para la ciencia cuando realiza tutoriales sobre los mismos, cuando organiza sesiones formativas o cuando difunde dichos contenidos en redes sociales, blogs o listas de información administradas por la biblioteca.

La biblioteca selecciona recursos, los recomienda y los difunde, pero además, elabora objetos de aprendizaje que sirven para conocer su manejo u organiza sesiones formativas espe-

cíficas para enseñar a emplearlos. Los investigadores se suscriben a los blogs y a las listas de las bibliotecas para conocer los recursos que se describen en dichas fuentes, que a menudo van más allá de las colecciones propias; pero también se inscriben en sus cursos de formación y son los principales usuarios de tutoriales sobre técnicas de documentación, ya no sólo orientados a la búsqueda en recursos sino ampliados a temas como la gestión bibliográfica, la obtención de índices de citas e indicios de calidad de publicaciones o el manejo de herramientas participativas para compartir proyectos de investigación.

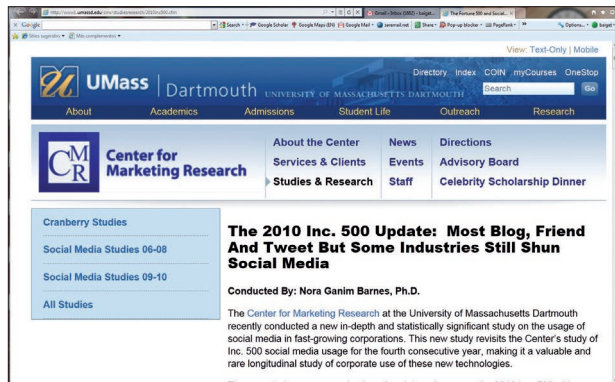
La biblioteca universitaria debe ser promotora de la investigación participativa, como administradora de contenidos informativos, productora de recursos de aprendizaje y capacitadora en servicios colaborativos. La ciencia 2.0 presenta a las bibliotecas la oportunidad de cumplir con su compromiso de servicio a la investigación. Rebiun ha sido consciente de la necesidad de fomentar el uso de las herramientas colaborativas en la investigación y, mediante su estudio sobre ciencia 2.0, ha querido contribuir con una guía práctica para facilitar y fomentar los servicios científicos basados en tecnologías participativas. Se trata de otro buen ejemplo de cómo las bibliotecas están sabiendo aplicar las tecnologías para cumplir y ampliar sus objetivos.

Referencias bibliográficas

1. *Rebiun. Ciencia 2.0: aplicación de la web social a la investigación*. Madrid: Rebiun, 2010.
<http://www.rebiun.org>
<http://eprints.rclis.org/handle/10760/3867>
2. **Cabezas-Clavijo, Álvaro; Torres-Salinas, Daniel; Delgado-López-Cózar, Emilio**. "Ciencia 2.0: catálogo de herramientas e implicaciones para la actividad investigadora". *El profesional de la información*, 2009, v. 18, n. 1, pp. 72-79. DOI: 10.3145/epi.2009.ene.1
3. *Research Information Network. If you build it, will they come? How researchers perceive and use web 2.0*. London: RIN, July, 2010.
<http://www.rin.ac.uk>

Informes

Inc. 500 2010: la mayoría de empresas usan blogs, hacen contactos y tuitean, pero algunas todavía evitan los medios sociales

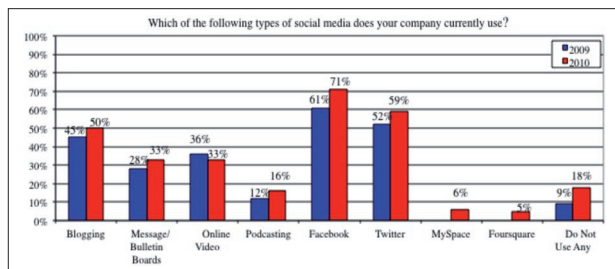


Ganim-Barnes, Nora. *The 2010 Inc. 500 update: Most blog, friend and tweet but some industries still shun social media.* Center for Marketing Research, University of Massachusetts Dartmouth, 2010, 11 pp.

Descargar el informe (324 KB):
<http://www1.umassd.edu/cmrl/studiesresearch/2010inc500.pdf>

El Center for Marketing Research de la University of Massachusetts, Dartmouth, EUA, llevó a cabo un estudio sobre el uso de los medios de comunicación social en empresas en rápido crecimiento. Se trata de la 5ª edición anual del mismo, obteniendo así un estudio longitudinal valioso y poco frecuente del uso corporativo de estas nuevas tecnologías.

Las Inc. 500 es una lista de las compañías privadas de mayor crecimiento en los EUA, elaborado anualmente por la revista Inc. Magazine.



En 2007, el primer estudio reveló ya que el grupo Inc. 500 dejó atrás a las tradicionales empresas de la lista Fortune 500 en el uso de los medios de comunicación social. Al igual que en los estudios anteriores, el estudio de 2010, bajo la dirección de **Nora Ganim-Barnes**, se realizó mediante una encuesta telefónica a las empresas de la lista Inc. 500. Todas las entrevistas se llevaron a cabo entre

octubre y noviembre de 2010. La lista de 2010 se publicó en la edición de septiembre de la revista. En esta cuarta versión participó el 34% (171) de las Inc. 500, por lo que la investigación es estadísticamente válida con un error de +/- 6%.

Según el informe, los medios de comunicación social han penetrado en el mundo de los negocios a una gran velocidad. También indica que la familiaridad de las empresas con el uso de los medios de comunicación social ha seguido creciendo en los últimos 12 meses. Las empresas encuestadas en este estudio, como en los estudios anteriores, son diversas en sector, tamaño y ubicación. Se incluyen dos de las primeras 10, cuatro de 25, y 28 de las 100 principales empresas de la lista Inc. 500.

Las empresas respondieron a preguntas detalladas sobre su uso y valoración de blogs, podcasts de vídeo online, redes sociales, foros y wikis.

Cambios a lo largo de los años: los wikis caen (se utilizan más como herramienta de colaboración que de participación de las comunicaciones), y varían las proporciones de uso de las redes sociales Twitter, Facebook, MySpace, LinkedIn y Foursquare.

La adopción de medios de comunicación social varía entre las industrias. A pesar de que el 83% de las empresas utilizan al menos una de las herramientas, la adopción es sesgada según la industria. Las empresas de servicios gubernamentales constituyen el 12% de la Inc. 500, pero representan el 27% de las que no utilizan medios sociales. Las compañías de energía constituyen un 3% de la lista, pero constituyen el 17% de los no usuarios. Las sociedades de servicios financieras siguen el mismo patrón: son el 5% de la lista pero el 10% de las que aún no han adoptado los medios de comunicación social.

El estudio también incluye preguntas sobre el uso de medios sociales para reclutar y evaluar a los empleados, así como para comunicarse con los consumidores y los stakeholders (personas con intereses en la empresa).

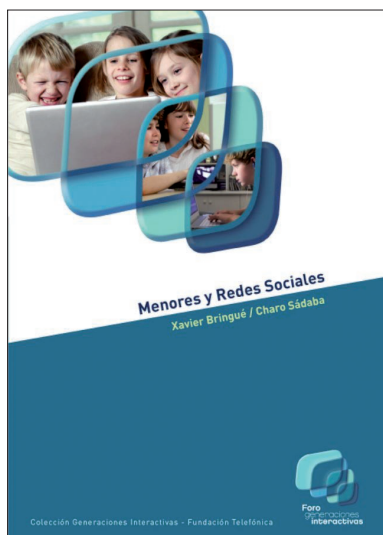
Según el informe, las redes sociales siguen señalando el camino a seguir. La plataforma más familiar para las 500 Inc. es Facebook, con 87% de los encuestados que afirman estar 'muy familiarizados' con ella. El blogging sigue siendo una herramienta importante: el 50% tiene un blog corporativo (era un 45% en 2009 y un 39% en 2008). Además, el informe señaló que las nuevas herramientas de comunicación están cambiando la manera de operar con éxito los negocios. En 2009 el 43% creía que los medios de comunicación social eran "muy importantes" para su negocio o estrategia de marketing, y ese número

aumentó a 56% en 2010. Además, los medios de comunicación social no sólo se utilizan para la comunicación entre empresas y consumidores, sino también para comunicarse con proveedores y socios.

Nota de prensa:

<http://www1.umassd.edu/cmrl/studiesresearch/2010inc500.cfm>

Menores y redes sociales



14,6 MB):

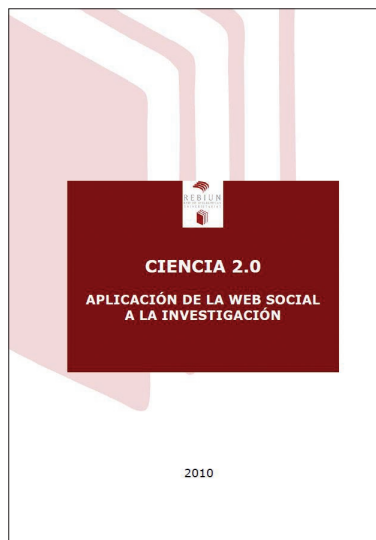
http://www.generacionesinteractivas.org/wp-content/uploads/2011/01/Libro-Menores-y-Redes-Sociales_Fin.pdf

Bringué-Sala, Xavier; Sádaba-Chalezquer, Charo. *Menores y redes sociales.* Colección / Foro generaciones interactivas, Fundación Telefónica, enero 2011, 329 pp.

ISBN: 84-8081-207-9

Descargar informe (pdf,

Ciencia 2.0: Aplicación de la web social a la investigación



Rebiun. *Ciencia 2.0: Aplicación de la web social a la investigación.* Madrid: CRUE - Red de Bibliotecas Universitarias Españolas, nov. 2010, 74 pp.

Descargar el informe (565 KB)
<http://eprints.rclis.org/19304/>

Estudio sobre de Ciencia 2.0 y las aplicaciones de la web social para la investigación. Se establecen tres categorías: compartir la investigación, compartir los recursos y compartir los resultados.

Se describen las aplicaciones y se seleccionan recursos de interés: redes sociales científicas, bases de datos de científicas, plataformas para la investigación, encuestas, mapas conceptuales, uso compartido de archivos, gestión bibliográfica, marcadores sociales, índices de citas, blogs y wikis, noticias científicas, acceso abierto. Los servicios se evalúan, describiendo su interés para las bibliotecas.

F. INDIZACIÓN Y RECUPERACIÓN

Informe de situación

Nueva normativa de catalogación: pasos hacia un futuro prometedor pero incierto

Por Assumpció Estivill-Rius

Estivill-Rius, Assumpció. "Nueva normativa de catalogación: pasos hacia un futuro prometedor pero incierto". *Anuario ThinkEPI*, 2011, v. 5, pp. 162-167.



Resumen: Se analizan las principales características de la norma RDA (*resource description and access*) publicada en 2010, diseñada para facilitar la captura, el almacenaje, la recuperación y la visualización de datos con las tecnologías más avanzadas de bases de datos, al tiempo que continúa siendo compatible con la tecnología más antigua aún en uso. Se presentan sus ventajas e inconvenientes, su relación con otras normas bibliográficas y algunas consecuencias de su implantación.

Palabras clave: RDA, Normas, Estándares, Catalogación, Descripción bibliográfica, Metadatos, Bibliotecas.

Title: *New cataloguing rules: steps toward a promising but uncertain future*

Abstract: The main features of the resource description and access (RDA) standard, published in 2010, are analyzed. Designed to facilitate the capture, storage, retrieval and visualization of data with the most advanced data base technologies, RDA remains compatible with older technology still in use. Its advantages and disadvantages, relationship to other bibliographic standards and some consequences of its implementation are discussed.

Keywords: RDA, Standards, Cataloguing, Bibliographic description, Metadata, Libraries

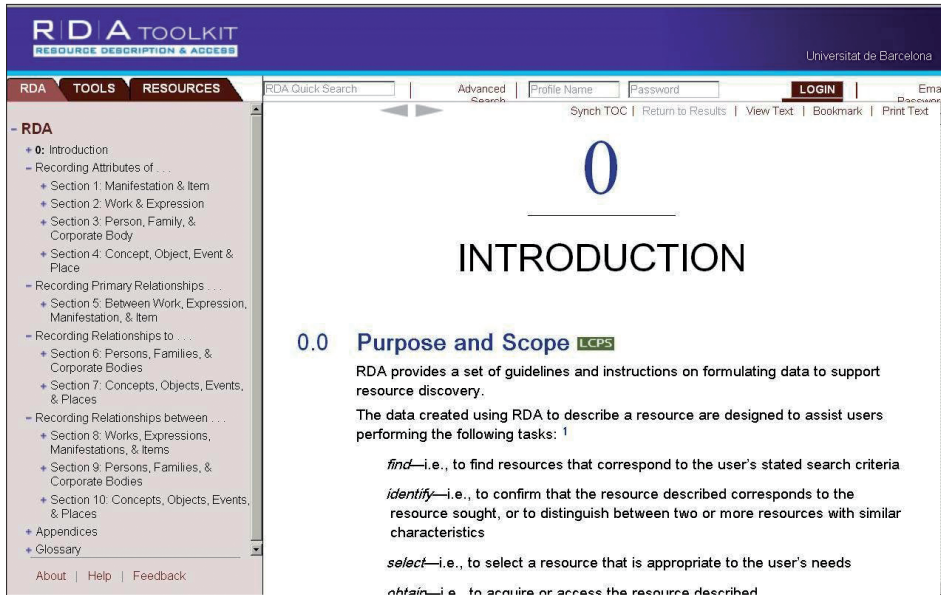
EN LOS ANALES DE LA CATALOGACIÓN, 2010 será el año de la publicación oficial del nuevo código RDA (*resource description and access*) y del RDA toolkit, la aplicación informática en línea que acoge la RDA.

También se actualizaron otros recursos útiles en la catalogación, como los *Library of Congress policy statements*, con las concreciones de la *Library of Congress (LC)* adaptadas a la RDA, las *Anglo-American cataloguing rules (AACR2)*, y algunos instrumentos desarrollados para usar con la RDA, como las correspondencias de los atributos y las relaciones definidas en el código con los formatos MODS y MARC 21 (bibliográfico y de autoridades), o diversos flujos de trabajo para describir algunos tipos de materiales (libros en una o más partes, publicaciones seriadas impresas, fondos y colecciones de archivo, etc.).

Sin embargo, tanto RDA como su *toolkit* están todavía en período de evaluación. Si todo marcha de la manera prevista las bibliotecas nacionales

de los Estados Unidos les darán el visto bueno una vez den por finalizada, a principios de abril, la evaluación del test que han llevado a cabo, con 23 participantes más, para determinar si los beneficios obtenidos con la implantación del nuevo código justifican el cambio.

El proceso hasta llegar al momento actual, que todavía está presidido por una cierta incógnita, ha sido largo y no exento de críticas, tanto de los sectores más ortodoxos, representativos de la catalogación tradicional, como de los más tecnológicos y más relacionados con el entorno del Dublin Core. De hecho, el test mencionado anteriormente es consecuencia directa de la recomendación del *Working Group on the Future of Bibliographic Control*, que reunió a diversos expertos convocados por la misma LC, de suspender la elaboración del código mientras no se solucionara una serie de cuestiones previas (*On the record*, 2008, p. 29). En concreto el grupo aconsejaba demostrar los beneficios del cambio;



Inicio del capítulo introductorio de RDA

mejorar el texto de RDA y la navegación; reconsiderar la discrepancia con las convenciones de la ISBD (*International standard bibliographic description*), y en todo caso argumentarlas, y trabajar en colaboración con la *Dublin Core Metadata Initiative (DCMI)* para elaborar un vocabulario de descripción bibliográfica basado en los modelos FRBR y DC y en la tecnología de la web semántica. Como veremos más adelante, esta última recomendación ha empezado a dar algunos frutos.

La versión impresa de RDA tiene 1.096 páginas. Así es que en esta presentación no se pretende hacer un análisis minucioso de sus características, sino sólo señalar aquellos aspectos más destacables que marcan las diferencias con el código todavía vigente de las AACR2.

Características principales de la RDA

La RDA, que se define como un conjunto de directrices e instrucciones para formular datos que ayuden en la búsqueda de recursos, se presenta como un marco flexible y extensible que permite la descripción de recursos digitales y tradicionales.

¿Demasiada flexibilidad?

La flexibilidad y la extensibilidad son características positivas porque facilitarán la adopción internacional del código tanto en el entorno bibliotecario como, quizás, en otros contextos. Sin embargo, también puede tener su lado negativo que puede conducir a la proliferación de prácticas locales y a una ausencia de normalización que dificulte el intercambio de datos. RDA define un conjunto de elementos básicos (*core elements*) que

son obligatorios para describir una obra, una expresión, una manifestación y un ítem, y para describir las entidades asociadas a estos recursos. Prescribe también elementos adicionales para diferenciar recursos con una información identificativa similar y entidades con el mismo nombre o título. La inclusión de otros elementos es opcional, y el texto indica que el centro catalogador puede establecer sus propias políticas locales o dejar la anotación de otros elementos al criterio de

la persona que crea los datos. A partir de estas directrices introductorias (0.6), se multiplican las indicaciones similares.

“Unas directrices tan laxas pueden dar como resultado registros poco uniformes que no convivirán fácilmente en una misma base de datos”

Además, a lo largo del texto abundan las instrucciones alternativas y las opciones de añadir u omitir datos concretos. Unas directrices e instrucciones tan laxas pueden dar como resultado tanto registros minimalistas como registros detallados hasta el exceso, y obligarán a innumerables concreciones locales que aumentarán la pesadez del texto (en número de páginas, como mínimo). Serán registros poco uniformes que no convivirán fácilmente en una misma base de datos.

Interoperabilidad

Entre las características de la RDA también destaca el hecho de que está diseñada para facilitar la captura, el almacenaje, la recuperación y la visualización de datos con las tecnologías más avanzadas de bases de datos, al tiempo que continúa siendo compatible con la tecnología más antigua aún en uso. En este sentido, y como iniciativa del *DCMI/RDA Task Group*, se han empezado a registrar los vocabularios usados en elementos de la RDA en el *Open Metadata Registry* usando los estándares de la web semántica –RDF

y SKOS (*simple knowledge organization system*), y se ha iniciado la elaboración de un perfil de aplicación del DC basado en RDA y en los modelos FRBR (*functional requirements for bibliographic records*) y FRAD (*functional requirements for authority data*), lo cual facilitará en el futuro la interoperabilidad de los metadatos RDA.

En esta misma dirección, hay que señalar que la LC ya mantiene desde los años 90 una serie de tablas de correspondencias entre el formato MARC 21 y estándares usados por otras comunidades (como el DC, el *Content standards for digital geospatial metadata* del *Federal Geographic Data Committee* y ONIX). Recientemente, OCLC también se ha interesado en la captura de datos a partir de ONIX, el estándar usado en el ámbito de las industrias del libro, y ha publicado unas tablas de correspondencias actualizadas (Godby, 2010). Así es que, de alguna manera, la capacidad para el intercambio de datos bibliográficos con formatos usados por otras comunidades existe desde hace años, pero de momento los resultados han sido bastante limitados.

“RDA está planteada como un estándar de contenido y no de presentación de datos”

El MARC usado en bibliotecas es un formato potente, pero rígido, que fue desarrollado hace casi 45 años y que sólo se ha adaptado al entorno web a base de formatos paralelos (como MARC XML y MODS).

Por tanto la característica de la RDA de adaptarse a las tecnologías web forma parte de un escenario futuro, ya que, de momento, los registros elaborados con este estándar continuarán codificados en primera instancia en formato MARC 21.

Por otro lado, la adaptabilidad de la RDA a la tecnología actual dependerá también de que los diseñadores de sistemas integrados de gestión bibliotecaria sepan interpretar de la manera

Name	Owner	Last Updated	Actions
PBCore pbcoreTitle/titleType	PBCore	2011-01-28	
PBCore pbcoreTitle/version	PBCore	2011-01-28	
PBCore publisherRole	PBCore	2011-01-28	
PBCore relationType	PBCore	2011-01-31	
Plant Types	Tracy Mehlh	2007-11-14	
RDA Applied Material	Metadata Management Associates	2008-05-23	
RDA Aspect Ratio	Metadata Management Associates	2009-03-31	
RDA Base Material	Metadata Management Associates	2008-05-23	
RDA Base Material for Microfilm, Microfiche, Photographic Negatives, and Motion Picture Film	Metadata Management Associates	2009-04-01	
RDA Book Format	Metadata Management Associates	2009-04-12	
RDA Broadcast Standard	Metadata Management Associates	2009-04-14	
RDA Carrier Type	Metadata Management Associates	2009-01-01	
RDA Choruses	Metadata Management Associates	2009-03-27	
RDA colour	Karen Coyle	2008-05-17	
RDA Colour of Moving Images	Metadata Management Associates	2009-05-27	
RDA Colour of Still Image	Metadata Management Associates	2009-05-27	
RDA Colour of Three-Dimensional Form	Metadata Management Associates	2009-11-11	

Open metadata registry con algunos de los vocabularios RDA ya registrados, <http://metadataregistry.org/vocabulary/list/page/4.html>

adecuada los modelos conceptuales en los que se sustenta el código.

No regula la presentación

Como consecuencia de la flexibilidad de RDA y de su adaptabilidad a la tecnología emergente, el texto está planteado como un estándar de contenido y no como un estándar de presentación de datos. Esto significa que las directrices prescriben únicamente los datos o elementos que hay que anotar en el registro, y no dan indicaciones sobre su presentación o visualización –es decir, no indican el orden de los elementos de la descripción o de los elementos de los puntos de acceso, ni la puntuación que los precede–.

Como ejemplos de formatos de presentación y codificación de los datos se citan ISBD y MARC 21, y se remite a los apéndices D y E, que incluyen tablas de correspondencias con los dos estándares. En general, los catálogos en línea nunca han usado el formato de presentación ISBD al pie de la letra, pero era un marco válido para estructurar la información bibliográfica y puntuarla según un sistema común que han adoptado todos los códigos de catalogación desde los años 70. La ISBD continúa vigente en el entorno internacional, y próximamente se publicará la edición consolidada definitiva del texto; también se está trabajando en la elaboración de un esquema ISBD-XML con el objetivo de que la normativa evolucione hacia las técnicas y servicios de la web semántica. Por ello es de suponer que los catálogos en línea continuarán basándose en este texto para presentar la información bibliográfica al usuario. Pero recordemos que ISBD es un estándar de presentación y, al mismo tiempo, es también un estándar de

contenido, y en este aspecto fundamental, son notables algunas discrepancias con las prescripciones de la RDA.

“Son notables algunas discrepancias entre las prescripciones de la RDA y la ISBD en relación con los contenidos”

Cerca de FRBR y FRAD, se aleja de AACR2

La RDA se declara conforme con los modelos FRBR y FRAD, y ésta afirmación tiene varias implicaciones.

– En primer lugar, el texto parte de las tareas que el usuario hace en el catálogo según los dos modelos: encontrar, identificar, seleccionar y obtener para los datos bibliográficos, y encontrar, identificar, clarificar y entender para los datos de autoridad.

– En segundo lugar, y tal y como queda implícito en las tareas anteriores, la RDA se aparta de las AACR2 en que no sólo incorpora disposiciones para la elaboración de registros bibliográficos, sino que también contempla instrucciones para la creación de registros de autoridad.

– En tercer lugar, contempla las entidades definidas en el modelo FRBR: los productos de la actividad intelectual o artística (la obra, la expresión, la manifestación y el ejemplar); los responsables del contenido intelectual o artístico de las entidades anteriores, su producción física, diseminación y custodia (persona, entidad corporativa y familia), y las entidades que pueden ser la materia de una obra (todas las anteriores además de concepto, objeto, acontecimiento y lugar). Así es que también se amplía el alcance de las AACR2 para incluir las materias, aunque, de momento, los capítulos dedicados a los atributos y las relaciones de concepto, objeto, acontecimiento y lugar están vacíos.

La adhesión a los modelos conceptuales de la *IFLA* impregna todo el texto, hasta el punto que la misma organización de las directrices e instrucciones se organiza alrededor de las entidades del modelo FRBR. La RDA se estructura en dos grandes apartados que comprenden 10 secciones y un total de 37 capítulos.

– Las primeras cuatro secciones están dedicadas a la anotación o registro de los atributos de manifestación e ítem, obra y expresión, persona, familia y entidad corporativa, y concepto, objeto, acontecimiento y lugar.

– Las seis secciones restantes se ocupan de la anotación de relaciones primarias (entre obra, expresión, manifestación e ítem); de las relaciones

a personas familias y entidades corporativas, y a conceptos, objetos, acontecimientos y lugares; y a las relaciones entre obras, expresiones, manifestaciones e ítems, entre personas, familias y entidades corporativas, y entre conceptos, objetos, acontecimientos y lugares.

El texto se complementa con doce apéndices y un glosario.

“La adhesión a los modelos conceptuales de la IFLA impregna todo el texto”

Conflicto con el formato MARC y la ISBD

La adopción de los modelos FRBR y FRAD tendrá repercusiones en la organización de la información bibliográfica y en la estructura de las bases de datos que se desarrollen en el futuro, y en principio facilitará la organización de los resultados según las entidades básicas.

Pero, como se ha anotado anteriormente, el formato MARC continúa siendo el formato predominante en el entorno bibliotecario, tanto en el proceso de introducción de datos, como en el intercambio. Así que mientras no haya un cambio sustancial en el formato y en los sistemas que lo gestionan, el catalogador tendrá que enfrentarse a una estructura del código realmente compleja –o como mínimo radicalmente diferente de la actual– que organiza la información bibliográfica de manera muy distinta al orden de entrada de datos en el registro MARC y también a su presentación en el catálogo.

“El catalogador tendrá que enfrentarse a una estructura del código realmente compleja –o como mínimo radicalmente diferente de la actual–”

En los registros que resultarán de la aplicación de la RDA los cambios no son tantos, y se concentran, excepto algunos casos, en la parte descriptiva del registro, en la anotación de atributos de manifestaciones e ítems. Es decir, son cambios que en la mayoría de los casos tendrán pocas consecuencias para el acceso.

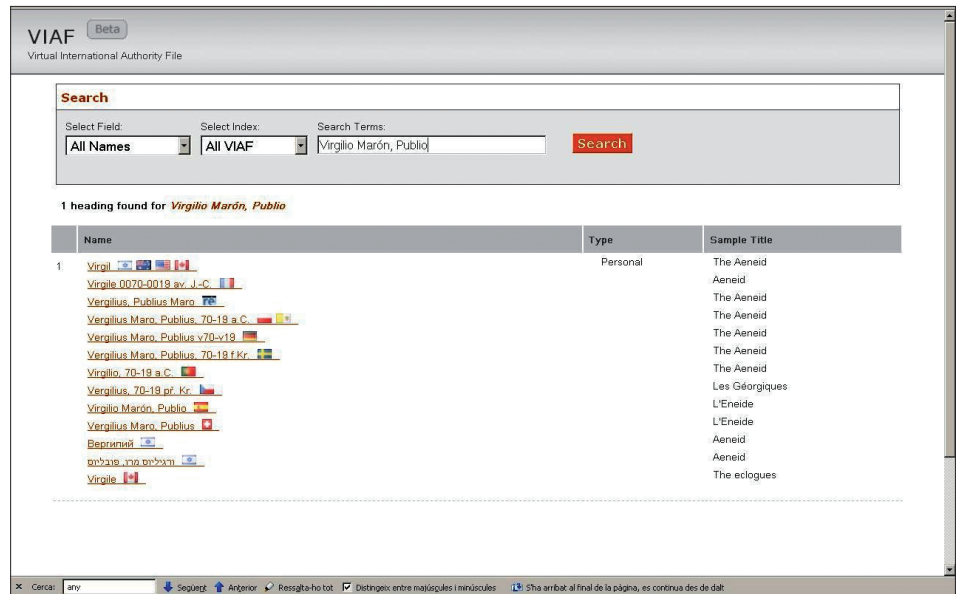
Sin embargo sí tendrán implicaciones en la copia y el intercambio de registros, ya que en general se apartan de las prescripciones de

la ISBD consolidada, y algunos consisten en la sustitución de abreviaturas normalizadas por indicaciones textuales en la lengua del centro catalogador (por ejemplo, *Lugar de publicación no identificado : editor no identificado*, en lugar de [S.l.] : [s.n.], o [y tres más] en lugar de [et al.] como se venía haciendo hasta hora y como aún dispone la ISBD consolidada).

Probablemente cuando RDA funcione en el marco de los estándares de la web semántica alcanzará su objetivo de facilitar la captura de metadatos creados por otras comunidades. Sin embargo, el hecho de que se aparte del estándar de la IFLA en diversos puntos, que elimine las abreviaturas y las sustituya por textos, y que los dos textos tampoco coincidan en los vocabularios que sustituyen al antiguo elemento *Designación general de material*, complicará la copia y el intercambio automático de registros procedentes del ámbito de las bibliotecas y hará más necesaria la revisión y edición de los registros copiados.

“La nueva normativa significará un coste muy alto para unas bibliotecas en tiempos de crisis”

La RDA aún tiene que demostrar sus beneficios respecto al código actual, pero en todo caso derivarán más bien de la aplicación de los estándares web que de las propias reglas de contenido. En este sentido, es probable que las AACR2 hubieran tenido un éxito similar si se hubieran adaptado al entorno de la web. En todo caso, la introducción de la nueva normativa significará un coste muy alto para unas bibliotecas en tiempos de crisis, y no sólo por la necesidad de adaptar los sistemas de gestión del catálogo a los nuevos modelos conceptuales o de la conversión de algunos de los registros de las base de datos a la RDA, sino sobre todo por la formación que será necesaria en el momento de implementarla y la elaboración de documentación adicional que requerirá su adopción.



Búsqueda de VIAF: la búsqueda por el nombre “Virgilio Marón, Publio” devuelve las autoridades equivalentes para este autor en las distintas bases de datos nacionales de autoridades <http://viaf.org>

Linked data

En la letra pequeña de los anales de la catalogación para el año 2010 constará toda una serie de actividades menos aparatosas que la publicación de la RDA, pero que tendrán su impacto para adaptar el entorno bibliotecario a la tecnología *linked data* y a los estándares web. Ya se ha hecho referencia al registro de los vocabularios RDA en el *Open Metadata Registry*, donde también se encuentran registrados vocabularios FRBR, FRAD e ISBD. Desde un punto de vista más práctico y operativo destacan, en primer lugar, los trabajos de la LC para ofrecer el acceso, a personas y a máquinas, a los *Library of Congress subject headings* y a otros vocabularios similares, como el *Thesaurus of graphic materials* o a las listas de lenguas, de países, de áreas geográficas, y de códigos de relación usados en el formato MARC 21. En la misma línea de adaptación a los nuevos entornos tecnológicos, un antiguo proyecto de la IFLA que actualmente gestiona OCLC, el *Virtual international authority file* (VIAF) usa la tecnología *linked data* para dar acceso y enlazar los datos de autoridad equivalentes de diversas bibliotecas nacionales y otros organismos.

Los proyectos se suceden uno tras otro y, así, *de mica en mica s’omple la pica* o, lo que es lo mismo, muchos pocos hacen un mucho. A la vuelta de la esquina quizá nos daremos cuenta que, de nuevo, las bibliotecas han aportado algo más que un grano de arena a la consolidación de esta web semántica que, hasta hace relativamente poco, era poco más que una entelequia acompañada de unos cuantos estándares.

Referencias bibliográficas

Byrne, Gillian; Goddard, Lisa. "The strongest link: libraries and linked data". *D-lib magazine*, 2010, v. 16, n. 11/12. DOI: 10.1045/november2010-byrne <http://www.dlib.org/dlib/november10/byrne/11byrne.html>

Godby, Jean. *Mapping ONIX to MARC*. Report and crosswalk produced by OCLC Research. OCLC, 2010. <http://www.oclc.org/research/publications/library/2010/2010-14.pdf>

Tabla de correspondencias:
<http://www.oclc.org/research/publications/library/2010/2010-14a.xls>

Library of Congress. Working Group on the Future of

Bibliographic Control. *On the record*. Library of Congress, 2008. <http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>

Open metadata registry
<http://metadataregistry.org>

RDA: Resource description and access. Chicago, Ill.: American Library Association, 2010-, 1096 pp. ISBN 978-0-8389-1093-1.

RDA toolkit: Resource description and access. American Library Association, cop. 2010. <http://www.rdatoolkit.org>

VIAF: *Virtual international authority file*. OCLC, cop. 2010. <http://viaf.org>

iralis®

International Registry for Authors:
Links to Identify Scientists

¡Ni un día más sin decidir tu firma!

Inscríbete en <http://iralis.org>

es:

- una **guía** para los autores hispanos para que firmen sus trabajos en el formato internacional usual
- una **base de datos** que registra las variantes de firma usadas por cada autor en diferentes épocas
- un **buscador** que usa automáticamente todas las variantes registradas

El formato de firma internacional

Con el apoyo de:

El profesional de la
información



F.1. Catalogación de objetos culturales y difusión digital del patrimonio

Por Tomás Saorín

14 octubre 2010

Saorín, Tomás. "Catalogación de objetos culturales y difusión del patrimonio por la Red".
Anuario ThinkEPI, 2011, v. 5, pp. 168-172.



Resumen: *Tras constatar el progreso de los museos en la informatización de sus colecciones, el autor advierte sobre la necesidad de asegurar que los esfuerzos de catalogación adoptan las soluciones tecnológicas adecuadas para producir valor en la web 2.0. Los metadatos deben poder definir objetos culturales complejos mucho más amplios que la descripción de una simple pieza o monumento, con más información contextual y del entorno. No se debe reducir la normalización documental al interior del museo.*

Palabras clave: *Patrimonio, Museos, Catalogación, Objetos culturales, Informatización, Digitalización, Metadatos, Web 2.0.*

Title: *Cataloguing cultural objects and digital dissemination of cultural heritage*

Abstract: *Having observed the progress of museums in the computerization of collections, the author warns about the need to ensure that the cataloguing efforts adopt appropriate technological solutions to produce value on Web 2.0. Metadata should be able to define complex cultural objects much more broadly than the simple description of a single piece or monument. Context and environment information should be added. Formal description processes should not be limited to inside the museum.*

Keywords: *Heritage, Museums, Cataloguing, Cultural objects, Computerization, Digitization, Metadata, Web 2.0*

VIVIMOS UN BUEN MOMENTO para los proyectos digitales. Si mezclamos las mejores características de cada uno de ellos, detectamos los puntos de innovación y exploramos analogías con otros servicios de éxito en la Red, quizá podamos decir que están alcanzando la masa crítica y la robustez tecnológica para dar el salto. ¿Hacia dónde?

Una respuesta oscura sería hacia un espacio semántico de objetos de información relevantes. Una respuesta más llana sería: crear el *Youtube* del patrimonio cultural.

¿Cuánta información sobre objetos culturales existe en la Red? ¿Alguien duda de la sobrea-bundancia de contenidos sobre patrimonio, arte o museos?

Sin embargo aún hay escasez de contenidos, tanto en cantidad como en calidad. Bastan unos pocos clicks para localizar unas fotos sobre el Teatro Romano de Cartagena¹, un artículo y varias menciones en *Wikipedia*, el sitio web específico del museo, una página didáctica en el canal de patrimonio del portal regional de contenidos digitales, unos vídeos en *Youtube*, la memoria de intervenciones arqueológicas, noticias de prensa

de su inauguración y musealización, y un largo etcétera.

“En la última década los museos españoles han dado el salto hacia la gestión de información documental de sus colecciones mediante el sistema integral *Domus*”

Con un poco más de esfuerzo accederíamos a bibliografía científica sobre el Teatro y su investigación arqueológica. Sin embargo el objeto informativo digital integrado “Teatro Romano de Cartagena” no existe.

¿Cómo sería tal objeto? Vamos a explicarlo para poner de relieve la importancia y el espacio de crecimiento que hay para la catalogación y creación de repositorios normalizados de objetos culturales.

Deliberadamente he optado por usar como ejemplo un sitio arqueológico que no pertenece



Figura 1. Fundación Teatro Romano de Cartagena
<http://www.teatroromanocartagena.org>

a la colección de un museo (aunque él mismo lo sea). Lo mismo ocurriría con monumentos, edificios, conjuntos históricos, etc., que pertenecen al ámbito de gestión del patrimonio cultural extra-museístico.

En la última década los museos españoles han dado el salto hacia la gestión de información documental de sus colecciones con la implementación del sistema integral *Domus*², que abarca todos los subsistemas informativos del museo, siendo de especial relevancia la catalogación o descripción de las piezas de las colecciones.

¿Cómo repercute en el público la mejora de la capacidad informativa de los museos? Mediante la publicación y consulta en línea de sus colecciones.

“Los modelos conceptuales para la descripción en bibliotecas, archivos y museos tienden a confluir”

El catálogo colectivo *Ceres*³ permite acceder a una información descriptiva y visual de las

colecciones de un número creciente de museos. Aquí nos estamos acercando a la existencia del objeto informativo digital de carácter cultural, dentro del museo, en sus colecciones. Este objeto informativo procedente del catálogo del museo está reutilizado en *Ceres*, en *Hispana* y en *Europeana*⁴. Dispone de un identificador OAI, y sus metadatos son adaptativos para usos en diferentes contextos. Pueden incorporarse a una exhibición virtual usando el software libre *Omeka*.

<http://omeka.org/>

Lo que sucede es que podemos hacer que vuelva a ser central la catalogación –y con ella la activación de oportunidades profesionales– en los proyectos de información cultural liderados por museólogos o historiadores del

arte. En realidad, catalogación con valor añadido tecnológico: asegurar que los esfuerzos de formalización adoptan las soluciones tecnológicas adecuadas para producir valor en la web 2.0, la web social o como quiera que la llamemos: la red que evoluciona e impacta.

¿Qué catalogación?

Primero decidir cuáles son los objetos, explicarlos individualmente y articularlos. Luego crear metadatos que permitan su uso contextualizado, estableciendo relaciones complejas de diferente naturaleza y nivel de análisis, entre colecciones propias y ajenas. Actualmente los modelos conceptuales para la descripción en bibliotecas, archivos y museos tienden a confluir.

Las características singulares de estas instituciones de la memoria, que gestionan las colecciones de objetos en los que permanece el conocimiento individual, institucional y colectivo, operan sobre una base común: su organización para permitir el conocimiento. En cada una de estas tres instituciones ha habido un ritmo de desarrollo diferente, en relación con los estándares de gestión de información sobre colecciones.

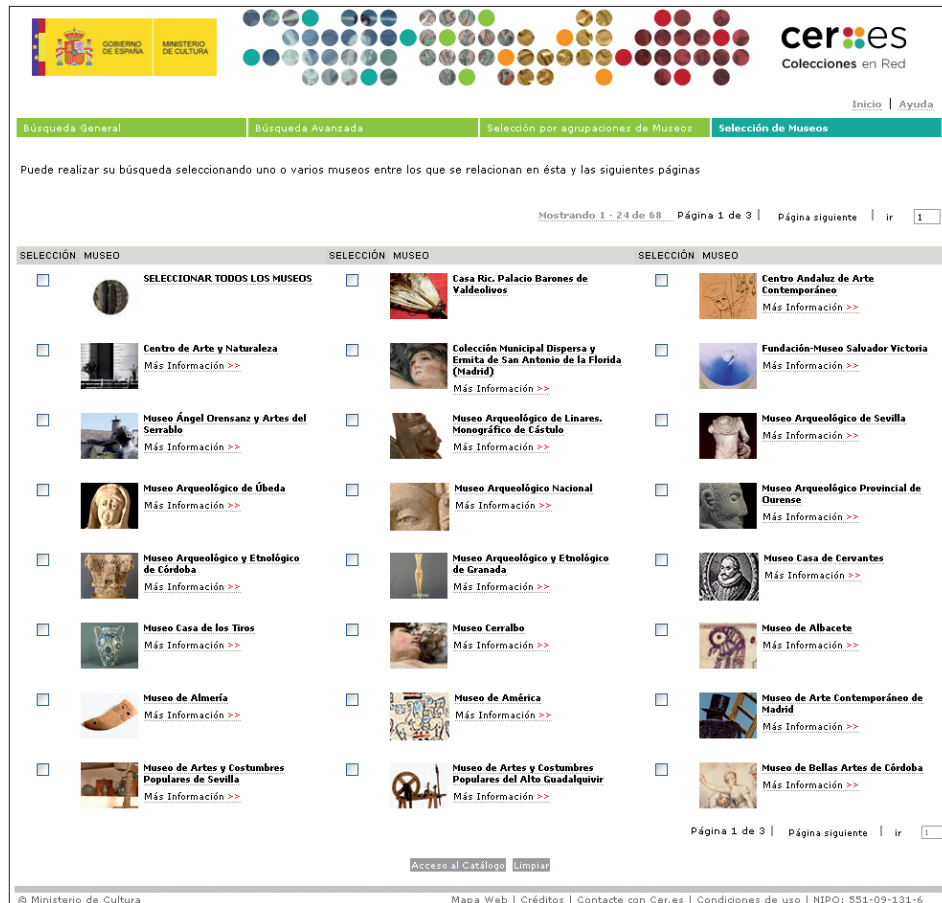


Figura 2. Red digital de colecciones de museos de España (Ceres) <http://ceres.mcu.es>

Las bibliotecas lideraron la normalización bibliográfica y la creación de instrumentos globalizados para compartir información: los catálogos colectivos.

En la última década hemos visto aumentar la eficacia que la normalización en el ámbito de archivos ha supuesto para el acceso a información digital sobre sus fondos, a partir de la familia de normas *Isad (G)* (*General international standard archival description*) o *Isaar (CPF)* (*International standard archival authority record for corporate bodies, persons and families*) y el concepto de descripción multinivel. Han impactado al mismo tiempo en los nuevos requerimientos funcionales para registros bibliográficos (*Frbr: Functional requirements for bibliographic records*) y de autoridades (*Franar: Functional requirements and numbering of authority records*).

En el campo de los museos y las colecciones culturales o patrimoniales, ya ha madurado la corriente de normalización e integración, cuyos fundamentos y posibilidades informativas podemos encontrar primero en la normalización de vocabularios controlados para arte y arquitectura y otros como *Nomenclature 3.0 for museum cataloguing*, *Art & architecture thesaurus (AAT)*⁵ o *Herein thesaurus project. European heritage network (Herein)*⁶, y posteriormente en la creación

de un esquema de metadatos como *VRA Core*⁷.

El sistema de información del museo conlleva una complejidad derivada de la propia diversidad de sus colecciones, servicios y tipología, asociada a las necesidades de documentación de la investigación, conservación, movimiento de colecciones y protección institucional del patrimonio cultural, sintetizado en el *Modelo conceptual de referencia del Cidoc (Gutiérrez-Usillos, 2010)*. Dentro de este sistema, la descripción de los contenidos adquiere una especificidad diferente del resto de aspectos de gestión, por lo que admite un enfoque globalizado que permite definir un estándar común.

Pero si reducimos la normalización documental al interior del museo, estaremos perdiendo una oportunidad de crear valor informativo duradero en todo el resto del patrimonio de valor cultural, que tanto tirón tiene en el sector económico del turismo cultural y la gestión cultural de la ciudad. ¿A qué objetos referencian las plataformas de realidad aumentada o las webs informativas basados en mapas? A *Wikipedia*.

diendo una oportunidad de crear valor informativo duradero en todo el resto del patrimonio de valor cultural, que tanto tirón tiene en el sector económico del turismo cultural y la gestión cultural de la ciudad. ¿A qué objetos referencian las plataformas de realidad aumentada o las webs informativas basados en mapas? A *Wikipedia*.

“El sistema de información del museo conlleva una complejidad derivada de la propia diversidad de sus colecciones, servicios y tipos”

El artículo de *Wikipedia* sobre un sitio cultural se ha convertido en el objeto informativo central, por delante del propio de las instituciones especializadas en cultura y patrimonio. ¿Es ésta la única vía? ¿Eclipsará *Wikipedia commons* a *Europeana*?

Difusión de las imágenes del patrimonio cultural

Una primera aproximación al concepto de “objeto cultural” nos sitúa más allá de las colec-

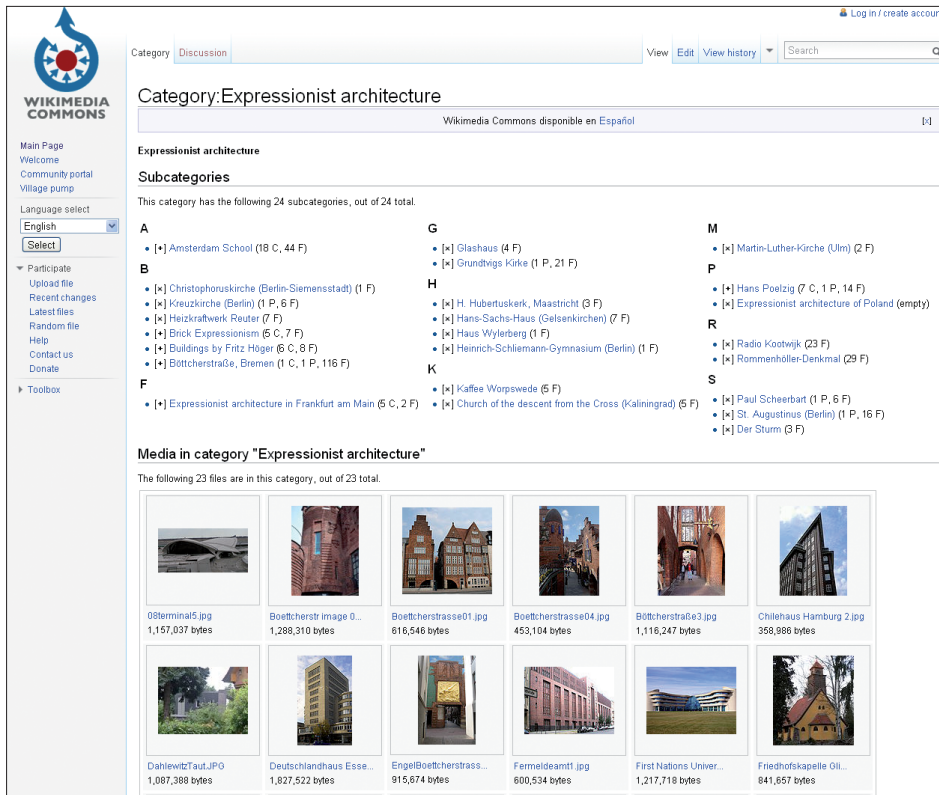


Figura 3. Arquitectura expresionista. Wikipedia commons http://commons.wikimedia.org/wiki/Category:Expressionist_architecture

ciones de los museos, aunque ellas constituyan el núcleo. No sólo las obras de arte, sino todos aquellos restos de la cultura material conservados en museos de todo tipo. Se consideran objetos culturales principalmente aquellos con un valor histórico-artístico, aunque bajo la definición de museos se acoja un buen número de piezas de valor científico (historia natural) o incluso espacios de creación contemporánea, o de exhibición de artes escénicas y musicales.

“El artículo de Wikipedia sobre un sitio cultural se ha convertido en el objeto informativo central, por delante del propio de las instituciones especializadas”

El objeto cultural nos remite al concepto de patrimonio cultural: edificios, monumentos, espacios monumentales y sitios arqueológicos. Se trata pues de objetos singulares, poseedores de un valor cultural establecido por las instituciones artístico-culturales, que son recurso primario para la investigación histórica y cultural, y que son contemplados desde la óptica de la museología contextual.

El objeto es portador de unos valores originales en un contexto histórico-cultural determina-

do, que los agentes que lo custodian y comunican tienen que conservar, fundamentalmente mediante información secundaria asociada. Además debe comunicarlo globalmente, en red, para cumplir su rol en la sociedad digital (Mateo-Rusillo, 2008).

El objeto cultural tiene pues unas características tangibles o directas, pero incluye diferentes capas de información intangible o contextual, que debe mantenerse unida a él para que mantenga a lo largo del tiempo su significado.

La American Library Association y la Visual Resource Association elaboraron *Cataloging cultural objects: a guide to describing cultural works and their images*,

que debería ser una lectura obligada no sólo para los equipos técnicos de los museos, sino para los responsables de difusión digital del patrimonio.

Existe información cultural, pero su naturaleza no formalizada limita su potencial de impacto, perjudicando a la calidad intrínseca que puedan tener. Al trabajar con el concepto de objeto cultural se permite delimitar una estructura de datos para describir de forma homogénea tanto un edificio histórico, una escultura monumental, un retablo, un cuadro, un tapiz, un sílex, un vestido, una urna funeraria, una inscripción o un yacimiento arqueológico. Y también un graffiti urbano, una escultura de rotonda, un barrio singular, un arco de triunfo, el Puente de los franceses, la torre Agbar, un parque minero o la ciudad de Toledo entera.

El objeto es además una entidad compleja, compuesta a menudo de partes con valor individual, del mismo o diferente autor o período, susceptibles de descripción independiente, que conformarán diferentes niveles de catalogación en objetos complejos, así como relaciones y agrupaciones apropiadas a la naturaleza de cada manifestación cultural. Piénsese en los diferentes elementos que pueden conformar la descripción de una catedral construida en diferentes fases, conforme a diferentes proyectos y estilos, que además alberga otras manifestaciones artísticas en su interior, articuladas a lo largo del tiempo como un todo cultural.

Se trata fundamentalmente de restos materiales, por lo que un elemento clave en ellos son sus representaciones gráficas, que constituyen un elemento descriptivo de alto valor, permitiendo un acceso digital al mismo, en cierto modo sustitutivo del acceso directo. Los bancos de imágenes procedentes de instituciones culturales son otro recurso sobre el que habría que reflexionar en términos de escasez, gestión de derechos, modelo de negocio y misión social.

**“Los sitios webs de los museos,
y muchos canales de patrimonio,
son islas informativas”**

¿Qué campo está abriéndose?

Cobrará fuerza la coexistencia en los sitios web de patrimonio de modos de difusión complementarios, apoyados en repositorios de contenidos digitales de objetos culturales, susceptibles de ser enlazados, remezclados y apropiados por los usuarios de la Red.

Las webs de los museos, y muchos canales de patrimonio, son islas informativas. Quizá vistosas, quizá adecuadas para sus fines inmediatos, pero al no ofrecer datos, desaprovechan un recurso de información cuya puesta a disposición de la sociedad generaría una inyección de valor, difusión y socialización diferente al impacto actual.

¿Cuántas webs de museos ofrecen código para embeber sus contenidos en blogs? ¿Cuántas proporcionan un url estable para sus piezas singulares? ¿Cuántas correrán el riesgo de aceptar etiquetado social?

Estas preguntas pueden parecer un llamamiento desde la web 2.0, pero por debajo de ellas están otras menos visibles:

¿Cuántas permiten navegar entre objetos de sus colecciones a otras colecciones? ¿Cuántas permiten descubrir relaciones interesantes? ¿Cuántas definen una licencia clara y amplia de reutilización de los metadatos y contenidos visuales? ¿Cuántas se ofrecen para *mashups*?

La reutilización de datos culturales pasa por la extensión de las prácticas de catalogación de objetos culturales más allá de los sistemas exclusivamente museísticos y su formalización mediante los recursos de la web semántica (RDF, skos).

También por la definición de condiciones claras y amplias de reutilización (*commons*, dominio público).

El objeto cultural será el nodo central para la organización y difusión no sólo de las colecciones museográficas sino, en un sentido más amplio, del patrimonio histórico-cultural. Este enfoque permite un espacio de conexión entre ambos ámbitos, rescatando el valor informativo del patrimonio que no forma parte directa de los fondos de museos, dibujando un amplio espacio para proyectos informativos culturales innovadores.

Referencias bibliográficas

Cataloging cultural objects: a guide to describing cultural works and their images. Chicago: American Library Association, 2006.

Gutiérrez-Usillos, Andrés. *Museología y documentación: criterios para la definición de un proyecto de documentación en museos*. Trea, 2010.

Mateo-Rusillo, Santos. *La comunicación global del patrimonio cultural: del marco teórico al estudio de casos*. Trea, 2008.

Nomenclature 3.0 for museum cataloguing. 3rd ed. of Robert G. Chenhall's *System for classifying man-made objects*. AASLH, 2009.

Recursos citados

1. *Teatro Romano de Cartagena*
<http://www.teatroromanocartagena.org>
2. *Domus (Documentación de los museos)*, Ministerio de Cultura.
<http://www.mcu.es/museos/CE/Funciones/Documentacion/DocumenatacionColecciones.html>
3. *Ceres (Colecciones en red)*. Red digital de colecciones de museos de España
<http://ceres.mcu.es>
4. *Europeana, Think Culture*
<http://www.europeana.org>
5. *Art & architecture thesaurus (AAT)*. J. Paul Getty Trust
http://www.getty.edu/research/conducting_research/vocabularies
6. *Herein thesaurus project*. European heritage network
<http://thesaurus.european-heritage.net>
7. *VRA Core metadata set v 4.0*
<http://www.vraweb.org/projects/vracore4>

F.2. Analítica de búsqueda: cómo y qué buscan los usuarios

Por **Jorge Serrano-Cobos**

8 noviembre 2010

Serrano-Cobos, Jorge. "Analítica de búsqueda: cómo y qué buscan los usuarios".
Anuario ThinkEPI, 2011, v. 5, pp. 173-176.



Resumen: Se comentan ejemplos de frases y términos de búsqueda utilizados por los usuarios del buscador Google, que pueden consultarse gratuitamente en su servicio Google Adwords. El análisis de qué y cómo busca el público puede ofrecer interesantes pautas para decidir los encabezamientos de materias de las bibliotecas, así como para la SEO (optimización de la posición de nuestra web en los resultados de los buscadores).

Palabras clave: Google Adwords, Términos de búsqueda, Hábitos, Mejora de la búsqueda.

Title: *Search analytics: how and what users are looking for*

Abstract: Some examples of search terms and phrases used by Google users, which are available free of charge on the Google Adwords service, are discussed. The analysis of why and how the public searches for information can offer interesting guidelines for deciding particular subject headings in libraries, and for search engine optimization (SEO) to improve the position of a website in the search engine results pages.

Keywords: Google Adwords, Search terms, Habits, Search improvement.

LAS QUEJAS SOBRE las carencias de los sistemas de búsqueda de los opacs son algo casi tradicional en nuestro entorno profesional¹ (si bien es verdad que además de trabajar en mejorar los algoritmos de recuperación y la presentación de los mismos a los usuarios, podrían mejorarse también los contenidos a recuperar²).

Cómo hacerlo es otra cuestión. Las iniciativas se han multiplicado en los últimos años³ y podemos encontrar, por ejemplo, que se usan:

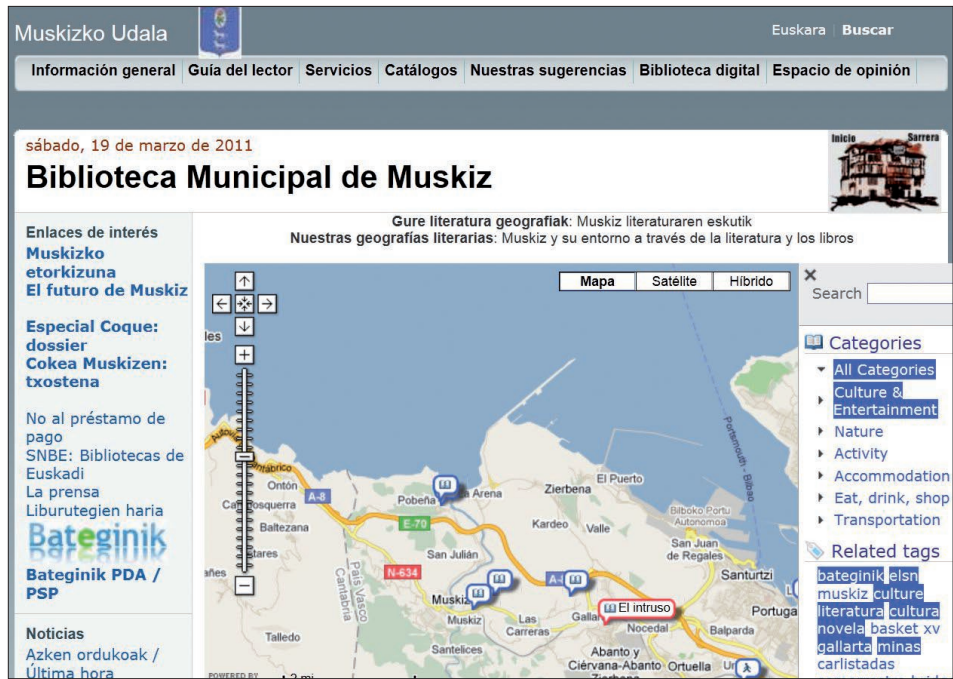
- *tags* para implantar la recuperación por ojeo o *browsing* en *LibraryThing*⁴ o *GoodReads*⁵;
- *rich results*⁶ en los resultados de la búsqueda de libros en *Google Books*, infiriendo qué busca el usuario en concreto⁷ mediante minería de uso o *web user mining*, con una tendencia a mostrar la última edición (probablemente para promover la compra online);
- *linked data* (dentro del movimiento de la web semántica) en rdf para enriquecer los contenidos de los libros, por ejemplo permitiendo recuperar libros de poemas por tipo de métrica⁸ o por caracteres de ficción⁹ e incluso libros que hablan en algún momento de un país o una ciudad¹⁰, lo que se puede hacer mediante *mashups* con mapas¹¹.

Pero a la hora de enriquecer los resultados de una búsqueda es importante entender cómo y

para qué buscan nuestros usuarios. Así, sabemos de los 3 tipos de intencionalidad en la búsqueda más conocidos¹²: el 75% de las búsquedas en la Web son informacionales, el 13% de navegación, y un 12% de transacción, aproximadamente¹³. Y también sabemos que en general la mayoría de los usuarios reformula sus búsquedas infructuosas mediante cambios en su contenido¹⁴, aunque harían falta estudios más actuales, una vez popularizados los últimos cambios en la presentación/facetación de resultados de los grandes buscadores de internet.

"A la hora de enriquecer los resultados de una búsqueda, es importante entender cómo y para qué buscan nuestros usuarios"

Otra cosa es que los usuarios y no usuarios de las bibliotecas actúen igual. En cuanto a los no usuarios que buscan en español y en España, mediante técnicas de analítica de búsqueda (*search analytics*¹⁵) podremos destacar algunos detalles curiosos de algunas de sus cadenas de



<http://www.muskiz-liburutegia.org>

búsqueda, en este caso simplemente consultando *Google Keywords (Adwords) Tool*¹⁶:

– Las búsquedas con errores gramaticales son muy comunes, tanto que en ocasiones se busca más por la suma de los posibles errores que por la palabra clave correcta. Pero *Google* hoy día casi elimina ese problema de las búsquedas en internet al corregir esos errores, y ciertos sistemas bibliotecarios integran software del tipo “quiso decir”.

– En otras ocasiones el usuario conoce la enorme variedad de contenidos que se puede encontrar y con su lenguaje natural intenta contextualizar y desambiguar el resultado que

busca. Por ejemplo, en búsquedas como “el caballero de la armadura oxidada libros”, “cien años de soledad libro”, o “don quijote de la mancha libro”. Es decir, el usuario faceta o filtra su búsqueda por formato, pero usando su lenguaje.

– En general se usan más los verbos en infinitivo (“comprar libros” más que “compra libros” o “compro libros”) pero hay que tener en cuenta que el español es un idioma que usa la forma activa y, al parecer, más aún en España.

“Es interesante analizar búsquedas como ‘autora harry potter’ o ‘romeo y julieta autor’”

– Dependiendo de lo que se busca, se utiliza más el plural que el singular, o viceversa. Por ejemplo, a la hora de recuperar información general o listas de elementos, se busca más en plural (12.100 veces al mes de media “lecturas para niños”, frente a 8.100 veces “lectura para niños”). Sin embargo en el caso de bibliotecas, para ahí comenzar

la búsqueda de los ítems que interesan, se usa más el singular, habitualmente acompañado de una localización para desambiguar (2.740.000 veces “biblioteca” frente a 450.000 de “bibliotecas”).

– Los sinónimos también deben ser tenidos en cuenta en nuestra búsqueda de la excelencia catalogadora: hemos de preguntarnos, por ejemplo, si los usuarios buscan lo mismo en el caso de “aprendizaje lectura” (2.900 búsquedas de media) que en “enseñanza lectura” (1.600).



<http://www.librarything.com>

– Deberíamos ser capaces de jugar con las cartas que da el desconocimiento de los usuarios de lo que buscan, que intentan dar un rodeo usando los datos que sí conocen. Así, podemos encontrar a usuarios que buscan “hogar del libro” en lugar de “casa del libro”, pero más interesantes –por el problema que pueden acarrear en la recuperación en un motor de búsqueda del catálogo– son las búsquedas como “autor harry potter” (1.900 búsquedas), “romeo y julieta autor” (1.600) o “autor de la eneida” (590). Si el usuario en estos casos lo que busca es una lista de libros de ese autor, o información biográfica del mismo, el opac puede que como mucho les aporte el título buscado; o ni eso si el sistema quiere encontrar todos los términos de la búsqueda.

– Por último, es interesante contrastar la poliseamia existente entre la intención dada al usar nuestros encabezamientos de materia con el que tiene quien realiza las búsquedas. Por ejemplo, cabe preguntarse cuántas de las 12.100 veces que se busca la expresión “escritores mexicanos” o de las 2.400 al mes que se busca “escritores hispanoamericanos” se hacen con la intención de encontrar uno o más libros que estudien a los escritores mexicanos o hispanoamericanos, o con la de encontrarse con listas de autores con sus obras asociadas.

Al parecer, según Google y su uso extenso e intensivo del *crowdsourcing*¹⁷, lo más probable es lo segundo¹⁸.

“Deberíamos ser capaces de jugar con las cartas que da el desconocimiento de los usuarios de lo que buscan”

Referencias bibliográficas

1. **Schneider, Karen G.** “How opacs suck, part 2: the checklist of shame”. *ALA TechSource*, 2006.

The screenshot shows the Goodreads website interface. At the top, there's a search bar and navigation links for 'home', 'my books', 'friends', and 'groups'. Below that, the 'listopia' logo and the text 'vote for your favorites' are visible. The main heading is 'Best Science Fiction Books'. A sub-heading reads: 'The best science fiction books of all time. Please only add books that are science fiction, not fantasy or horror. 1700 books on the list, 6493 voters, list created: May 29th, 2008'. There are tags for 'sci-fi, sciencefiction y sf' and buttons for 'all votes' and 'add books to this list'. The list contains four items:

Rank	Book Title	Author	Avg Rating	Number of Ratings	Score	Number of Voters
1	Ender's Game (Ender's Saga, #1)	Orson Scott Card	4.17	139625	213257	2154
2	Dune (Dune Chronicles, #1)	Frank Herbert	3.93	112734	164686	1672
3	The Hitchhiker's Guide to the Galaxy (Hitchhiker's Guide, #1)	Douglas Adams	4.02	184235	156277	1593
4	1984	George Orwell	4.00	342377	115806	1188

<http://www.goodreads.com>

<http://www.alatechsource.org/blog/2006/04/how-opacs-suck-part-2-the-checklist-of-shame.html>

2. **Castillo-Vidal, Jesús.** “Descenso del número de visitas a las webs de bibliotecas y opacs”. *IweTel*, 29 sept. 2010.

<http://www.mail-archive.com/iwetel@listserv.rediris.es/msg04422.html>

3. **Serrano-Cobos, Jorge; Sellés, Alicia.** “Catálogos online y portales bibliotecarios: ¿un reto para la integración?”. *Mi biblioteca*, 2009, n. 19, pp. 70-75.

4. LibraryThing

<http://www.librarything.com/work/8653840>

5. Goodreads

<http://www.goodreads.com/user/new?remember=true>

6. **Catacchio, Chad.** “Google Books to get ‘rich results’ starting today”. *The next web (TNW)*, 1 Nov. 2010.

<http://thenextweb.com/google/2010/11/01/google-books-to-get-rich-results-starting-today>

7. **Madrigal, Alexis.** “Inside the Google Books algorithm”. *The Atlantic*, 1 Nov. 2010.

<http://www.theatlantic.com/technology/archive/10/11/inside-the-google-books-algorithm/65422/#>

8. Freebase. Poetic verse form

http://www.freebase.com/view/book/poetic_verse_form

9. Freebase. Book character

http://www.freebase.com/view/book/book_character

10. Open Library

http://openlibrary.org/subjects/place:new_york

11. Biblioteca Municipal de Muskiz.
<http://www.muskiz-liburutegia.org/mapalit.html>
12. **Broder, Andrei**. "A taxonomy of web search". IBM research.
<http://www.sigir.org/forum/F2002/broder.pdf>
13. **Jansen, Jim**. "Classifying the user content of web queries using k-means clustering". *Web search*, 1 Nov. 2010.
<http://jimjansen.blogspot.com/2010/11/classifying-user-intent-of-web-queries.html>
14. **Young Rieh, Soo; Hong (Iris) Xie**. "Patterns and sequences of multiple query reformulations in Web searching: a preliminary study (2001)". En: *Proc. of the 64th Asist annual meeting*, Washington DC, 2001, pp. 246-255.
http://rieh.people.si.umich.edu/papers/rieh_asist2001.pdf
15. **Serrano-Cobos, Jorge**. *Search analytics* [presentación].
<http://www.slideshare.net/jorgeserranocobos/search-analytics-2219355>
16. Google Adwords
<https://adwords.google.com/select/KeywordToolExternal>
17. Google "define:crowdsourcing"
<http://bit.ly/aZHVXU>
18. Google "escritores hispanoamericanos"
<http://bit.ly/aTbEW3>



RedIRIS IWETEL

Foro para profesionales de bibliotecas y documentación

<http://www.rediris.es/list/info/iwetel.html>

***IweTel*, foro de información y debate de la biblioteconomía y la documentación**

Fundada por Tomàs Baiget en 1993, *IweTel* es la lista pionera en español de los profesionales de las bibliotecas, documentación, bases de datos y sistemas de información en general.

Al principio se alojó en *Sarenet* y en 1998 pasó a *RedIRIS*. Posteriormente se han ido creado otras listas más especializadas como *Arxiforum* (archivos), *Bib-Med* (información bio-médica), *Bescolar* (bibliotecas escolares), *Incyt* (indicadores científicos), etc., pero *IweTel*, con más de 5.000 miembros, es la lista de referencia, el medio de comunicación básico y central para los profesionales de la información.

En la lista se cumple la conocida regla del 80/20 (el 80% de los mensajes los genera el 20% de los inscritos), o su reciente reformulación a 90, 9, 1%: el 90% de los inscritos son pasivos, casi nunca envían nada, el 9% (unos 360) participa alguna vez, y existe un 1% (50 personas) que genera la mayoría de mensajes.

Con el aumento de inscritos y el número de mensajes (algunas semanas se distribuyen más de 100) fue necesario hacer la lista moderada, y en ello estamos los 4 firmantes, intentando aplicar nuestro sentido común para decidir cuáles se aprueban y cuáles no, y evitando los mensajes repetidos. Rechazamos alrededor de un 15-20%, lo cual a veces provoca quejas de sus autores, y para dirimir las dudas se creó un Consejo Asesor formado por veteranos de la lista, a quienes los moderadores pedimos consejo.

La lista cumple los dos objetivos básicos típicos: tablón de anuncios (conferencias, cursos, publicaciones, noticias) y foro de debates. Además se usa como sistema abierto de evaluación por pares (*open peer review*) de las notas que los miembros del think tank *ThinkEPI* envían periódicamente a la lista para su pública crítica y discusión. Esas notas y los principales mensajes que generan se publican cada año reeditados en el *Anuario ThinkEPI* de la editorial *EPI*.

Con los cambios tecnológicos habidos a lo largo de estos años y, más recientemente, con las nuevas plataformas web 2.0, se ha planteado muchas veces si las listas de correo se han hecho "obsoletas". La verdad es que pensamos que una lista sigue siendo el medio ideal de comunicación de una comunidad profesional: rápida, limpia, discreta y eficaz, lejos de la faramalla de las redes sociales, también muy interesantes y útiles pero para otras cosas.

Más información e inscripciones:
<http://www.rediris.es/list/info/iwetel.html>

Javier Leiva-Aguilera (*Catorze.com*), **Paco López-Hernández** (*Universidad Carlos III de Madrid*), **Isabel Olea** (*Universidad de León*) y **Tomàs Baiget** (*El profesional de la información*).

F.3. Google: el buscador que era algo más

Por José-Antonio Millán

11 enero 2011

Millán, José-Antonio. "Google: el buscador que era algo más".
Anuario ThinkEPI, 2011, v. 5, pp. 177-179.



Resumen: Se comenta que los cambios que el buscador Google ha ido introduciendo en los resultados de las búsquedas para "adivinar" las preferencias de los usuarios, y acercarse más a sus pretendidos objetivos de búsqueda (y también en aras de la publicidad) pueden entorpecer algunos usos del buscador. Por ejemplo, a diferencia de cuando era menos "inteligente", ahora da resultados sesgados hacia lo local próximo al usuario, o corrige el deletreo de términos frecuentes cuando se buscan los infrecuentes.

Palabras clave: Buscadores, Conocimiento del usuario, Ordenación de los resultados, Sesgo, Búsqueda inteligente.

Title: *The search engine that was something more*

Abstract: The changes that Google have introduced in the search results for "guessing" user's preferences, and trying to give results closer to their intended search targets (and also for the sake of advertising), can hinder some uses of the search engine. For example, unlike when it was less "intelligent", now Google gives biased results toward local information deemed to be nearer to the user, or corrects the spelling of common terms when looking for the unusual ones.

Keywords: Search engines, User knowledge, Results ranking, Bias, Intelligent search.

DURANTE MUCHOS AÑOS hemos estado utilizando Google, hasta tal extremo que ya formaba parte de nosotros: eran las antenas con las que palpábamos el universo, desde el circundante hasta el remoto.

Su fiabilidad global –no tanto en la ordenación de los resultados como en el hecho de que nos llevara a apariciones de lo que buscábamos, dondequiera que se encontraran– lo convertía en una herramienta realmente útil, y no sólo para fines de búsqueda estricta. Por ejemplo, los estudios sobre presencia de lenguas en las redes hacían uso de la consulta de Google, convertido entonces en una auténtica "base de datos" de ocurrencias de términos. Es sabido cómo, utilizando la metodología de buscar términos exclusivos de una lengua, *Funredes* trazó un mapa de la presencia relativa de las lenguas neolatinas en relación con el inglés, en los primeros años de la pasada década.

<http://funredes.org>

Hacia 2007 sin embargo los datos del buscador empezaron a presentar discrepancias. Las mismas consultas hechas por distintas personas en diferentes lugares daban distintos resultados. La diferencia concernía no sólo a la ordenación de las repuestas sino también (y sobre todo) a la cantidad de resultados y a su procedencia. Todo

parecía indicar que se estaban utilizando para la ponderación y quizás para la misma configuración del universo de búsqueda:

- a) anteriores consultas del usuario, posiblemente con análisis de qué resultados del buscador le habían llevado a un clic, y
- b) el entorno geográfico desde donde se hacía la consulta.

Además, y como el conjunto de servicios de Google ya afectaba prácticamente a todas las dimensiones de la acción personal (desde búsquedas y visionado de vídeos a correo, búsqueda en libros y mapas, y suscripción a rss), la suma de experiencias que podía tener en cuenta la respuesta a una simple pregunta en el buscador era enorme. Como resultado de todo esto, Google apuntaba cada vez más al entorno inmediato del usuario.

"Ni siquiera con la 'búsqueda avanzada' y utilizando restricciones podemos estar seguros de que se va a respetar la literalidad de nuestra consulta"

La razón parece clara: al buscar la optimización del sistema de anuncios que acompañan a los resultados, todo sesgo hacia accesos, bienes y servicios próximos al usuario podía estar acompañado más eficazmente por la propia publicidad de éstos. Y además –y quizás para una amplia base de usuarios– el servicio del buscador estaba siendo realmente “mejor”, en el sentido de que llevaba a resultados más directamente relacionados con su entorno. Así, ante la consulta de un nombre propio sin más identificación, *Google* puede proponer (muchas veces, insistamos, acertadamente) una calle con ese nombre en la ciudad desde la que se ha preguntado.

La expansión semántica y morfológica de las consultas hace también que el universo de respuestas se expanda no siempre en la dirección que uno desearía. Un nombre propio inusual se puede interpretar como la variante mal escrita de otro que el buscador supone más verosímil, dado nuestro historial y localización. Una palabra en una lengua extranjera puede correr idéntica suerte. Y la cuestión es que ni siquiera echando mano de la “búsqueda avanzada” y utilizando todo tipo de restricciones podemos estar seguros de que se va a respetar la literalidad de nuestra consulta, y no podremos obtener lo que *Google* nos daba hace no demasiado tiempo.

El problema –ya lo estamos viendo– se plantea básicamente para las personas que hacemos un uso de *Google* para el que quizá no estaba destinado, pidiéndole un servicio de cobertura universal de apariciones en la Web, pero es que hubo una época en que prestó este servicio y, de momento, no tiene un sustituto claro.

Quienes hemos utilizado *Google* para rastrear piezas de bibliografía arcana, que al final hemos detectado en el sitio de una universidad checa o en una página personal americana; quienes hemos usado el buscador a la pesca de neologismos en las lenguas, o de usos creativos o paródicos de las palabras, estamos privados en gran medida de lo que teníamos. Y muchos usuarios menos especializados están también privados de algo muy importante: la posibilidad de toparse con algo que no buscaban, lo que se conoce como serendipia.



Este servicio de búsqueda que se ha trocado de golpe en algo más local y provinciano creo que, entre otras cosas, le hace un flaco servicio a la extensa comunidad de hispanohablantes, con su enorme dispersión geográfica, que había confiado (tal vez ingenuamente) en la posibilidad de que el común digital, y las herramientas que lo servían, fuera un elemento que favoreciera la mutua interrelación y comprensión.

* * *

Ya no confío en *Google*

Por **Lara Rey**

Desde 1998 utilizo habitualmente *Google*. Lo tengo como página de inicio en mis navegadores. En los últimos años, para mí, era más rápido encontrar una página a través de *Google* que en mi carpeta de favoritos.

Últimamente, desde hace quizá casi dos años, *Google* ha cambiado. Supongo que es la misma sensación que tienen los padres cuando son plenamente conscientes de que sus hijos ya no son esos niños pequeños e ingenuos. No sé explicar punto por punto en qué ha cambiado, pero sé a ciencia cierta que no es el que era.

No me interesa ver en *Google* las últimas actualizaciones de hace 10 segundos en las redes sociales o en *Twitter*. Creo que para ese tipo de búsquedas hay otros buscadores u otras opciones. No me interesan las búsquedas tan locales y tan mediatizadas por donde te encuentras geográficamente. A mí me gustaba *Google* cuando era “mundial”. Me gustaría resetear mi historial de búsquedas, perfil y preferencias que en estos años *Google* ha ido acumulando sobre mí y ser capaz de empezar de cero.

En mi trabajo varias personas utilizamos *Google Alert*. En algunos casos lo configuramos el mismo día y con las mismas palabras clave. Y es curioso porque no nos llegan los mismos resultados...Y ya no sólo eso. Una misma búsqueda en dos ordenadores ofrece dos resultados.

No me acaba de gustar cómo está evolucionando *Google* como buscador. Pero no tengo a este nivel una alternativa mejor.

Hace tiempo **Juan Freire** escribió un post en la ya no disponible *Generacionred*, que se titulaba "¿Hasta cuándo podremos confiar en *Google*?". Y mi respuesta es: yo ya no confío en *Google*.

<http://www.atinachile.cl/content/view/39061/Hasta-cuando-podremos-confiar-en-Google.html>

"Los analfabetos digitales, de cualquier edad y condición, consideran a *Google* más próximo al demonio.

Los inmigrantes digitales tienen una extraña relación con este antiguo buscador convertido en cerebro universal y dios virtual.

Pero los nativos digitales (que salvo algún pionero precoz, rara vez han cumplido los 30 años), tienen otro tipo de relación social, intelectual y afectiva con *Google*. Los nativos usan *Google* pero odian a *Google*. ¿Paradójico?"

http://nomada.blogs.com/jfreire/2006/06/google_ha_muert.html

Lo que más me llama la atención es que cuanto más sabe *Google* de mí, de mis búsquedas, de mi entorno..., menos útil me resulta.
lararey@ferroatlantica.es

Soluciones alternativas a comportamientos extraños en *Google*

Por **Isidro F. Aguillo**

Estimados "agentes en nómina de *Microsoft*": Es cierto que *Google* está siendo objeto de investigación por ciertas prácticas, que sus algoritmos son secretos y que el buscador sigue evolucionando de una manera más bien opaca para el usuario final. Sin embargo es una herramienta formidable y con el adecuado conocimiento se le puede sacar mucho provecho. En concreto citáis algunos comportamientos extraños o nuevos que tienen explicación y/o solución alternativa:

– Resultados diferentes en dos ordenadores uno al lado de otro o en el mismo ordenador apenas pasados unos segundos: *Google* no puede responder los millones de peticiones desde un único sitio, por lo que utiliza un sistema distribuido. Son los llamados *Google data centers*, cada uno con una o varias IP numéricas diferentes. Y además al parecer con bases de datos actualizadas de diferente manera. Eso significa que cuando utilizas la

dirección te puede responder el *data center* más próximo o el menos saturado, que no tiene por qué ser el mismo que te contestará momentos después o a tu compañero en otro ordenador al lado. ¿Y con resultados distintos? Se puede evitar esto utilizando siempre la misma IP que se puede obtener de las listas que aparecen en la Web.

– Personalización "geográfica": los algoritmos de *Google* son muy complejos, incluyendo los de localización y personalización. Mucha gente cree que dependen de utilizar diferentes dominios (*google.com*, *google.es*) pero en realidad parece ser que el filtro más importante es el idioma de la interfaz. Es decir, para tener resultados más "internacionales" hay que utilizar la interfaz en inglés (ojo, no buscar en inglés o filtrar por idioma).

Espero que os sirva y recuperéis el cariño por *Google*.

Dominar la herramienta de trabajo

Por **Isidro F. Aguillo**

Debemos abordar este tema de manera profesional. *Google* es una empresa comercial cuyo principal objetivo es ganar dinero.

Personalmente me puede interesar o no la trayectoria y problemas legales, económicos, éticos o morales de *Google*, pero como documentalista en ejercicio mi trabajo consiste en conocer y dominar la herramienta que de forma mayoritaria utilizan los usuarios de internet. Es decir, cuando estoy en mi puesto de trabajo no utilizo *Google* como lo haría mi sobrino de 7 años, sino que aprendo a sacarle partido, a entender cómo funciona y a conocer sus debilidades y fortalezas.

Por ejemplo:

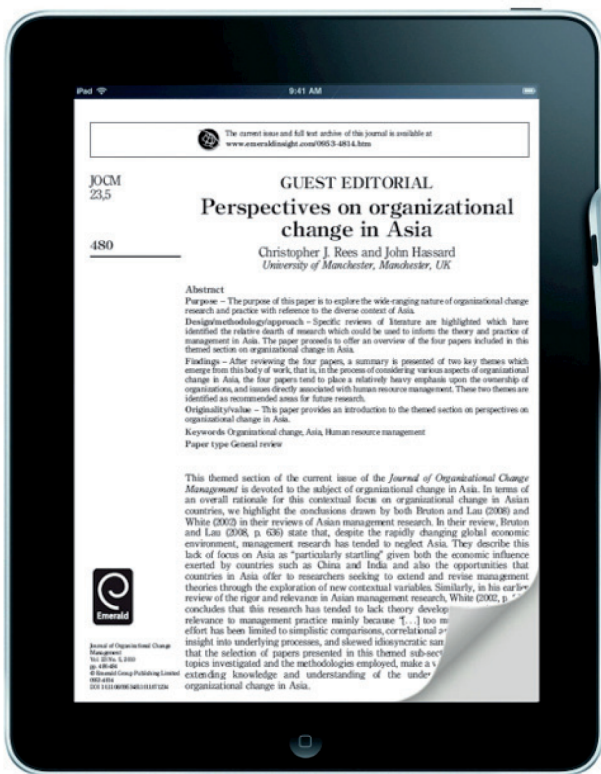
Millán se pregunta cómo recuperar los resultados de "siempre", no los personalizados geográficamente. Fácil, no depende de *google.com*, ni de *google.es*. Es tan simple y "extraño" como cambiar el idioma de la interfaz. No es obvio, ni intuitivo, pero no me acuerdo de la madre del ingeniero jefe de *Google* sino que busco (y encuentro) una solución.

Otro ejemplo:

Mucha gente se pregunta por qué los resultados de *Google* varían tanto incluso en una misma máquina. Un poco de exploración me permite descubrir que hay cientos de *Googles* y que me pueden responder según su disponibilidad. No me interesa ni la arquitectura distribuida ni los vericuetos de los protocolos, sólo me hace falta saber que si siempre utilizo la misma ip los resultados no son tan irregulares.

isidro.aguillo@cchs.csic.es

Una biblioteca al alcance de su mano



Emerald eBookSeries



A partir de ahora el acceso en línea a las dos colecciones de libros digitales de Emerald está disponible también en formato EPUB, además de en HTML y PDF. El formato EPUB amplía enormemente las

funcionalidades de acceso de los libros digitales de Emerald a través de dispositivos móviles, tales como lectores de libros electrónicos, teléfonos móviles, iPads, etc.

Las colecciones de libros electrónicos de Emerald constituyen un recurso excelente para mejorar y complementar los fondos actuales de su biblioteca.

Están organizadas en dos colecciones:

- **La colección de Negocios, Gestión y Economía**, que cubre, entre otras, las disciplinas siguientes: *Estrategia, Economía, Contabilidad y finanzas, y Gestión de recursos humanos.*
- **La colección de Ciencias Sociales**, que cubre, entre otras, las disciplinas siguientes: *Educación, Gestión ambiental/Medio ambiente, Gestión sanitaria/Servicios de salud, Lenguaje y lingüística, y Sociología y política pública.*

Si está interesado en obtener más información sobre ellas y sus opciones de suscripción o compra, póngase en contacto con Jordi Caralt, Director Comercial de Emerald Group Publishing para España, en jcaralt@emeraldinsight.com o en el teléfono 00 44 1274 785126.

G. PUBLICACIÓN CIENTÍFICA

Informe de situación

2010 a vista de pájaro: publicación científica, OA, indicadores e informetría

Por Isidro F. Aguillo

Aguillo, Isidro F. "2010 a vista de pájaro: publicación científica, OA, indicadores e informetría".
Anuario ThinkEPI, 2011, v. 5, pp. 181-186.



Resumen: *Revisión de la situación de la publicación científica europea y sobre todo española. Entre otros temas se trata de informetría, indicadores, rankings universitarios, bases de datos WoS y Scopus, acceso abierto y repositorios.*

Palabras clave: *Publicación científica, Rankings, Bibliometría, Indicadores, Bases de datos de citas, Acceso abierto.*

Title: *2010 bird's eye view: scientific publishing, OA, indicators and informetrics*

Abstract: *Review of the scientific publication situation in Europe and especially in Spain. Among other issues, infometrics, indicators, university rankings, databases, WoS and Scopus, and open access repositories are discussed.*

Keywords: *Scientific publication, Rankings, Bibliometrics, Indicators, Citation databases, Open access.*

REDUCIR LOS ACONTECIMIENTOS MÁS RELEVANTES de un año a una mera sinopsis no es tarea fácil, puesto que hay que reflejar tanto hitos señeros, cuya importancia es difícil de valorar en el corto plazo, como tendencias y evoluciones, cuyo desarrollo excede el marco temporal propuesto.

Por ello nos centraremos en una enumeración, ni completa ni objetiva, de algunos sectores cuyas actividades o resultados merecen ser destacados y de algunas noticias, relevantes o no, que merecen nuestro comentario.

También resulta complicado delimitar unas temáticas heterogéneas entre sí y que incluso, en algún caso, dan lugar a discrepancias respecto a su denominación. Por ejemplo, el término *informetría* se ha asociado a veces a la descripción matemática de las distribuciones, la parte más técnica, mientras que ahora parece que ya ha triunfado una definición más amplia, que incluye a todas las disciplinas cuantitativas. Batalla ganada desde el momento en que el *Journal of informetrics*¹ se ha convertido, en apenas cuatro años de vida, en el estandarte de bibliómetras y cienciómetros. En 2009 su factor de impacto ha

sido de 3,379, una cifra nunca alcanzada por las revistas más prestigiosas de esta disciplina.

Acceso a WoK y Scopus

El acontecimiento local más relevante se produjo al final de año, con la renovación de la licencia nacional de las bases de datos de *ISI/Thomson (Web of Knowledge)* y la contratación de una licencia similar con *Elsevier* para el acceso a la base *Scopus*. En época de crisis económica se agradece el esfuerzo de *Fecyt* para garantizar la disponibilidad de estas dos herramientas fundamentales, máxime si como se ufana la propia *Fecyt* se ha logrado un ahorro de 4 millones de euros.

Sin embargo, ha de hacerse notar que se está pagando por una serie de prestaciones cuya rentabilidad habría que considerar más detalladamente. Como instrumentos de recuperación de información bibliográfica no ofrecen ventaja alguna sobre otros productos, ya que por ejemplo, los gratuitos *Google Scholar* y *Dialnet* ofrecen mayor cobertura.

Respecto a su uso como herramienta de evaluación, hay que recordar que en este país todavía se utiliza el factor de impacto (se miden citas esperadas en vez de citas reales, ¡qué barbaridad!) que sólo requiere contratar el muchísimo más económico *Journal Citation Reports (JCR)* [el equivalente *Scimago Journal Rank (SJR)*, que se basa en datos de *Scopus*, una fuente todavía fríamente recibida por nuestros gestores, es gratuito].

“España es el 9º país en papers y el 11º en citas recibidas (2000-2010)”

Queda su uso para la medida de la producción, puesto que proporciona la afiliación de todos los autores, algo que obviamente no requiere que se pague el acceso para comprobación y vanidad a todos y cada uno de los científicos españoles. No se trata de una *boutade*, para realizar el sensacional *In-Recs*² sólo se necesita una licencia y un pequeño y entregado grupo de investigación.

España en el mundo

El balance global de la producción científica española en revistas de impacto nos sitúa en la 9ª posición mundial³ teniendo en cuenta el número

Rank	Country	Papers	Citations	Cites per paper
1	USA	2,967,957	46,796,090	15.77
2	JAPAN	770,252	7,877,699	10.23
3	GERMANY	762,599	9,960,100	13.06
4	PEOPLES R CHINA	719,971	4,227,779	5.87
5	ENGLAND	679,394	9,979,737	14.69
6	FRANCE	542,293	6,660,630	12.28
7	CANADA	430,856	5,619,293	13.04
8	ITALY	409,232	4,770,753	11.66
9	SPAIN	315,420	3,256,075	10.32
10	AUSTRALIA	284,250	3,359,748	11.82
11	RUSSIA	267,319	1,243,711	4.65
12	INDIA	266,230	1,497,065	5.62
13	SOUTH KOREA	254,599	1,767,799	6.94
14	NETHERLANDS	239,892	3,687,829	15.37
15	BRAZIL	190,801	1,197,953	6.28
16	SWEDEN	174,052	2,548,046	14.64
17	SWITZERLAND	172,904	2,873,881	16.62
18	TAIWAN	162,197	1,115,524	6.88
19	POLAND	144,559	954,220	6.6
20	TURKEY	138,345	687,389	4.97

SOURCE: ESSENTIAL SCIENCE INDICATORSSM FROM THOMSON REUTERS. TIME PERIOD: 2000-AUGUST 31, 2010 (FOURTH BIMONTHLY PERIOD OF 2010).

Total de papers publicados y citas recibidas, por países, enero 2000-agosto 2010³

de artículos publicados, aunque como es conocido el balance del impacto obtenido no es tan halagüeño (11ª por número de citas). El incremento de la colaboración internacional no es menoscabo para estas cifras globales, aunque debería enfriar el entusiasmo que ciertos sectores proclaman al realizar series temporales. Los porcentajes de la citada colaboración podrían, conveniente y correctamente ajustados, hundir las generosas cifras de crecimiento en la producción española, y más aún las de la latinoamericana, situación que muchos quieren ocultar por su gravedad.

La incorporación de nuevas revistas fuente de nuestro país, que ya suman 167 según el último balance⁴, también deja un sabor agridulce, pues es probable que dichas revistas locales penalicen severamente nuestra visibilidad global (como bien señala el genial, a veces, **Anthony Van Raan**⁵), pues los artículos que no están en inglés bajan la posición nacional en el ranking mundial. Las polémicas actuaciones de *Fecyt*, en este caso defendiendo las revistas nacionales, pueden tener estas consecuencias.

“Los artículos que no están en inglés bajan la posición de los países en el ranking mundial”

El año 2010 no permite todavía realizar una estimación del impacto real de la crisis económica mundial en la actividad científica, aunque las políticas expansivas de gasto en I+D de países como Francia y Alemania o la fortaleza de las economías emergentes (BRIC = Brasil, Rusia, India y China) nos hacen albergar cierto nivel de optimismo. Las dudas sobre el comportamiento ético de ciertas instituciones chinas con una producción abultada⁶ no debe hacernos pensar que el potencial chino es menor. No olvidemos el alto número de autores chinos residentes en EUA que firman con afiliación estadounidense.

Los datos de *Scopus*⁷ muestran un crecimiento continuado de la producción española, que en 2009 casi alcanzó los 59.000 entradas y que los datos provisionales de 2010 la estiman en más de 64.000, lo que supone un incremento anual del orden del 10%. El grupo *Scimago*⁸ ha logrado con sus productos bibliométricos colocarse como referente mundial en la explotación de los datos de *Scopus*⁹. Esta base de datos con citas, aunque sigue recibiendo críticas por la opacidad de sus criterios de selección, ya ha roto de hecho el monopolio de *ISI/Thomson*.

A ello también ayuda, aunque todavía modestamente, *Scholar*¹⁰, la gratuita base de datos de

citas del gigante Google. Entre los bibliómetros, *Scholar* se ha vuelto muy popular gracias al programa *Publish or Perish*¹¹, que ha facilitado de forma notable la explotación de los resultados. En la actualidad el programa, que para uso personal se distribuye gratuitamente, ya va por la versión 3.1, y sigue incorporando nuevas funciones.

“El grupo español Scimago, creador del SJR, es un referente mundial en la explotación de los datos de Scopus”

Un hito destacable ha sido la celebración de la primera *European Summer School for Scientometrics*¹², celebrada en Berlín en junio de 2010. El éxito, al que no es ajeno como eficaz organizador nuestro hombre en la *Universidad de Viena*, **Juan Gorraiz**, ha permitido que la iniciativa tenga continuidad y haya convocada una nueva edición en septiembre de 2011, en Viena.

Rankings institucionales

A nivel institucional, cabría recurrir a los rankings de universidades y centros de investigación, pero el año 2010 no ha sido especialmente generoso en esta área debido a la falta de profesionalidad de algunos de sus editores y la igualmente sorprendente impericia de muchos gestores académicos en la interpretación de los datos publicados.

La primera edición tras romper el acuerdo con la empresa QS¹³ del tan esperado y publicitado ranking de la revista inglesa *Times Higher Education (THE)*¹⁴ (nada que ver con el prestigioso periódico *The times*) ha estado plagada de fallos catastróficos, ausencias injustificadas y malas prácticas. La citada empresa británica QS, envuelta en polémica desde hace varios años y con serias sospechas de sesgos intencionados, desbanca a *Harvard* en favor de *Cambridge* (británica como QS) en la primera posición de su ranking. Signo de los tiempos, el ranking *THE* está disponible como

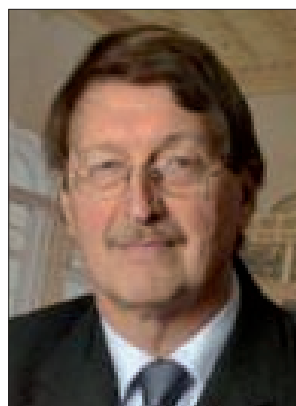
aplicación del *iPhone* (pagando, claro).

Se ha vislumbrado también lo que será el nuevo ranking del *Centre for Higher Education Policy Studies (Cheps)*, de la *Universidad de Twente*, Holanda, financiado generosamente por la *Comisión Europea* y que parece una colección de emoticones de colores.

Ni el *Centre of Science and Technology Studies (CWTS)* de la *Universidad de Leiden*¹⁵ ni el grupo *Scimago*¹⁶ se atreven finalmente a publicar auténticos rankings únicos, aunque nos dejan con unas interesantes colecciones de datos, que en el caso de *Scimago* superan tanto en calidad como en cobertura a los chapuceros *Essential Science Indicators* de *ISI/Thomson*.

Otro indicador protagonista ha sido el índice h, o mejor dicho la familia de indicadores derivada de él, que no hace sino crecer, aunque sea de forma desordenada. Es posible que ésta sea una moda pasajera, no sólo por la irrelevancia de ciertas propuestas, sino por las importantes

dudas que genera el propio índice h respecto a su consistencia matemática¹⁷.



Grandes bibliómetros holandeses: Loet Leydesdorff, Henk Moed, Anthony Van Raan y Paul Wouters

Sobre el CWTS

El CWTS, posiblemente el centro de investigación bibliométrica líder en Europa, ha sido un importante protagonista de este año. Además del cambio de liderazgo tras la jubilación de **Van Raan**, que había marginado de su sucesión a **Henk Moed** (uno de los bibliométricos más ortodoxos de su generación, finalmente contratado por Elsevier) en favor de **Paul Wouters** (con un perfil como mínimo “distinto”), el CWTS ha visto peligrar su corona al ser objeto de un agrio debate académico (y personal) su propuesta de *Crown indicator*. Este indicador utilizado generosamente en la evaluación de instituciones, a menudo de forma opaca, fue desenmascarado empírica y teóricamente por **Loet Leydesdorff**, de la *Amsterdam School of Communications Research (ASCoR)*, y sus colaboradores. **Van Raan** (al parecer todavía activo) reaccionó presentando un nuevo indicador, reconociendo implícitamente las limitaciones del anterior, pero sin confirmar aún, de forma definitiva, su validez. Merece la pena revisar el epílogo –hasta el momento– de este debate en las páginas de las actas de la *11th Intl conf on science and technology indicators, Universidad de Leiden, sept. 2010*¹⁹, y del *Journal of informetrics*¹⁹.

Acceso abierto (open access, OA)

Las iniciativas de OA se multiplican por todo el mundo, incluyendo España donde el número de repositorios institucionales se ha disparado. Los bibliotecarios de las universidades catalanas pueden estar más que satisfechos, puesto que tanto la *Autònoma de*

Barcelona (Dipòsit Digital de Documents) como la *Politécnica (UPCommons)* colocan sus repositorios entre los 50 primeros del mundo²⁰ (*Digital/CSIC* ocupa el puesto 38 y *UPCommons* el 41) en una lista de 1.184 repositorios.

*Dialnet*²¹, desde sus modestos orígenes, se ha convertido en el referente nacional de las bases de datos bibliográficas y ya indiza 2,4 millones de trabajos. Pero es mucho más, un portal de revistas, un repositorio de tesis y una emergente herramienta bibliométrica, entre otros servicios y nuevas promesas. Los acuerdos en marcha para canibalizar sus ruinosos competidores no son

Ranking Web de Repositorios del Mundo
Enero 2011

inicio UK

> inicio > Top repositorios

Datos

- Acerca de Nosotros
- Acerca del Ranking
- Top Repositorios
- Top USA y Canadá
- Top Europa
- Top Institucionales
- Top Portales
- Directorio
- Metodología
- Buenas Prácticas

Información

- Glosario
- Blog
- Enlaces
- Contacta con Nosotros
- Mapa del Sitio
- Aviso Legal

Buscar en Repositorios:

Categoría:

Todas

Top Repositorios
Primero | Previo | Siguiente | Último | Repositorios 1 a 50 de 1184

RANKING MUNDIAL	REPOSITORIO	POSICIÓN				
		PAÍS	TAMAÑO	VISIBILIDAD	FICHEROS RICOS SCHOLAR	
1	Social Science Research Network	USA	3	2	1	5
2	ArXiv.org e-Print Archive	USA	5	3	3	3
3	CiteSeerX	USA	2	1	924	2
4	Research Papers in Economics	EU	1	5	201	4
5	CERN Document Server	CH	4	20	2	8
6	Kyoto University Research Information Repository	JP	13	8	6	22
7	Smithsonian/NASA Astrophysics Data System	USA	15	4	839	1
8	National Taiwan University Repository	TW	37	9	32	7
9	Hokkaido University Collection of Scholarly and Academic Papers	JP	41	12	49	24
10	HAL Sciences de l'Homme et de la Société	FR	57	11	58	36
11	Munich Personal RepEc Archive	DE	49	26	17	29
12	Digital Library and Archives Virginia Tech University	USA	56	21	15	38
13	University of Michigan Deep Blue	USA	30	43	11	30
14	Universidade de São Paulo Biblioteca Digital de Teses e Dissertações	BR	34	49	13	9
15	MIT Dspace	USA	26	72	7	12
16	HAL Institut National de Recherche en Informatique et en Automatique Archive Ouverte	FR	35	19	131	26
17	University of Queensland Espace	AU	27	24	78	43
18	University of Twente Repository	NL	116	16	27	72
19	Igitur Archive Universiteit Utrecht	NL	53	59	5	23
20	École Polytechnique Federale de Lausanne Infoscience	CH	8	36	41	220
21	Universitat Autònoma de Barcelona Dipòsit Digital de Documents	ES	17	74	4	78
22	Biblioteca Digital Jurídica do Superior Tribunal de Justiça	BR	69	23	70	103
23	Universidade do Minho Repositorium	PT	32	44	66	49
24	Dissertations of the University of Groningen	NL	55	63	12	47
25	Universitat Stuttgartelektronische Hochschulschriften	DE	191	10	111	299

Cybermetrics Lab

M.A.V.I.R.
www.mavir.net

OpenAIRE
Open Access Infrastructure for Research in Europe

Primeros 25 repositorios según el “Ranking web de repositorios del mundo”, Laboratorio de Cibermetría, CCHS, CSIC²⁰.



sino excelentes noticias para el contribuyente patrio.

A nivel internacional algunos rubicones se han cruzado. *PLoS One*²², la controvertida revista de acceso abierto que cobra (y mucho) a sus autores, se ha convertido en la revista más grande del mundo (¡6.749 artículos!).

La *Comisión Europea*, a través del proyecto *OpenAire*²³ del VII Programa Marco, ha decidido por fin promover una plataforma de acceso abierto a la ingente producción científica que ha financiado. La decisión, guiada por una viabilidad percibida de que se alimente de los repositorios pre-existentes, será puesta a prueba en los próximos años.

Empiezan a aparecer con fuerza las propuestas de depósitos de datos científicos. Estas iniciativas están todavía en su infancia, si excluimos los campos científicos de la llamada "Big science" o los usuarios habituales de grids (parrillas o redes de ordenadores que unen sus capacidades de computación), donde la necesidad de compartir grandes (gigantescos) bancos de datos han tenido efectos positivos también en la mejora de las infraestructuras.

Algunos nubarrones preocupan a la comunidad OA, generalmente ocasionados por la presencia de importantes editores de revistas en proyectos y propuestas relacionadas con el futuro de la publicación científica. Sus intromisiones tienen unos fuertes tintes cuantitativos, negando evidencias favorables al acceso abierto o proponiendo indicadores de consumo claramente sesgados hacia los títulos adquiridos en compras consorciadas.

Es necesario indicar que la iniciativa del *European Reference Index for the Humanities (ERIH)*²⁴, una propuesta para poner en valor las revistas europeas de humanidades, apenas se preocupe de considerar los aspectos relacionados con el acceso abierto. Esperemos que no ocurra como con *Latindex*, que con su obstinada y obsoleta preocupación sólo por los aspectos formales ha

retrasado en décadas la evolución de las revistas latinoamericanas.

A la sombra de la bibliometría

Mi último comentario tiene un carácter profesional, relacionado con el papel del bibliómetro (sensu lato) en el diseño y gestión de la política científica española en la segunda década del

tercer milenio. En los últimos años la comunidad bibliométrica española ha conseguido situarse entre la élite mundial de la disciplina, como demuestra el crecimiento de las publicaciones en revistas de primer cuartil, el peso de nuestros ponentes en las conferencias internacionales y la presencia de expertos en comités multinacionales.

"A la sombra de la élite bibliométrica española se ha desarrollado una comunidad dada al estudio puntual, obsesionada con un puntillismo inútil"

A la sombra de dicha élite se ha desarrollado una comunidad imperfecta, dada al estudio local o puntual, obsesionada con un puntillismo metodológico inútil, alejada de las comunidades que mide, con una preocupante carencia de rudimentos matemáticos, ni hablar por supuesto del uso de las técnicas más modernas y empecinada en resolver problemas irrelevantes o necesidades obsoletas.

En el entorno de la revista *El profesional de la información* se ha llamado la atención sobre esta deriva, cosa que algunos han aprovechado para lanzar sus puyas sobre la disciplina. Pero se equivocan ya que la realidad acabará por desenmascarar a holgazanes e ineptos cortos de miras, que así liberarán los preciosos recursos que absorben para que otros puedan dedicarlos a proyectos, contratos, infraestructuras, divulgación y publicaciones más serios.

Referencias

1. <http://www.elsevier.com/locatel/joi>
2. <http://ec3.ugr.es/in-recs>

3. <http://sciencewatch.com/drl/cou/2010/10decALL>
4. <http://www.accesowok.fecyt.es/factor>
5. **Van Raan, Ton; Van-Leeuwen, Thed; Visser, Martijn.** "Non-English papers decrease rankings". *Nature*, 2011, n. 469, p. 34, 06 Jan 2011. DOI: 10.1038/469034a <http://www.nature.com/nature/journal/v469/n7328/full/469034a.html>
6. "Scientific fraud: action needed in China". *The lancet*, v. 375, n. 9709, p. 94, 9 January 2010. DOI: 10.1016/S0140-6736(10)60030-X [http://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(10\)60030-X/fulltext](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(10)60030-X/fulltext)
7. <http://www.scimagojr.com/countryrank.php>
8. <http://www.scimago.es>
9. <http://www.scopus.com>
10. <http://scholar.google.com>
11. <http://www.harzing.com/pop.htm>
12. <http://www.scientometrics-school.eu>
13. <http://www.topuniversities.com>
14. <http://www.timeshighereducation.co.uk/world-university-rankings>
15. <http://www.socialsciences.leiden.edu/cwts/products-services/leiden-ranking-2010-cwts.html>
16. <http://www.scimagoir.com>
17. **Waltman, Ludo; Van-Eck, Nees-Jan.** *A taxonomy of bibliometric performance indicators based on the property of consistency*. Erasmus Research Institute of Management (ERIM), ERIM Report series research in management, 12 March 2009. <http://repub.eur.nl/res/pub/15182>
18. http://www.cwts.nl/pdf/BookofAbstracts2010_version_15072010.pdf
19. **Leydesdorff, Loet; Opthof, Tobias.** "Remaining problems with the 'New Crown Indicator'" (MNCS) of the CWTS [letter to the editor]. *Journal of informetrics*, 2011, v. 5, n. 1, Jan., pp. 224-225.
20. <http://repositories.webometrics.info>
Método
Tamaño (S). Número de páginas recuperadas desde los 4 motores más grandes: *Google, Yahoo, Live Search y Exalead*.
Visibilidad (V). Número total de enlaces externos únicos recibidos (inlinks). Sólo puede obtenerse de forma consistente desde *Yahoo* y *Exalead*.
Ficheros ricos (R). Número de archivos en formato pdf (según *Google* y *Yahoo*).
Scholar (Sc). Mediante *Google Scholar* se calcula la media del número total normalizado de artículos y de aquellos (recientes) publicados entre 2001 y 2008. El puesto en el ranking se calcula con la fórmula:
 $P = 0,2 S + 0,5 V + 0,15 R + 0,15 Sc$
21. <http://dialnet.unirioja.es>
22. <http://www.plosone.org/home.action>
23. <http://www.openaire.eu>
24. <http://www.esf.org/research-areas/humanities/erih-european-reference-index-for-the-humanities.html>

G.1. La alegría va por barrios: revistas españolas como publicaciones fuente en la *Web of Science*

Por **Luis Rodríguez-Yunta**

26 abril 2010

Rodríguez-Yunta, Luis. "La alegría va por barrios: revistas españolas como publicaciones fuente en la *Web of Science*". *Anuario ThinkEPI*, 2011, v. 5, pp. 187-191.



Resumen: Se analiza la situación de las publicaciones españolas en la lista de revistas fuente de los índices de citas de ISI Thomson Reuters. Los cambios experimentados desde 2007 influirán sobre los futuros resultados de los indicadores bibliométricos de los *Journal Citation Reports*. La presencia de las revistas españolas ha aumentado de forma notable, especialmente en ciencias sociales, en donde se partía de una situación de casi invisibilidad. Sin embargo, se realizan dos importantes matizaciones a este éxito: el peso porcentual de las publicaciones españolas sigue siendo muy escaso y no todos los campos científicos se van a beneficiar por igual de este progreso.

Palabras clave: Revistas científicas, Índices de citas, *Web of Science*, WoS, España.

Title: *Happiness depends on the neighborhood: Spanish journals as source publications in Web of science*

Abstract: This paper analyzes the situation of Spanish journals in the master journal list of ISI Thomson Reuters' citation indexes. The changes since 2007 will influence future results of bibliometric indicators of *Journal Citation Reports*. The presence of Spanish journals has increased significantly, especially in Social Sciences, where it started from a situation of invisibility. However, there are two important constraints in this success: the percentage of Spanish publications remains very limited and not all scientific fields will benefit equally from this progress.

Keywords: Scientific journals, Citation indexes, *Web of Science*, WoS, Spain

LOS CAMBIOS EXPERIMENTADOS desde 2007 en la lista de revistas fuente que toma en consideración ISI Thomson Reuters para elaborar sus índices de citas siguen mereciendo la atención no sólo en los foros científicos, sino también entre bibliotecarios y documentalistas.

A pesar de las controversias que despierta el Factor de impacto, no se puede negar su relevancia sobre los procesos de valoración de la actividad científica y la gestión de colecciones de publicaciones periódicas en bibliotecas académicas y científicas. La aparición de nuevos indicadores alternativos provoca dudas sobre si generan rankings similares o complementarios (Torres-Salinas; Jiménez-Contreras, 2010), pero en esta guerra abierta frente a *Scopus* también deben contemplarse las variaciones que pueden producirse a corto plazo por la expansión en la selección de publicaciones para elaborar los índices de citas.

Los datos sobre las revistas "fuente" de la *Web of Science* (WoS) están en libre acceso en la Web,

donde se publica la *Master journal list* (MJL). Sin embargo, resulta mucho más trabajoso extraer valoraciones cuantitativas sobre esta web, frente a las facilidades que ofrece *Scopus*, que sí permite la descarga de un fichero *Excel* con todos los datos ya tabulados.

<http://science.thomsonreuters.com/mjl>

"El crecimiento de la presencia de las revistas españolas en la *Master journal list* de la *WoS* ha sido espectacular"

Otra dificultad para abordar el análisis de estos cambios radica en el hecho de que se trata de un proceso vivo, y por tanto no es posible sacar conclusiones definitivas sin correr el riesgo de que queden obsoletas en poco tiempo. Por este motivo, cualquier momento puede ser oportuno para plantear un estado de la cuestión y conocer

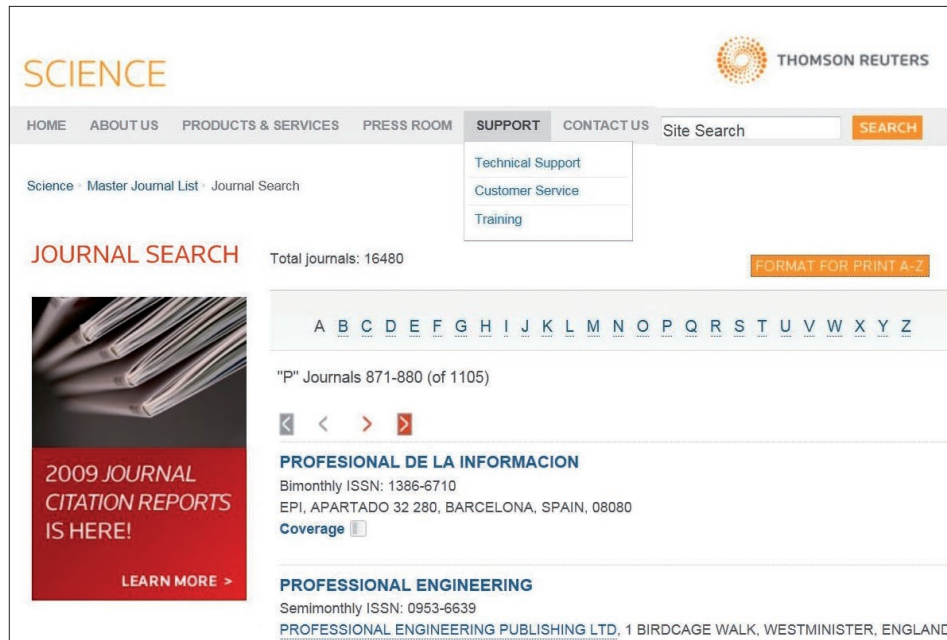


Figura 1. Master journal list. En febrero de 2011 incluye 16.480 revistas indizadas por las bdds de Thomson Reuters. De ellas, en la WoS hay algo más de 10.000. <http://science.thomsonreuters.com/mjl>

mejor en qué situación se encuentran las revistas españolas e iberoamericanas en los índices de citas, y así se ha realizado para una comunicación en el VII Seminario hispano mexicano con los datos publicados en enero de 2010 (Rodríguez-Yunta, 2010).

Destacada mejora de España

Sin duda, el aumento de la presencia de revistas españolas en la *MJL* de la *WoS* ha sido espectacular. En la actualidad se recogen 165 publicaciones editadas en España. Evidentemente esta cifra queda muy por debajo de la aportación de Estados Unidos, Reino Unido, Países Bajos y Alemania, que continúan en los primeros puestos de este listado. Si se considera la presencia por países en la *MJL* como un indicador del prestigio como país en las diferentes disciplinas, hay que destacar que España ha pasado al quinto puesto en el *Social science citation index (SSCI)*, al octavo en el *Arts & humanities citation index (A&HCI)* y al decimocuarto en el *Science citation index expanded (SCIE)*.

Esta situación es especialmente exitosa para las ciencias sociales, en donde se partía de una situación de casi invisibilidad ya que la selección de revistas españolas en el *SSCI* se limitaba anteriormente al área de psicología. Debe felicitarse a todos los que han contribuido a este progreso, pero también es necesario matizar dos aspectos:

1. El peso porcentual de las publicaciones españolas sigue siendo muy escaso. Tan sólo se ha pasado de la miseria a la pobreza. A pesar de

situarse en una muy digna posición en un *ranking* por países editores de las revistas incluidas en la *MJL*, las publicaciones editadas en España tan sólo representan el 1,0% en el *SCIE*, el 1,8% en el *SSCI* y el 3,2% en el *A&HCI*. Los cuatro países principales (Estados Unidos, Reino Unido, Países Bajos y Alemania) siguen acaparando el mayor número de fuentes de datos en los índices de citas.

2. La alegría va por barrios. Como si se tratase de una lotería, no todos los campos científicos se van a beneficiar por igual de este progreso. Frecuentemente

se evalúan las revistas no por su mera presencia dentro de estos listados sino por la posición en los primeros cuartiles dentro de las tablas de los *Journal citation reports*. Y estos datos se establecen según el esquema de clasificación propio del *ISI*, para cada uno de sus 255 epígrafes. Por ello, las posibilidades de recibir citas crecen considerablemente para aquellas revistas que cuentan dentro de la *MJL* con publicaciones relacionadas con la misma comunidad de usuarios (autores-lectores) dentro de un mismo campo científico. Y por el contrario, es mucho más difícil que reciba citas una revista de un área científica en la cual no hay ninguna publicación que incluya trabajos de la misma comunidad o red científica.

“El peso porcentual de las publicaciones españolas sigue siendo muy escaso: se ha pasado de la miseria a la pobreza”

La situación es ésta: de los 255 epígrafes, la presencia de revistas españolas se distribuye del siguiente modo:

– En el *SCIE* hay 78 publicaciones españolas encuadradas en 60 categorías de las 173 presentes en la clasificación de la *WoS*. Sólo en 21 de ellas se recoge más de 1 revista editada en España, mientras que no existe ninguna en 113 (65,3%).

– En el *SSCI* figuran 52 revistas españolas, presentes en 25 disciplinas de 55 posibles, aunque sólo en 15 con más de un título. No existe ninguna fuente española en 30 (54,5%).

– En el *A&HCI* hay 48 publicaciones de España, que están incluidas en 16 categorías de 27 existentes. Sólo en 9 de ellas hay más de un título. No hay revistas españolas en 11 (40,7%).

La probable incidencia de una revista dentro de la categoría o categorías en las que ha quedado seleccionada es muy irregular, ya que el número de publicaciones en cada epígrafe es muy variable. Por ello, aunque la categoría con mayor número de revistas en *A&HCI* es "History", con 12 publicaciones, por peso porcentual esta misma categoría, pero en el *SSCI*, es la que tiene mayor probabilidad de incidencia sobre los indicadores, pese a tener tan sólo 4 títulos (10,8%).

En los cuadros adjuntos se incluye la lista de categorías que tienen revistas españolas. Los títulos concretos seleccionados en la *MJL* de la *WoS* pueden consultarse en su web, o bien a través del directorio *Latindex* y los directorios de revistas de las bases de datos del *CSIC*. Para las revistas de ciencias humanas y sociales puede consultarse también el portal *Dice* del grupo *Epuc* (*Grupo de Investigación de Evaluación de Publicaciones Científicas*).

<http://dice.cindoc.csic.es>

"Ya no puede afirmarse que sea necesario publicar revistas en inglés para tener presencia en el *WoS*"

El estudio realizado también muestra que los cambios producidos están derrumbando algunos tópicos que hasta ahora podían ser creencias comunes en la comunidad científica:

– Ya no puede afirmarse que sea necesario publicar revistas en inglés para tener presencia en la *WoS*. De las revistas aceptadas, la mayoría publican artículos sólo en español.

– El cumplimiento de buenas prácticas en la edición científica tampoco es imprescindible para

D I C E		Difusión y Calidad Editorial de las Revistas Españolas de Humanidades y Ciencias Sociales y Jurídicas	
<input type="checkbox"/> Inicio <input type="checkbox"/> Cómo buscar <input checked="" type="checkbox"/> Búsqueda <input type="checkbox"/> Metodología <input type="checkbox"/> Grupo de trabajo <input type="checkbox"/> Contacto <input type="checkbox"/> Enlaces <input type="checkbox"/> Mapa			
Datos de la revista: El Profesional de la Información			
Título abreviado	Prof. Inf. ^C		
ISSN	1386-6710		
Año comienzo-fin	1998-		
Periodicidad	Bimestral		
Editor	Tomás Balget EPI SCP		
Lugar de edición	Barcelona		
Soporte	Impresa		
Historia	Desde 1992 hasta 1997 Information World en español. AC-F: 1992-1997		
URL	http://www.elprofesionaldelainformacion.com/		
Presencia en Internet	Texto completo		
Bases de datos que la incluyen	FRANCIS ; INSPEC ; ISTA ; LISA ; LISTA ; PASCAL ; SSCI ; ACADEMIC SEARCH COMPLETE ; ISOC ; SCOPUS ; ACADEMIC SEARCH PREMIER		
Área temática ISOC	Bibliotecología y Documentación		
Área de conocimiento	Bibliotecología y Documentación		
Clasificación UNESCO	Documentación		
Criterios Latindex cumplidos	33 (Criterios Latindex de revistas impresas o electrónicas)		
Evaluadores externos	Sí		
Cumplimiento periodicidad	Sí		
Apertura exterior del consejo de redacción	Sí		
Apertura exterior de los autores	Sí		
Fecha de actualización	20/04/2010		

Figura 2. Ejemplo de registro del Directorio Dice <http://dice.cindoc.csic.es>

superar el proceso de evaluación de *ISI Thomson Reuters*. En la selección figuran indistintamente publicaciones que están o no en el catálogo de *Latindex*.

Nota

– La categoría "History" en *Arts & humanities citation index* tiene 12 títulos españoles, más que ninguna otra de la *WoS*, pero compite en un conjunto de 250 revistas, por lo que sólo representa el 4,8%.

– La categoría "History" en *Social sciences citation index* tiene sólo 4 títulos, pero compite en un conjunto de sólo 37 revistas. Por ello aunque haya menos publicaciones españolas, en esta categoría es donde se alcanza un mejor porcentaje (10,8%), el más alto para España de todas las categorías que se contemplan en los *Journal citation reports* para hacer *rankings* de revistas.

Ese párrafo pretendía resaltar que para valorar la probable incidencia de las publicaciones españolas sobre el factor de impacto en una categoría concreta, no basta con dar el dato del número de revistas, ya que el número total de publicaciones en cada categoría es muy variable.

Una revista en un grupo de 20 da lo mismo que 10 sobre 200. Y claro, también podría añadirse que su capacidad de incidencia dependerá igualmente del número de artículos que contenga. Si tenemos pocas revistas y además con escasos artículos al año, en el cálculo del factor de impacto en cualquier categoría siempre van a tener mayor incidencia las publicaciones norteamericanas e inglesas, no sólo más numerosas en las listas de *ISI Thomson Reuters*, sino quizá también con mayor cantidad de artículos al año como media por título de revista.

Categorías de la WoS con más de 1 revista española	Base de datos	Rev. españolas	Total revistas	%
Allergy	SCIE	2	23	8,7
Archaeology	A&HCI	3	74	4,1
Art	A&HCI	3	69	4,3
Business	SSCI	2	100	2,0
Clinical neurology	SCIE	2	185	1,1
Communication	SSCI	3	72	4,2
Construction & building technology	SCIE	2	51	3,9
Economics	SSCI	5	295	1,7
Education & educational research	SSCI	8	191	4,2
Engineering, multidisciplinary	SCIE	2	90	2,2
Gastroenterology & hepatology	SCIE	2	69	2,9
Geography	SSCI	2	74	2,7
Geology	SCIE	3	53	5,7
History	A&HCI	12	250	4,8
History	SSCI	4	37	10,8
History & philosophy of science	A&HCI	3	51	5,9
Humanities, multidisciplinary	A&HCI	3	121	2,5
Immunology	SCIE	3	138	2,2
Infectious diseases	SCIE	3	63	4,8
Information science & library science	SSCI	2	79	2,5
Language & linguistics	A&HCI	7	160	4,4
Law	SSCI	2	141	1,4
Linguistics	SSCI	6	148	4,1
Literature, romance	A&HCI	6	61	9,8
Marine & freshwater biology	SCIE	3	94	3,2
Mathematics	SCIE	5	272	1,8
Mathematics, applied	SCIE	2	224	0,9
Medicine, general & internal	SCIE	4	164	2,4
Microbiology	SCIE	2	106	1,9
NeuroSciences	SCIE	2	240	0,8
Nutrition & dietetics	SCIE	2	74	2,7
Operations research & management science	SCIE	2	75	2,7
Pediatrics	SCIE	2	105	1,9
Pharmacology & pharmacy	SCIE	7	261	2,7
Philosophy	A&HCI	7	154	4,5
Plant sciences	SCIE	2	192	0,5
Political science	SSCI	2	128	1,6
Psychology, clinical	SSCI	2	103	1,9
Psychology, educational	SSCI	2	48	4,2
Psychology, multidisciplinary	SSCI	4	134	3,0
Public, environm. & occup health	SSCI	2	118	1,7
Religion	A&HCI	5	121	4,1
Sociology	SSCI	2	136	1,5
Statistics & probability	SCIE	2	109	1,8
Surgery	SCIE	2	187	1,1
Urology & nephrology	SCIE	2	67	3,0

Tabla 1

Referencias bibliográficas

Rodríguez-Yunta, Luis. "Las revistas iberoamericanas en *Web of Science* y *Scopus*: visibilidad internacional e indicadores de calidad". En: Memoria del VII Seminario Hispano-Mexicano de Investigación en Bibliotecología y Documentación, Ciudad de México, 7-9 de abril de 2010, pp. 347-363.

<http://eprints.rclis.org/18343>

Torres-Salinas, Daniel; Jiménez-Contreras, Evaristo. "Introducción y estudio comparativo de los nuevos indicadores de citación sobre revistas científicas en *Journal citation reports* y *Scopus*". *El profesional de la información*, 2010, marzo-abril, v. 19, n. 2, pp. 201-207.

Categorías de la WoS con sólo 1 revista española	Base de datos	Total revistas	%
Agricultural economics & policy	SCIE	13	7,7
Agriculture, multidisciplinary	SCIE	59	1,7
Anthropology	SSCI	94	1,1
Architecture	A&HCI	36	2,8
Automation & control systems	SCIE	61	1,6
Behavioral sciences	SCIE	48	2,1
Biotechnology & applied microbiology	SCIE	157	0,6
Business, finance	SSCI	83	1,2
Cardiac & cardiovascular systems	SCIE	110	0,9
Cell biology	SCIE	171	0,6
Chemistry, applied	SCIE	73	1,4
Chemistry, multidisciplinary	SCIE	158	0,6
Classics	A&HCI	41	2,4
Computer science, interdiscipl applics	SCIE	96	1,0
Dentistry, oral surgery & medicine	SCIE	74	1,4
Developmental biology	SCIE	69	2,6
Engineering, civil	SCIE	117	0,9
Entomology	SCIE	84	1,2
Film, radio, television	A&HCI	24	4,2
Folklore	A&HCI	18	5,6
Food science & technology	SCIE	123	0,8
Forestry	SCIE	57	1,8
Health care sciences & services	SCIE	69	1,4
History & Philosophy of science	SCIE	48	2,1
History & philosophy of science	SSCI	36	2,8
International relations	SSCI	73	1,4
Limnology	SCIE	19	5,3
Literary reviews	A&HCI	57	1,8
Literary theory & criticism	A&HCI	21	4,8
Materials science, ceramics	SCIE	25	4,0
Materials science, multidisciplinary	SCIE	219	0,5
Mathematics, interdiscipl applics	SCIE	88	1,1
Medieval & renaissance studies	A&HCI	33	3,0
Metallurgy & metallurgical engineering	SCIE	76	1,3
Micology	SCIE	22	4,5
Obstetrics & gynecology	SCIE	77	1,3
Oceanography	SCIE	60	1,7
Oncology	SCIE	181	0,6
Ornithology	SCIE	21	4,8
Pathology	SCIE	84	1,2
Physiology	SCIE	85	1,2
Psychiatry	SSCI	121	0,8
Psychiatry	SCIE	130	0,8
Psychology	SCIE	77	1,3
Psychology, applied	SSCI	76	1,3
Psychology, developmental	SSCI	65	1,5
Psychology, experimental	SSCI	77	1,3
Psychology, social	SSCI	58	1,7
Public, environm & occupational health	SCIE	138	0,7
Respiratory system	SCIE	46	2,2
Robotics	SCIE	17	5,9
Sport sciences	SCIE	85	1,2
Substance abuse	SSCI	26	3,8
Substance abuse	SCIE	13	7,7
Veterinary sciences	SCIE	153	0,7
Water resources	SCIE	72	1,4

Tabla 2

G.2. Library analytics: métricas online en servicios de información web

Por Jorge Serrano-Cobos y Ana Casado-Sánchez

28 septiembre 2010

Serrano-Cobos, Jorge; Casado-Sánchez, Ana. "Library analytics: métricas online en servicios de información web". *Anuario ThinkEPI*, 2011, v. 5, pp. 192-194.



Resumen: *Presentación del concepto de analítica web y su aplicación a los portales bibliotecarios. Se propone establecer los objetivos que quiere alcanzar la biblioteca y luego definir unos indicadores para medir su grado de consecución, como, por ejemplo, visitas a la web, tiempo de permanencia en la web, conversión a préstamos, visitas físicas a la biblioteca, fidelización y adecuación de los descriptores al vocabulario empleado por los usuarios.*

Palabras clave: *Analítica web, Objetivos, Indicadores bibliotecarios, Vocabulario.*

Title: *Library analytics: online metrics in web information services*

Abstract: *Introduction to the web analytics concept and its application to library portals. It is proposed to establish the goals the library wants to reach and then to define indicators to measure their achievement, such as number of visits, time connected to the web, conversion to book loans, physical visits to the library, fidelity and suitability of the descriptors to the vocabulary used by users.*

Keywords: *Web analytics, Library analytics, Objectives, KPI, Key performance indicators, Vocabulary.*

LIBRARY ANALYTICS es un concepto traducible aproximadamente por "analítica web de bibliotecas". Hace referencia al análisis mediante herramientas de analítica web de portales bibliotecarios, y podría hacerse extensible a servicios de información web en general.

Debemos entender el portal bibliotecario como un todo, que englobaría el sistema de consulta del catálogo (opac) y otras secciones del portal (agenda, página de servicios, noticias, etc.). Aunque obviamente no es imprescindible, para medir lo que ocurre en un portal se utiliza mucho *Google Analytics*, que en muchos casos permite también estudiar el buscador interno del portal bibliotecario y, en cierta medida, el del catálogo.

Este análisis debería permitirnos saber mejor cómo y cuánto está aportando el portal bibliotecario a los objetivos de la institución. En primer lugar por tanto, deberemos tener claras las prioridades del portal y explicitarlas por escrito, por ejemplo:

- atraer no-usuarios físicos;
- atraer no-usuarios de cualquier parte del mundo (con un universo de estudio totalmente distinto, usuarios físicos más usuarios web);

- fidelizar a los usuarios físicos;
- potenciar el uso de los servicios físicos de la biblioteca;
- potenciar el uso de los servicios online de la biblioteca;
- etc.

Después, en función de los objetivos o prioridades, debemos determinar cómo medimos el éxito de nuestras acciones promocionales mediante los denominados *KPIs* (*key performance indicators* o indicadores clave de desempeño), que nos permiten entender si estamos haciéndolo suficientemente bien o no (y por qué) para conseguir los objetivos que nos hayamos fijado.

"El análisis de visitas y transacciones en un portal bibliotecario se conoce en inglés por *library analytics*"

Algunos ejemplos típicos serían el número de páginas vistas, el número de visitantes únicos al día, el tiempo medio por visita que los usuarios están en el portal, etc.



Dashboard (tablero de instrumentos) o estadísticas básicas que presenta Google Analytics

Más allá de estos KPIs generalistas, aplicables a cualquier portal, podemos buscar mediciones más concretas y adecuadas al contexto de nuestros objetivos. Algún ejemplo:

- Ratio de conversión: ¿cuántos usuarios que llegan al portal terminan reservando un ítem online? ¿Cuántos se registran en un boletín semanal de novedades, o en el RSS de las noticias de nuestra institución, o quieren hacerse socios?

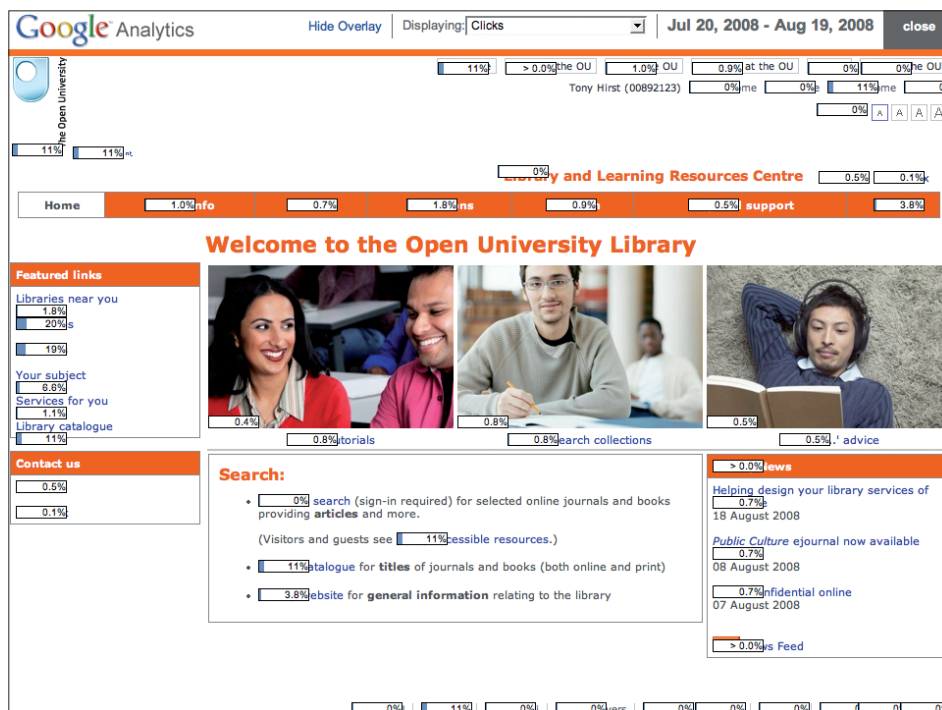
- Días/horas de la semana con más visitas. Lo interesante aquí es encontrar el porqué, más allá del mero dato. ¿Hay una interrelación causal entre el envío de un boletín, y un aumento de visitas? Si realizamos campañas con *Twitter*, ¿aumentan mucho las visitas en ciertas horas, o aumentan indepen-

der cómo buscan y qué palabras clave usan. Si nuestro sistema de recuperación no responde a esas expresiones de búsqueda, los usuarios no

dientemente de las horas debido a cierto tipo de avisos temáticos (es decir, si miran nuestra información –sea la que sea– a tal hora, o si las visitas están relacionadas con cierto tema que interesó más que otros)?

- Usuarios nuevos versus fidelizados: nos permitiría saber si un usuario vuelve a utilizar más servicios del portal tras su primera visita, qué servicios son los que fidelizan más, etc.

- Tasa de correspondencia entre *key words* buscadas y lenguaje controlado: si queremos satisfacer las necesidades de recuperación de información de los usuarios que han llegado a través de internet, es lógico intentar entender



Resultados de Google Analytics puestos sobre la home de la Open University Library, señalando el porcentaje de clicks recibidos por cada botón

encontrarán la información, aunque ésta exista en nuestro catálogo.

Por ejemplo, una expresión bastante buscada en Google por usuarios españoles en idioma español dentro del entorno temático de una biblioteca pública podría ser "autores españoles". Si es coincidente con una materia usada por los catalogadores, el sistema al menos podrá recuperarlo y ofrecer recursos de información específicos a esos usuarios que utilizan esa forma específica de expresar esa necesidad.

Referencias bibliográficas

Hirst, Tony. "Library analytics (part 1)". *OUseful.Info*,

the blog..., 20 agosto 2008.

<http://blog.ouseful.info/2008/08/20/library-analytics-part-1/>

Inan, Hurol. "Resetting your website's priorities". *Hurol Inan*, 16 abril 2009.

<http://hurolinan.com/index.php/2009/04/16/resetting-your-websites-priorities/>

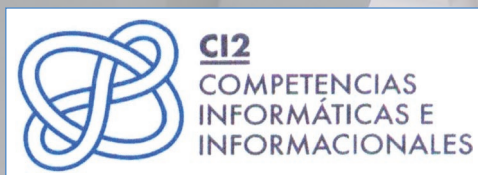
Kaushik, Avinash. "Six web metrics / key performance indicators to die for". *Web analytics blog: Occam's razor by Avinash Kaushik*, 16 septiembre 2008.

<http://www.kaushik.net/avinash/2008/09/rules-choosing-web-analytics-key-performance-indicators.html>

Mi, Jia; Weng, Cathy. "Utilizing opac search logs and Google Analytics: assessing opac effectiveness and user search behavior". En: *VALE Users'INJLA CUSINJ ACRL conference*. November, 2009.

<http://valenews.files.wordpress.com/2009/01/utilizing-opac-search-logs.ppt>

IX Jornadas CRAI



16-17 de junio de 2011

Santiago de Compostela

Por cuarto año consecutivo, dos comisiones sectoriales de la *Conferencia de Rectores de las Universidades Españolas (Crue)*,

- *Red de Bibliotecas Universitarias Españolas (Rebiun)*, y
- *Tecnologías de la información y las Comunicaciones (TIC)*

organizan conjuntamente las **IX Jornadas CRAI**, dedicadas a difundir el concepto de Centro de Recursos para el Aprendizaje y la Investigación.

Acogidas por la *Universidad de Santiago de Compostela* estarán dedicadas a la:

Evaluación y acreditación de las competencias informáticas e informacionales

Las jornadas se dirigen a toda la comunidad universitaria pero de manera especial todos aquellos que desarrollan su labor en el ámbito de los servicios de apoyo al aprendizaje, la docencia y la investigación.

Ven a Santiago y aprovecha para visitar la espectacular recién estrenada Cidade da Cultura

G.3. Archivos centralizados vs institucionales: ¿estamos equivocando el camino con la fragmentación de repositorios?

Por **Luis Rodríguez-Yunta**

4 octubre 2010

Rodríguez-Yunta, Luis. "Archivos centralizados vs institucionales: ¿estamos equivocando el camino con la fragmentación de repositorios?". *Anuario ThinkEPI*, 2011, v. 5, pp. 195-202.



Resumen: Se defienden las ventajas de los repositorios centralizados frente a la dispersión de proyectos institucionales. La aparición constante de nuevos recursos, archivos abiertos y bibliotecas digitales, afecta tanto al medio académico como a las instituciones de fondo antiguo. La "cultura del corte de cinta" debería ir acompañada de garantías para la preservación y la prestación de buenos servicios. Se citan algunos servicios cooperativos en el ámbito académico y un ejemplo concreto de digitalización de fondo antiguo sobre revistas españolas en la Universidad de Connecticut.

Palabras clave: Repositorios, Políticas, Archivos abiertos, Bibliotecas digitales, Proyectos.

Title: *Centralized vs institutional archives: are we going the wrong way with the fragmentation of repositories?*

Abstract: *This paper advocates the advantages of centralized repositories against the spread of institutional projects. The continuing emergence of new resources, open archives and digital libraries affects both academia and institutions that house ancient collections. The current "ribbon-cutting culture" should be subject to safeguards for the preservation and provision of good services. This paper lists some cooperative services in the academic world and a concrete example of digitization of an early collection of Spanish journals at the University of Connecticut.*

Keywords: *Repositories, Policies, Open archives, Digital libraries, Projects.*

EN LOS ÚLTIMOS años estamos asistiendo a una auténtica fiebre de creación de archivos digitales institucionales. El directorio español *BuscaRepositorios* identifica 66 recursos de este tipo en el medio académico a fecha de 1 de octubre de 2010.

Por su parte en el recolector *Hispana* se relacionan en la misma fecha 503 proyectos y 129 repositorios participantes. Parece que cada institución desea ofrecer sus propios recursos, manteniendo el control sobre la versión digital de sus fondos o sobre los documentos generados por sus investigadores. Es un esfuerzo loable pero, ¿es el mejor camino para llegar a los usuarios y garantizar la sostenibilidad de los recursos?

Esta dinámica tiene sus riesgos: no todos los organismos tienen el mismo potencial para asegurar la preservación de sus archivos digitales o para adaptarse con facilidad a las nuevas demandas tecnológicas que vayan surgiendo. Es muy probable que los actuales formatos de edición o versión digital queden obsoletos en poco tiempo y tengan que ser transformados. Y ya es necesari-

rio instalar versiones adecuadas para los nuevos dispositivos de lectura. Otro inconveniente es la dispersión de fuentes para localizar información y la posible duplicación de proyectos.

"Recursos como *Hispana* o *Recolecta*, al excluir la consulta del texto completo, están muy lejos de competir con buscadores como *Google*"

En esta dinámica de trabajo disperso, los recursos centralizados como *Hispana* o *Recolecta* se limitan a la integración y recuperación de los metadatos, manteniendo meros enlaces a los documentos, y limitando generalmente la capacidad de recuperación de contenidos. Al excluir la consulta del texto completo se está muy lejos de competir con la capacidad de buscadores como *Google*.

La alternativa a este funcionamiento es la explotación de repositorios colaborativos con un servicio técnico que garantice todas las prestaciones deseables. Con ello, la centralización de los recursos es plena y los documentos generados por una institución se integran en un servicio global, sin perder el carácter de colección identificable en sus páginas de navegación. Esta alternativa afecta tanto al ámbito de los archivos abiertos de carácter científico, en el que se han multiplicado los proyectos universitarios y de centros de investigación, como a las bibliotecas digitales de fondo antiguo, a las que se suman las iniciativas de centros e instituciones culturales de cualquier tipo.

A la inflación de repositorios académicos se han referido **Chris Armbruster** y **Laurent Romary** en varias contribuciones (**Armbruster; Romary, 2009; Romary; Armbruster, 2009**). Hay que alabar la aportación de los proyectos institucionales en la promoción del autoarchivo entre los investigadores, especialmente cuando existe un mandato institucional de depósito. Sin embargo, la mayoría de estos recursos no alcanzan un tamaño suficiente, no ofrecen servicios de calidad a la comunidad académica y no logran un buen posicionamiento en el *Ranking web of world repositories*.

Frente a la fragmentación de proyectos institucionales, los archivos temáticos como *arXiv* o *PubMed Central* pueden ocupar un lugar destacado en la comunicación internacional para una disciplina. Los investigadores están interesados en suscribirse a las alertas de nuevos trabajos



Figura 1. Hispana. Directorio y recolector de recursos digitales, <http://roai.mcu.es>



Figura 2. HAL, <http://hal.archives-ouvertes.fr>

“En España, el principal ejemplo de funcionamiento centralizado es Recercat”

depositados sobre un tema y consecuentemente en incluir sus propios *preprints*, mientras que usualmente no desean recibir este tipo de datos

sobre una única institución. **Romary y Armbruster** (2009) destacan además dos ejemplos que marcan una nueva tendencia hacia los esfuerzos colaborativos: el sistema *HAL* en Francia y *eSciDoc* en Alemania. Ambos son recursos centralizados en los que participan varias instituciones con el objetivo de garantizar la preservación y ofrecer servicios de valor añadido: estadísticas, páginas de autor, portales temáticos y enlaces a recursos disciplinares, múltiples formatos de exportación y sistemas de búsqueda avanzada que incluyen la interrogación del texto completo.

En España, el principal ejemplo de funcionamiento centralizado es *Recercat*, un proyecto coordinado por el *Consorci de Biblioteques Universitàries de Catalunya (Cbuc)* y alojado en el *Centre de Supercomputació de Catalunya*. En suma, red institucional y servicio informático de garantía. Las universidades participantes pueden tener sus propios archivos abiertos, pero la existencia de un repositorio colectivo es una gran ventaja sobre todo para integrar a las instituciones científicas de menor tamaño, organismos interinstitucionales e instituciones culturales.

Respecto a los proyectos de digitalización de fondo antiguo, hay que destacar un ejemplo concreto de la *Universidad de Connecticut* (**Ramos; Bennett, 2010**), que muestra este modo alternativo de trabajar frente a la fragmentación de sitios. El *Thomas J. Dodd Research Center* de esta universidad posee desde la década de 1970 una rica colección de revistas y periódicos españoles que perteneció a Juan Pérez de Guzmán y Boza, Duque de T'Serclaes, con publicaciones desde el siglo XVIII hasta principios del siglo XX, en especial editadas en Andalucía. Se trata por tanto de un recurso de alto interés que forma parte del patrimonio cultural español, depositado fuera de nuestro país y probablemente al margen de las fuentes recopiladas en *Europeana*.

El servicio de colecciones especiales de la biblioteca universitaria decidió abordar la digitalización de una parte de esta colección, en concreto la formada por las publicaciones feme-

ninas, de gran interés para la historia de las mentalidades, los estudios de género y también la literatura o la moda. No se incluyeron títulos al azar, previamente se cotejó que no estuvieran presentes en los recursos electrónicos españoles. De los 18 títulos que se han digitalizado, 17 son fondos únicos y uno, *El correo de las damas*, estaba incluido en los recursos electrónicos de la *Hemeroteca de la Biblioteca Nacional de España (BNE)*, pero los ejemplares de Connecticut incluían dibujos originales de las modas parisinas que la *BNE* no poseía, de modo que completaba lo que ya se había hecho.

“La existencia de repositorios consorciados permitiría abordar con mayor agilidad las sucesivas renovaciones tecnológicas”

La digitalización de este tipo de fondos es prioritaria por razones de conservación, precisamente para poder ofrecer su consulta a cualquier usuario sin restricciones por su eventual mal estado. La existencia de un acceso libre para la consulta del texto completo en internet permite evitar su uso físico y además llegar a un gran número de

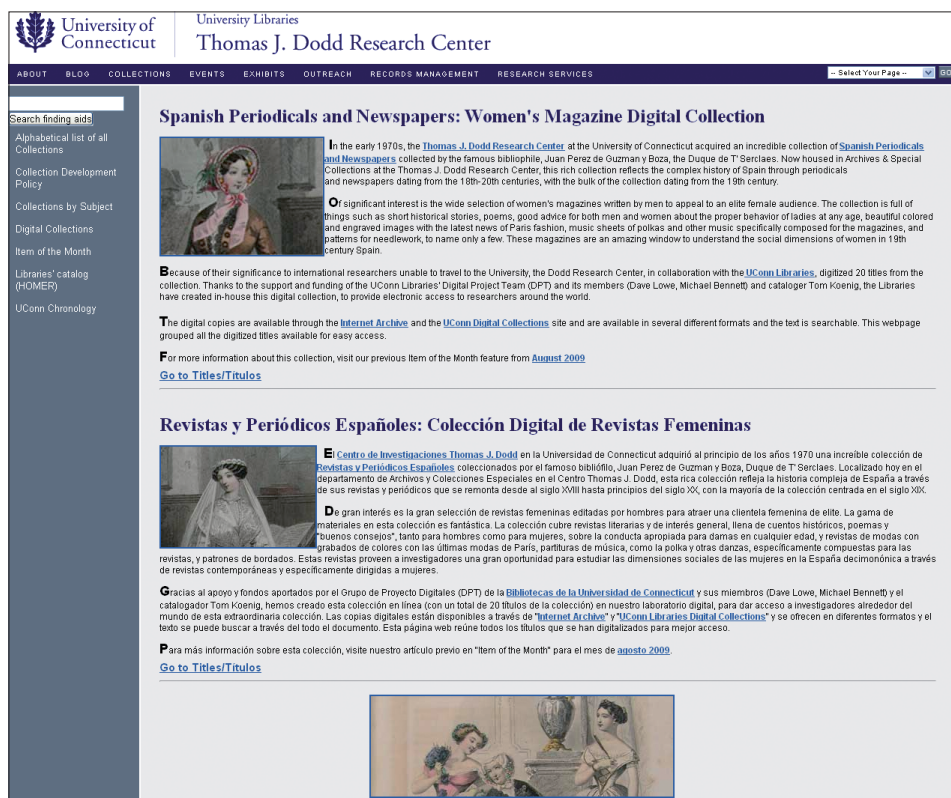


Figura 3. Revistas y periódicos españoles: Colección digital de revistas femeninas. Thomas J. Dodd Research Center. University of Connecticut. <http://doddcenter.uconn.edu/collections/spanwomen.htm>



Figura 4. Acceso a la Colección digital de revistas femeninas desde Internet Archive. Ejemplo: Las hijas de Eva (1880), <http://www.archive.org/details/lashijasdeeva00merc>

posibles lectores muy alejados geográficamente. Desde que se puso en marcha el proyecto la mayor parte de los accesos en línea han llegado precisamente desde España.

La web de la biblioteca incluye una página específica sobre el proyecto y una relación de los ejemplares incluidos. Desde aquí se enlaza con los documentos electrónicos existentes en *Internet Archive*, un proyecto cooperativo sin ánimo de lucro que garantiza el libre acceso a un repositorio digital permanente.

La *Universidad de Connecticut* actúa como suministrador de contenidos en un depósito colectivo, sumando sus fondos directamente a un recurso de carácter global.

¿Cuál es la ventaja de volcar los ficheros digitales y no los metadatos en un archivo común?

Sin duda, si el recurso es de alto nivel, las garantías de preservación y además la calidad de la oferta tecnológica. Basta entrar en los ejemplos de los registros de estas revistas españolas en *Internet Archive* para verlo: el acceso no se limita a un fichero pdf o una imagen escaneada; también puede descargarse en otros formatos, como *ePub*, *Djvu* o *Mobi*, destinados a los nuevos dispositivos lectores de libros electrónicos.

La existencia de repositorios consorciados permitiría abordar con mayor agilidad las sucesivas renovaciones tecnológicas. La lectura desde nuevos dispositivos es ya una realidad, y sus usuarios no pueden quedarse al margen de los servicios que se ofertan. Pronto vendrán otros formatos de exportación o de navegación y los sistemas de archivo tendrán que adaptarse a ello. La dispersión de repositorios no favorece este proceso.

Una comparación similar podría hacerse para las revistas digitales: ¿hemerotecas institucionales

o proyectos centralizados?

En España hay un riesgo claro de dispersión por la sólida implantación de la "cultura del corte de cinta": hay que inaugurar y presentar nuevos productos, nuevos proyectos, nunca hablar de sostener y financiar lo ya existente.

¿Cómo revertir esta tendencia? ¿Por qué no aplicar en la propia institución aquello que más se valora como usuarios?

Referencias bibliográficas

Armbruster, Chris; Romary, Laurent. *Comparing repository types. Challenges and barriers for subject-based repositories, research repositories, national repository systems and institutional repositories in serving scholarly communication.* Working Paper Series, 2009. <http://ssrn.com/abstract=1506905>

Ramos, Marisol; Bennett, Michael J. "Mujeres, damas y señoritas: el mundo de las revistas femeninas españolas del siglo XIX al alcance de la mano; The women's magazine digital collection at The Thomas J. Dodd research center". *UConn libraries presentations.* Paper 26, 2010. http://digitalcommons.uconn.edu/libr_pres/26

Romary, Laurent; Armbruster, Chris. *Beyond institutional repositories.* Working paper series, 2009. <http://ssrn.com/abstract=1425692>

Recursos citados

arXiv
<http://arxiv.org>

BuscaRepositorios
<http://www.accesoabierto.net/repositorios>

eSciDoc
<http://escidoc.org>

Europeana
<http://www.europeana.eu/portal>

Hyper articles en ligne (HAL)
<http://hal.archives-ouvertes.fr>

Hemeroteca digital
<http://www.bne.es/es/Catalogos/HemerotecaDigital>

Hispana, directorio y recolector de recursos digitales
<http://hispana.mcu.es>

Internet Archive
<http://www.archive.org>

PubMed Central
<http://www.ncbi.nlm.nih.gov/pmc>

Ranking web of world repositories
<http://repositories.webometrics.info>

Recercat
<http://www.recercat.net>

Recolecta
<http://www.recolecta.net/buscador>

University of Connecticut:
– Thomas J. Dodd Research Center
<http://doddcenter.uconn.edu>

– Spanish periodicals and newspapers
<http://doddcenter.uconn.edu/collections/spanper.htm>

– Women's magazine digital collection
<http://doddcenter.uconn.edu/collections/spanwomen.htm>

* * *

Centralizados o distribuidos: depende

Por Remedios Melero
[Lista OS Repositorios]



No se puede generalizar, y lo que hay que barajar, como en cualquier “proyecto/negocio”, es la versatilidad, la sostenibilidad y los servicios que se prestan.

El título del artículo quiere dar a entender que parece un error la construcción de repositorios: la palabra fragmentación induce a “división”, y ésta a pérdida de fuerza; sin embargo, en un sistema de repositorios de lo que se habla es de distribución. Las garantías a las que alude **Rodríguez-Yunta** son válidas para todos los proyectos, si bien los de digitalización tienen una función científico-académica y de uso común, sus objetivos y funciones no son las mismas que las de un repositorio institucional de acceso abierto, por lo tanto deben tener un trato diferente.

Hispana y *Recolecta* son *harvesters* o recolectores de metadatos según el protocolo OAI-PMH; obviamente si los proveedores de datos (los repositorios) se eligen con diferente criterio no pueden ofrecer los mismos resultados.

La preservación es un factor que debería tenerse en cuenta antes de poner en marcha cualquier iniciativa con contenidos digitales, aunque la realidad es distinta. En eso sí que estoy de acuerdo con la centralización, entre otras cosas porque la preservación es costosa y porque la centralización

permitiría estándares comunes de migración, transformación, etc.

Puede haber servicios colaborativos y repositorios institucionales, ambos son compatibles; de hecho en eso se basaba *Driver* y ahora *Coar*.

“Puede haber servicios colaborativos y repositorios institucionales, ambos son compatibles”

Ofrecer servicios es un argumento que he esgrimido muchas veces para que un repositorio tenga éxito; donde se sitúe es cuestión de criterios.

Es cierto que los repositorios temáticos *arXiv* o *PubMed Central* son un éxito, pero hay que ver por qué. Del primero, huelga decirlo, por su antigüedad y las materias que trata en las que es habitual el uso de preprints. Del segundo hay que recordar que la medicina y ciencias afines son áreas inmensamente ricas en producción científica, que la mayoría de entidades financiadoras con mandatos OA tienen que ver con éstas y que, además, las editoriales facilitan el depósito directo a *PubMed*.

Respecto al *HAL* en Francia y *eSciDoc* en Alemania, el primero sí es totalmente multidisciplinar, nacional y centralizado pero el segundo es de *Max Planck*, equivalente a *Digital CSIC*.

Sobre los proyectos de digitalización de fondo antiguo, estoy segura de que la duplicación de fuentes primarias digitalizadas es importante, ¿cuántos “Don Quijote” habrá digitalizados en el mundo? Lo que hay que hacer es localizar lo que ya existe para no duplicarse, en este caso los recolectores harían un gran papel.

Dispersión y distribución no son sinónimas, si hablamos de islas es una cosa y si hablamos de sistemas interoperables es otra.

Estoy de acuerdo en que crear algo nuevo cuando ya existe es una pérdida de tiempo y de recursos económicos y humanos, pero eso no es exclusivo de los repositorios o revistas, es propio del que se quiere poner medallas. Por cierto, supongo que al hablar de proyectos centralizados de revistas el autor se refiere a portales de revistas, porque si no llegaríamos a una enciclopedia de babel seriada.

Quería ser más breve pero no lo he conseguido, terminaré como he empezado, centralizados o distribuidos, “depende”. Cuando solía ejercer de química, las ventajas o inconvenientes de un método u otro se tenían que demostrar y avalar con resultados, no bastaba con creer que podría

ser de una manera u otra, creo que eso no es exclusivo de un área de conocimiento.

La prioridad absoluta es poblar los repositorios

Por **Isidro F. Aguillo**

[Lista OS Repositorios]



Creo que **Rodríguez-Yunta** mezcla en su mensaje dos propósitos que, aunque complementarios, son distintos y requieren diferentes aproximaciones. El objetivo primario de los repositorios es facilitar el acceso al conocimiento generado y aunque es relevante su función

de preservación de dicho conocimiento, poner el énfasis en este último aspecto puede hacer peligrar la viabilidad de las iniciativas *open access*, dividiendo esfuerzos, haciendo de proyectos que deben ser sencillos unas tremendas pesadillas de complejidad y burocracia, que al final puede alejarnos del *mainstream* de los procesos de comunicación científica.

Dejaré de lado el aspecto preservacionista en el resto de mis argumentos.

Puede que en el futuro la mayor parte de la producción científica esté disponible en abierto, pero ahora que esa cifra raramente roza el 15%, me parece inadecuado discutir si son galgos o podencos. La prioridad absoluta es poblar los repositorios, convencer a los autores de las ventajas de depositar y facilitarles la tarea en la medida de lo posible.

Por eso veo con sorpresa que en nuestra comunidad se plantean estrategias que a menudo ignoran al usuario clave de todo esto: el autor (científico, investigador, profesor, doctorando, etc.).

Y como autor me llama la atención que los recolectores me roben el posicionamiento de mis artículos en las búsquedas en *Google* cuando se trata de intermediarios de escaso valor añadido que además diluyen la identidad institucional de la organización que financia dichos trabajos.

“En nuestra comunidad se plantean estrategias que a menudo ignoran al usuario clave: el autor”

Soy un firme defensor de los repositorios institucionales ya que creo que son el camino más rápido para promover el acceso abierto. Para los

autores su repositorio institucional es tanto el palo (mandatos) como la zanahoria (prestigio), pero si se les ningunea, al final el autor acabará por no contribuir y la pirámide caerá por su base.

No, no me niego a la colaboración. Sólo digo que priorizarlo ahora es canibalizar los esfuerzos que deben hacerse desde las instituciones productoras de contenidos.

Compartir esfuerzos en redes

Por **Luis Rodríguez-Yunta**

Melero y Aguillo aportan sin duda ideas muy relevantes y atinan de lleno en el tema, lo cual siempre es de agradecer pues frecuentemente los debates en *IweTel* se dispersan con gran facilidad.

Mi intención desde luego no es pontificar, sino sólo plantear la duda sobre lo que tenemos, porque me pongo en la posición del usuario y ver la lista actual de repositorios creo que marea a cualquiera y temo que los recolectores no son capaces de ofrecer un servicio de valor añadido idéntico al de los repositorios.

Y sobre todo me asombra descubrir que algunos repositorios son de tamaño ridículamente pequeño. Y bueno, lo de juntar bibliotecas digitales de fondo antiguo con repositorios de universidades y centros de investigación, no es una idea original mía, es algo que está ocurriendo ya en algunos sistemas de búsqueda que utilizan OAI, luego la relación es evidente aunque su naturaleza sea distinta.

“No todas las instituciones necesitan crear un repositorio; para entidades pequeñas o proyectos de digitalización muy concretos puede ser más rentable asociarse”

No pretendo ningunear a los repositorios institucionales. Como apunta **Aguillo**, pueden ser indispensables hoy en día. Pero también veo grandes ventajas en la centralización o al menos en compartir esfuerzos en redes. Simplemente, no todas las instituciones necesitan crear un repositorio o una biblioteca digital; para entidades pequeñas o para proyectos de digitalización muy concretos puede ser mucho más rentable asociarse.

Respecto a la corrección que envía **Melero** sobre *eSciDoc*, lo que hago en mi texto es reseñar lo que expresan **Romary** y **Armbruster** en el artículo que cito. Allí estos autores se refieren

a eSciDoc como un sistema multidisciplinar en el que colaboran *Max Planck Society* y *FIZ Karlsruhe*. ¿No he recogido bien la idea de estos autores? Finalmente, sobre la frase final de **Melero** de que no basta con creer que algo puede funcionar, que hay que demostrarlo, por eso puse ejemplos de sistemas que considero que funcionan y otros citados por **Romary** y **Armbruster**, que ellos ponen de modelo. Son sistemas existentes, ejemplos válidos, no creencias.

La aventura del intento de autoarchivar

Por **Remedios Melero**
[Lista OS Repositorios]

En el boletín de **Peter Suber** de octubre 2010 habla de su “aventura en el intento de autoarchivar” (interpretación propia), y entre muchas de las cosas que menciona, hay un párrafo que viene al hilo de esta discusión:

<http://www.earlham.edu/~peters/fof/newsletter/10-02-10.htm>

“Hay grandes ventajas en que los autores depositen sus trabajos en su propio repositorio institucional. Ayuda a las instituciones a compartir, analizar y evaluar los resultados de la investigación propia. Añade incentivos locales para los mandatos de los financiadores de producir y premiar la participación de los autores. Se añade robustez a la conservación por el principio de *Lockss* (cuantas más copias existan, más seguridad de conservación) mediante la distribución de la bibliografía en torno a una gran red. Se asegura que el sistema vaya adaptándose al crecimiento de la investigación publicada, simplemente por el hecho de que las redes de distribución son más espaciales que cualquier nodo individual. Sobre todo, nutre las culturas locales de auto-archivo en todas las universidades, lo que beneficiará tanto a la investigación no financiada como a la financiada pero sin mandato de los donantes”.

Agregación de contenidos oportunos

Por **Pablo De-Castro**
[Lista OS Repositorios]



Abundando en el argumento de **Melero**, si uno trata de listar mentalmente los repositorios temáticos o cooperativos que hay disponibles en las proximidades, no encuentra precisamente el cuerno de la abundancia.

Está *Recercat*, ya citado por **Rodríguez-Yunta**, y de

ahí hay que pasar a *arXiv* (EUA), *PubMed Central* (EUA) y *Ukpmc* (UK).

Entonces, si la disyuntiva es entre un esquema de repositorios distribuidos que –con sus posibles puntos débiles a nivel conceptual– progresa con ritmo sostenido, y unos repositorios centralizados de los que no se tiene excesivas noticias salvo en lo que atañe a sus dificultades de financiación, creo que la alternativa es clara (si es que fuera preciso escoger, que sinceramente creo que no).

Uno diría que **Romary** (*Inria*) y **Armbruster** (*MPS*), estrechos colaboradores en el proyecto *PEER*, están argumentando en realidad a favor de crear un único repositorio centralizado donde se recojan todas las publicaciones resultantes de proyectos financiados por la UE (y afectados por su mandato de archivo abierto de resultados de investigación).

Esto, teniendo detrás toda la infraestructura *OpenAIRE*, puede ser una excelente idea, pero no veo la necesidad de oponer los dos modelos. Claro que probablemente pudiera también recopilarse este gran conjunto de producción científica –bien que más lentamente– por agregación de los contenidos oportunos de los repositorios institucionales distribuidos, y entonces ¿qué sería de toda esa infraestructura potencialmente redundante?

Directrices de OpenAIRE

Por **Eloy Rodrigues**
[Lista OS Repositorios]



Si me permiten la intrusión desde Portugal en vuestra interesante discusión, me gustaría, además de compartir los argumentos de **Melero**, **Aguillo** y **De-Castro** sobre la importancia y las ventajas para el progreso del *open access* del archivo en repositorios

institucionales, clarificar la estrategia del proyecto *OpenAIRE*.

Siguiendo lo definido en el proyecto-piloto de *open access* de la *Comisión Europea* (que afortunadamente no exige el archivo en un repositorio específico –como lo han hecho los *NIH* con *PubMed*– sino el archivo en cualquier repositorio de acceso abierto), la idea no es construir una infraestructura para el archivo centralizado. Al contrario, la estrategia de *OpenAIRE* es recolectar las publicaciones de los centenares de repositorios institucionales (o temáticos) que existen en los países europeos.

Por eso ha escrito unas “Directrices” para garantizar la conformidad de los repositorios institucionales con los requisitos de *OpenAIRE*.

http://www.openaire.eu/index.php?option=com_attachments&task=download&id=43&lang=es

Solamente para los autores que no tengan un repositorio institucional o temático donde archivar sus publicaciones, *OpenAIRE* ofrece un repositorio central, que en el proyecto llamamos, por eso mismo, repositorio huérfano (pero que sería mejor llamar de “homeless”, o sea, para las publicaciones que no tienen un “hogar” local).

Los repositorios deben demostrar su utilidad

Por José-Manuel Barrueco



Es interesante comprobar que el viejo tema de centralizados *versus* institucionales sigue provocando debates. Desde mi experiencia de trabajar con ambos creo que la solución está en la convivencia y la cooperación.

El papel del repositorio institucional (RI) es almacenar la producción de la institución y proporcionar los mejores metadatos posibles para que los recolectores puedan a su vez proporcionar buenos servicios de valor añadido.

El problema viene cuando el RI no cumple con su función: tiene contenidos cuestionables (la cultura de corte de cinta hace que se llenen a toda costa) y, lo que es peor, pésimos metadatos. Sobre esa base los recolectores no pueden hacer nada. Como se suele decir, de donde no hay...

Sobre lo de “juntar bibliotecas digitales de fondo antiguo con repositorios de universidades y centros de investigación”, la única relación que existe es que utilizan un mismo protocolo. Yo no veo qué servicios de valor añadido comunes a un manuscrito y a un artículo de física nuclear se puedan prestar. Más allá de tener una ventana de búsqueda, que es precisamente en lo que se quedan estas iniciativas.

Echo en falta en la lista de **Rodríguez-Yunta** otro sistema que funciona: *RePEc* (*Research*

papers in economics). *RePEc* es la suma de más de 1.200 “repositorios institucionales” agregados para formar un *dataset* de 900.000 documentos. Con un crecimiento mensual de 18.000 nuevos *papers* y más de 600.000 descargas “reales” de documentos al mes.

RePEc funciona sin necesidad de mandatos, por el mero hecho de que es una iniciativa útil para el investigador, porque estar en él proporciona a los autores visibilidad y una serie de valores añadidos:

- páginas personales;
- alertas temáticas;
- análisis de citas;
- rankings de autores e instituciones;
- datos de uso: descargas y vistas de sus documentos;
- directorio de economistas;
- y también: motores de búsqueda.

“Cuando los repositorios consigamos demostrar nuestra utilidad, los autores vendrán a nosotros”

[How you can use RePEc](#) | [RePEc information for participants](#) | [Major participants and activity](#)

RePEc (**R**esearch **P**apers in **E**conomics) is a collaborative effort of hundreds of volunteers in [73 countries](#) to enhance the dissemination of research in economics. The heart of the project is a decentralized database of working papers, journal articles and software components. All RePEc material is freely available. Participation in RePEc as a provider only involves the cost of your time in preparing and maintaining metadata describing your publications.

You may add your own materials to RePEc through a department or institutional archive. All institutions, nonprofit and commercial, are welcome to join and [contribute their materials](#) by establishing and maintaining their own RePEc archive. If your institution does not yet participate in RePEc, you may submit your own papers to [MPRA](#) (the Munich Personal RePEc Archive), and they will automatically be included in RePEc. RePEc does not support personal archives: only institutional archives.

<http://www.repec.org>

Cuando los repositorios consigamos demostrar nuestra utilidad, entonces los autores vendrán a nosotros. Para ello tenemos que empezar por construir repositorios hacia afuera, pensando más en los recolectores que en nuestra institución. Teniendo en cuenta que no somos un fin en nosotros mismos, sino un camino para conseguir entre todos una alternativa válida de comunicación.

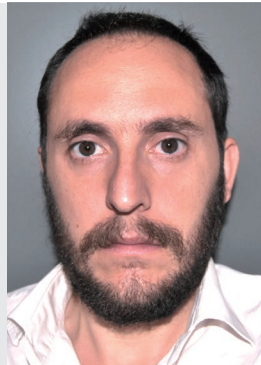
Esos son los deberes pendientes que tenemos los bibliotecarios...

G.4. Book citation index: nueva historia sobre *big science* y *little science*

Por Elea Giménez-Toledo y Daniel Torres-Salinas

13 enero 2011

Giménez-Toledo, Elea; Torres-Salinas, Daniel. "Book citation index: nueva historia sobre *big science* y *little science*". *Anuario ThinkEPI*, 2011, v. 5, pp. 203-205.



Resumen: Se describe el nuevo producto bibliométrico *Book citation index* anunciado por Thomson Reuters y se reflexiona sobre cómo esta poderosa empresa puede llevar a cabo ese gran proyecto (*big science*) –y también mejorar sus otros productos– beneficiándose de los muchos estudios que muchos pequeños equipos de investigación bibliométrica han realizado previamente (*little science*). Se augura éxito al nuevo índice puesto que los libros son el principal vehículo de comunicación en las humanidades.

Palabras clave: Citas de libros, *Book citation index*, BCI, *Science citation index*, Thomson Reuters, Humanities.

Title: *Book citation index: a new take on "big science" and "little science"*

Abstract: The new bibliometric product announced by Thomson Reuters, *Book citation index*, is described. The authors discuss how this powerful company can carry out this big project (*big science*) and also improve its other products, taking advantage of the many studies undertaken by small bibliometric research teams (*little science*) in previous years. The new index promises to be successful because books are the main communication vehicle in the humanities.

Keywords: *Book citations*, *Book citation index*, BCI, *Science citation index*, Thomson Reuters, Humanities.

DE-SOLLA-PRICE¹ empleó en 1963 los términos "big science" y "little science" para describir la evolución de la actividad científica en el conjunto de la sociedad, los cambios que se habían producido en las escalas de la ciencia y, fundamentalmente, el crecimiento exponencial de investigadores y de publicaciones.



Derek-John De-Solla-Price frente a una calculadora Antikythera, de la antigua Grecia, que él analizó
http://es.wikipedia.org/wiki/Derek_John_de_Solla_Price

En concreto "big science" se refiere a los grandes proyectos con alto presupuesto, y

"little science" a los estudios llevados a cabo por un solo investigador o un equipo pequeño. Releyendo su obra hemos creído que esos términos estaban plenamente vigentes, no estrictamente en el sentido que él les daba sino en una interpretación particular y personal que nos permitimos hacer sobre lo que ocurre hoy. Quizá sea algo atrevida pero creemos no estar demasiado lejos de la realidad.

Desde que la *Web of science* (*WoS*) y los *Journal citation reports* (*JCR*), de la empresa Thomson Reuters, se empezaron a emplear extensivamente en muchos países del mundo como herramienta fundamental para la evaluación de la producción científica de investigadores, instituciones y países, han sido numerosos los estudios bibliométricos que han demostrado sus sesgos y limitaciones para evaluar a determinados países y sobre todo a determinadas disciplinas.

Gran parte de los congresos sobre estudios de la ciencia, bibliometría o edición científica están consagrados o ampliamente dedicados a analizar datos provenientes de *WoS* o a mostrar sus limi-

taciones. Lo mismo ocurre con las publicaciones especializadas en el área. Hasta tal punto es así que se puede afirmar sin temor a errar que los grupos de investigación de todo el mundo que han ido demostrando minuciosamente todas y cada una de las limitaciones de *WoS*, han ido construyendo un corpus teórico e incluso metodológico como respuesta a las deficiencias de los sistemas de evaluación basados únicamente en esta fuente.

“En algunos países se han hecho esfuerzos para suplementar los aspectos no cubiertos por *WoS*”

De esta forma se han puesto de manifiesto públicamente los caminos a seguir para lograr una evaluación científica más ajustada a las características de las disciplinas humanísticas y sociales. Esto es particularmente evidente en el caso de países cuyas revistas están poco representadas en *WoS* y entre los grupos de investigación que trabajan específicamente en la evaluación de las ciencias humanas y sociales.

Así por ejemplo se fueron analizando datos de citación e impacto y hábitos de publicación y citación entre investigadores de ciencias humanas y sociales llegando a la conclusión de que las ventanas de citación de dos años empleadas por *JCR* eran insuficientes y poco significativas. Al mismo tiempo se propuso su ampliación para que pudieran ser aplicables y útiles a algunas materias. Desde *Thomson Reuters* dieron respuesta mejorando su producto (*JCR*) mediante la incorporación de ventanas de citación más amplias y ofreciendo así una respuesta más completa, precisa y adecuada a científicos sociales y humanistas²... Eso sí, a partir no sólo de su propia investigación sino también de los resultados de pequeños y medianos grupos de todo el mundo que no han tenido ni tendrán la repercusión científica y social que tiene la empresa *Thomson Reuters*.

“El principal vehículo de comunicación en las humanidades es el libro”

Otro claro ejemplo ha sido la reivindicación de los humanistas de que las monografías fueran protagonistas de los sistemas de la evaluación

de la actividad científica, avalada por decenas de trabajos bibliométricos que muestran el alto porcentaje de citas a monografías y la relevancia del libro como medio de comunicación científica en esas áreas.

Todos los grupos de investigación que han trabajado en este tema –independientes a priori de los grupos de poder–, desde australianos a españoles, pasando por holandeses, italianos, etc., han propuesto y aplicado procedimientos de evaluación de monografías y/o editoriales de monografías para cubrir ese hueco y permitir así tener elementos de valoración del principal vehículo de comunicación en las humanidades: el libro.

“Previsiblemente los países de todo el mundo comprarán el acceso a los índices de citas de libros”

Esos intentos siempre han quedado circunscritos a proyectos de investigación con fecha de finalización o a experimentos puntuales. Las razones por las que no han continuado han sido diversas, pero no han tenido que ver con la falta de consistencia, validez o utilidad de los resultados.

Durante la feria del libro de Frankfurt de 2010 se anunció *Book citation index*³, el nuevo proyecto de análisis de citas aplicado a libros que *Thomson Reuters* presentará a lo largo de 2011. No cabe ninguna duda de que esta empresa ofertará a la comunidad científica internacional un estupendo producto de información.

Book citation index, inspirado en los *citation indexes* tradicionales de revistas, incorporará de partida a las mejores editoriales científicas mundiales e invitará a otras a que participen (dicho sea de paso, seguramente las editoriales abrirán más fácilmente las puertas a *Thomson Reuters* que a cualquier otro grupo de investigación con pretensiones estrictamente científicas).

Previsiblemente países de todo el mundo ampliarán sus ya enormes inversiones para tener acceso a estos índices de citas y seguiremos siendo dependientes de los productos de información generados por *Thomson Reuters*, empresa que se irá haciendo cada vez más grande y poderosa. Nadie hablará de la base científica de ese proyecto: la investigación invisible o casi invisible realizada por grupos de todo el mundo que durante años han estado poniendo en bandeja a dicha empresa el desarrollo de un producto que sólo alguien como ella puede llevar a cabo.



<http://wokinfo.com>

Podrían mencionarse dos o tres proyectos españoles que con una inversión mínima están dando respuestas sólidas a las necesidades de los evaluadores de la actividad científica, pero volveríamos a hablar de *little science* y *big science* o de David contra Goliat.

Es innegable la labor de los *citation indexes*, la extraordinaria idea de crearlo y la investigación que la propia empresa –primero *ISI* y luego *Thomson*– ha llevado a cabo durante estos años. Pero también es innegable el esfuerzo que se ha hecho en algunos países para cubrir aquellos aspectos no tratados por *WoS* y para proporcionar indicadores y pautas más cercanas y útiles a las humanidades y a las ciencias sociales, además de introducir pluralidad en los sistemas de evaluación de la actividad científica.

Rememorando dos ideas del artículo de prensa de **Miguel Delibes-de-Castro** a partir de uno previo de **Laurent Ségalat**, sobre la analogía entre burbuja científica y burbuja financiera^{4, 5}, es peligroso depender de un único sistema y además

“a menudo confundimos en ciencia el éxito y el brillo con el mérito real” (**Ségalat**).

Sirva este pequeño texto para reflexionar sobre ciencia y mercado.

Notas

1. **De-Solla-Price, Derek J.** *Little science, big science*. New York: Columbia University Press, 1963.

2. Mientras tanto, por cierto, en España ya se calculaban índices de impacto con ventanas de citación de tres a cinco años tanto para humanidades como para ciencias sociales (*Resh e In-Recs*).

3. **Herther, Nancy.** “*Thomson Reuters announces Book citation indexes –scheduled for release in 2011*”. *Information today*, 2010, 25 October. <http://newsbreaks.infotoday.com/NewsBreaks/Thomson-Reuters-Announces-Book-Citation-IndexScheduled-for-Release-in-Q--70969.asp>

4. **Delibes-de-Castro, Miguel.** “¿Burbuja científica?”. *Público.es*, 26 junio 2010. <http://blogs.publico.es/ciencias/general/9461%C2%BFburbuja-cientifica>

5. **Ségalat, Laurent.** “System crash. Science and finance: same symptoms, same dangers?”. *European Molecular Biology Organization (EMBO) reports*, 2010, v. 11, n. 2, pp. 86-89. DOI: 10.1038/embor.2009.278 <http://www.nature.com/embor/journal/v11/n2/pdf/embor2009278.pdf>

G.5. Efecto Aneca: producción española en comunicación en el *Social science citation index*

Por **Pere Masip**

11 enero 2011

Masip, Pere. "Efecto Aneca: producción española en comunicación en el *Social science citation index*". *Anuario ThinkEPI*, 2011, v. 5, pp. 206-210.



Resumen: Se describen los efectos que los criterios usados por las agencias de evaluación tienen en la producción española en comunicación. Por un lado se multiplica el número de revistas académicas especializadas, que realizan un esfuerzo importante por mejorar su calidad, y por otro aumenta de forma espectacular el número de artículos de autores españoles en revistas indexadas en el *Social science citation index*.

Palabras clave: Comunicación, Investigación, Efecto Aneca, Producción científica, España.

Title: *The Aneca effect in the Social science citation index: Spain's productivity in communication research*

Abstract: Effects of the evaluation criteria used by assessment agencies on Spanish productivity in communication research. On the one hand, the number of academic journals has increased and their quality has improved; on the other hand the production of Spanish authors in journals indexed in the *Social Science Citation Index* has spectacularly increased in the last years.

Keywords: Communication, Research, Aneca effect, Scholarly production, Spain.

EN EL ÚLTIMO número de 2010 de *El profesional de la información*, Torres-Salinas et al ofrecen una valiosa propuesta de clasificación integrada de revistas científicas especializadas en ciencias sociales y humanas¹.

Como los propios autores apuntan, el principal objetivo de la herramienta es permitir la obtención de indicadores bibliométricos sobre la calidad de las publicaciones para distintos niveles de agregación. Su consolidación ofrecería a los investigadores un recurso unificado al que atenderse a la hora de decidir en qué revistas publicar para obtener el reconocimiento de las agencias de evaluación.

En el área de la comunicación, en la que nos centramos a lo largo de esta nota, tanto *Aneca* (*Agencia Nacional de Evaluación de la Calidad y Acreditación*)² como *Cneai* (*Comisión Nacional Evaluadora de la Actividad Investigadora*)³, así como las diversas agencias autonómicas, priorizan y privilegian en el proceso de evaluación los artículos publicados en las revistas indexadas en la base de datos *Social science citation index* (*Ssci*). Si bien también pueden valorarse los artículos aparecidos en revistas listadas en otras bases de datos (como *Scopus*) y clasificaciones nacionales o internacionales (como *Erih*, *In-recs*, *Latindex*,

Dice-ledcyt) el peso asignado al *Ssci* es abrumador.

"A pesar del impulso de las revistas, es significativo que ninguna de las que se publican en el ámbito de la comunicación sea una iniciativa comercial"

A través de los criterios y principios de evaluación establecidos, las agencias de evaluación se han convertido en auténticas *vedettes* del sistema universitario español y condicionan de forma más o menos consciente la investigación que se realiza. **Jaume Soriano** denomina "efecto Aneca" a la influencia que los procedimientos de acreditación y evaluación científicos oficiales tienen en la investigación española⁴. En opinión de **Soriano**, la comunidad académica ha modificado la forma de difundir los resultados de su actividad investigadora y los jóvenes investigadores trazan sus carreras sobre los objetos de estudio que mejor se ajusten a los criterios definidos por las agencias.

(Re)surgir de las revistas científicas de comunicación

Los investigadores planifican su actividad en función de esos criterios. Si las agencias priorizan inequívocamente la difusión de los resultados, y entre ellos muy especialmente la publicación en revistas, publíquese en revistas. Para los investigadores es mucho más rentable publicar un artículo que publicar un libro.

A aquellos que todavía se atreven a intentar publicar un libro les espera un segundo reto que afrontar: encontrar una editorial que los acoja. La mayoría evitan publicar monografías fruto de resultados de investigación alegando la falta de mercado para este tipo de trabajos. Prefieren manuales y trabajos de divulgación que puedan convertirse en lecturas obligatorias. Los que aceptan el desafío, con frecuencia lo hacen pertrechados de la sólida red de seguridad que constituye el compromiso del autor de comprar en bloque varias decenas de ejemplares o, directamente, el pago de su edición.

La alternativa a las editoriales comerciales son las universitarias, que han sido tradicionalmente ninguneadas. Un libro publicado por una editorial vinculada a un centro universitario no era valorado por algunas agencias ni por la *Cneai*. La discriminación se mantuvo hasta diciembre de 2009, fecha en que se publicaron los nuevos criterios de evaluación, donde se equiparaban las editoriales universitarias con las comerciales.

Sin lugar a dudas, las grandes beneficiadas de los criterios de evaluación han sido las revistas científicas. Los títulos se han multiplicado y han trabajado para mejorar la calidad, cuando menos la formal.

En la última actualización de *In-recs*⁵ –de diciembre de 2010, que incluye el índice de impacto de 2009–, en el área de comunicación se incluyen ya 23 títulos, 2 más que en la edición de 2008 y 6 más que en 2000. A éstos se les podrían añadir otras publicaciones, algunas de larga tradición como *Quaderns del CAC* y otras de aparición más reciente, como *Textual & visual media*, *Miguel Hernández communication journal*, *Revista mediterránea de comunicación* o *Fonseca journal of communication*, entre otras.

Una cifra nada desdeñable, pero que en opinión de algunos todavía es insuficiente en la medida que “toda facultad ha de tener, como mínimo, una cabecera (...), para evitar la endogamia editorial” [sic]⁶. Una afirmación que contrasta con la escasa incidencia de algunos títulos, tal como se desprende de *In-recs* o como concluyen algunos estudios⁷.

A pesar del impulso de las revistas, es significativo que ninguna de las que se publican en el ámbito de la comunicación sea una iniciativa

comercial. La mayoría son editadas por universidades, algunas por asociaciones científicas o fundaciones, y no faltan las que pueden catalogarse abiertamente de auténticas iniciativas personales. Aunque, como veremos a continuación, las revistas españolas en comunicación han realizado grandes avances desde el punto de vista cualitativo, la aparición de iniciativas comerciales, similares a la de *El profesional de la información* en el ámbito de la biblioteconomía y la documentación, contribuiría a la profesionalización del sector, a erradicar algunas prácticas poco recomendables y a aumentar el nivel de los artículos.

La carrera por mejorar el factor de impacto (FI) y atraer más y mejores artículos ha tenido sus efectos en las revistas. Los editores están cada vez más preocupados por cumplir los criterios de calidad de catálogos como *Latindex* o *Dice* y mejorar su índice de impacto en *In-recs*. Como los grandes cocineros cuando se otorgan las estrellas *Michelin*, están pendientes de las actualizaciones de los índices de impacto⁸ y algunos no dudan en poner en marcha verdaderas estrategias de ingeniería que faciliten la obtención de un índice mayor⁹.

“El número de artículos publicados por las revistas españolas de comunicación aumenta año tras año”

Algunas revistas dirigieron su mirada al *Social science citation index* y ahora empiezan a recoger los primeros frutos. En 2 años se ha pasado de no disponer de ninguna revista española –ni en una lengua distinta al inglés– en el área de comunicación del *Scsi* a poder encontrar hasta 3. Primero fue *Comunicar* (FI 2009: 0,022) y ahora *Comunicación y sociedad* y *Estudios sobre el mensaje periódico*. De las 67 publicaciones que constituyen el *Journal master list* de comunicación del *Scsi*, España aporta 3, las mismas que los Países Bajos y sólo está superada por Estados Unidos (29) y Gran Bretaña (28). Alemania, Australia, Eslovenia y Suiza contribuyen con 1 título cada uno¹⁰.

Internacionalización de la investigación española en comunicación

Daniel Jones¹¹ denunciaba que la producción española era excesiva, repetitiva y con escasa proyección internacional. Ciertamente, la presencia de autores españoles en revistas internacionales ha sido limitada^{12, 13}.

Entre 1994 y 2004 la contribución española en las publicaciones de comunicación del *Ssci* se redujo a 43 artículos. España se situaba en la octava posición del ranking europeo, superada por Francia (48 artículos publicados), Bélgica (56), Finlandia (64), Suecia (69), Alemania (186), Países Bajos (201) y Gran Bretaña (681). En el contexto mundial, se ubicaba en el decimosexto lugar¹³.

El promedio era de 3,9 artículos por año, siendo 1997 el más prolífico, con 7, seguido por 1998 y 2000 con 6. *Public understanding of science* fue la publicación que acogió más artículos con 6, seguida con 4 artículos por *European journal of communication*, *Javnost-The public*, *Public relations review* y *Telecommunications policy*. En total 16 de los 35 títulos que constituían la muestra publicaron algún artículo con la participación de al menos un autor vinculado a una institución española.

Las cifras obtenidas en 2005 para el período 1994-2004 han sido recientemente actualizadas¹⁴ y muestran un cambio de tendencia que no puede desvincularse del efecto *Aneca*.

Entre 2005 y 2009 la aportación española a las publicaciones de la sección de comunicación del *Ssci*¹⁵ ascendió a 64 artículos, 28 de los cuales aparecieron en 2009. Esto es, un incremento del 166% respecto al período 2000-2004 y del 237% para el 1994-1999.

Este incremento se traduce también en un importante salto en la clasificación de naciones europeas más productivas. España asciende hasta la cuarta posición, sólo superada por Alemania (105 artículos), los Países Bajos (171) y Gran Bretaña (327). En términos relativos la contribución española supone ahora el 6,36% de la aportación europea al *Ssci*, muy por encima del 2,6% del período anterior.

Las revistas preferidas para publicar son similares a las apuntadas más arriba: *Public relations review* (15 artículos), *Telecommunications policy* (12), *Public understanding of science* (4), a las que hay que sumar *Journal of advertising* (4) y *Discourse & society* (4). En total 17 de las 35 revistas analizadas recogen artículos de autores españoles.

A modo de reflexión final

Más allá de la discusión sobre la bondad o no del sistema de evaluación de las enseñanzas superiores, las actividades docentes y de investigación del profesorado universitario que se ha instaurado en España, lo cierto es que este modelo ha incidido inequívocamente en los diversos actores que intervienen en la investigación. En primer lugar –obviamente– en los propios investigadores, que pueden condicionar su actividad a los criterios de evaluación utilizados por las agencias. Por ejemplo es sabido que algunas revistas prefieren ciertas aproximaciones metodológicas y que algunos temas son más difíciles de “colocar” en revistas indizadas en el *Ssci* que otros. Si se quiere publicar en esos títulos, el investigador tendrá que adecuar sus trabajos a esos requisitos, la mayor parte de las veces no escritos, pero conocidos por todos los que trabajan en el área.

Conscientes del valor que las agencias atribuyen a los artículos de carácter internacional, en particular del *Ssci*, los autores parecen haberse volcado en esas revistas. Ello tiene ventajas: evaluaciones positivas, mayor visibilidad potencial de la investigación y los autores españoles, etc., aunque no faltan consecuencias menos favorables.

Giménez-Toledo y **Alcaín**⁷ apuntan por ejemplo que el peso asignado a las revistas *ISI* hace que los autores tiendan a publicar en

<http://www.aneca.es>

esas revistas trabajos sin que sean el destino más oportuno. Ello se produciría en detrimento de las cabeceras españolas, más adecuadas para aceptar investigaciones que abordan temas de carácter local.

También es bien conocido el alcance parcial y el sesgo geográfico y lingüístico del *Scsi*⁸. El ámbito de la comunicación no es una excepción. La edición de 2009 incluye por ejemplo 3 revistas específicas de publicidad (*Journal of advertising*, *International journal of advertising* y *Journal of advertising research*), 2 de comunicación para la salud (*Communication health* y *Journal of health communication*), o de comunicación científica (*Public understanding of science* y *Science communication*), y una sola dedicada inequívocamente al periodismo: *Journalism & mass communication quarterly*, con marcado acento norteamericano¹⁶.

No incluye por ejemplo, las 3 revistas que con toda seguridad se pueden considerar de referencia en periodismo: *Journalism studies* (Routledge), *Journalism practice* (Routledge) y *Journalism: theory, practice & criticism* (Sage).

Este déficit de revistas de periodismo puede condicionar la estrategia de publicación de los autores de la disciplina. Por un lado se ven perjudicados por no encontrar títulos de su materia en el *Scsi* y, por otro, presionados por los criterios de las agencias, pueden mandar originales a revistas incluidas en el *Scsi* pero que no constituyen el "núcleo duro" de la especialidad, lo que puede dificultar la visibilidad de esos artículos entre los investigadores del área¹⁷.

Además de los autores, las revistas también se han visto condicionadas, y en su caso también beneficiadas, por el nuevo rol otorgado por las agencias de evaluación. Ciertamente, las revistas españolas han realizado un esfuerzo por cumplir los criterios de calidad, y la mejora es sin duda sustancial, pero en general todavía adolecen de un *amateurismo* poco acorde con las exigencias de un sistema que las ha convertido en el centro gravitacional.

El número de artículos publicados por las revistas españolas de comunicación aumenta año tras año⁹, si bien parece identificarse un grupo bien definido de revistas que concentran las preferen-

ISI Web of KnowledgeSM

Web of Science Additional Resources

Search Cited Reference Search Advanced Search Search History Marked List (0)

Web of Science[®] – with Conference Proceedings

<< Back to previous

Results Topic=(communication) AND Address=(Spain OR Barcelona OR Madrid OR Valencia OR Bilbao OR Seville)
Refined by: Subject Areas=(COMMUNICATION)
Timespan=All Years Databases=SSCI.

Results: 140 Page 1 of 14 Go Sort by: Latest Date

Print E-mail Add to Marked List Save to EndNote Web Analyze Results
Save to EndNote, RefMan, ProCite more options Create Citation Report

Refine Results

Search within results for Search

Subject Areas Refine

- COMMUNICATION (140)
- BUSINESS (11)
- SOCIOLOGY (8)
- LINGUISTICS (7)
- LANGUAGE & LINGUISTICS (5)

more options / values...

Document Types Refine

- ARTICLE (117)
- BOOK REVIEW (11)
- EDITORIAL MATERIAL (5)
- PROCEEDINGS PAPER (4)
- NEWS ITEM (1)

more options / values...

Authors

- Title: Ideas wich lead to writing. The beginnings of José Jimenez Lozano in the newspaper (1956-1969)
Author(s): Bobillo MM
Source: **COMUNICACION Y SOCIEDAD** Volume: 23 Issue: 2 Pages: 149-178 Published: DEC 2010
Times Cited: 0
- Title: Making things happen: the role of communication in strategic management A case study on banking industry
Author(s): Garcia E, Garraza TS
Source: **COMUNICACION Y SOCIEDAD** Volume: 23 Issue: 2 Pages: 179-201 Published: DEC 2010
Times Cited: 0
- Title: Research in Communication. Bibliometric Study in Journals of Communication in Spain
Author(s): Castillo A, Carreton MC
Source: **COMUNICACION Y SOCIEDAD** Volume: 23 Issue: 2 Pages: 289-327 Published: DEC 2010
Times Cited: 0
- Title: The early stages of the integration of the internet in EU newrooms
Author(s): Sarrica M, Fortunati L, O'Sullivan J, et al.
Source: **EUROPEAN JOURNAL OF COMMUNICATION** Volume: 25 Issue: 4 Pages: 413-422 Published: DEC 2010
Times Cited: 0
- Title: Influence of Mass Media on Body Image and Eating Disordered Attitudes and Behaviors in Females: A Review of Effects and Processes
Author(s): Lopez-Guimera G, Levine MP, Sanchez-Carracedo D, et al.
Source: **MEDIA PSYCHOLOGY** Volume: 13 Issue: 4 Pages: 387-416 Published: 2010

Web of science (Scsi).

cias de los autores tanto a la hora de publicar como de citar^{7, 9}.

Giménez-Toledo y Alcaín-Partearroyo subrayan los riesgos de este sistema. Por un lado, revistas españolas se ven privadas de artículos de calidad al preferir sus autores remitirlas a las revistas del *Scsi*, mientras que en España las consideradas mejores aumentarán sus tasas de rechazo y algunas se verán condenadas a la desaparición al no recibir originales suficientes. Ciertamente que estas últimas pueden acabar nutriéndose de los trabajos rechazados por las primeras, pero los más que pausados ritmos de revisión y publicación adoptados por algunas revistas consideradas "buenas" condenan a algunas investigaciones a la obsolescencia antes de su publicación.

Por último no puede pasarse por alto la presencia de los autores españoles en las revistas del *Scsi*. Los criterios de evaluación sin duda han obligado a los investigadores en comunicación a mirar al exterior. De tener una presencia poco más que testimonial, se ha pasado a que España sea el cuarto país europeo más productivo (como se ha dicho antes). Y cabe esperar que ese peso aumente en los próximos años gracias a la incorporación de 3 publicaciones españolas al *Scsi*.

Desde la óptica de los procesos de evaluación, los investigadores españoles tenderán a publicar en las revistas que priorizan las agencias de acreditación. Otra cosa distinta será ver si esos trabajos ganan visibilidad internacional y si su presencia en el *Scsi* se traduce también en un aumento de las citas recibidas, lo que aportaría un indicador más sobre la mejora cualitativa de la producción española en comunicación.

Notas y referencias bibliográficas

1. **Torres-Salinas, Daniel; Bordons, María; Giménez-Toledo, Elea; Delgado-López-Cózar, Emilio; Jiménez-Contreras, Evaristo; Sanz-Casado, Elías.** "Clasificación integrada de revistas científicas (CIRC): propuesta de categorización de las revistas en ciencias sociales y humanas". *El profesional de la información*, 2010, noviembre-diciembre, v. 19, n. 6, pp. 675-683. DOI: 10.3145/epi.2010.nov.1.
2. **Aneca.** *Principios y orientaciones para la aplicación de los criterios de evaluación. Programa de evaluación de profesorado para la contratación*, 2008. http://www.aneca.es/var/media/551398/pep_2010_07_ppios_070515.pdf
y:
Principios y orientaciones para la aplicación de los criterios de evaluación. Programa Academia, 2008. http://www.aneca.es/var/media/557171/academia_ppiosyorientaciones_100616.pdf
3. *Resolución de 18 de noviembre de 2009*, de la presidencia de la *Comisión Nacional Evaluadora de la Actividad Investigadora (CNAI)*, por la que se establecen los criterios específicos en cada uno de los campos de evaluación.
4. **Soriano, Jaume.** "El efecto Aneca". En: *Investigar la comunicación. Asociación Española de Investigación de la Comunicación*. Santiago de Compostela, 30 de enero a 1 de febrero, 2008.
5. *In-recs*
<http://lec3.ugr.es/in-recs/iii/Comunicacion-fecha-2009.ht>
6. **De-Pablos, José-Manuel.** "Análisis de las revistas españolas de comunicación, tras la actualización del índice de impacto de 2009 (octubre de 2010)". En: *II Congreso internacional latino de comunicación social*. La Laguna, Noviembre, 2010. http://www.revistalatinacs.org/10SLCS/actas_2010/217_De_Pablos.html
7. **Giménez-Toledo, Elea; Alcaín-Partearroyo, María-Dolores.** "Estudio de las revistas españolas de periodismo". *Comunicación y sociedad*, 2006, v. XIX, n. 2, pp. 107-131. http://www.unav.es/fcom/comunicacionysociedad/es/articulo.php?art_id=62
8. Ver, por ejemplo: **Archambault, Eric;** et al. "Benchmarking scientific output in the social sciences and humanities: the limits of existing databases". *Scientometrics*, 2006, v. 68, n. 3, pp. 329-342. DOI: 10.1007/s11192-006-0115-z.
9. **Fernández-Quijada, David.** "El perfil de las revistas españolas de comunicación (2007-2008)". *Revista española de documentación científica*, 2010, v. 33, n. 4, pp. 551-579. DOI:10.3989/redc.2010.4.756. <http://redc.revistas.csic.es/index.php/redc/article/view/671>
10. Puede encontrarse un análisis reciente sobre la presencia de revistas españolas en la *Web of science* en el siguiente thinkepi: **Rodríguez-Yunta, Luis.** "La alegría va por barrios: revistas españolas como publicaciones fuente en la *Web of Science*". *Anuario ThinkEPI*, 2011, v. 5, pp. 187-191. <http://www.thinkepi.net/la-alegria-va-por-barrios>
11. **Jones, Daniel E.** "La comunicación en el escape-rate". En: Díaz-Nosty, Bernardo (dir.). *Tendencias'07. Medios de comunicación: el escenario iberoamericano*. Ariel: Barcelona; Madrid: Fundación Telefónica, 2007, pp. 395-408. http://www.infoamerica.org/primeralanuario_medios.pdf
12. **Lauf, Edmund.** "National diversity of major international journals in the field of communication". *Journal of communication*, 2005, v. 55, n. 1, pp. 139-151.
13. **Masip, Pere.** "European research in communication during the years 1994-2004: a bibliometric approach". En: *1st European communication conference*. European Consortium for Communication Research, Amsterdam, 24-26 November 2005.
14. **Masip, Pere.** "Mapping communication research in Europe (1994-2009)". En: *3rd European communication conference*. European Communication Research and Education Association, Hamburg, 15-12 October 2010.
15. Para el análisis realizado en 2010 se tuvieron en cuenta las mismas 35 publicaciones que se analizaron en 2005.
16. Editada por la norteamericana *Association for Education in Journalism and Mass Communication*. De los 1.519 artículos publicados entre 1999 y 2004, sólo 12 eran de investigadores europeos.
17. La carencia de revistas de periodismo en el *Ssci* se verá parcialmente corregida en la próxima edición con la incorporación de *Journalism: theory, practice and criticism*, además de *Ecquid novi-African journalism studies* (que como su propio título indica, está dedicada específicamente al estudio del periodismo africano), y la española *Estudios sobre el mensaje periodístico*.

G.6. Google Scholar: no es oro todo lo que reluce

Por Isidro F. Aguillo

11 enero 2011

Aguillo, Isidro F. "Google Scholar: no es oro todo lo que reluce".
Anuario ThinkEPI, 2011, v. 5, pp. 211-215.



Resumen: *Luego de una corta perspectiva histórica de las bases de datos bibliográficas con citas, que permiten realizar estudios bibliométricos, se comentan las características de Google Scholar como posible base de datos para la misma utilización. Si bien Scholar es gratuito y más exhaustivo que WoS y Scopus, no se aconseja su uso para realizar análisis bibliométricos con fines de evaluación de personas e instituciones.*

Palabras clave: *Google Scholar, Google Académico, WoS, Web of science, Scopus, Bases de datos bibliográficas, Citas, Bibliometría.*

Title: **Google Scholar: all that glitters is not gold**

Abstract: *After a short historical perspective of bibliographic databases with citations, which allow bibliometric studies, the features of Google Scholar as a possible database for the same purpose are discussed. While Scholar is free and more comprehensive than WoS and Scopus, its use is not recommended for bibliometric analysis, especially for the evaluation of individuals and institutions.*

Keywords: *Google Scholar, WoS, Web of science, Scopus, Bibliographic databases, Appointments, Bibliometrics.*

Introducción

EL FACTOR LIMITANTE en los estudios de la actividad científica, especialmente los que utilizan técnicas cuantitativas, ha sido la disponibilidad de bases de datos.

La bibliometría de las últimas décadas no hubiera sido posible sin las bases de *ISI/Thomson (citation indexes)*¹. La explosión de la patentometría coincide con el acceso en abierto de los servicios web de las organizaciones de patentes europeas, estadounidenses y japonesas y, en fin, la cibermetría existe en buena medida por las bondades de los motores de búsqueda comerciales.

En muchos casos dichas bases de datos no habían sido diseñadas específicamente para la actividad bibliométrica y fue necesario (y todavía lo es) realizar un importante esfuerzo de selección, limpieza, organización y normalización de los resultados, antes de comenzar cualquier análisis.

Los costes eran enormes (acceso vía *Dialog*², adquisición de versiones en cd-rom) y lo siguen siendo (licencias nacionales *WoK*³ y *Scopus*), y, además de ciertas limitaciones legales, estaban las de carácter técnico. Éstas son relevantes para entender la tipología y profundidad de los traba-

jos bibliométricos realizados en los 80 y los 90. Era difícil exportar grandes cantidades de registros, ciertos campos tenían múltiples valores difíciles de segregar (autores, direcciones, citas), había que repasar errores y normalizar entradas, era complejo hacer correspondencias entre autores y sus direcciones cuando varios tenían la misma afiliación institucional.

La imposibilidad práctica de corresponder referencias con artículos generalizó el uso de las citas "esperadas" (el *infame* factor de impacto, por el que se supone a cada artículo particular el valor de la revista), en vez de utilizar las "observadas" o reales de cada uno. Otras consecuencias fueron el desprecio hacia los recuentos fraccionados de los cada día más frecuentes trabajos multiautorados o el insólito filtrado temático por categorías disciplinares de revistas o por selección de palabras clave (¡en bases de datos sin auténtica indización!). Todo ello motivado por las limitaciones de contenido y estructura de las bases de datos, pero también por un sistema de gestión intencionadamente capado que impedía una adecuada automatización de ciertos procesos.

La consecuencia directa es que el usuario final del trabajo del bibliómetro (otros colegas, fundamentalmente aquéllos objeto de análisis, y los gestores de instituciones y de políticas científicas)

apenas se reconoce en los resultados, que pueden pecar tanto de excesiva sencillez (plenos de errores) como de inaguantable profundidad (tablas densísimas, sin ninguna utilidad práctica).

“La dificultad de indizar la llamada internet invisible motivó la elaboración de un producto que no dependiera de los robots automáticos”

Hubo y sigue habiendo (cada vez menos, eso sí) trabajos mediocres, pero quizá la principal carencia es la ausencia de escenarios generales, con históricos de datos correctamente organizados y que evitara la continua reinención de la rueda a la que nos tiene acostumbrada esta disciplina en nuestro país (aunque en este caso la culpa es compartida por la inaudita ausencia de un manual de calidad, actualizado en castellano, del conjunto de las disciplinas cuantitativas).

Obviamente esta nota no es la primera que llama la atención sobre el cuidado extremo que se ha de tener tanto a la hora de seleccionar las fuentes bibliográficas como en el diseño de la extracción y utilización de los datos correspondientes. Y es posible que vuelva a caer en saco roto.

Google Scholar

Para los afortunados que trabajan en instituciones que se pueden permitir el indecente dispendio de tener contratadas las dos grandes bases de datos de citas (*Web of science* y *Scopus*), la labor bibliométrica se hizo un poco más compleja con la aparición de éste segundo, el nuevo producto de *Elsevier*. No sólo las bases de datos eran diferentes (*Scopus* es ligeramente mayor y con menor sesgo anglosajón), sino que las herramientas de consulta y extracción y los indicadores (externos en el caso de *Scopus*) eran también distintos.

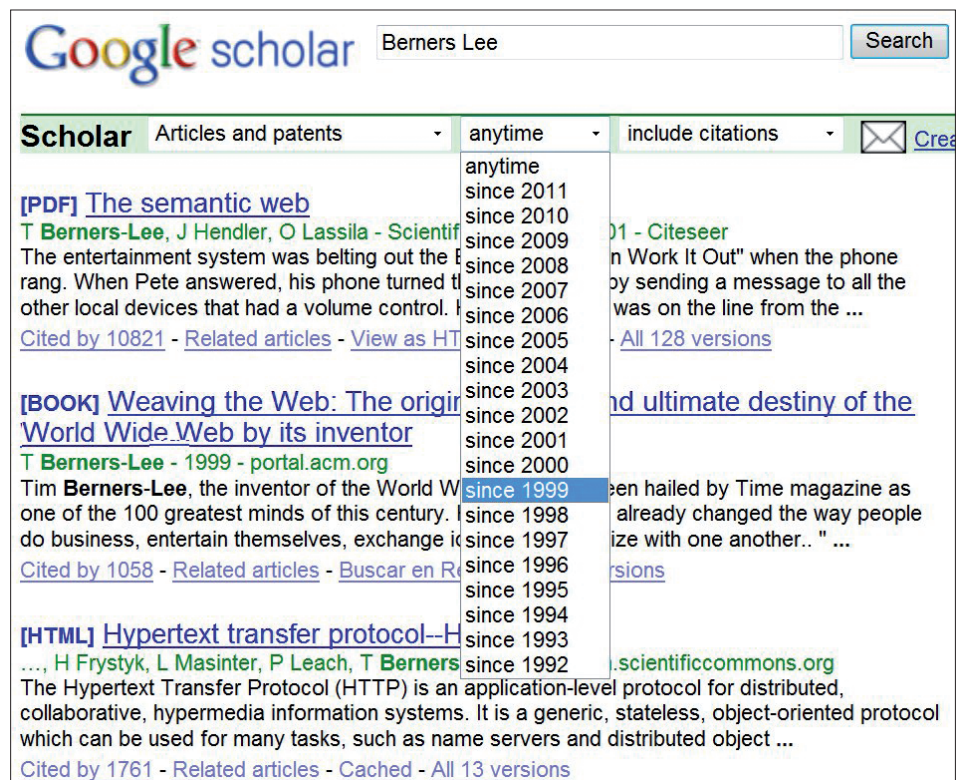
Sin embargo las ventajas se impusieron, ya que la competencia mejoró las prestaciones del *WoS* (ventanas de citación más amplias, nuevos indicadores) y su cobertura geográfica (con un cierto número de revistas no anglosajonas, sobre todo de ciencias sociales y humanas –que al parecer tienen un significativo menor impacto–). A medio plazo, trabajos de fusión de ambas bases de datos proporcionarán una mejor idea de las bondades y limitaciones de cada una de ellas, pero mientras tanto, cabe esperar la multiplicación de estudios disciplinares y/o temporales que remeden los ya realizados previamente con ayuda de *WoS*.

En ese contexto apareció un nuevo e interesante actor, *Google Scholar*, la base de datos académica del famoso buscador⁴.

Dentro de la estrategia global de *Google* de recolectar toda la información posible⁵, la dificultad de indizar la llamada internet invisible motivó la elaboración de un producto que no dependiera de los robots automáticos.

“La opacidad de Google respecto a las fuentes que utiliza ha dificultado el análisis global del buscador académico”

Scholar se nutre de una serie de acuerdos con productores y distribuidores de bases de datos



académicas y científicas de todo el mundo que ceden sus registros bajo distintas condiciones (tanto la lista de suministradores como los detalles de los contratos son secretos comerciales de Google).

Google proporciona ciertos valores añadidos (citas, enlaces, etiquetas) además de añadir la gigantesca sección académica de la web visible que aparece en el buscador general.

El resultado es una gran base de datos bibliográfica multidisciplinar que incluye citas a los diferentes artículos (fundamentalmente como ayuda a la recuperación). Es decir, es el tercer gran sistema de citas junto con WoS y Scopus, con la ventaja de su mayor tamaño y el hecho fundamental de ser de acceso gratuito. Se trata de un producto todavía en versión beta (¡desde 2004!), cuyo futuro no está garantizado y que al parecer es mantenido por un equipo muy reducido. Todo ello podría explicar la falta de normalización documental, muy necesaria en un producto multifuente tan heterogéneo formal y sustantivamente.

A pesar de los distintos problemas documentales de Google Scholar, la reciente aparición del software gratuito *Publish or perish*⁶, que permite la captura directa de los registros y calcula automáticamente diversos indicadores (incluyendo distintas variantes del índice h), ha renovado y generalizado el interés por Scholar en la comunidad bibliométrica.

En la bibliografía de esta nota figura una selección de artículos que tratan fundamentalmente dos áreas: la comparación directa de Google Scholar con las otras grandes bases de datos de citas (WoS y Scopus), y la utilización de registros de Scholar para la realización de estudios bibliométricos.

Los árboles no dejan ver el bosque

La opacidad de Google respecto a las fuentes que utiliza (y la evolución temporal de dicha cobertura, que parece se incrementó significativamente en los últimos años) ha dificultado el análisis global del buscador académico. De hecho, el diseño de muchos estudios comparativos implicaba utilizar básicamente instituciones y autores de reconocido prestigio, para los que se obtenía una cierta equivalencia con los resultados obtenidos en los productos de "calidad contrastada" (basados más o menos en núcleos de Bradford). Las diferencias en los estudios disciplinares se atribuían a diferencias de cobertura y, en fin, otras discrepancias se atribuían a problemas y limitaciones técnicas que se trataban de describir y evaluar o simplemente se citaban sin más, como pretexto.

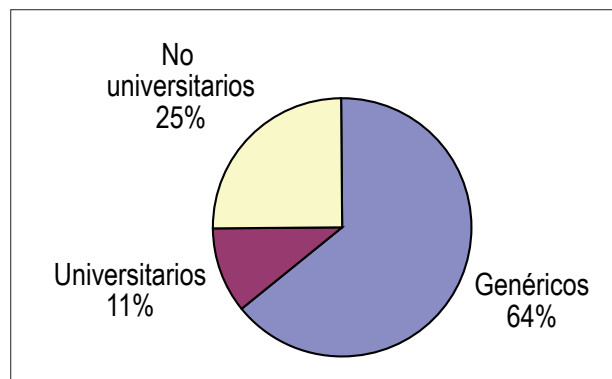
En el curso de un estudio cibernético sobre la distribución institucional de los contenidos recogidos en Google Scholar, descubrimos que las discrepancias son mayores de lo que se estimaba y que de hecho esta base no es comparable a WoS o Scopus, y su uso bibliométrico puede estar desaconsejado como norma general.

Se recogieron los registros totales (al menos con resumen) que aparecen en Scholar para dos grupos de dominios: 225 top level domains (incluyendo dominios nacionales como .es, .fr o .it, y los genéricos tales como .com, .org o .net) y 10.442 dominios universitarios (por ejemplo: ucm.es, harvard.edu u ox.ac.uk).

"Google Scholar es el tercer gran sistema de citas junto con WoS y Scopus, con la ventaja de su mayor tamaño y ser gratuito"

De la primera población se obtuvo un total de 86 millones de registros, de los que 55 millones (el 64%) correspondían a dominios genéricos, lo que cabría esperar de productores y distribuidores comerciales (.com) u organizaciones sin ánimo de lucro fuertemente presentes en este "mercado" (.org). Hay que tener en cuenta que Google Scholar muestra registros únicos, que "unifica" duplicados, es decir registros que pueden aparecer en repositorios institucionales o páginas personales pero que están también recogidos en distribuidores comerciales.

El segundo grupo (universidades) proporcionó 9 millones de registros, que supone un 10,6% del total obtenido en la estimación global de 86 millones, lo que implica que hay alrededor de un cuarto de los contenidos que bajo bandera nacional (dominio propio) son provistos desde instituciones no universitarias (productores locales, centros de investigación, portales, bibliotecas y repositorios digitales).



Results							
Papers:	97	Cites/paper:	2.84	h-index:	7	AWCR:	23.23
Citations:	275	Cites/author:	260.06	g-index:	16	AW-index:	4.82
Years:	41	Papers/author:	52.14	hc-index:	5	AWCRpA:	21.50
Cites/year:	6.71	Authors/paper:	2.70	hI-index:	6.13	e-index:	13.82
				hI,norm:	6	hm-index:	6.50

Cites	Per year	Rank	Authors	Title	Year
101	8.42	1	L Codina	Evaluación de recursos digitales en l...	2000
65	5.42	2	L Codina	El libro digital y la WWW	2000
29	1.81	3	L Codina	El libre digital	1996
18	1.00	4	L Codina	Modelo conceptual de un sistema d...	1994
11	1.83	5	L Codina	El libro digital y el territorio de la lec...	2006
8	0.47	7	LL CODINA	Teoría de recuperación de informaci...	1995
8	0.00	8	L Codina...	Web y cine: análisis comparativo de...	
5	0.00	13	L Codina	El nou sector emergent dels bancs ...	
3	0.16	16	L CODINA	Software a prueba: Windows perso...	1993
3	0.14	15	LL Codina	Introducción a las bases de datos d...	1991

Resultado parcial del análisis de la producción del autor Lluís Codina mediante el programa Publish or Perish, <http://www.harzing.com>

La muestra universitaria puede utilizarse para un análisis en más profundidad, aunque hay que advertir que en muchos casos se trata de producción hospedada, es decir, además de trabajos publicados por personal de la institución se pueden encontrar contribuciones de terceros, tales como presentaciones en congresos celebrados en la universidad hospedadora o material didáctico producido por otros autores pero puesto a disposición (posiblemente sin cobertura contractual) por el profesorado propio.

En dicho análisis aparecen las sorpresas, ya que tras EUA, los siguientes países mejor representados son respectivamente España, Brasil y Taiwán (por delante de Japón, Alemania, Canadá y Reino Unido). Entre los veinte primeros aparecen también Costa Rica, México e Indonesia.

Descendiendo a nivel institucional, tras *Harvard* (base de datos de astronomía) se encuentran *Pennsylvania State University (CiteSeerX)*, la *Universidad de La Rioja (Dialnet)*, *Johns Hopkins University (MUSE)*, *Catie* (Costa Rica, base de datos de agronomía), *Universidad Complutense de Madrid (CompluDoc)* o la *Universidad Autónoma del Estado de México (Redalyc)*.

Es decir, de acuerdo con las actuales políticas institucionales, sus páginas web buscan reflejar no sólo la producción de "excelencia" de la universidad, sino todos los resultados independientemente de su calidad y tipo, e incluso hospedando producción de terceros, ya sea puntualmente o exhaustivamente como parte de consorcios amplios. *Google Scholar* está recogiendo y reflejando todo ello (y cada vez más, a medida que las iniciativas *open access* van triunfando, aunque sea lentamente).

En resumen, *Google Scholar* es una interesante herramienta de recuperación de información, con limitaciones derivadas de su falta de control documental, que se pueden soslayar dado su

tamaño y el hecho de ser gratuita. La oferta de citas bibliográficas claramente incrementa su valor, pero la evolución reciente la aleja cada día más de aquellas que filtran contenidos de acuerdo con criterios de calidad (¿o impacto?). Este ruido extra desaconseja su uso liberal en los estudios bibliométricos, especialmente aquellos que tengan fines evaluativos.

“Scholar es una interesante herramienta de recuperación de información, pero se desaconseja su uso liberal en los estudios bibliométricos”

Notas

1. *ISI (Institute for Scientific Information)* es la empresa que en 1960 fundó **Eugene Garfield**, creador de las 3 bases de datos *Science Citation Index*. Fue comprada por *Thomson Reuters* en 1992.
2. *Dialog*, fundada por **Roger K. Summit** en 1980, fue comprada por *Thomson* en 2000, y revendida a *ProQuest* en 2008.
3. *WoK (Web of knowledge)* es el nombre comercial de un paquete de bases de datos de *Thomson Reuters* que incluye *WoS* (los 3 *citation indexes*), *Journal Citation Reports (JCR)*, *Biosis*, *Derwent*, y otras.
4. <http://scholar.google.com>
5. <http://www.google.com/corporate>
6. <http://www.harzing.com/pop.htm>

Referencias bibliográficas

Bar-Ilan, Judit. "A closer look at the sources of informetric research". *Cybermetrics*, 2009, v. 13, paper 4. <http://www.cindoc.csic.es/cybermetrics/articles/v13i1p4.pdf>

Bar-Ilan, Judit. "Citations to the 'Introduction to informetrics' indexed by *WoS*, *Scopus* and *Google Scholar*". *Scientometrics*, 2010, v. 82, n. 3, pp. 495-506. DOI: 10.1007/s11192-010-0185-9.

Bar-Ilan, Judit. "Which h-index? A comparison of *WoS*,

Scopus and Google Scholar". *Scientometrics*, 2008, v. 74, n. 2, pp. 257–271. DOI: 10.1007/s11192-008-0216-y. <http://sci2s.ugr.es/index/pdf/Bar-Ilan2008.pdf>

Beel, Joeran; Gipp, Bela. "Academic search engine spam and Google Scholar's resilience against it". *Journal of electronic publishing*, 2010, v. 13, n. 3. DOI: 10.3998/3336451.0013.305.

<http://quod.lib.umich.edu/cgilt/text/text-idx?c=jep;view=text;rgn=main;idno=3336451.0013.305>

García-Pérez, Miguel A. "Accuracy and completeness of publication and citation records in the Web of Science, PsycInfo, and Google scholar: a case study for the computation of h indices in psychology". *Journal of the American Society for Information Science and Technology*, 2010, v. 61, n. 10, pp. 2070-2085. DOI: 10.1002/asi.21372.

Harzing, Anne-Wil; Van-der-Wal, Ron. "A Google Scholar h-index for journals: an alternative metric to measure journal impact in economics and business". *Journal of the American Society for Information Science and Technology*, 2008, v. 60, n. 1, pp. 41-46.

Harzing, Anne-Wil; Van-der-Wal, Ron. "Google Scholar as a new source for citation analysis". *Ethics in science and environmental politics*, 2008, v. 8, n. 1, pp. 61-73. DOI: 10.3354/esep00076.

<http://www.int-res.com/articles/esep2008/8/e008p061.pdf>

Jacsó, Peter. "Google Scholar revisited". *Online information review*, 2008, v. 32, n. 1, pp. 102-114.

<http://www.cs.unibo.it/~cianca/wwwpages/dd/08Jacso.pdf>

Jacsó, Peter. "Savvy searching. Pragmatic issues in calculating and comparing the quantity and quality of research through rating and ranking of researchers based on peer reviews and bibliometric indicators from Web of Science, Scopus and Google Scholar". *Online information review*, 2010, v. 34, n. 6, pp. 972-982.

Kousha, Kayvan; Thelwall, Mike. "Sources of Google Scholar citations outside the Science Citation Index: a comparison between four science disciplines". *Scientometrics*, 2008, v. 74, n. 2, pp. 273-294. DOI: 10.1007/s11192-008-0217-x.

Li, Jie; Burnham, Judy F.; Lemley, Trey; Britton, Robert M. "Citation analysis: comparison of Web of Science, Scopus, SciFinder, and Google Scholar". *Journal of electronic resources in medical libraries*, 2010, v. 7, n. 3, pp. 196-217. DOI: 10.1080/15424065.2010.505518.

Mayr, Phillip; Walter, Anne-Kathrin. "An exploratory study of Google Scholar". *Online information review*, 2007, v. 31, n. 6, pp. 814-830.

<http://www.ib.hu-berlin.de/~mayr/arbeiten/OIR-Mayr-Walter-2007.pdf>

Meho, Lokman I.; Yang, Kiduk. "Impact of data sources on citation counts and rankings of LIS faculty: Web of Science vs. Scopus and Google Scholar". *Journal of the American Society for Information Science and Technology*, 2007, v. 58, n. 13, pp. 2105-25. DOI: 10.1002/asi.v58:13.

Mikki, Susanne. "Comparing Google Scholar and ISI Web of Science for earth sciences". *Scientometrics*, 2010, v. 82, n. 2, pp. 321-331. DOI: 10.1007/s11192-009-0038-6.

Torres-Salinas, Daniel; Ruiz-Pérez, Rafael; Delgado-López-Cózar, Emilio. "Google Scholar como herramienta para la evaluación científica". *El profesional de la información*, 2008, v. 18, n.5, pp. 501-510. DOI: 10.3145/epi.2009.sep.03.

White, Bruce. "Examining the claims of Google Scholar as a serious information source". *New Zealand library & information management journal*, 2006, v. 50, n. 1, pp. 11-24.

<http://muir.massey.ac.nz/bitstream/10179/571/5/GoogleScholar.pdf>