



Observatorio

Tecnologías, investigación y futuro de la profesión
Isidro F. Aguillo

Artículos

Future trends in information technologies for information units
Marshall Breeding

Creación automática de sistemas de búsqueda de respuestas en dominios restringidos
Katia Vila, José-Norberto Mazón y Antonio Ferrández

Uso de tecnologías semánticas y análisis visual para la anotación automática de imágenes y su recuperación
Sergio Rodríguez-Vaamonde, Pilar Ruiz y Marta González

Futuro de las revistas científicas de comunicación en España
Antonio J. Baladrón y Beatriz Correyero

World shares of publications of the USA, EU-27, and China compared and predicted using the new *Web of Science* interface vs *Scopus*
Loet Leydesdorff

Valoración de editoriales especializadas en Comunicación, Biblioteconomía y Documentación: encuesta a profesores e investigadores
Elea Giménez-Toledo y Carlos Tejada-Artigas

Uso y credibilidad de fuentes periodísticas 2.0 en Portugal y España
João Canavilhas y Begoña Ivars-Nicolás

La precariedad te hace dócil. Problemas que afectan a la profesión periodística
Mònica Figueras, Marcel Mauri, Salvador Alsius y Francesc Salgado

Análisis

El futuro de la Web ante la neutralidad de la Red: estado de la cuestión en la Unión Europea
Cristina Cullell-March

El libro electrónico: invarianzas y transformaciones
José-Antonio Cordón y Carlos A. Lopes



Revista internacional científica y profesional sobre documentación, comunicación, bibliotecas, sistemas y tecnologías de la información.



FUTURO WEB DE LA BIBLIOTECAS

00:12 AM -DRIVING/CONNECTED TO THE CLOUD



El profesional de la

información

Revista bimestral fundada en 1992 por
Tomàs Baiget y Francisca García-Sicília

El profesional de la información es una revista de
la editorial EPI SCP

Apartado 32.280 - 08080 Barcelona

+34-609 352 954

<http://www.elprofesionaldelainformacion.com>

Redacción

El profesional de la información

Apartado 32.280

08080 Barcelona

+34-934 250 029

epi@elprofesionaldelainformacion.com

Publicidad

Tel.: +34-609 352 954

publici@elprofesionaldelainformacion.com

Suscripciones

El profesional de la información

Apartado 32.280

08080 Barcelona, España

suscripciones@elprofesionaldelainformacion.com

<http://www.elprofesionaldelainformacion.com/suscripciones.html>

Servicios online

María T. Moreno

mt.moreno@ono.com

Diseño

MASmedios, <http://www.masmedios.com>

Director artístico: Moisés Mañas

Maquetación

SA de Litografía

Producción e Impresión

SA de Litografía

Ramon Casas, 2

08911 Badalona, Barcelona

Tel. +34 - 933 847 676

<http://www.sadelitografia.com/>

Distribución online

MetaPress, Birmingham, Alabama, EUA

<http://elprofesionaldelainformacion.metapress.com>

Depósito legal: B. 12.303-1997

Los trabajos publicados en EPI son aprobados según el sistema tradicional "peer review" en doble ciego: son revisados al menos por dos expertos en el tema, del Consejo Asesor de la revista y/o externos.

Para conseguir que los trabajos no pierdan actualidad, la dirección y los evaluadores de esta revista ponen especial esfuerzo en revisar los artículos con gran rapidez, consiguiendo un tiempo medio de aceptación o rechazo de los trabajos de sólo unas pocas semanas.

DIRECCIÓN EDITORIAL

Tomàs Baiget

EPI SCP

<http://www.baiget.com>

SUBDIRECTOR

Javier Guallar

Univ. de Barcelona / Univ. Ramon Llull / Univ. Oberta de Catalunya

<http://sites.google.com/site/sitiodeljavierguallar/>

COORDINADOR EDITORIAL

Carlos Tejada-Artigas

Universidad Complutense de Madrid

tejada@ccdoc.ucm.es

REDACTORA JEFE

Isabel Olea

Universidad de León

isabel.iolea@gmail.com

REDACCIÓN

Natalia Arroyo-Vázquez

Fundación Germán Sánchez Ruipérez

narroyo@fundaciogsr.es

Lluís Codina

Universitat Pompeu Fabra

<http://www.lluiscodina.com>

Ricardo Eito-Brun

Grupo GMV

reito@gmv.es

Elea Giménez-Toledo

Inst. de Estud. Document. sobre Ciencia y Tecnología

elea.gimenez@cchs.csic.es

Javier Leiva-Aguilera

Catorze.com

<http://www.javierleiva.info>

Toon Lowette

Grid Electronic Publishing

toon@grid.be

Roser Lozano

CRAI Universitat Rovira i Virgili

roser.lozano@urv.cat

José-Antonio Millán

Libros y bitios

<http://jamillan.com>

Fernanda Peset

Universidad Politécnica de Valencia

mpesetm@upv.es

Jorge Serrano-Cobos

MASmedios

jorgeserrano@gmail.com

Daniel Torres-Salinas

Universidad de Navarra

torressalinas@gmail.com

REVISIÓN DE LENGUA INGLESA

Elaine M. Lilly

Writer's First Aid

elaine@writersfirstaid.com

CONSEJO ASESOR

Ernest Abadal

Universitat de Barcelona, Barcelona.

Isidro F. Aguillo

Centro de CC Humanas y Sociales, CSIC, Madrid.

Ramon Alberch

Generalitat de Catalunya, Barcelona.

Adela d'Alòs-Moner

Doc6, Barcelona.

Ricardo Baeza-Yates

Univ. de Chile, Santiago, Chile. Yahoo! Research, Barcelona.

Carlos B. Amat

Inst. Agroquím. y Tecn. Alimentos, CSIC, Valencia.

Jesús Bustamante

Biblioteca, Cedefop, Salónica, Grecia.

Carlota Bustelo-Ruesta

Consultora, Madrid.

Emilio Delgado-López-Cózar

Universidad de Granada, Granada.

Javier Díaz-Noci

Universitat Pompeu Fabra, Barcelona.

Assumpció Estivill

Universitat de Barcelona, Barcelona.

Antonia Ferrer-Sapena

Univ. Politécnica de Valencia, Valencia.

António Fidalgo

Universidade da Beira Interior, Portugal.

Francisco-Javier García-Marco

Universidad de Zaragoza, Zaragoza.

Paola Gargiulo

Caspar, Roma, Italia.

Johannes Keizer

Food and Agriculture Org. (FAO). Roma, Italia.

Thomas Krichel

Palmer School of Libr. & Inform. Sci. LIU, NY, USA.

Victoria Manglano

Ovid Technologies, Madrid.

Mari-Carmen Marcos

Universitat Pompeu Fabra, Barcelona.

Pere Masip

Blanquerna, Univ. Ramon Llull, Barcelona.

Charles McCallieNeville

Opera Software, Oslo, Norway.

Marcos Palacios

Universidade Federal da Bahia, Brasil.

Joan Roca

Minnesota State University, Mankato, USA.

Ramón Salaverría

Universidad de Navarra, Pamplona.

Robert Seal

Loyola Univ. Chicago, Evanston, Illinois, USA.

Ernesto Spinak

Consultor, Montevideo, Uruguay.

Jesús Tramullas

Universidad de Zaragoza, Zaragoza.

La revista **El profesional de la información** está referenciada en:

BASES DE DATOS

Academic search premier (Ebsco)

<http://www.ebscohost.com/academic/academic-search-premier>

Compludoc (Universidad Complutense de Madrid)

<http://europa.sim.ucm.es/compludoc/>

Dialnet (Universidad de La Rioja)

http://dialnet.unirioja.es/servlet/revista?&clave_revista=469

Economía y negocios (Ebsco)

<http://www.ebscohost.com/academic/economia-y-negocios>

Francis (Inist)

<http://www.inist.fr/spip.php?article23>

ISI Social science citation index, Social SCI, WoS (Thomson Reuters)

http://wokinfo.com/products_tools/multidisciplinary/webofscience/ssci/

Impact Factor 2010 = 0,375

Inspec, Information services in physics, electronics and computing (IET, The Institution of Engineering and Technology)

<http://www.theiet.org/publishing/inspec/>

ISOC, Índice español de ciencias sociales y humanidades (Iedcyt)

<http://bddoc.csic.es:8080/ver/ISOC/revi/0721.html>

ISTA, Information science and technology abstracts (Ebsco)

<http://www.ebscohost.com/public/information-science-technology-abstracts>

Lisa, Library and information science abstracts (CSA)

<http://www.csa.com/factsheets/lisa-set-c.php>

Lista, Library, information science & technology abstracts (Ebsco)

<http://www.libraryresearch.com>

Pascal (Inist)

<http://www.inist.fr/spip.php?article22>

Scopus (Elsevier) Scimago Journal Rank 2010 = 0,030

<http://www.scopus.com>

CATÁLOGOS Y SERVICIOS BIBLIOTECARIOS

Argos-Bolsum (GVA, Generalitat Valenciana)

http://www.pre.gva.es/argos/es/contenido_general/recursos/bolsum/

Catálogo de Latindex

<http://www.latindex.unam.mx/buscador/ficRev.html?folio=6772>

In-Recs, Revistas españolas de ciencias sociales Grupo EC3, Evaluación de la Ciencia y de la Comunicación Científica, Universidad de Granada

<http://ec3.ugr.es/in-recs/iii/Documentacion-fecha-2009.htm>

Registros bibliográficos para bibliotecas públicas españolas (Rebeca)

<http://www.mcu.es/bibliotecas/MC/Rebeca/>

ACCESO A LOS TEXTOS COMPLETOS

MetaPress (2000-)

<http://elprofesionaldelainformacion.metapress.com/>

ALPSP Learned journals collection (ALJC) (2011-)

<http://aljic.swets.com>

Ebscohost Electronic Journals Service (2000-embargo 1 año)

<http://ejournals.ebsco.com/direct.asp?JournalID=105302>

Library, information science & technology abstracts with full text (2000-embargo 1 año)

<http://www.ebscohost.com/thisTopic.php?marketID=1&topicID=584>

Academic search complete (2000-embargo 1 año)

<http://www.ebscohost.com/thisTopic.php?marketID=1&topicID=633>

Business source complete (2000-embargo 1 año)

<http://www.ebscohost.com/academic/business-source-complete>

El profesional de la información (1992-embargo 2 años)

<http://elprofesionaldelainformacion.com/contenidos.html>

SwetsWise (2000-)

<https://www.swetswise.com>

EPI EN FACEBOOK

<http://www.facebook.com/elprofesionaldelainformacion>

EPI EN TWITTER

http://twitter.com/revista_EPI

PLATAFORMA DE PRODUCCIÓN OJS

Recyt, Repositorio español de ciencia y tecnología (Fecyt)

<http://recyt.fecyt.es/index.php/EPI>

Sumario **EPI, 2012, v. 21, n. 1**

Tema central: El futuro de la Web

OBSERVATORIO

- 5 **Tecnologías, investigación y futuro de la profesión**
Isidro F. Aguillo

ARTÍCULOS

- 9 **Current and future trends in information technologies for information units**
Marshall Breeding
- 16 **Creación automática de sistemas de búsqueda de respuestas en dominios restringidos**
Katia Vila, José-Norberto Mazón y Antonio Ferrández
- 27 **Uso combinado de tecnologías semánticas y análisis visual para la anotación automática de imágenes y su recuperación**
Sergio Rodríguez-Vaamonde, Pilar Ruiz-Ibáñez y Marta González-Rodríguez
- 34 **Futuro de las revistas científicas de comunicación en España**
Antonio J. Baladrón-Pazos y Beatriz Correyero-Ruiz
- 43 **World shares of publications of the USA, EU-27, and China compared and predicted using the new Web of Science interface versus Scopus**
Loet Leydesdorff
- 50 **Valoración de editoriales especializadas en Comunicación, Biblioteconomía y Documentación: encuesta a profesores e investigadores**
Elea Giménez-Toledo y Carlos Tejada-Artigas
- 63 **Uso y credibilidad de fuentes periodísticas 2.0 en Portugal y España**
João Canavilhas y Begoña Ivars-Nicolás
- 70 **La precariedad te hace dócil. Problemas que afectan a la profesión periodística**
Mónica Figueras-Maz, Marcel Mauri-Ríos, Salvador Alsius-Clavera y Francesc Salgado-De-Dios

ANÁLISIS

- 77 **El futuro de la Web ante la neutralidad de la Red: estado de la cuestión en la Unión Europea**
Cristina Cullell-March
- 83 **El libro electrónico: invarianzas y transformaciones**
José-Antonio Cordón-García y Carlos A. Lopes
- 91 **Gestión de objetos digitales con ContentDm: estudio de caso en el Consorci de Biblioteques Universitàries de Catalunya y el Institut Cartogràfic de Catalunya**
Rafael Roset, Sandra Reoyo-Tudó y Robert Torre-Marín
- 98 **Registro, almacenamiento y acceso simultáneo de vídeo streaming sobre protocolo UDP. El sistema MReclPTV de la Universitat Pompeu Fabra**
Frederic Guerrero-Solé e Hibai López-González

SOFTWARE DOCUMENTAL

- 105 **Carrot2: búsqueda y visualización de la información**
Audilio Gonzales-Aguilar y María Ramírez-Posada

INDICADORES

- 113 **Sistema de indicadores culturales local (Siclo)**
Vicente Coll-Serrano, Salvador Carrasco-Arroyo, Olga Blasco-Blasco y Luis Vila-Lladosa

AGENDA

120 INFORMACIÓN PARA LOS AUTORES



Creación de Archivos y Bibliotecas Virtuales

Desde la digitalización de materiales bibliográficos hasta la asignación de metadatos y su implementación en la red, conforme a la normativa internacional.

Productos para crear Bibliotecas Digitales y Virtuales

DIGIBIB 7.0

Solución avanzada para la creación de Bibliotecas Digitales y la Gestión Bibliotecaria Multilingüe.

DIGIARCH 2.0

Sistema digital de descripción y gestión archivística. Descripción en ISAD(G) y EAD 2.0.

DIGIDIR 2.1

Directorio para Archivos, Bibliotecas y Museos con generación automática de estadísticas y sistemas de información geográfica (GIS).

OASIS-PMH 2.0

Sistema integrado de recolección de diversos esquemas de metadatos:

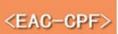
- DCMI sin cualificar
- MARC 21
- SWAP
- EAD
- mod_OAI
- DRIVER 2.0.

ADAPTACIÓN A EUROPEANA (FASE DANUBIO)

Implementación del esquema ESE 3.4 (Europeana Semantic Elements) y EDM 5.2.2 (Europeana Data Model) Adaptado a la Agenda Digital Europea 2020.

DIGITALIZACIÓN AVANZADA

Con asignación dinámica de metadatos.



- Recolección en la Web para Entidades e Instituciones de Memoria en OAI-PMH y Dublin Core cualificado con ESE 3.4
- Consultoría y mappings a EDM 5.2.2 (Europeana Data Model)
- Bibliotecas digitales que permiten la creación, recuperación y recolección de metadatos (MARCXML, DCMI y RDF y RDFs)
- Archivos Web que facilitan la creación, recuperación y recolección de metadatos (EAD 2.0 y EAC 2010)
- Implementación de la Europeaana OpenSearch API
- Adaptación del repositorio OAI para la transmisión de instancias RDF según ORE
- Repositorios Institucionales DIGIPRESV para Preservación Digital a largo plazo mediante PREMIS 2.1 y OAIS ISO 14721
- Intercambio de metadatos en METS 1.9 (diferentes Profiles) integrando todos los esquemas de metadatos
- Creación de METSRights para el control de los derechos de autor
- Reconocimiento Óptico de Caracteres OCR y generación dinámica de ALTO (Analyzed Layout and Text Object)
- Generación e integración de registros SKOS mediante MARC21(Up.12)/RDA
- Creación de eBooks o libros digitales en formatos: ePub y Mobipocket.
- Adaptación de DIGIBIB a Linked Open Data



Validación en el Data Providers de la Open Archives Initiative. Genera un Sitemap para Google.



Nº 6002188



C/ Claudio Coello, 123. Madrid. Tel.: 915 81 20 01. digibis@digibis.com

www.digibis.com



Isidro F. Aguillo



Isidro F. Aguillo coordina el *Laboratorio de Cibermetría* del *Centro de Ciencias Sociales y Humanas (CCHS)* del *Consejo Superior de Investigaciones Científicas (CSIC)*. Sus líneas de trabajo incluyen el desarrollo de indicadores de la actividad investigadora en la Web, la evaluación y análisis de las iniciativas de acceso abierto y los procesos de comunicación científica a través de revistas electrónicas y repositorios institucionales. Dirige o participa en varios proyectos de I+D de la Unión Europea (*Acumen*, *OpenAire*) y del *Plan Nacional de Investigación Científica*. Es el responsable de la publicación de los rankings web de: *Universidades* (desde 2004), *Centros de Investigación*, *Hospitales*, *Escuelas de Negocios* y *Repositorios del Mundo*, que reciben más de 5 millones de visitas por año. Ha sido miembro de la *Oficina Española de Ciencia y Tecnología (SOST)* en Bruselas

y *Metcalf* visitor professor en la *University of New South Wales* (Sidney, Australia). Ha publicado más de un centenar de artículos, impartido más de 400 conferencias y seminarios en más de 80 universidades españolas y extranjeras. Es el editor de la revista *Cybermetrics* desde 1997, la primera revista electrónica del CSIC. Participa en distintos comités organizadores y científicos de congresos, en grupos de trabajo de la *Comisión Europea* y en la evaluación y revisión de proyectos y artículos científicos. Es licenciado en biología por la *Universidad Complutense* y máster en información y documentación por la *Universidad Carlos III de Madrid*. En 2009 fue nombrado doctor *honoris causa* por la *University of Indonesia*, distinción que comparte, entre otros, con 5 jefes de estado. Su número de Erdős es 4.

Consejo Superior de Investigaciones Científicas (CSIC)

Centro de Ciencias Humanas y Sociales (CCHS), Laboratorio de Cibermetría

Albasanz, 26-28. 28037 Madrid

isidro.aguillo@cchs.csic.es

<http://www.cchs.csic.es/es/personal/isidro.aguillo>

Resumen

El éxito de los profesionales de la información en los próximos años dependerá de su compromiso con la investigación y desarrollo en las áreas de identificación, acceso, recuperación, almacenamiento, descripción y puesta en valor de la ingente cantidad de recursos disponibles. Este esfuerzo de investigación debe ser coordinado con los informáticos, no ya para recabar su colaboración en tareas técnicas y desarrollos puntuales, sino para realizar avances significativos en áreas de frontera de la ciencia de los contenidos. Se propone una mayor participación en proyectos, especialmente en consorcios europeos o internacionales, prestando especial atención a la aplicación de estándares, defensa del acceso abierto, explotación de nuevas interfaces, interconexión de datos y la preservación digital. El proceso debe incluir aspectos éticos, involucrar usuarios finales, promover plataformas de integración de recursos y extender la cobertura de formatos más allá del documento textual.

Palabras clave

Observatorio, Investigación en documentación, Acceso abierto, Web social, Interconexión de datos, *Linked data*, Metadatos.

Title: Technologies, research and future of the profession

Abstract

The success of information professionals in the coming years will depend on their commitment to research and development in the areas of identification, access, retrieval, storage, description and assessment of the vast amount of resources available. This research effort should be coordinated with computer scientists, not to seek their cooperation in specific technical tasks and developments, but to make significant progress in frontier areas of content sciences. Greater project participation by librarians is proposed, especially in European or international consortia, paying particular attention to the implementation of standards, advocacy of open access, exploitation of new interfaces, data networking and digitization. The process should include ethical issues, involve end users, promote resource integration platforms and extend formats coverage beyond text documents.

Keywords

Observatory, Information science research, Open access, Social web, Data networking, Linked data, Metadata.

Aguillo, Isidro F. "Tecnologías, investigación y futuro de la profesión". *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 5-7.

<http://dx.doi.org/10.3145/epi.2012.ene.01>

Introducción

EPI es posiblemente la revista profesional de documentación en español más volcada con los aspectos tecnológicos, cubriendo muchos de sus artículos las principales novedades e innovaciones que se producen en nuestro campo. Si de algo peca es, quizá, de presentar demasiados avances genéricos o implementaciones internacionales y pocos casos de estudio derivados de experiencias propias.

Pretendo dedicar este observatorio a identificar algunas tendencias, a sugerir líneas de trabajo presentes y futuras y a reclamar una mayor colaboración con los expertos informáticos cuyo papel, a veces por soberbia, otras por dejarlo relegado a misiones secundarias, o incluso por indolencia en ciertos casos, no ha alcanzado aún las cuotas de esfuerzo cooperativo que se requieren para un salto cualitativo en este sector específico, pero muy horizontal de nuestra disciplina.

Las interfaces de acceso a la información evolucionarán hacia el reconocimiento del habla, la comprensión del lenguaje natural y la traducción multilingüe instantánea

Entorno tecnológico

Empezaré señalando la importancia de la "fuerza bruta" en el desarrollo tecnológico, algo que a menudo no se valora suficientemente en las perspectivas que se realizan en nuestra profesión y que a la luz de la ley de Moore (ver la *Wikipedia*, ¡por supuesto!) seguirá proporcionando muchas novedades: me refiero a la reducción de tamaño, menor consumo y sustancial incremento de la capacidad de proceso. Si el boom de las tabletas fue propiciado por microprocesadores más potentes y memoria más rápida, mi bola de cristal muestra que las interfaces de acceso a la información evolucionarán hacia el reconocimiento del habla, la comprensión del lenguaje natural y la traducción multilingüe "instantánea". Todo ello es perfectamente posible ahora, pero está lastrado por la falta de potencia de la actual generación de "aparatos". *Apple* ya lo permite en sus nuevos *iPhones*, *Microsoft* lleva mucho tiempo trabajando en ello y seguro que es una de las futuras evoluciones que *Google* incluirá en próximas generaciones de *Android*.

Obviamente, aunque podemos adelantar la implantación de alguna de estas nuevas tecnologías, nuestra contribución a su desarrollo ha de ser limitada, por lo que estas predicciones parecen tener poco valor práctico en nuestro caso particular. ¡Nada más lejos de la realidad!, pues las interfaces están pensadas para acceder a la información, y la evolución descrita no hará sino hacer más difícil la gestión de volú-

menes ingentes de recursos. La información en la Web va a seguir creciendo a ritmos exponenciales, y sus formatos y estructuras se hacen cada vez más complejos, lo que podría aumentar las posibilidades de que nuestros profesionales fueran expulsados de ciertas áreas, como lamentablemente ya estamos viendo que ocurre en bastantes casos. Sirva como ejemplo la escasa presencia de aplicaciones (apps) en cualquiera de las tiendas online realizadas o inspiradas por documentalistas. Y eso que muchos desarrolladores buscan con ahínco recursos que implementar.

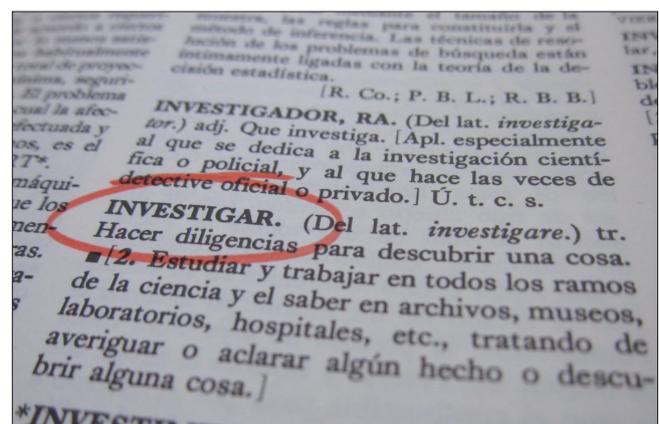
Integrar informática y documentación

Sin remontarnos demasiado atrás, tenemos un claro ejemplo en la construcción de los llamados "mashups" bibliotecarios, aplicaciones web relativamente simples que, salvo contadas excepciones, apenas se implantaron en nuestro país. Era más fácil apostar por sistemas bibliotecarios obsoletos, con interfaces poco amigables y nullos servicios de valor añadido. En buena medida esto ocurría por la escasa colaboración con los informáticos, opción esta al parecer no muy potenciada por ninguno de ambos colectivos.

Ciertamente, existen casos de éxito, donde esa colaboración ha sido fructífera y es de esperar que se perpetúe, aunque para ello hacen falta acciones específicas. Por ejemplo, queda mucho por andar en la explotación no ya de los documentos individuales, sino de los datos embebidos en los mismos, promoviendo su portabilidad y reutilización.

Queda mucho trabajo por hacer para explotar los datos embebidos en los documentos, promoviendo su portabilidad y reutilización

Por su lado los informáticos deben eliminar de lo más alto de su agenda la prioridad por la seguridad o los temas de privacidad. Esto abriría la puerta a una mejor explotación





de datos, donde nuestra labor debe ser imponer estándares abiertos y el libre acceso a la información y construir ricas colecciones de metadatos. El viejo sueño de la web semántica, nunca cercano pese a lo que se insista en las aulas universitarias, podría entonces estar más cerca.

Sin embargo, no basta una mera aproximación técnica donde unos y otros resuelvan problemas prácticos y desarrollos a medida demasiado personalizados.

La exigencia a la que me refiero tiene más que ver con la, ahora recortada, I+D. Sorprende ver la escasa participación de nuestras bibliotecas y centros de documentación en proyectos puros de investigación, donde el socio informático no sea un mero desarrollador contratado para resolver problemas que básicamente le son ajenos. En los *Programas marco* de la Unión Europea existe una amplia gama de áreas –aparte de las puramente bibliotecarias– donde nuestra contribución puede ser relevante si se coordina con departamentos universitarios de informática en la resolución de retos globales.

En qué colaborar

¿Cuáles serían las áreas más interesantes para este tipo de colaboraciones? En mi modesta opinión y sin ánimo de ser exhaustivo:

Archivo digital. La Web se renueva continuamente a costa de renacer de sus cenizas, unos contenidos volátiles que no se preservan. Son necesarios estándares, posiblemente legislación, pero sobre todo herramientas de recogida, etiquetado y preservación que garanticen un alto nivel de automatización del proceso y una escalabilidad que permita abordar proyectos de nivel nacional.

Web semántica. La interoperabilidad parece bastante resuelta (¡viva el XML, el RDF y las ontologías!), pero los estándares para los *open linked data* deben ser más diversos y la libertad de mercado (oferta/demanda) debe ser regulada más estrictamente para impedir monopolios. Mi preferencia es obviamente que los recursos se generen de forma cooperativa y se coloquen en el dominio público. La lingüística computacional, o como quiera que llamemos a esta disciplina, debe entrar en el curriculum de nuestros másteres. Y no sólo para enseñar html 5.

Acceso abierto. El diálogo entre los agentes involucrados (administraciones, instituciones, investigadores, bibliotecas) debe ser más amplio y profundo. Es inaceptable que los bibliotecarios hayan marginado a los usuarios finales (au-

tores de artículos científicos) de los repositorios, abriendo una brecha que quizá tenga más una solución técnica que política. Ampliaré este comentario en un trabajo que espero publicar próximamente.

Web 2.0 o Web social. La dictadura de ciertas herramientas muy populares (me refiero fundamentalmente a *Facebook*) que nos obligan a ciertos modelos de comunicación, es posible por una falta de propuestas más específicamente documentales para el desarrollo de soluciones personalizadas. Es significativo el poco uso de los APIs en la mayoría de nuestras páginas web a pesar de la multiplicación de los servicios ofertados. La solución no es el uso, es la integración.

Multimedia. El vídeo es un formato perfectamente catalogable, con herramientas informáticas capaces de describir contenidos, separar secuencias, transcribir diálogos o identificar música. Las folksonomías y las contribuciones de los usuarios pueden ayudar, pero el rubicón a cruzar es fundamentalmente tecnológico. Esa tecnología ya está en marcha y los informáticos necesitan corpus más grandes y anotados para proseguir sus desarrollos.

Web móvil. No veo la necesidad de invertir más esfuerzos en una web específica y distinta para los dispositivos móviles. Pero, por favor, ¿no habéis visto nunca una tableta en funcionamiento? Al parecer por ahí hay una que se llama *iPad*...

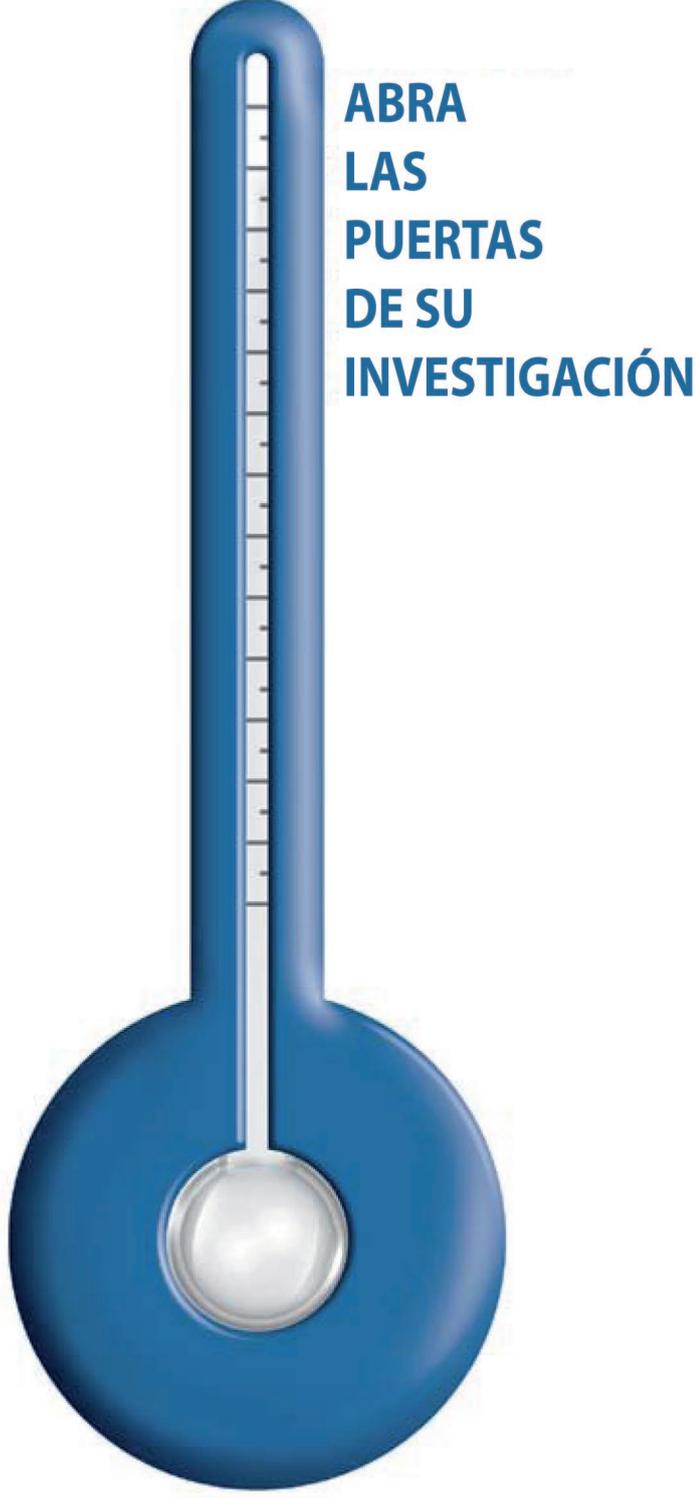
Evaluación de calidad. Un potencial para el cual estamos plenamente preparados, pero que aplicamos con dificultad y hasta con recelo. Me refiero fundamentalmente a la identificación y explotación de recursos de calidad, aplicando criterios que deben ser menos automáticos y más rigurosos. Ello incluye también aspectos éticos y comerciales, y citaríá como ejemplo palmario el uso de fuentes (carísimas) explotadas en régimen de monopolio y con reconocidos sesgos (*WoS/ISI-Thomson*, *Scopus/Elsevier*) cuando existen alternativas gratuitas (*Google Scholar*, *Microsoft Academic Search*) y con cobertura más amplia. En muchos casos dichas fuentes se ofrecen vía API para su mejor adaptación y más general utilización.

Nube. Hay dos posibles perspectivas respecto al uso de la nube, una institucional guiada por motivos económicos, pero lastrada por cuestiones de seguridad, y otra personal. Como tal opción personal y a la vista de los problemas de privacidad identificados, diría que de entrada ¡no gracias! (yo no tengo cuenta de *Gmail*, por ejemplo), aunque tal postura tiene fecha de caducidad.

“Necesitamos un mayor esfuerzo investigador para aplicar tecnología a los nuevos retos que afrontamos”

Colofón

Nuestro modelo de explotación de la información no está obsoleto, pero necesitamos un mayor esfuerzo investigador para aplicar tecnología a los nuevos retos que afrontamos. En ese contexto aún tenemos mucho que aportar. Utilizando una frase que he leído recientemente en una de nuestras revistas: “Los contenidos son el rey, pero los metadatos construyen imperios”.



La llave para maximizar su potencial de investigación

SciVal es una pionera suite de herramientas de búsqueda desarrollada por Elsevier que le permite explotar todo su potencial de investigación. A través de sus extensas bases de datos, las soluciones de SciVal le proporcionan la visión estratégica que necesita para permanecer a la vanguardia de la investigación.

SciVal le ayuda a tomar decisiones seguras de inversión en materia de investigación en su lucha por destacar en un mercado tan competitivo, permitiéndole para que pueda analizar minuciosamente su rendimiento en investigación interdisciplinar, evaluar sus estrategias de investigación y hallar su talento natural.

Descubra las nuevas oportunidades que SciVal puede ofrecerle en www.scival.com/unlock
Se mire como se mire, tiene sentido.



CURRENT AND FUTURE TRENDS IN INFORMATION TECHNOLOGIES FOR INFORMATION UNITS



Marshall Breeding



Marshall Breeding is director for *Innovative Technology and Research* at *Vanderbilt University Library* and the *Executive Director of the Vanderbilt Television News Archive*. He is founder and publisher of *Library Technology Guides*. He earned a Bachelor of Arts in Philosophy (1980) and a Masters of Arts (1982) from *Colorado State University* and a Masters of Arts from *Vanderbilt University* (1985). He serves as editor for *Smart Libraries Newsletter* published by *ALA TechSource*, writes *Systems Librarian* column for *Computers in Libraries* published by *Information Today*, and has authored the *Automation Marketplace* feature for *Library Journal* since 2002.

Vanderbilt University Library
110 Twenty-first Avenue South, Suite 704
Nashville, TN 37203, USA
<http://www.librarytechnology.org>
marshall.breeding@librarytechnology.org

Abstract

With the increasing dominance of electronic content and digital collections in academic libraries, the capabilities lacking in the current slate of automation systems has increasingly become an obstacle to progress. A new generation of digital services platforms for libraries is emerging, designed to provide a more comprehensive approach to the management and access to all formats of library materials: print, electronic and digital. These new systems involve a modernization of technology, embracing service-oriented architectures, availability for APIs to facilitate interoperability with external systems, support for task workflows more aligned with current operational realities, and more tightly coupled discovery interfaces that deliver access to library collections and services in a more unified and comprehensive manner. These new products, emerging in this era of cloud computing, have been designed for deployment through software as a service and rely on highly shared data models.

Keywords

Libraries, Library management systems, Integrated library systems, Web based services, Infrastructure, Digital collections, Digitization, Library collections, Discovery software, eBooks, eJournals, Software as a service, Cloud computing.

Título: Tendencias actuales y futuras en tecnologías de la información para unidades de información

Resumen

Con la fuerte irrupción de las bibliotecas en el acceso a los recursos electrónicos y la gestión de colecciones digitales, la oferta actual de sistemas de automatización se ha ido convirtiendo en un obstáculo para el progreso. Una nueva generación de plataformas de servicios digitales para bibliotecas está emergiendo, diseñada para proporcionar un apoyo integral para la gestión y el acceso de todos los formatos de materiales de la biblioteca: impresos, electrónicos y digitales. Estos nuevos sistemas implican una modernización de las arquitecturas orientadas a servicios, con un mayor desarrollo del concepto de "software como servicio" y de otros modelos basados en la "nube".

Palabras clave

Bibliotecas, Sistemas de gestión de bibliotecas, Sistemas integrados de bibliotecas, Servicios basados en la web, Infraestructura, Colecciones digitales, Digitalización, Software de descubrimiento, Libros electrónicos, Revistas electrónicas, Software como servicio, Computación en la nube.

Breeding, Marshall. "Current and future trends in information technologies for information units". *El profesional de la información*, 2011, v. 21, n. 1, pp. 9-15.

<http://dx.doi.org/10.3145/epi.2012.ene.02>

Nota: Este artículo puede leerse traducido al español en: <http://elprofesionaldelainformacion.com/breeding.html>

Artículo recibido el 25-01-12

Aceptación definitiva: 28-01-12

1. Introduction

This paper is based on a presentation made at the *IX Conferencia Internacional sobre Bibliotecas Universitarias*, organized by *Dirección General de Bibliotecas* of the *Universidad Nacional Autónoma de México (UNAM)*, México DF, October 26th-28th 2011.

In 1996 the *UNAM's General Directorate of Libraries* celebrated its thirtieth anniversary with a series of lectures and panel discussions in which the central idea was to imagine the features of the library of the future at that time. In this paper we intend to analyze how the library of the future is being built and how it would be in 15 years from now, in 2026.

Academic libraries today find themselves at a critical juncture between the automation systems of the past and a new genre of emerging technologies. This essay aims to look forward to anticipate the progression of technologies and automation systems. Five-year projections can extend the current trends to estimate the progress by 2016 of developments and trends in place today. Following the charge of the panel discussion on which this essay is based, these trajectories are extended another decade to the year 2026, but must be considered more speculative.

How libraries develop and implement their technology strategies will be informed by opportunities or limitations defined by the larger scholarly publishing landscape, paradigm shifts in technology architectures, and the legal and business environment related to intellectual property and copyright

A variety of factors will shape the state of libraries of the future, both at the 5 and 15 year horizons. How libraries develop and implement their technology strategies will be informed by opportunities or limitations defined by the larger scholarly publishing landscape, paradigm shifts in technology architectures, and the legal and business environment related to intellectual property and copyright. Progress for libraries will also relate to their economic situation and their access to related resources, including network and computing infrastructure, Internet bandwidth, and the level of interest and ability for their users to take advantage of computer and web-based services and many other factors. Painting trends with a broad brush may not account for the finer details that reflect the realities of libraries in different international regions or specialized environments.

The *Technological Resources Panel* was asked to analyze the prospective of technological infrastructure and its implementation in library services and activities, imagining technical infrastructure in academic libraries in 2026, including:

- Projections made on trends underway today.
- Disruptions can produce radically different long-term outcomes.

- Present an optimistic view that libraries will maintain key role in academic institutions.
- Acquire, manage, and deliver access to information in support of teaching and research.

2. Assumptions regarding library collections

2.1. 2011: Transition toward electronic and digital content

Today, organizations involved in the creation of software for libraries have entered a new phase of competition. Following a period of research and development that aimed to provide alternatives to libraries, both in back-end automation and end user discovery, a slate of new solutions have emerged, often representing quite different conceptual models. In a continued trend, librarians seek solutions that immediately improve the experiences of their users, especially via discovery products.

These new automation products arise in the context of the realities that shape the work of academic libraries, especially related to the ever increasing proportions of electronic and digital materials that comprise library collections. The transition from print to electronic formats has largely run to completion with scholarly journals. The vast majority of journal articles are now available in electronic form, with many libraries shifting away entirely from printed subscriptions, even when available. An increasing number of scholarly journals follow an open access model where articles are freely available, mostly due to up-front fees paid by the authors. Monographs also see great changes in format options. E-books have been available for many years, though usually to supplement print copies, and embedded within aggregated databases of specialized materials. Today we are approaching a tipping point where interest in e-books has risen dramatically in public libraries in tandem with the proliferation of a large array of e-reading devices. E-book lending now ranks as one of the top concerns for public libraries; academic libraries will not be far behind and also need significantly better tools for managing this growing component of their collections. Today most academic libraries routinely create and provide access to collection of digital objects, including documents, photographs and other images, video content, and sound recordings.

2.2. 2016: All new content produced electronically

By 2016 we should expect significant movement toward the digital realm on many fronts. Electronic scholarly articles, possibly disaggregated from specific e-journals will reach near 100 percent penetration, with ever higher amounts of open access materials. Though the proportions of open access content will have continued to increase, traditional subscriptions will remain dominant by this date. E-books may achieve majority status in academic libraries, and any new monographic acquired in print form will be exceptionally rare. Academic and research libraries with large legacy collections of print monographs will continue to maintain those collections, though the majority of these collections will be relegated to remote storage facilities, available upon request. I also anticipate that most library procurements will be based on patron-driven acquisitions. The combina-

tion of limited resources and the options to acquire digital editions with near immediate delivery alleviates the need for librarians purchase large amounts of resources that may never see use. By this time, libraries will also supplement their legacy print collections with global resources of digital books from the *HathiTrust*, *Google Books*, the *Open Library*, and other initiatives specific to geographic regions and languages. Access to digital resources will increasingly be globalized through initiatives such as *Europeana* and the *Digital Public Library of America*. By this time all new content will be routinely produced electronically.

2.3. 2026: No technical obstacles but likely legal constraints

Looking further forward to 2026 we can anticipate these trends to be even more fully realized. Academic libraries will operate in a thoroughly digital arena. All new content will be acquired, increasingly collaboratively, in some kind of digital form. The technical capability will long since have been achieved to create digital representations of all books, articles, manuscripts, archives, maps, and other materials of interest to academic libraries. With the intervening decade since new content shifted to digital creation, current research materials can naturally be expected to be readily available for incorporation into a library's digital collections. Libraries will expect discovery services that truly represent the entire body of scholarly content available to their users across a wide range of media and formats, including materials available through resource sharing arrangements.

While the technical obstacles for digital dominance by 2026 will be entirely obliterated, events on the business and legal front will stand as the most likely constraining factors. It's not guaranteed that the processes to digitize the remaining analog books not in the public domain will continue an ever more severe intellectual property environment. The outcome of legal battles taking place today among libraries, publishers, author's groups and other content stakeholders stand to make an enormous impact on the long-term digital future of libraries. In recent days, for example, the United States has seen proposed legislation that would undo major accomplishments in open access, such as the mandatory posting of articles derived from research funded by the United States federal government. Constant vigilance and advocacy will be essential to ensure that libraries maintain a strong position in the digitally dominant future. But barring legal obstacles, technologies will be available in 2026 to deliver powerful access to the entire body of scholarly content.

3. Impact on library management systems

Advancing from the present to the future, libraries will also undergo major transformations regarding their needs for management systems. We can anticipate a transition from the integrated library systems (ILS) in place today to a new genre of library services platforms, embracing major changes in the management, access, and fulfillment of library resources consistent with the transformations described above and with other changes underway in the realm of scholarly publishing and scientific research.

It's safe to say that in 2011 integrated library systems dominate the scene. This model of library automation, based on modules of cataloging, circulation, acquisitions, serials management, and online catalog was established as early in the 1970's when libraries dealt primarily with print materials and has evolved too little since. The ILS tends to operate in a self-contained way, communicating with external systems primarily through library-specific protocols such as Z 39.50, SIP or NCIP, ISO ILL and based on library specific data formatting conventions such as MARC21 or Unimarc.

The traditional ILS model does not lend itself to the management of electronic resources and has led to the proliferation of supplemental products to address this increasingly strategic aspect of library operations. Today, academic libraries, in addition to their ILS, routinely implement OpenURL-based link resolution services, specialized electronic resource management systems, discovery interfaces, digital asset management platforms, institutional repositories, proxy servers, and other components addressing one aspect or another of a library's digital operations.

‘ In addition to their ILS, academic libraries routinely implement many other software components, resulting in a very complex environment for the library to manage ’

Each of these applications tends to require separate implementation and management processes and hardware platform, resulting in a very complex environment for the library to manage, exceeding the resource capacity of many libraries. These systems often fail to interoperate with each other effectively, due to isolated data models and a lack of robust APIs.

3.1. Library services platforms

2011 marks the beginning of a new cycle of library automation that will take hold by 2016, with full maturity by 2026. Following two or three years of research and development, a new slate of products launched in 2011-12, that I term *library services platforms*, aim to take a much more comprehensive approach to managing content resources.

While each brings its own distinctive features and design characteristics to the table, these library services platforms generally aim to manage all types of resources, offering specialized workflows depending on whether the item is analog or digital, local or remote, purchased or licensed. They mostly incorporate highly shared data models, following a metadata lifecycle approach that favors processing of batches or streams of metadata from external feeds, enriching records as possible through automated and manual processes. These platforms also take a more agnostic approach relative to metadata formats, supporting major library and non-library standards, including MARC21, Unimarc, Dublin Core, METS, MODS, MARC/XML, ONIX, and others. The availability of a robust suite of application programming interfaces, or APIs, will stand as one of the key defining characteristics

of this new genre of library services platforms. The “platform” aspect of this genre provides the ability for library programmers to access the data and functionality of these systems to create new services beyond those delivered, to construct new interfaces and widgets, and to dynamically connect with external systems within the library’s own enterprise network, that of its parent institution, or with external partners or providers.

“The new *library services platforms* will provide the ability for library programmers to access the data and functionality to create new services beyond those delivered”

This brief summary gives only a glimpse of this new model of library automation that takes a radically divergent approach to the ILS model. Figure 1 presents something of a visual representation of a general schematic of the orientation of these new products.

2011 and 2012 saw, or will see the launch of several of these new platforms, including: *OCLC’s WorldShare Management Services*, *Alma* from *Ex Libris*, *Serials Solutions Intota*, *Sierra* from *Innovative Interfaces*, and the open source *Kuali OLE* project. But these are the systems of the future more than the present with vast majority of academic libraries today operating traditional ILS products such as *Millennium*, *Voyager*, *Aleph 500*, *SirsiDynix Symphony*, *Horizon*, *VTLS Virtua*, *Vubis Smart*, and others that vary by global region.

3.2. Full maturity of LSP by 2026

By 2016, we can expect to see implementation of the new library service platforms in significant proportions. But a timeframe of only five years from now is not necessarily sufficient to complete generational changes. Given that even today, previous ILS more than a decade past the point when they have been actively developed, such as *Geac Advance* or *Dynix Classic*, remain in use in at least some libraries, reflecting the reality that some libraries take a very long time to acquire and implement new products. Libraries with limited resources often have little choice but to make the best of outdated automation systems. Historically, the full completion of a cycle from existing legacy products to new flagship offerings spans over a decade. Turning the calendar forward to 2026, we can speculate the realm of library services platforms to reach full maturity and near universal deployment. In the same way that almost any academic library today has implemented an integrated library system of some sort, by 2026 we can expect routine use of platforms capable of comprehensive resource management across print, electronic, and digital formats. Many of the automation products used in

academic libraries back in 2011 will be extinct, especially those offered by organizations offering new platforms.

I also expect that at least some of the products in use today to survive though 2026, though the intervening years of evolution will render them almost unrecognizable relative to their current form, and they will have taken on many of the traits of the library services platforms launched in 2012. The library automation environment has always been favorable to an evolutionary approach to product development, even though the cycle underway today introduces a number of revolutionary alternatives. Several vendors have been able to successfully navigate through a long series of technology cycles, including *Innovative Interfaces, Inc.*, *VTLS*, *SirsiDynix* (and its predecessor companies), and *Ex Libris*, and we can anticipate that these, and others, will continue to evolve existing products or create new ones through the next era of library automation.

“Library automation technologies have historically seen cycles of about 10 years”

The emerging technologies of today will inevitably turn into the legacy systems of the future. Given that library automation technologies have historically seen cycles of about 10 years, we should anticipate that beginning in about 2020 sufficient change will have transpired in academic libraries, technologies, and scholarly publishing to spark the next wave of innovation and product development. 2026 may well be a time when the systems we see today as the wave of the future may be due for redevelopment or replacement. The technology and business cycles of library automation churn relentlessly.

4. Cloud computing

In 2011 cloud computing has entered the mainstream. Though the majority of libraries today operate local servers, an increasing number are shifting to vendor hosting ar-

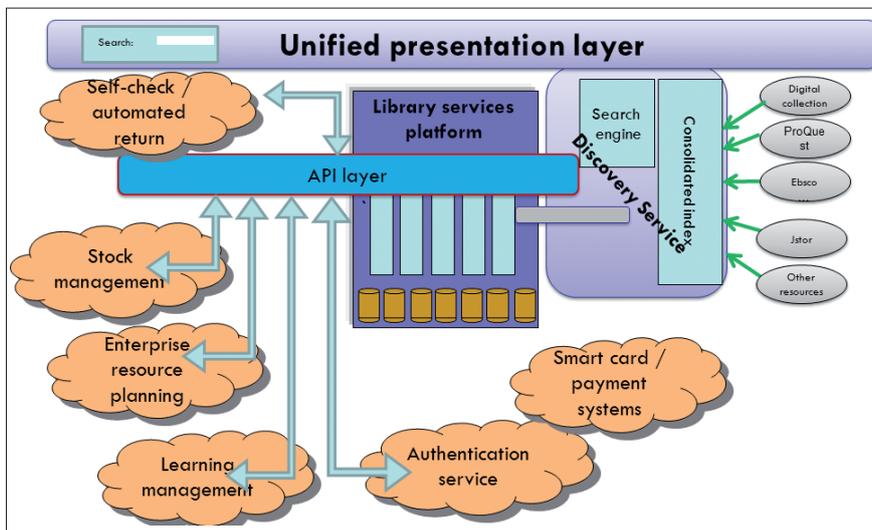


Figure 1. General schema of the library services platforms

rangements or software as a service (saas). Today's legacy systems are moving rapidly to vendor-hosted arrangements. New products are generally being design for multi-tenant software as a service deployment. Today we can see that the wheels are in motion for a steady advance toward cloud computing. Internationally, libraries in some regions do not find cloud computing as a viable option due to the lack of adequate internet bandwidth.

We can expect cloud computing to play a much more dominant role by 2016. The new slate of library services platforms will natively embrace cloud computing. But even those libraries continuing with traditional integrated library systems will shift from local installations to software as a service. By 2016, locally managed servers for library automation systems will be a minority approach. Internet bandwidth should be sufficient in almost all international regions to support cloud computing.

By 2026, should the trajectory of the current trends remain reasonably on course, we might anticipate that almost all computer services will be deployed through cloud computing models, including the core automation and discovery services implemented by academic libraries. By this date the needed infrastructure of robust internet bandwidth should be present ubiquitously, including both high-speed connections to data centers and wireless communications to individual users. Mobile devices will far outnumber desktop and laptop computers.

“The automation systems themselves likewise will be offered as open source software routinely by 2026”

5. Open source software

We can expect that open source software see steady growth over the next fifteen years. Today we see a vigorous competition between library automation systems offered through proprietary licenses and those available as open source software. Over time, library automation products will include increasing proportions of open source components. The automation systems themselves likewise will be offered as open source software routinely by 2026.

If the trend toward cloud computing that we noted above comes to pass, the license models of the automation products becomes at least somewhat less relevant. In the multi-tenant software as a service model, the concept of individual libraries modifying the source code does not really apply. The dynamics of a cloud-based automation environment blur many of the distinctions seen with locally deployed software. In general, both open source and proprietary systems will involve a business model based on subscriptions to services.

Given the general orientation of libraries to prefer open systems, over the course of the next 15 years I anticipate a shift accordingly towards open source licenses, even by the organizations that today offer proprietary software. In the future, points of differentiation for all products –includ-

ing both open source and proprietary– will focus on characteristics such as quality of support, the functionality of the system, the relative power of the APIs offered, as well as competitive pricing of the subscription and service fees.

“If the trend toward cloud computing comes to pass, the license models of the automation products becomes less relevant”

6. Discovery services

Beginning in about 2005, one of the key trends in library automation involved the emergence of discovery interfaces that offered an improved way of accessing library collections than the online catalog modules that were delivered as part of their ILS. These discovery interfaces offered a more modern search experience, including features such as relevancy-ranked results, faceted browsing, and more intuitive and attractive page designs. Beginning about 2010 a set of discovery services, oriented primarily oriented to academic libraries, came out with an expansive vision of search that included access to the articles and other information available through a library's subscriptions to electronic resources. This slate of discovery services included *Summon* from *Serials Solutions*, *Primo Central* from *Ex Libris*, and *Ebsco Discovery Service*. In 2012 an increasing number of academic libraries will be shifting away from the online catalog module that was delivered with their ILS to a more broad-based discovery service.

Discovery products will dominate by 2016. Traditional library catalogs will serve mostly as specialized tools for researchers needing to work closely with the library's physical materials and manuscript collections. We can anticipate that the discovery services available by this time will offer much more sophisticated access to library services and collections than those in place today. The generalized keyword-oriented search technologies will have been extended to include more effective means of ranking and selecting results. We might anticipate better use of social and associative metadata to improve search results. A decade after the start of the technology cycle of discovery products, we can expect that this genre to reach a state of full maturity and near universal deployment among academic libraries.

“By 2016 we can expect discovery products to reach a state of full maturity and near universal deployment among academic libraries”

Discovery services will have advanced significantly by 2026. It seems reasonable to believe that most of the shortcomings of content discovery will be resolved. Collections will be fully represented in the indexes of the discovery services and at a highly granular level. Indexes will routinely incorporate the full text of books, articles, newspapers and other

textual material; audio will be represented through both transcriptions and speech-based searching. High-quality metadata will enhance search capabilities. We can reasonably expect that library discovery services will provide access to the body of library materials, spanning all disciplines and media types, with at least the breadth and sophistication that search engines such as *Google* offer for searching the Web.

7. Metadata and progress toward the semantic web

In 2011, the semantic web and the realm of open linked data made no significant impact on library automation products or discovery services. New initiatives such as RDA (resource description and access) were launched to bring at least some of the characteristics of linked data to the ways that libraries describe their materials. Other efforts with even more potential to propel libraries toward more interoperable metadata practices include the announcement that the *Library of Congress* will begin investigating alternatives to MARC as the carrier for its bibliographic records.

By 2016 we might expect that RDA will overtake AACR2 as new bibliographic metadata is created, but that it will make only incremental improvements in the ways that automation systems or discovery services make more meaningful use of metadata. Yet, interest in bringing semantic web principles may spark other initiatives in how metadata can be structured.

One might optimistically hope that by 2026 the semantic web will constitute a major part of the metadata fabric that underlies library services platforms and discovery services. Any success in operationalizing open linked data should result in advances in how libraries and library users gain access to metadata and content in ways that reduce the dependence on commercial environments that currently restrict their availability.

By 2026 there will be a fewer number of organizations producing library automation products, each with very broad global reach and extensive product development and support capacity

8. General forecast

The cycles of technology take a toll on industry players and products. Some companies may lag behind in innovation and research and development needed to move forward successfully. Each phase of the ongoing cycles of technology tends to narrow the field of competitors. Companies not able to successfully develop and market products consistent with current technology architectures and library expectations become vulnerable to the inevitable industry consolidation produced through business cycles of mergers and acquisitions. 2026 might bring a business environment of a much fewer number of organizations producing library automation products, each with very broad global reach

and extensive product development and support capacity. Globalization will have a major impact on the library automation industry, though I expect that some national and regional companies to continue to prosper and to offer significant competition.

While the library automation products available will achieve maturity and technical sophistication relative to most library demands and requirements apparent today, we can also anticipate that new challenges will arise that will in turn require ever new dimensions of technology support.

While it's tempting to imagine a world by 2026 that solves all of the problems that impede libraries today, it would be naïve to hold an overly optimistic view. Throughout the history of library automation, the demands of libraries have consistently outpaced the development of technology products that deliver appropriate support. Even though technology advances rapidly, the forces that challenge and complicate the future of libraries seem to move even faster. Surely other issues will emerge that will disrupt some of the current trends and will demand technology and automation systems to turn in other directions other than the ones addressed in this essay.

This exercise in which we speculate on the state of libraries and library technologies by 2026 must be taken with a healthy dose of skepticism. While it's interesting to think about the trajectories currently in motion, we must recognize that no crystal ball can give a true look that far into the future and that any variety of unanticipated factors and events could propel libraries of the future into quite different directions.

9. Bibliography

Arriola-Navarrete, Óscar; Tecuatli-Quechol, Graciela; González-Herrera, Guadalupe. "Software propietario vs software libre: una evaluación de sistemas integrales para la automatización de bibliotecas". *Investigación bibliotecológica*, 2011, v. 25, n. 54. pp. 37-70.

<http://www.ojs.unam.mx/index.php/ibi/article/view/27480>

Breeding, Marshall; Yelton, Andromeda. Librarians' assessments of automation systems: Survey results, 2007–2010, *Library technology reports*, American Library Association, 2011, May-June, v. 47, n. 4, pp. 5-8. ISSN 0024-2586 (Print). <http://www.alatechsource.org/taxonomy/term/106/librarians-assessments-of-automation-systems-survey-results-2007-2010>

Breeding, Marshall. "A cloudy forecast for libraries". *Computers in libraries*, 2011, v. 31, n. 7, pp. 32-34.

Breeding, Marshall. "Transformations in academic libraries demand transformed automation support". *Computers in libraries*, 2011, v. 31, n. 4, pp. 27-29.

Cho, Jane. "Study on a saas-based library management system for the Korean library network". *The electronic library*, 2011, v. 29, n. 3, pp. 379-393. <http://dx.doi.org/10.1108/02640471111141115>

Fronk, Elizabeth. "A review of 'next-gen library catalogs'". *Journal of web librarianship*, 2011, v. 5, n. 1, pp. 74-75. <http://dx.doi.org/10.1080/19322909.2011.544564>

Gerolimos, Michalis; Konsta, Rania. Services for academic libraries in the new era. *D-lib magazine*, 2011, v. 17, n. 7-8. <http://www.dlib.org/dlib/july11/gerolimos/07gerolimos.html>
<http://dx.doi.org/10.1045/july2011-gerolimos>

Gross, Julia; Sheridan, Lutie. "Web scale discovery: the user experience". *New library world*, 2011, v. 112, n. 5/6, pp. 236-247.
<http://dx.doi.org/10.1108/03074801111136275>

Hane, Paula; Hawkins, Don. "CIL: Focus and value for library communities". *Information today*, 2011, May, v. 28, n. 5, pp. 20-21.

Jackson, Darla W. "Thinking about technology...: Watson, answer me this: will you make librarians obsolete or can I use free and open source software and cloud computing to ensure a bright future?". *Law library journal*, 2011, n. 103, p. 497.

Liu, Jia. "Digital library and digital reference service. Integration and mutual complementarity". Educational futures rethinking theory and practice. In: R. Rikowski (ed.) *Digitisation perspectives*, 2011, v. 46, part VI, pp. 191-212.
http://dx.doi.org/10.1007/978-94-6091-299-3_12

Nesta, Frederick; Mi, Jia. "Library 2.0 or library III: returning to leadership". *Library management*, 2011, v. 32, n. 1/2, pp. 85-97.
<http://dx.doi.org/10.1108/01435121111102601>

Yang, Sharon Q.; Hofmann, Melissa A. "Next generation or current generation?: A study of the opacs of 260 academic libraries in the USA and Canada". *Library Hi Tech*, 2011, v. 29, n. 2, pp. 266-300.

Wynnea, Susan C.; Hanscoma, Martha J. "The effect of next-

generation catalogs on catalogers and cataloging functions in academic libraries". *Cataloging & classification quarterly*, 2011, v. 49, n. 3, pp. 179-207.
<http://dx.doi.org/10.1080/01639374.2011.559899>

Žumer, Maja. "Next-gen library catalogs". *Program: electronic library and information systems*, 2011, v. 45, n. 1, pp. 122-123.

10. Other recent bibliography by Mr. Marshall Breeding

(recommended by the editor for further reading)

Breeding, Marshall. "A smart approach to RFID technologies". *Smart libraries newsletter*. American Library Association, 2010, November, v. XXX, n. 11.
<http://alatechsource.metapress.com/content/I72006156702>

Breeding, Marshall. "The state of the art in library discovery 2010". *Computers in libraries*, 2010, Jan.-Feb., v. 30, n. 1, pp. 31-34.

Breeding, Marshall. "Leveraging technology for success in a challenging economy". *Computers in libraries*, 2010, April, v. 30, n. 3, pp. 24-26.

Breeding, Marshall. "Using technology to enhance a library as place". *Computers in libraries*, 2011, v. 31, n. 3, pp. 29-31.

Breeding, Marshall. "Discovering Harry Pottery barn". *Computers in libraries*, 2011, v. 31, n. 2, pp. 21-23.

Breeding, Marshall. "Preparing for the long-term digital future of libraries". *Computers in libraries*, 2011, v. 31, n. 1, pp. 24-26.

Próximos temas centrales

Marzo 2012	Publicación científica y acceso abierto
Mayo 2012	Organización del conocimiento
Julio 2012	Comunicación digital
Septiembre 2012	Información y derecho
Noviembre 2012	Bibliotecas académicas
Enero 2013	Soportes digitales
Marzo 2013	Educación y biblioteca
Mayo 2013	Bibliotecas y documentación de museos

Los interesados pueden remitir notas, artículos, propuestas, publicidad, comentarios, etc., sobre estos temas a: <http://recyt.fecyt.es/index.php/EPI/index>



CREACIÓN AUTOMÁTICA DE SISTEMAS DE BÚSQUEDA DE RESPUESTAS EN DOMINIOS RESTRINGIDOS



Katia Vila, José-Norberto Mazón y Antonio Ferrández



Katia Vila es doctora en ingeniería informática por la *Universidad de Alicante*. Ha sido profesora asistente en el *Departamento de Informática* de la *Universidad de Matanzas* en Cuba, impartiendo docencia en ingeniería informática en asignaturas relacionadas con la inteligencia artificial, la programación descriptiva, y la recuperación de información. Es miembro del *Grupo de investigación en Procesamiento del Lenguaje y Sistemas de Información (GPLSI)* del *Departamento de Lenguaje y Sistemas Informáticos* de la *Universidad de Alicante*. Ha publicado diferentes trabajos sobre sistemas de búsqueda de respuestas (BR). Disfrutó de una beca *MAEC-AECID* (España) para un proyecto post-doctoral sobre sistemas de BR.

Universidad de Alicante
Departamento de Lenguajes y Sistemas Informáticos
Apartado de Correos 99. 03080 Alicante
kvila@dlsi.ua.es



José-Norberto Mazón es profesor ayudante doctor en el *Departamento de Lenguajes y Sistemas Informáticos* de la *Universidad de Alicante*. Ha publicado varios trabajos en revistas y conferencias especializadas de carácter internacional (como *ER, ICCS, JISBD, DSS, DKE*, etc.). Ha realizado estancias de investigación en la *Universidad de Münster* (Alemania) y la *Universidad de Toronto* (Canadá). Su investigación se centra principalmente en inteligencia de negocio y desarrollo de software dirigido por modelos. Ha participado en la organización de varias ediciones de los talleres "*Business intelligence and the Web*" (*Beweb*) y "*Web and requirement engineering*" (*WeRE*).

Universidad de Alicante
Departamento de Lenguajes y Sistemas Informáticos
Apartado de Correos, 99. 03080 Alicante
jnmazon@dlsi.ua.es



Antonio Ferrández es profesor en el *Departamento de Lenguajes y Sistemas Informáticos* de la *Universidad de Alicante*. Obtuvo su doctorado en informática en la misma universidad. Sus intereses de investigación son: procesamiento del lenguaje natural, resolución de problemas lingüísticos (anáfora o elipsis), extracción de información, sistemas de recuperación de información y búsqueda de respuestas. Ha participado en proyectos, acuerdos con empresas privadas y organismos públicos relacionados con sus temas de investigación. Ha dirigido tesis doctorales y es autor de artículos de revista y ponencias de congreso.

Universidad de Alicante
Departamento de Lenguajes y Sistemas Informáticos
Apartado de Correos, 99. 03080 Alicante
<http://www.dlsi.ua.es/~antonio/antonio.html>
antonio@dlsi.ua.es

Resumen

Los sistemas de búsqueda de respuestas (BR) se pueden considerar como potenciales sucesores de los buscadores tradicionales de información en la Web. Para que sean precisos deben adaptarse a dominios concretos mediante el uso de recursos semánticos adecuados. La adaptación no es una tarea trivial, ya que deben integrarse e incorporarse a sistemas de BR existentes varios recursos heterogéneos relacionados con un dominio restringido. Se presenta la herramienta *Maraqá*, cuya novedad radica en el uso de técnicas de ingeniería del software, como el desarrollo dirigido por modelos, para automatizar dicho proceso de adaptación a dominios restringidos. Se ha evaluado *Maraqá* mediante una serie de experimentos (sobre el dominio agrícola) que demuestran su viabilidad, mejorando en un 29,5% la precisión del sistema adaptado.

Palabras clave

Búsqueda de respuestas, Recuperación de información, Dominios restringidos, Desarrollo dirigido por modelos.

Title: Automatic generation of question answering systems in restricted domains

Artículo recibido el 10-11-11
Aceptación definitiva: 11-01-12

Abstract

Question answering (QA) applications can be considered as the potential successors to the traditional information retrieval on the Web. However, QA systems should be adapted to restricted domains for the sake of precision. This adaptation is not a trivial task, since several heterogeneous resources related to a restricted domain must be integrated into existing QA systems. This paper presents the *Maraqqa* tool, whose novelty is the use of software engineering techniques such as model driven development to automate the adaptation process. It is worth noting that *Maraqqa* has been evaluated through a set of experiments (within the agricultural domain) that demonstrate its applicability: the precision of the adapted QA system showed 29.5% improvement.

Keywords

Question answering, Information retrieval, Restricted domains, Model driven development.

Vila, Katia; Mazón, José-Norberto; Ferrández, Antonio. "Creación automática de sistemas de búsqueda de respuestas en dominios restringidos". *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 16-26.

<http://dx.doi.org/10.3145/epi.2012.ene.03>

Introducción y motivación

Con la sobreabundancia de información y la heterogeneidad en los formatos de acceso es complicado localizar informaciones concretas. Los usuarios necesitan sistemas que les permitan acceder a información precisa desde los diferentes recursos existentes de una forma transparente y simple.

La primera solución aportada por la comunidad científica para conseguir un acceso sencillo y rápido a la inconmensurable cantidad de información digital accesible, fue la recuperación de información (RI) o *information retrieval* (Baeza-Yates; Ribeiro-Neto, 1999). La RI consiste en la selección en un depósito de documentos de los que tengan mayor relevancia para una consulta realizada por un usuario.

La BR tiene como objetivo obtener respuestas a preguntas precisas formuladas en lenguaje natural

Los sistemas de RI más conocidos son los que actúan sobre internet y localizan información en la Web, por ejemplo los motores de búsqueda como *Google* o *Yahoo*. Según Russell y Norvig (2003), este modelo de RI corresponde casi totalmente a palabras, ya que acepta una cantidad mínima de sintaxis (palabras que deben aparecer una junto a la otra), y un papel diminuto de clases semánticas (en forma de listas de sinónimos). Por ello los resultados son en muchas ocasiones documentos que contienen los términos de la consulta, pero no son la información deseada. Por otro lado, la salida es una lista de documentos ordenada en función de medidas de similitud con la pregunta y medidas indirectas de la visibilidad y prestigio de la información (por ejemplo *page rank* de la web), lo cual puede convertir a esos motores en sistemas manipulables. Luego resta una ardua tarea ya que el usuario debe revisar y leer cada documento de la lista obtenida, lo primero para ver si en realidad está relacionado con los requerimientos solicitados y lo segundo para localizar en su interior la información puntual que se desea.

<http://www.google.com>

<http://www.yahoo.com>

Los inconvenientes mencionados impulsaron la investigación en sistemas de búsqueda de respuesta (BR) o *question answering*, que tienen como objetivo la obtención de respuestas concretas a preguntas precisas indicadas por el usuario directamente en lenguaje natural.

En las figuras 1 y 2 se muestran las salidas para la pregunta *What year did Wilt Chamberlain score 100 points?* de un sistema de RI como *Google* y de un sistema de BR como *NSIR* respectivamente. En el caso de la RI se aprecia, en contraste con la BR, que el usuario ha de buscar dentro de los documentos devueltos.

<http://clair.si.umich.edu/clair/NSIR/html/nsir.cgi>

Los sistemas de BR deben adquirir un nivel de comprensión del texto muy superior al alcanzado por la RI, por lo que es habitual realizar un análisis léxico, sintáctico y semántico tanto de la pregunta como de los documentos. Este tipo de análisis conlleva un coste computacional muy elevado, el cual se supera mediante el filtrado de los documentos a través de la RI, es decir, entre los millones de documentos de partida, la BR realiza el análisis y búsqueda de la respuesta únicamente sobre unos cientos de pasajes (conjunto de oraciones de un documento).

Por último, se analizan las diferencias entre BR y RI desde el punto de vista de la propia pregunta que realiza el usuario. En primer lugar, habitualmente la pregunta en RI no es más que una serie de palabras ("invención teléfono"), mientras que en BR ésta ha de ser una pregunta formulada correctamente en lenguaje natural ("¿Cuándo fue inventado el teléfono?"). En segundo lugar, la RI se centra en la propia pregunta, mientras que la BR se centra en la respuesta a dicha pregunta. Por ejemplo, si planteamos la pregunta "¿Cuándo fue inventado el teléfono?" a un sistema de RI, buscará el documento que más veces contenga "teléfono", mientras que la BR buscará el que contenga la fecha concreta de la invención del teléfono.

La investigación en BR se ha visto incentivada a partir de 1999 por conferencias como *TREC* (*Text retrieval conference*) desde su octava edición, *CLEF* (*Cross-language evaluation forum*) y *Ntcir* (*NII Test collection for IR systems*). En estas conferencias compiten sistemas de BR desarrollados tanto por instituciones académicas como por empresas. Su

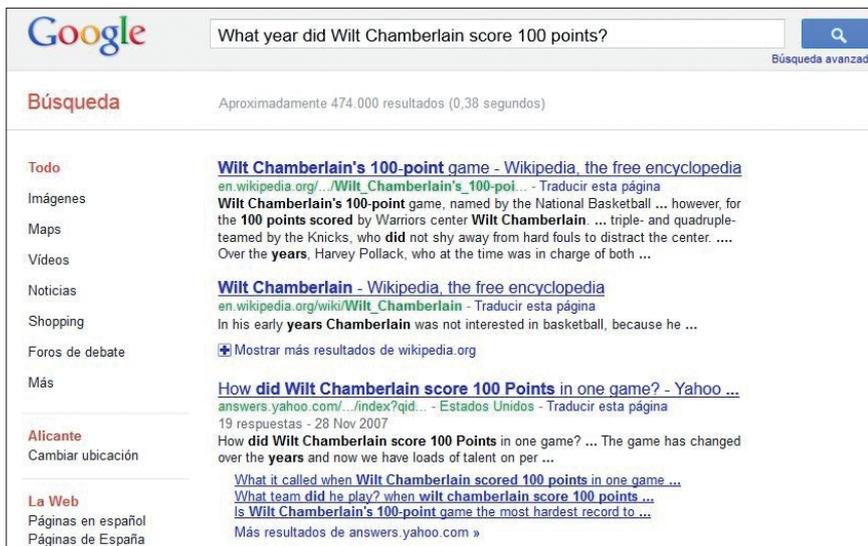


Figura 1. Salida de un sistema de RI para una pregunta en lenguaje natural

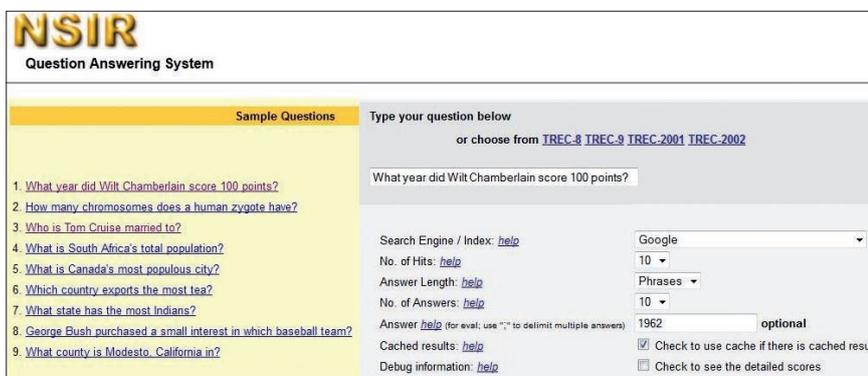


Figura 2. Salida de un sistema de BR para una pregunta en lenguaje natural

funcionamiento general (dependiendo del tipo de sub tarea sobre la que se compite) consiste en que los participantes disponen de unas colecciones de documentos y una serie de preguntas, cuya respuesta es analizada por un comité que ordena cada sistema por una serie de medidas. Estas colecciones de documentos suelen ser de un tamaño considerable y de contenidos diversos conocidos como de “dominio abierto” (DA).

<http://trec.nist.gov>

<http://www.clef-initiative.eu>

<http://research.nii.ac.jp/ntcir/index-en.html>

Algunos ejemplos de sistemas de BR con una dilatada trayectoria son *Start* (*SynTactic analysis using reversible transformations*) y *Wolfram Alpha*.

Start, operativo desde diciembre de 1993, fue realizado por **Boris Katz** y los componentes del *InfoLab Group* del *MIT Computer Science and Artificial Intelligence Laboratory* (**Katz**, 1997). Está basado en el etiquetado previo de segmentos de texto según diferentes granularidades (**Katz et al.**, 2006), lo cual facilita el acceso a los tipos de respuesta predefinidos, pero presenta la desventaja de limitar los tipos de pregunta a realizar al sistema.

<http://start.csail.mit.edu>

Siguiendo la misma estrategia de *Start*, *Wolfram Alpha* (anunciado en marzo de 2009) –de la compañía *Wolfram Research*, presidida por **Stephen Wolfram**–, extrae respues-

tas de una base de datos estructurada. Este sistema indexa una gran cantidad de respuestas, sobre las cuales realiza un proceso de inferencia cuando se plantea una pregunta. Es decir, ha construido una gran base de datos de conocimiento que utiliza para este proceso de inferencia. De un modo similar funciona *Cyc*, el cual aúna una ontología y una base de datos de conocimiento general para llevar a cabo razonamientos de tipo *humano*. Este tipo de planteamientos adoptados tradicionalmente tienen limitaciones de extensibilidad debido a la estructuración predefinida de los tipos de respuestas.

<http://www.wolframalpha.com>

<http://www.cyc.com>

A pesar de los amplios esfuerzos investigadores realizados, las carencias de los sistemas de RI actuales y el acuerdo general de que los sistemas de BR son sus potenciales sucesores, no se ha generalizado su uso. El motivo principal es que la BR no es lo suficientemente precisa sobre un dominio abierto para el coste computacional que supone, siendo necesarios muchos recursos semánticos y una capacidad de razonamiento no alcanzada de momento.

Estos objetivos sí se podrían alcanzar sobre “dominios restringidos” (DR) tales como el dominio médico o legal, en los que se disponga de recursos semánticos suficientes.

“ A pesar de las investigaciones, las carencias de los sistemas de RI y la creencia de que los sistemas de BR son sus sucesores, no se ha generalizado su uso ”

Nuestro trabajo se centra en la adaptación automática de sistemas de BR en DA a dominios restringidos. Se presenta la herramienta *Maraqqa* (*Model-driven adaptation for restricted-domain question answering*), que automatiza el proceso de adaptación y por tanto, puede usarse como un recurso que otorgará a todo sistema de BR en DA la capacidad para enfrentar el cambio, para introducir y extraer información de diferentes dominios restringidos independientemente del idioma y formato de sus recursos de conocimiento. Para ello, nos basaremos en el desarrollo dirigido por modelos o *model-driven development* (**Bézivin**, 2005).

Arquitectura de un sistema de BR

Después de estudiar varias de las aproximaciones de BR presentadas en conferencias como *TREC* y *CLEF*, se ha observado que la mayoría de ellas tiene una arquitectura común que se muestra en la figura 3.

El módulo de indexación incluye los procesos aplicados a la colección de documentos de manera offline con el objetivo de acelerar el proceso de BR:

- indexación de documentos para el sistema de RI, que usa información estadística (se asignan pesos a las palabras);
- indexación de documentos para el proceso de BR, que usa información léxica, sintáctica y semántica.

Por otro lado el módulo de búsqueda es el encargado de realizar 3 procesos para responder a la pregunta del usuario:

1. Análisis de la pregunta, en el cual
 - se clasifica la pregunta [ej.: en función del tipo de respuesta esperada (TRE), que nos indica el tipo semántico de la información a buscar, como “tipo persona” para la pregunta “¿Quién es el presidente de España?”];
 - se extraen las palabras clave de la pregunta para usarlas como consulta en el siguiente proceso de RI (ej.: para “¿Cuál es el correo del *Ministerio de Economía y Competitividad*?”, es habitual descartar el término “correo” porque éste no suele aparecer en la respuesta y su inclusión en la pregunta de RI podría devolver pasajes que contengan dicho término en lugar de la información buscada. Esto es lo que ocurre en la figura 1, en la que se muestran pasajes con las apariciones del término *year* para la pregunta *What year did Wilt Chamberlain score 100 points?*).
2. Recuperación de documentos o pasajes relevantes, donde se utilizan las palabras clave previamente detectadas y el indexado de los documentos para la RI realizado con anterioridad, como entradas del sistema de RI. La salida serán los documentos o pasajes relevantes para la pregunta del usuario, reduciendo así el espacio de búsqueda para la BR.
3. Análisis del conjunto de documentos o pasajes relevantes usando el indexado para la BR previamente realizado, con la finalidad de encontrar la respuesta esperada por el usuario.

Problemas en la adaptación de un sistema de BR de DA a DR

Si analizamos la arquitectura de un sistema de BR es obvio que las fases de análisis de la pregunta y extracción de la respuesta son dependientes del conocimiento, usualmente incluido en patrones, que se tenga del dominio de aplicación. En el contexto de este trabajo, se entiende por “patrones” (Vila, 2010) todas las posibles estrategias para la detección de las relaciones entre los elementos de la pregunta y la respuesta (ej.: formas lógicas, expresiones regulares, relaciones sintácticas, relaciones de dependencia, etc.). Además se puede destacar la repercusión que tiene la fase de análisis de la pregunta en el resto de fases de la arquitectura típica de un sistema de BR y por tanto, en la eficiencia del sistema.

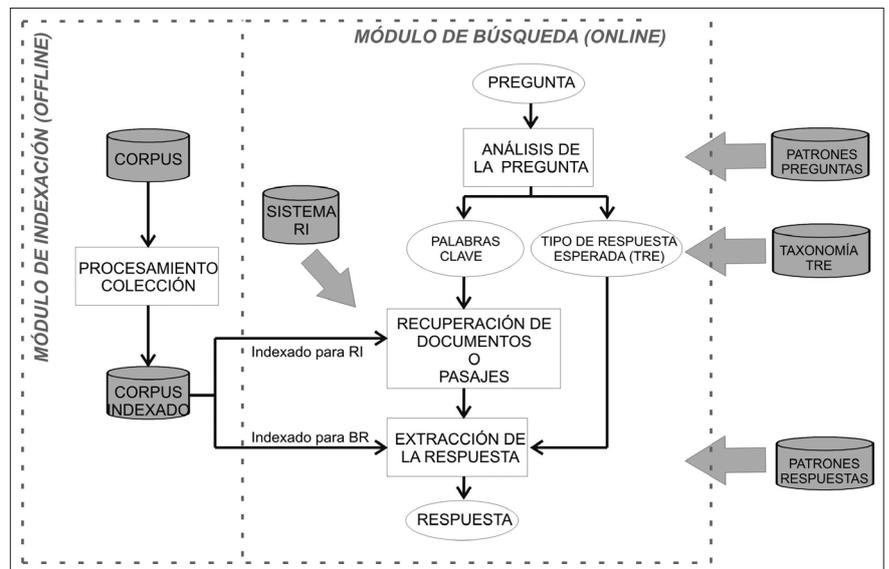


Figura 3. Arquitectura típica de un sistema de BR

En esta fase es crucial la determinación del tipo semántico de la respuesta que espera el usuario o TRE, por medio de una taxonomía de TRE [también conocida como jerarquía de preguntas (Li; Roth, 2006) u ontología de preguntas (Metzler; Croft, 2005)] predefinida. Una correcta especificación de la taxonomía de TRE implica una efectiva detección del TRE de la pregunta del usuario, reduciendo así el espacio de búsqueda de las respuestas candidatas y brindando una respuesta más precisa (Li; Roth, 2006; Hovy; Hermjakob; Ravichandran, 2002). Un ejemplo ilustrativo sería para la pregunta del usuario “¿Quién es el presidente de España?”, el TRE sería “persona” y las respuestas candidatas serían nombres propios como “Mariano Rajoy”.

Para el funcionamiento de un sistema de BR de DA en dominios restringidos es necesaria la adaptación de sus patrones de preguntas y respuestas y de la taxonomía de TRE que se utilice

Teniendo en cuenta estas consideraciones, se puede afirmar que para lograr el correcto funcionamiento de un sistema de BR de DA en dominios restringidos es necesaria la adaptación de:

- los patrones de preguntas y respuestas del sistema de BR;
- la taxonomía de TRE que se utilice.

Sin embargo, en las aproximaciones actuales para alcanzar dicha adaptación persisten dos problemas fundamentales:

- adaptación manual de los patrones de BR (Ferrés; Rodríguez, 2006; Peñas et al., 2009) y las taxonomías de TRE (Sekine; Sudo; Nobata, 2002; Hovy; Hermjakob; Ravichandran, 2002; Metzler; Croft, 2005; Li; Roth, 2006) a dominios restringidos, lo cual requiere de un esfuerzo enorme en función del tiempo y del coste propios de la complejidad inherente de los conceptos provenientes de esos dominios (Mollá; Vicedo, 2007);

– definición de los patrones de BR y las taxonomías de TRE para dominios restringidos por medio del análisis de corpus de preguntas-respuestas potenciales (Terol; Martínez-Barco; Palomar, 2006; Kosseim; Yousefi, 2008), lo que no es una situación realista ya que es muy difícil y compleja la adquisición de este tipo de corpus en dominios restringidos.

Después de analizar la situación actual de adaptación de sistemas de BR en DA a dominios restringidos, parece indiscutible la necesidad de diseñar estrategias que faciliten este proceso. Principalmente, las propuestas deben:

- elevar el grado de automatización de la adaptación, evitando que sea un proceso tedioso y complejo;
- explotar los recursos de conocimiento disponibles para un dominio concreto independientemente de su esquema de representación;
- utilizar el corpus textual como punto de partida y fuente de información principal, en detrimento del (poco realista en dominios restringidos) uso de corpus de preguntas;
- aumentar la precisión de la BR.

Para alcanzar estos objetivos, presentamos *Maraqqa*, una herramienta que permite una adaptación sistemática de los sistemas de BR en DA a nuevos dominios restringidos basada en el paradigma de ingeniería del software conocido como desarrollo dirigido por modelos (*model driven development, MDD*), que se basan en la representación de los conceptos y actividades que rigen un área concreta de conocimiento en lugar de usar terminología informática o codificación de algoritmos en lenguajes de programación complejos (Mellor et al., 2003). La creación de estos modelos debe basarse en una sintaxis gráfica, mediante reglas y normas establecidas en un metamodelo (Kleppe et al., 2003). Además, mediante el uso de *MDD* se posibilita la generación automática de la aplicación software a partir de modelos (Mellor; Clark; Futagami, 2003). Por lo tanto, *MDD* enfatiza dos aspectos clave: los modelos y las transformaciones entre ellos para llegar a obtener el código fuente correspondiente del sistema software.

Maraqqa

Ha sido desarrollado siguiendo el paradigma *MDD* mediante la plataforma *Eclipse*. Permite adaptar fácilmente sistemas de BR en DA a dominios restringidos, haciendo a estos sistemas útiles para su uso en dominios técnicos y no triviales que necesitan de una alta precisión como medicina, farmacia, derecho, etc. Por ejemplo, en el dominio agrícola una pregunta tipo podría ser “¿Qué enzima incrementa la digestibilidad del fósforo orgánico en los animales?”. La pregunta es sobre un dominio restringido por lo que su terminología es más específica y por lo tanto es incorrectamente respondida por sistemas de BR en DA. En el trabajo de Vila et al. (2011) se realizó una serie de experimentos para comparar la precisión de un sistema de BR de DA desarrollado en la Universidad de Alicante, llamado *AliQAn* (Roger et al., 2008) y su adaptación al dominio agrícola mediante *Maraqqa*. Los resultados indicaron que la precisión del sistema de BR *AliQAn* fue de 28,8% comparado con la media del sistema en un dominio abierto que es alrededor del 43% de precisión. Una vez *AliQAn* fue adaptado usando *Maraqqa*, la precisión

alcanzó un 58,3% (Vila; Mazón; Ferrández, 2011). En el apartado de evaluación de *Maraqqa* se describen con mayor profundidad los experimentos realizados.

<http://www.eclipse.org>

“*Maraqqa* permite adaptar sistemas de BR en dominios abiertos a dominios restringidos”

Para adaptar sistemas de BR a DR, *Maraqqa* requiere los siguientes recursos:

- a) Conocimiento no estructurado: en este caso se refiere a toda la documentación no estructurada disponible en el dominio restringido al cual se desea adaptar el sistema de BR. Esta documentación es conocida como colección de documentos o corpus.
- b) Conocimiento semi-estructurado o estructurado: cualquier sistema de organización del conocimiento (SOC) –llamados en inglés *knowledge organization systems, KOS* (Hodge, 2000)–, disponible en el dominio. Estos recursos de conocimiento incluyen una variedad de esquemas para organizar, manejar y recuperar información. El término SOC pretende abarcar cualquier tipo de esquema que sirva para gestionar el conocimiento, como diccionarios, taxonomías, tesauros, ontologías, etc. Según las áreas de conocimiento a las que se refiera un SOC, existen dos tipos: SOC genérico como *WordNet*, *EuroWordNet*, *SUMO*; o SOC de dominio, tales como el tesoro *Agrovoc* para el dominio agrícola, el tesoro multilingüe *EuroVoc* que abarca la terminología de los ámbitos de actividad de la Unión Europea, el metatesoro *UMLS* para el dominio médico, etc.
<http://wordnet.princeton.edu>
<http://www.illc.uva.nl/EuroWordNet>
<http://www.ontologyportal.org>
<http://www.fao.org/agrovoc>
<http://eurovoc.europa.eu>
http://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus
- c) Código fuente de los patrones de preguntas y respuestas del sistema de BR en DA que se desea adaptar.

Cabe resaltar que *Maraqqa* es independiente del idioma de los recursos que emplee, pero para su correcto funcionamiento todos estos recursos deben estar en el mismo idioma.

A partir de estos recursos, *Maraqqa* obtiene:

- taxonomía de TRE refinada para el dominio en cuestión;
- código fuente de los patrones de preguntas adaptados al dominio;
- código fuente de los patrones de respuesta adaptados al dominio.

Para explicar la herramienta utilizaremos como caso de estudio la pregunta “¿Qué glicósidos tienen un efecto defaunante en el rumen?”. Nuestro objetivo sería adaptar *AliQAn* a este dominio agrícola. Para ello se necesita, en primer lugar, un corpus o conocimiento no estructurado, sobre el

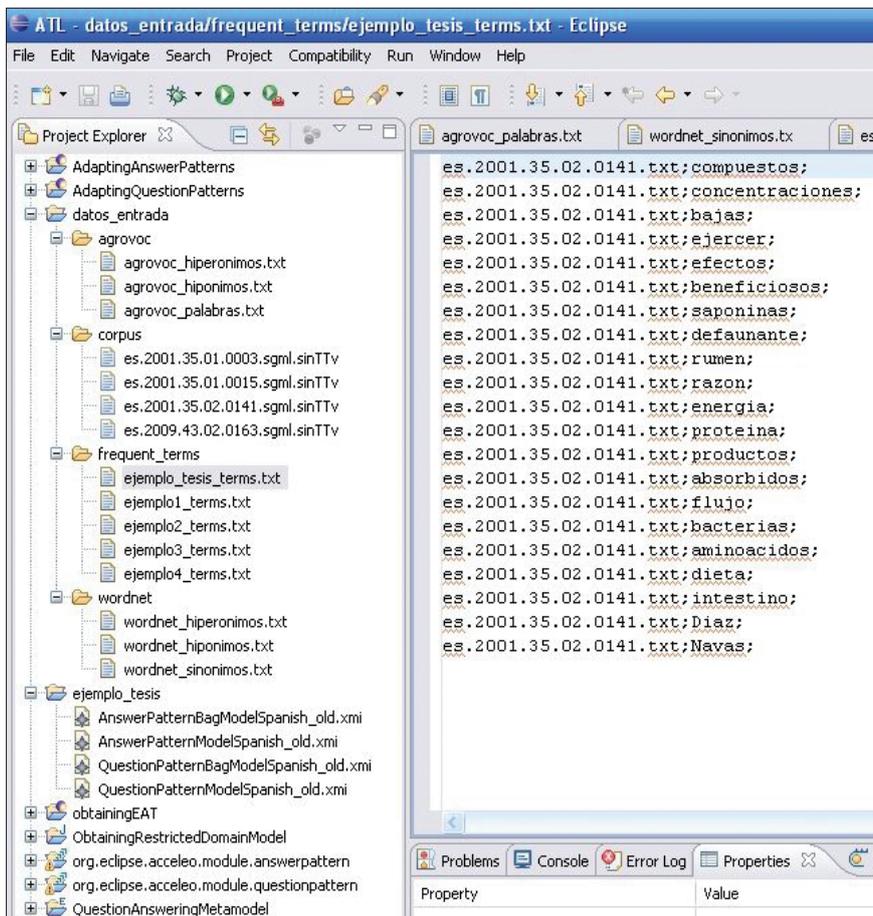


Figura 4. Obtención de términos más relevantes

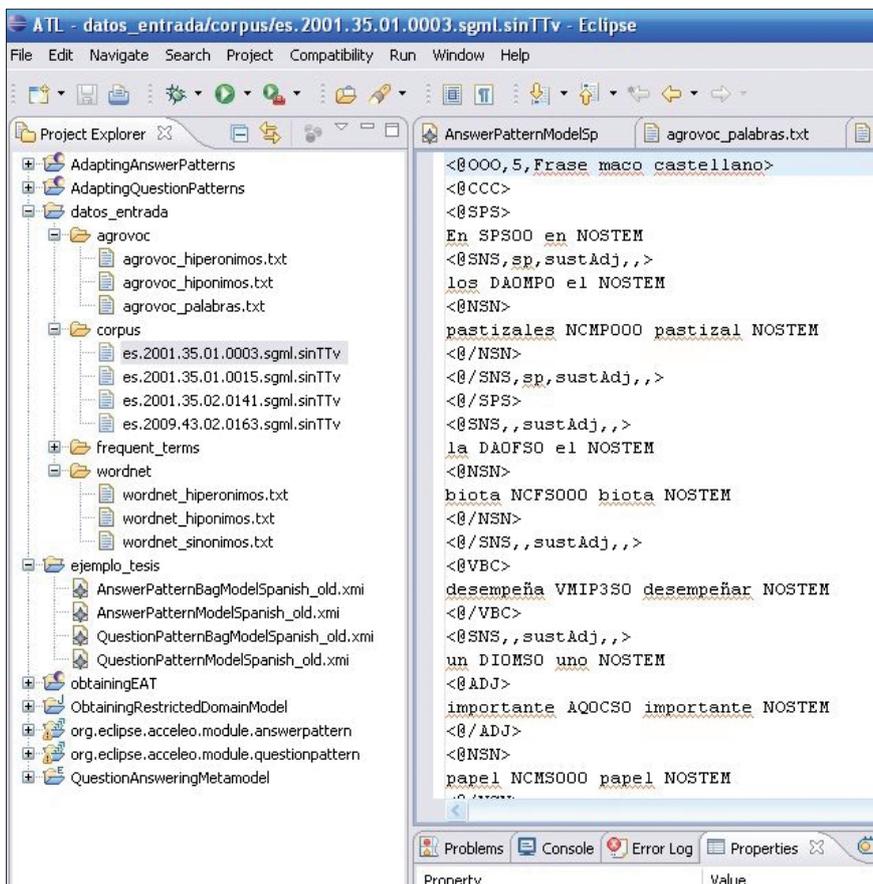


Figura 5. Ejemplo de relaciones sintácticas entre términos relevantes.

dominio. En este caso nuestro corpus es un conjunto de artículos de investigación de la *Revista cubana de ciencia agrícola (RCCA)*. En este corpus se encuentra el fragmento de texto que contiene una respuesta candidata: “[...] Sin embargo, en ocasiones, estos compuestos, en concentraciones bajas, pueden ejercer efectos beneficiosos. Por ejemplo, las saponinas tienen un efecto defaunante en el rumen, lo que puede contribuir al aumento de la razón energía: proteína de los productos absorbidos debido al incremento del flujo de bacterias y de aminoácidos de la dieta hacia el intestino [...]”. <http://www.ica.inf.cu/revista-cubana-de-ciencia-agricola/index.php>
<http://redalyc.uaemex.mx/src/>

El primer paso realizado por *Maraqá* consiste en determinar los términos más relevantes del dominio a partir de la indexación del corpus (para RI y BR según la figura 3), tomando los términos que sean sustantivos, verbos y adjetivos y que tengan una determinada frecuencia de aparición en el corpus. Para nuestro caso, los términos más relevantes se pueden ver en la captura de pantalla de *Maraqá* de la figura 4 precedidos por el documento en el que aparecen (ej.: “compuestos” que aparece en el documento “es.2001.35.02.0141.txt”).

Se obtienen las relaciones entre estos términos más relevantes usando el análisis sintáctico del corpus realizado con anterioridad en el indexado para BR (ver fase offline de figura 3); por medio de herramientas de procesamiento del lenguaje natural: primero se utilizó el etiquetador léxico o *PoS tagger MACO* (Acebo et al., 1994) y luego el analizador sintáctico *Supar* (Ferrández; Palomar; Moreno, 1999). Por ejemplo, si un sustantivo puede ser sujeto de un verbo o qué adjetivos modifican a un sustantivo. En la figura 5 se observa un ejemplo de este análisis con *Maraqá* (la relación adjetivo-sustantivo entre los términos “importante-papel”, o la relación sujeto-verbo entre “biota-desempeña”).

Además, se debe tener en cuenta el conocimiento estructurado tanto del dominio agrícola como del dominio abierto. En este caso para el primero *Maraqá* usa el tesoro *Agrovoc* (en la figura 6 aparece cada término del te-

sauro precedido por su identificador), mientras que se usa *WordNet* como recurso de dominio abierto (figura 7).

Luego *Maraqqa* debe obtener los patrones existentes de *AliQAn* tanto de pregunta (figura 8) como de respuesta (figura 9), con el fin de adaptarlos al dominio agrícola. En la figura 8 se aprecia cómo se extrae la estructura del patrón de pregunta “patronEOS” con:

- las expresiones “ptdtEO5” (con valor “qué”) y “snsEO5”, relacionado con el concepto “objeto_inanimado” y todos sus hipónimos, como “edificio” o “instrumento_musical”;
- la asociación “ptdtEO5-snsEO5” que sirve para relacionar ambas expresiones;
- el TRE “entidad_objeto” que sirve para identificar preguntas del tipo: “¿Qué instrumento musical tocaba Beethoven?”

En la figura 9 se aprecia cómo se extrae la estructura del patrón de respuesta “patronRespEO” con:

- TRE “entidad_objeto”;
- los conceptos asociados de “objeto_inanimado” y sus hipónimos; el cual permite encontrar la respuesta a la pregunta anterior en el siguiente texto “[...] Beethoven, último gran representante del clasicismo vienés, tocaba el piano desde joven [...]”.

Una vez se tienen todos estos recursos, *Maraqqa* obtiene un modelo del dominio restringido donde se determina la equivalencia entre los términos del corpus y los conceptos del dominio restringido. Después obtiene todos los conceptos relacionados usando SOC de dominios y genéricos. La figura 10 muestra el modelo generado por *Maraqqa* donde se observa que el término “saponinas” tiene un concepto equivalente en el SOC de dominio (*Agrovoc*) con código “6795”. Luego se obtienen los conceptos relacionados con el mismo en *Agrovoc* (como los hipónimos “glicoalcaloides” y “ginsenósidos” y los hiperónimos “glicósidos”, “carbohidratos” y “compuestos_organicos”) y en *WordNet* (como los hiperónimos “compuesto_químico”, “sustancia”, “objeto_inanimado” y “entidad”).

Una vez obtenido el modelo de dominio restringido, *Maraqqa* genera una taxonomía de TRE propia del dominio

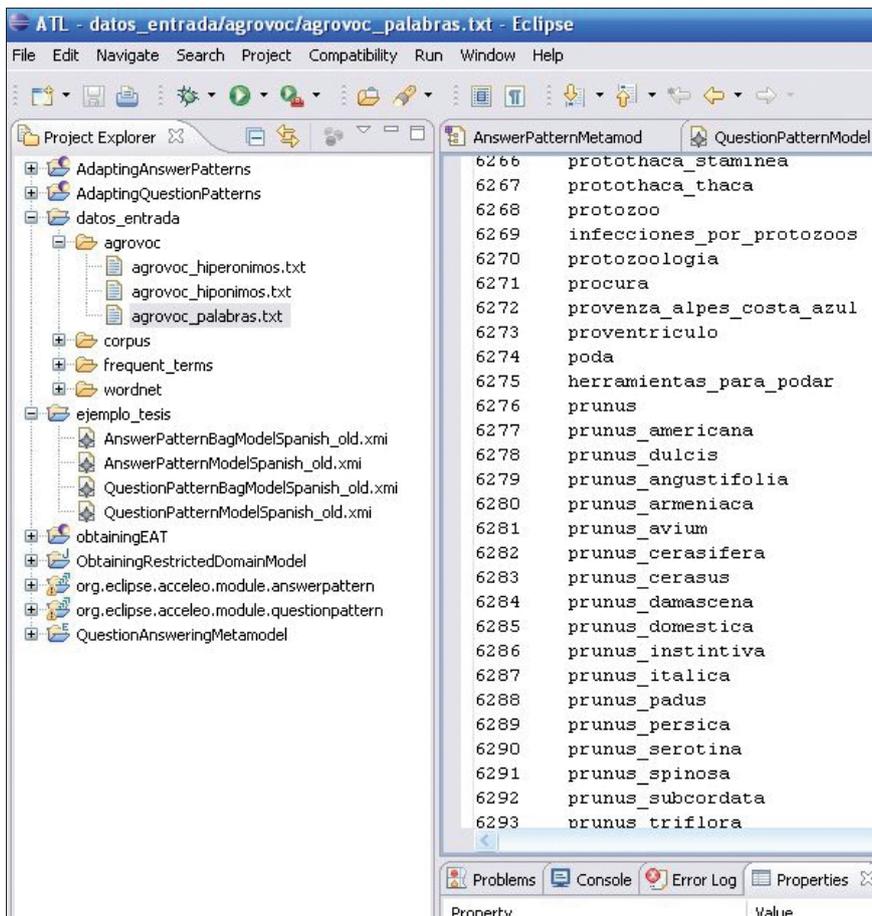


Figura 6. Ejemplo del uso de *Agrovoc* por *Maraqqa*

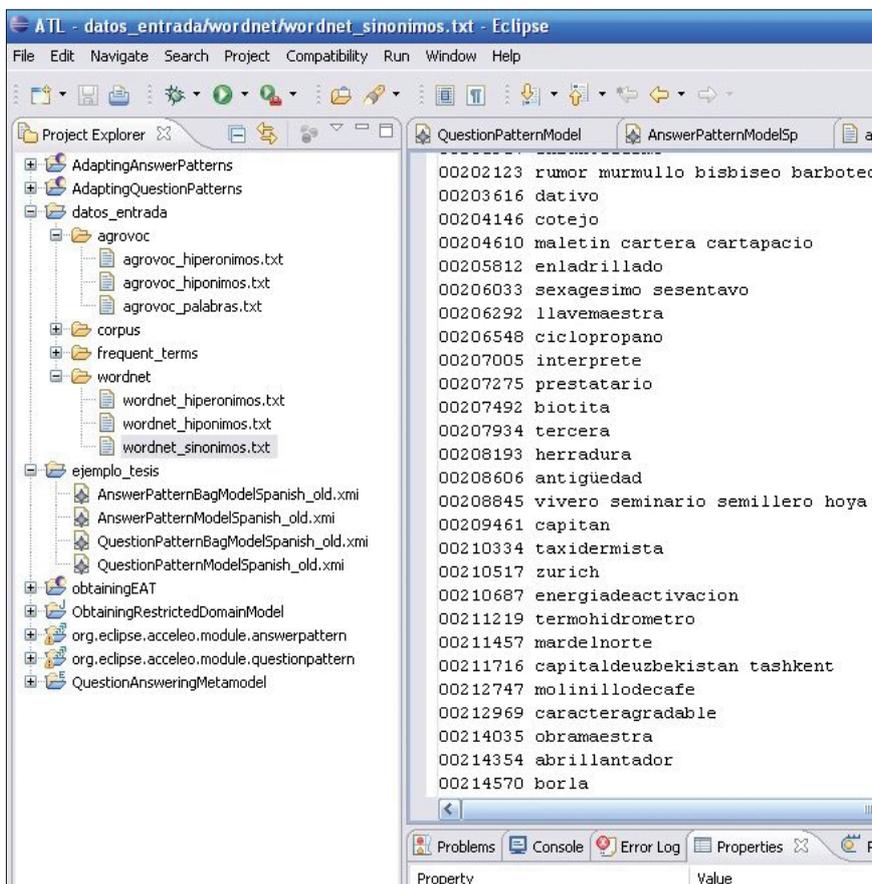


Figura 7. Ejemplo del uso de *WordNet* por *Maraqqa*

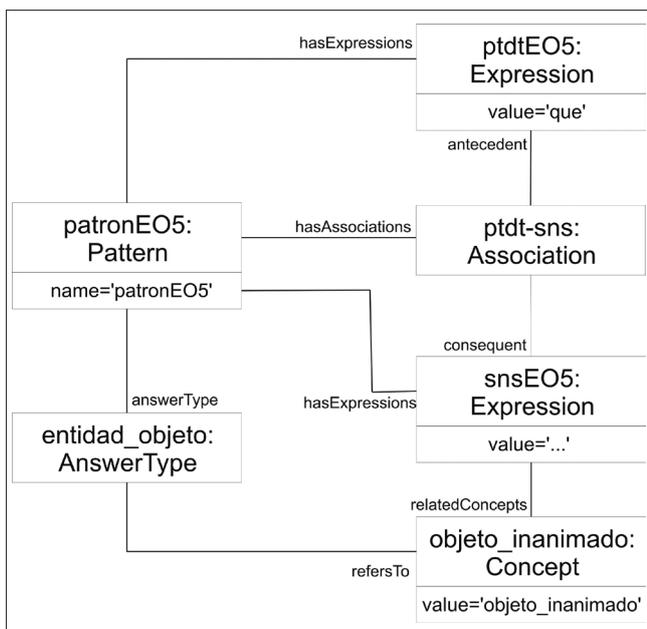


Figura 8. Ejemplo de patrones de pregunta de DA extraídos por *Maraqa*

restringido que estamos tratando. La figura 11 muestra la correspondiente para nuestro caso de estudio, donde se observa que el concepto “glicósidos” tiene como hiperónimo a “compuesto orgánico”. Siguiendo esta taxonomía, la pregunta ejemplo tendría como TRE “glicósidos” y el espacio de búsqueda estaría sólo restringido a tipos de “glicósidos” (como “saponinas”) los cuales podían ser aceptados como respuestas correctas.

Finalmente, con la taxonomía de TRE para el dominio específico, *Maraqa* adapta los patrones de pregunta y respuesta existentes al dominio (en este caso agrícola) y genera el código correspondiente a los mismos para ser adicionado al sistema de BR de DA. Un ejemplo de patrón de pregunta adaptado se muestra en la figura 12 y su correspondiente código en C++ en la figura 13. Dicho patrón permite dar respuesta a la pregunta que se ha empleado de ejemplo “¿Qué glicósidos tienen un efecto defaunante en el rumen?”, la cual sería “saponinas”

En la siguiente sección explicaremos cómo fueron realizados los experimentos y expondremos los resultados obtenidos.

Evaluación de *Maraqa* en el dominio agrícola

Para realizar la evaluación empleamos los primeros 43 volúmenes de la *Revista cubana de ciencia agrícola (RCCA)* como dominio restringido agrícola. Es una publicación del *Instituto de Ciencia Animal*, perteneciente al *Ministerio de Educación Superior* de la República de Cuba, editada en español e inglés desde 1967. Cada uno de los volúmenes tiene un promedio de tres o cuatro números, lo que hace un total de 140 números y 2.000 artículos (28,65 MB

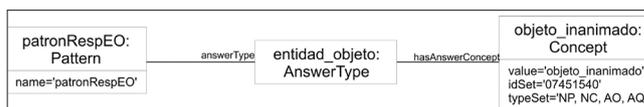


Figura 9. Ejemplo de patrones de respuesta de DA extraídos por *Maraqa*

como ficheros pdf). En concreto, el corpus textual *RCCA* está formado por un conjunto de ficheros de texto plano creados a partir de los archivos originales (con formato de texto enriquecido) de la *RCCA*, usando la utilidad de *Linux pdftotext*. De esta manera se evitaron los problemas con la codificación de los caracteres o con la maquetación (formato de doble columna, encabezados y pies de página, conversión de tablas y fórmulas, etc.). Para llevar a cabo los experimentos se contó con un conjunto de 330 preguntas *RCCA* elaboradas por expertos del dominio. La inmensa mayoría de las preguntas son de tipo factual y están basadas en hechos, preguntando por nombres de plantas, microorganismos, sustancias, fármacos, compuestos químicos, alimentos, fauna, personas, localización, día en el que sucedió algún hecho, etc. Algunos ejemplos de preguntas son: “¿Qué es la necrosis cerebrocortical?” o “¿Qué produce la cytophaga?”.

Después de usar la herramienta *Maraqa* y los recursos de conocimiento disponibles (corpus *RCCA*, *Agrovoc* y *WordNet*), se obtuvo un total de 9.022 términos relevantes, 921 conceptos en la taxonomía de TRE para el dominio agrícola de la *RCCA*, y 2.600 y 325 patrones de preguntas y respuestas, respectivamente. De esta manera obtuvimos el sistema de BR para el dominio agrícola de la *RCCA* (Vila; Mazón; Ferrández, 2011).

Para verificar la eficiencia del sistema de BR para el dominio agrícola, obtenido a partir de *AliQAn* y usando *Maraqa*, se llevaron a cabo dos experimentos (Vila, 2010; Vila; Mazón; Ferrández, 2011). Ambos tenían como objetivo medir la precisión de los dos sistemas de BR sobre el dominio agrícola: el primer experimento usando el sistema de BR de DA *AliQAn* (SBR-DA *AliQAn*) y el segundo empleando el sistema de BR de DR agrícola (SBR-DR *Agrícola*). Los resultados de ambos experimentos (figura 14) demostraron que la precisión del sistema de BR de DR *agrícola* era mayor que la precisión

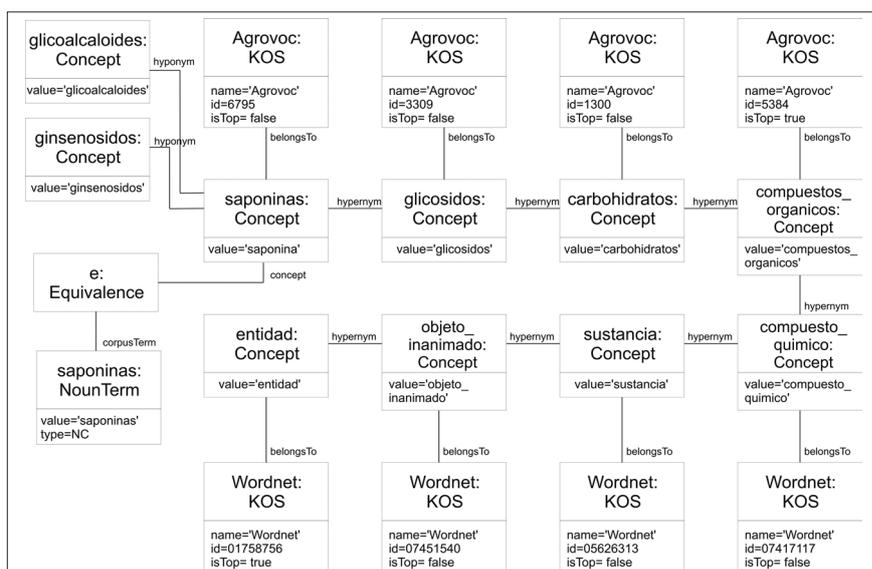


Figura 10. Ejemplo del modelo de dominio restringido y representación del concepto “saponinas” por *Maraqa*

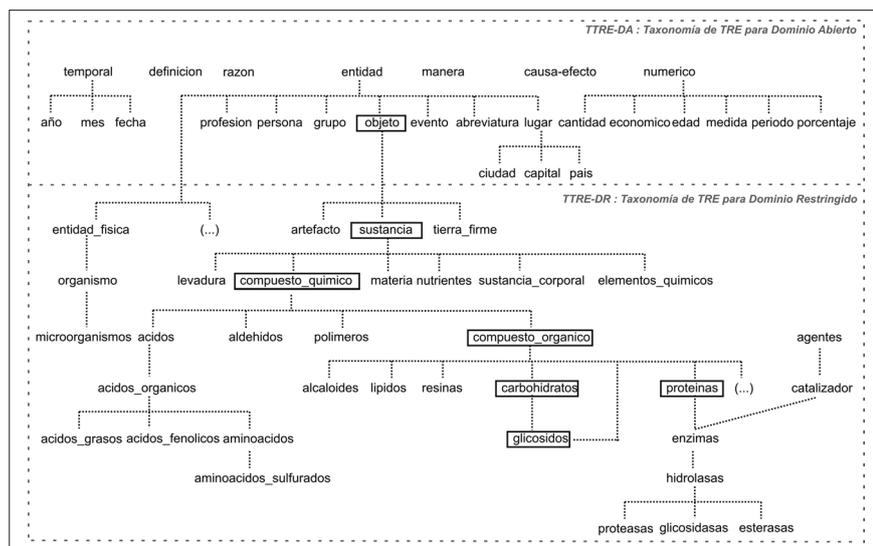


Figura 11. Ejemplo de taxonomía de TRE obtenida por *Maraqqa*

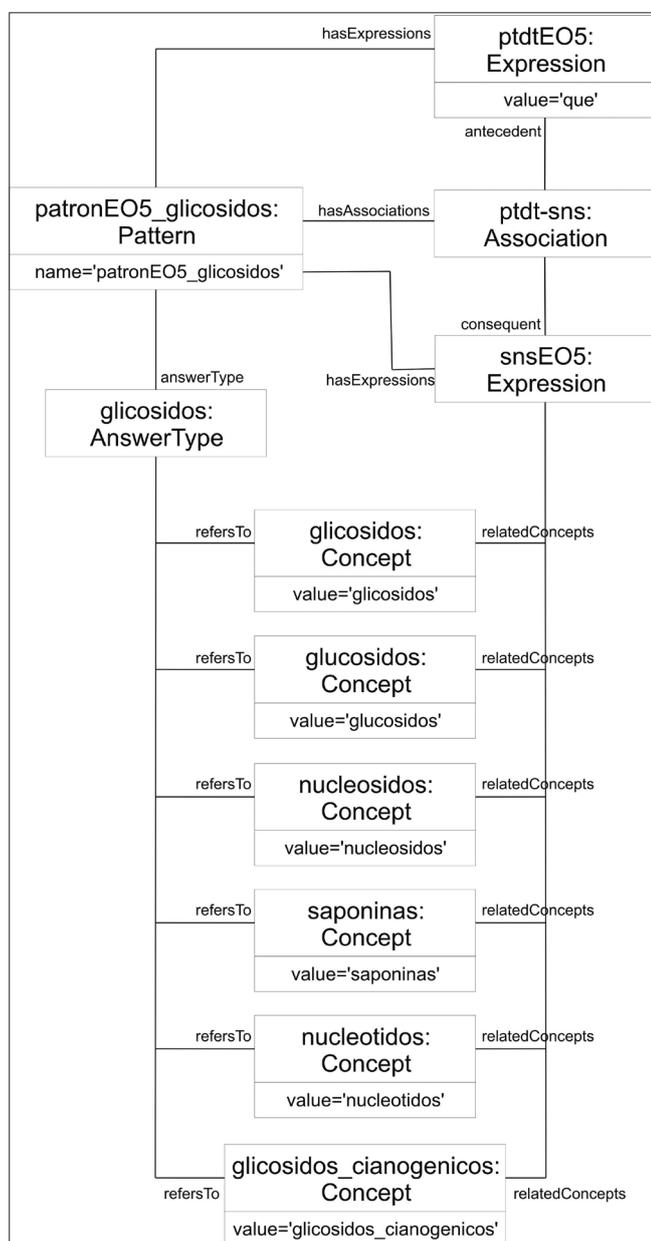


Figura 12. Ejemplo de modelo de patrón de respuesta adaptado al dominio por *Maraqqa*

del sistema de BR de DA *AliQAn* sobre el mismo dominio de aplicación. Por ejemplo, para las primeras respuestas del sistema *baseline AliQAn* sobre el dominio agrícola *RCCA* fue de 28,8% y la del sistema de BR para el dominio agrícola fue de 58,3%. De esta manera nuestro sistema tuvo una mejora del 29,5% de precisión en las respuestas. Estos resultados demuestran la efectividad de nuestro método que supera la media del sistema *AliQAn* (43% de precisión) en todas sus actuaciones pasadas sobre dominios abiertos (Roger et al., 2008).

Conclusiones

Se ha presentado el sistema BR *Maraqqa*, realizado siguiendo el paradigma del desarrollo dirigido por modelos mediante la plataforma *Eclipse*. Permite adaptar fácilmente sistemas de BR en DA a dominios restringidos, integrando los recursos de conocimiento anteriormente utilizados con los nuevos recursos del DR, y generando el código fuente del sistema de BR adaptado a dicho dominio.

Para ello requiere como entrada la colección de documentos sobre la que se realizará la búsqueda de información, los recursos de conocimiento del DR independientemente del formato que tengan, y el código fuente de los patrones de preguntas y respuestas del sistema de BR en DA que se desea adaptar. A partir de estos recursos, genera la taxonomía de TRE refinada para el dominio en cuestión, y el código fuente de los patrones de preguntas y respuestas adaptados al dominio. Todo el proceso de funcionamiento de la herramienta ha sido ilustrado mediante un caso de estudio en el dominio agrícola.

Con *Maraqqa* se supera el problema de la baja precisión alcanzada cuando un sistema de BR de DA trabaja en un dominio restringido (pasando de una precisión del 28,8% al 58,3%). Además se facilita la integración de los recursos de conocimiento de dicho dominio al sistema de BR, y todo ello a partir del corpus textual como punto de partida y fuente de información principal, en detrimento del uso de corpus de preguntas de otras propuestas, lo que consideramos poco realista en DR. La principal contribución es la adaptación automática de sistemas de BR existentes (específicamente de sus patrones de pregunta-respuesta y las taxonomías de TRE) a diferentes dominios restringidos, sin requerir ningún esfuerzo manual como el resto de propuestas previas existentes (Sekine; Sudo; Nobata, 2002; Hovy; Hermjakob; Ravichandran, 2002; Metzler; Croft, 2005; Li; Roth, 2006; Ferrés; Rodríguez, 2006; Terol; Martínez-Barco; Palomar, 2006; Kosseim; Yousefi, 2008; Peñas et al., 2009).

Es posible la obtención de recursos de conocimiento (con información semántica) del dominio restringido, necesarios para aplicar *Maraqqa*, a partir de la utilización de sistemas de alineamiento automático entre diferentes SOC's (Soergel et al., 2004; Van-Hage et al., 2010); lo que planteamos como un interesante trabajo futuro.

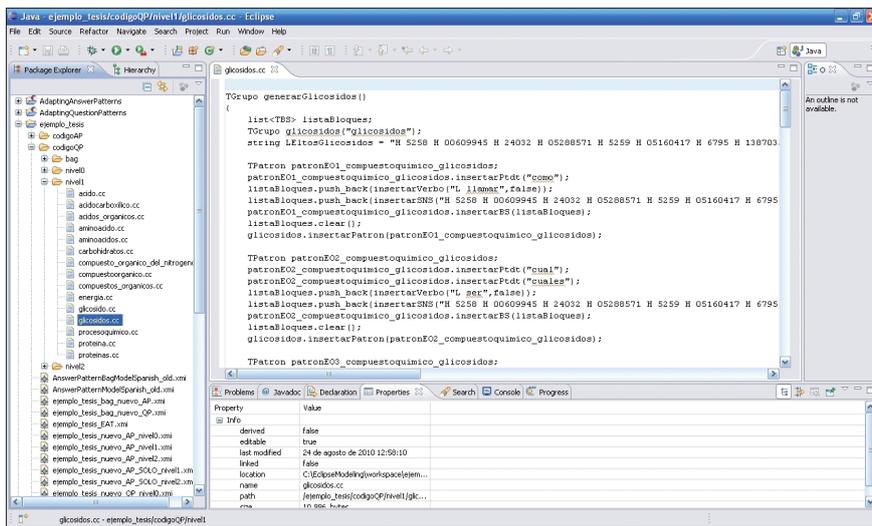


Figura 13. Ejemplo del código del patrón de pregunta adaptado al dominio por Maraqa

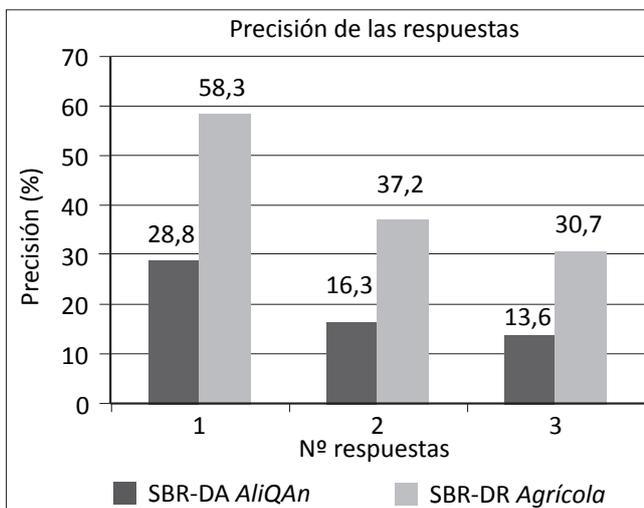


Figura 14. Precisión del sistema de BR para el dominio agrícola (desarrollado con Maraqa)

Agradecimientos

Esta investigación ha sido parcialmente financiada por el gobierno español bajo el proyecto *TIN2009-13391-C04-01*; y el gobierno valenciano bajo el proyecto *Prometeo/2009/119*. La investigación de **Katia Vila** ha sido financiada por una beca *MAEC-Aecid* del gobierno español.

Bibliografía

Acebo, S.; Ageno, Alicia; Climent, Salvador; Farreres, Javier; Padró, Lluís; Placer, Roberto; Rodríguez, Horacio; Taulé, Mariona; Turmo, Jordi. "MACO: Morphological analyzer corpus-oriented". *Esprit BRA-7315 Aquilex II. Working paper 31*, 1994.

Baeza-Yates, Ricardo; Ribeiro-Neto, Berthier. "Modern information retrieval". *ACM Press/Addison-Wesley*, 1999. ISBN: 978 0321416919

Bézivin, Jean. "On the unification power of models. Software and system modeling", 2005, v. 4, n. 2, pp. 171-188. <http://atlanmod.emn.fr/www/papers/OnTheUnificationPowerOfModels.pdf>

Ferrández, Antonio; Palomar, Manuel; Moreno, Lidia. "An empirical approach to Spanish anaphora resolution". *Machine translation*, 1999, v. 14, n. 3-4, pp. 191-216. ftp://dlsi.ua.es/people/antonio/ART_MUL5.pdf

Ferrés, Daniel; Rodríguez, Horacio. "Experiments adapting an open-domain question answering system to the geographical domain using scope-based resources". En: *MLQA'06: Procs of the Workshop on multilingual question answering. Association for Computational Linguistics*, 2006, pp 69-76. <http://dl.acm.org/citation.cfm?id=1708097.1708111>

Hodge, Gail. *Systems of knowledge organization for digital libraries: beyond traditional authority files*. The Digital Library Federation Council on Library and Information Resources, 2000. ISBN: 1887334769 <http://www.clir.org/pubs/reports/pub91/pub91.pdf>

Hovy, Eduard; Hermjakob, Ulf; Ravichandran, Deepak. "A question/answer typology with surface text patterns". En: *Procs of the 2nd intl conf on human language technology research*, Morgan Kaufmann Publishers Inc., 2002, pp. 247-251. <http://www.isi.edu/natural-language/projects/webclope/dia/pubs/02hlt.pdf>

Katz, Boris. "From sentence processing to information access on the World wide web". *AAAI Spring symposium on natural language processing for the World wide web*, 1997, pp. 77-94. <http://aaai.org/Papers/Symposia/Spring/1997/SS-97-02/SS97-02-010.pdf>

Katz, Boris; Borchardt, Gary; Felshin, Sue. "Natural language annotations for question answering". En: *Procs of the 19th intl Flairs conf (Flairs 2006)*, May 2006. <http://groups.csail.mit.edu/infolab/publications/FLAIRS0601KatzB.pdf>

Kleppe, Anneke; Warmer, Jos; Bast, Wim. *MDA explained. The practice and promise of the model driven architecture*. Addison Wesley, 2003. ISBN: 978 0321194428

Kosseim, Leila; Yousefi, Jamileh. "Improving the performance of question answering with semantically equivalent answer patterns". *Data and knowledge engineering*, 2008, v. 66, n. 1, pp. 53-67. <http://dx.doi.org/10.1016/j.datak.2007.07.010>

Li, Xin; Roth, Dan. "Learning question classifiers: the role of semantic information". *Natural language engineering*, 2006, v. 12, n. 3, pp. 229-249. <http://dx.doi.org/10.1017/S1351324905003955>

Mellor, Stephen J.; Clark, Anthony N.; Futagami, Takao. "Model-driven development - Guest editor's introduction". *IEEE software*, 2003, v. 20, n. 5, pp. 14-18. <http://dx.doi.org/10.1109/MS.2003.1231145>

Metzler, Donald; Croft, W. Bruce. "Analysis of statistical question classification for fact-based questions". *Informa-*

tion retrieval, 2005, v. 8, n. 3, pp. 481-504.
<http://ciir.cs.umass.edu/pubfiles/ir-323.pdf>
<http://dx.doi.org/10.1007/s10791-005-6995-3>

Mollá, Diego; Vicedo, José-Luis. "Question answering in restricted domains: an overview". *Computational linguistics*, 2007, v. 33, n. 1, pp. 41-61.
<http://www.mitpressjournals.org/doi/pdf/10.1162/coli.2007.33.1.41>
<http://dx.doi.org/10.1162/coli.2007.33.1.41>

Peñas, Anselmo; Forner, Pamela; Sutcliffe, Richard; Rodrigo, Álvaro; Forascu, Corina; Alegria, Iñaki; Giampiccolo, Danilo; Moreau, Nicolas; Osenova, Petya. "Overview of ResPubliQA 2009: question answering evaluation over European legislation". En: *Working notes of Cross language evaluation forum (CLEF)*, 2009.
http://clef.isti.cnr.it/2009/working_notes/ResPubliQA-overview.pdf

Roger, Sandra; Vila, Katia; Ferrández, Antonio; Pardiño, María; Gómez, José-Manuel; Puchol-Blasco, Marcel; Peral, Jesús. "Using AliQAn in monolingual QA@Clef 2008". *9th Workshop of the Cross-language evaluation forum. Lecture notes in computer science*, 2008, v. 5706, pp. 333-336.
http://dx.doi.org/10.1007/978-3-642-04447-2_38

Russell, Stuart; Norvig, Peter. "Artificial intelligence: a modern approach". *Pearson education*, 2nd ed., 2003. ISBN: 978 0137903955

Sekine, Sstoshi; Sudo, Kiyoshi; Nobata, Chikashi. "Extended named entity hierarchy". En: *Procs of 3rd Intl conf on lan-*

guage resources and evaluation (LREC'02), 2002, pp. 1818-1824.
<http://nlp.cs.nyu.edu/pubs/papers/sekine-lrec02.pdf>

Soergel, Dagobert; Lauser, Boris; Liang, Anita; Fisseha, Frehiwot; Keizer, Johannes; Katz, Stephen. "Reengineering thesauri for new applications: the Agrovoc example". *Journal of digital information*, 2004, v. 4, n. 4.
<http://journals.tdl.org/jodi/article/viewArticle/112/111>

Terol, Rafael M.; Martínez-Barco, Patricio; Palomar, Manuel. "Aplicación de técnicas basadas en PLN al tratamiento de preguntas médicas en búsqueda de respuestas". *Procesamiento del lenguaje natural*, 2006, n. 36, pp. 17-24.
<http://www.sepln.org/revistaSEPLN/revista/36/02.pdf>

Van-Hage, Willem-Robert; Sini, Margherita; Finch, Lori; Kolb, Hap; Schreiber, Guus. "The OAEI food task: an analysis of a thesaurus alignment task". *Applied ontology*, 2010, v. 5, n. 1, pp. 1-28.
<http://dx.doi.org/10.3233/AO-2010-0072>

Vila, Katia. *Búsqueda de respuestas en dominios restringidos: aplicación sobre el dominio agrícola*. Tesis doctoral, Universidad de Alicante, 2010.
http://rua.ua.es/dspace/bitstream/10045/18329/1/tesis_vila.pdf

Vila, Katia; Mazón, José-Norberto; Ferrández, Antonio. "Model-driven development for adapting question answering systems to restricted domains". *Journal of research and practice in information technology*, May 2011, v. 43, n. 2, pp. 23-40.
<http://dx.doi.org/10.1145/1966883.1966893>

Helena Martín Rodero



¿Te apuntas?
Ya somos
más de 2.000



3 documentos en E-LIS

Exit ID: 1177
IraLIS: No encontrado ¿Qué es?
Institución: [Facultad de Medicina](#)
Dirección: Alfonso X el Sabio, s/n
Campus Miguel de Unamuno
Código postal: 37007
Ciudad: Salamanca
País: ES - España
Teléfono: +34-923 294 500 ext. 1846
Fax: +34-923 294 519
Correo-e: helena@usal.es

Correo-e personal: anina.helena@gmail.com

Web institucional: <http://sabus.usal.es>
Pagerank 7/10

Web personal: <http://www.usalbiomedica.com>
Pagerank 5/10

Especialidades: Biblioteca digital; Biblioteca universitaria;
Información biomédica;
Recuperación de información y búsquedas;
Revistas electrónicas

Para titulados con más de 1 año de experiencia, que hayan publicado algún artículo o ponencia o puedan dar clase más de 1 hora.



USO COMBINADO DE TECNOLOGÍAS SEMÁNTICAS Y ANÁLISIS VISUAL PARA LA ANOTACIÓN AUTOMÁTICA DE IMÁGENES Y SU RECUPERACIÓN



Sergio Rodríguez-Vaamonde, Pilar Ruiz-Ibáñez y Marta González-Rodríguez



Sergio Rodríguez-Vaamonde es ingeniero de telecomunicaciones por la *Universidad del País Vasco* y diploma de estudios avanzados en ingeniería telemática. Actualmente es doctorando, centrando sus investigaciones en la anotación automática de imágenes. Es investigador en el centro tecnológico *Tecnalia*, donde lleva a cabo trabajos de aplicación de las tecnologías de visión artificial a diferentes sectores como las TICs o el sector industrial.

*Unidad de Sistemas de Información e Interacción
División ICT-European Software Institute, Tecnalia
Parque Tecnológico de Bizkaia
Ibaizabal Bidea, Edif. 202. 48170 Zamudio, España
sergio.rodriguez@tecnalia.com*



Pilar Ruiz-Ibáñez, licenciada en geografía e historia por la *Universidad de Granada*, es investigadora en la *Unidad de Sistemas de información e Interacción* de *Tecnalia*. Antigua responsable del Centro de Documentación de *Robotiker*. Directora del proyecto de fusión de gestión del conocimiento de *Tecnalia R&I*.

*Unidad de Sistemas de Información e Interacción
División ICT-European Software Institute, Tecnalia
Parque Tecnológico de Bizkaia
Ibaizabal Bidea, Edif. 202. 48170 Zamudio, España
pilar.ruiz@tecnalia.com*



Marta González-Rodríguez, licenciada en informática por la *Universidad del País Vasco* y máster en tecnologías avanzadas de fabricación por la misma universidad, es gerente de *Tecnologías Semánticas* de *Tecnalia*. Participa en proyectos de I+D de ámbito nacional y europeo en el campo de las tecnologías semánticas e investiga su aplicación a diferentes sectores.

*Unidad de Sistemas de Información e Interacción
División ICT-European Software Institute, Tecnalia
Parque Tecnológico de Bizkaia
Ibaizabal Bidea, Edif. 202. 48170 Zamudio, España
marta.gonzalez@tecnalia.com*

Resumen

Se analizan los actuales sistemas semánticos para la indexación y recuperación de la información no textual en internet. El uso combinado de técnicas de visión artificial y el contenido textual en el que se enmarcan las imágenes, conforman los algoritmos más efectivos para conseguir que los motores de búsqueda en la nube puedan llegar a generar los mejores resultados en la recuperación de la información pertinente. El futuro inmediato de la Web pasa por lograr una contextualización automática de las imágenes que permita establecer similitudes entre los contenidos no textuales, y poder recuperarlos de forma efectiva.

Palabras clave

Anotación automática de imágenes, Visión artificial, Tecnologías semánticas, Procesamiento de lenguaje natural, Contextualización de imágenes, *Content-based image retrieval (cbir)*.

Title: Joint use of semantics and visual analysis for automatic image annotation and retrieval

Abstract

The present stage of development of semantic systems used in the indexing and retrieval of non-text information on the internet is described. The most effective algorithms for retrieval of non-text information are the combined use of computer vision and content analysis of the text associated with the images. These techniques can lead to the best results in the retrieval of relevant information. The Web's immediate future lies in automatic contextualisation of images to establish similarities between them and to be able to effectively retrieve non-text content.

Artículo recibido el 16-11-11
Aceptación definitiva: 16-12-11

Keywords

Automatic image annotation, Computer vision, Semantic technologies, Natural language processing, Image contextualisation, Content-based image retrieval (cbir).

Rodríguez-Vaamonde, Sergio; Ruiz-Ibáñez, Pilar; González-Rodríguez, Marta. “Uso combinado de tecnologías semánticas y análisis visual para la anotación automática de imágenes y su recuperación”. *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 27-33.

<http://dx.doi.org/10.3145/epi.2012.ene.04>

1. Introducción

La mayoría de las aplicaciones y tecnologías que están surgiendo en internet se centran en permitir a los usuarios utilizar diferentes tipos de contenidos multimedia, como vídeos o imágenes. Por ejemplo, *Facebook* publicaba a principios de 2011 que en el último fin de semana del año anterior se subieron 750 millones de fotografías (*Facebook*, 2011). Esta tendencia crece de forma imparable, ya que *Youtube* reportaba una subida de 8 horas de vídeo por minuto en 2007 mientras que en mayo de 2011 mostraba unas estadísticas de 48 horas por minuto (*Youtube*, 2011).

“ El futuro de la Web pasa por gestionar toda la información multimedia ”

El futuro de la Web pasa por gestionar toda esa información multimedia, y sobre todo por hacer posible su acceso de forma inteligente desde cualquier parte del mundo, teniendo en cuenta que un contenido que no se puede encontrar, *no existe*. Las tecnologías involucradas son numerosas: desde el manejo de grandes cantidades de datos, hasta las redes de telecomunicaciones que soportan el movimiento de estas grandes cantidades de información, pasando por bases de datos distribuidas y de rápida accesibilidad para mejorar los tiempos de consulta. Es fundamental permitir a los usuarios el acceso eficiente a la información multimedia que deseen. El objetivo es favorecer que los diferentes motores de búsqueda sean capaces de encontrar la información multimedia en función de su contenido, lo que se denomina *content based information retrieval (cbir)*.

El caso de las imágenes es un ejemplo claro: si un usuario desea buscar una fotografía del edificio del museo *Guggenheim* de Bilbao, el sistema deberá analizar las existentes en internet y saber que una imagen muestra el *Guggenheim* de Bilbao y no el de New York, Venecia, Berlín o Las Vegas. En la actualidad esto dista de estar resuelto. Los grandes motores, como *Google* o *Bing* en su versión de búsqueda de imágenes, realizan dicha búsqueda en función de las etiquetas html que incluyen los usuarios, del propio nombre de la imagen o del pie de foto de la misma. Esto deriva en búsquedas no siempre adaptadas a las necesidades reales de los usuarios.

Estas necesidades latentes han provocado que tanto los motores de búsqueda como los investigadores de todo el mundo, estén analizando nuevas formas de llevar a cabo búsquedas efectivas sobre grandes conjuntos de contenidos multimedia y, específicamente, sobre imágenes estáticas.

De todas ellas, el modelo más utilizado, es la anotación automática de imágenes (**Makadia; Pavlovic; Kumar**, 2010). En este modelo, dada una imagen, un sistema automático se encarga de analizarla y etiquetarla en función de varios parámetros, de tal forma que al realizar la búsqueda los motores analicen las etiquetas textuales de forma rápida y arrojen los resultados más pertinentes.

El presente artículo se centra en analizar las tendencias actuales a la hora de anotar una imagen de forma eficiente, en base a su contenido visual y a la información que aporta el propio sitio web, lo que se denominaría el “contexto de la imagen”. Así mismo, se describirá brevemente el estado actual de los buscadores de imágenes, así como las necesidades planteadas de cara a encontrar lo que los usuarios realmente quieren.

“ Es fundamental permitir a los usuarios el acceso eficiente a la información multimedia que desean ”

2. Anotación semántica de imágenes

El primer paso a dar para realizar la búsqueda de imágenes es anotar su contenido. Es primordial analizar la propia imagen mediante tecnologías de visión artificial, con lo que se logrará conocer el contenido de la misma. Estas técnicas permiten obtener su representación matemática de tal forma que posteriormente es necesario interpretar esta información de bajo nivel para encontrar su correspondencia con conceptos semánticos de alto nivel. Este es el principal problema que encuentran los investigadores en la actualidad para poder realizar una anotación de imágenes en función de su contenido y se le denomina *semantic gap*.

Desde el comienzo de la anotación de imágenes no estaba claro si era posible atravesar el *semantic gap*. Diferentes estudios del cerebro humano han investigado cómo se representan los objetos en el córtex visual (**Serre et al.**, 2007; **Dicarlo**, 2009), pero no se podía saber cómo se realiza el cambio a una representación semántica de la imagen formada en la retina (**Tousch; Herbin; Audibert**, 2011). En cuanto a la posibilidad de realizar la interpretación semántica por parte de sistemas automáticos de procesamiento, no ha sido hasta 2011 cuando se ha demostrado que cuanto más similitudes visuales tengan dos imágenes, mayor será su similitud semántica (**Deselaers; Ferrari**, 2011).

La barrera del *semantic gap* es crítica a la hora de realizar la anotación y se ha abordado desde diferentes puntos:

- utilizando técnicas estadísticas para mapear características de bajo nivel a vocabularios no estructurados de alto nivel (Csurka *et al.*, 2004);
- mediante representaciones intermedias compartidas entre varios objetos (Torresani; Szummer; Fitzgibbon, 2010);
- usando jerarquías semánticas de vocabularios limitados que aportan un mayor grado de riqueza semántica (Gao; Chia; Cheng, 2010);
- mediante la utilización de técnicas complejas de procesamiento de lenguaje natural para definir, no sólo el contenido de las imágenes sino también las relaciones existentes entre el propio contenido (Lee; Trauman, 2010).

Es primordial analizar la propia imagen mediante tecnologías de visión artificial para conocer su contenido

El contenido de la imagen es muy importante a la hora de anotarla para su posterior recuperación. Pero hay situaciones en las que la información visual no aporta todo el contenido de la propia imagen y es evidente que la combinación de la información visual con otra información textual, es beneficiosa para los sistemas de anotación (García-Serrano, 2011). Para lograr una mayor riqueza de anotación, destacamos una línea de investigación a la que cada vez se están dedicando más esfuerzos y será muy relevante en el futuro: la “contextualización de la imagen”. Generalmente, una imagen estará localizada dentro de una o varias webs que tratarán sobre temas relacionados con ella. De esta forma, la contextualización de la imagen analiza las propias webs, en vez de la imagen, y busca obtener un mayor conocimiento del contenido de la imagen además del que se pueda extraer de la misma. Por ejemplo, si se presenta en una web una imagen del cuadro de *Las meninas*, el sistema de procesamiento visual sabrá que hay un perro, un espejo, varias personas, incluso podría llegar a saber quiénes son las personas. Pero, del texto que rodea la imagen se puede extraer información sobre quién fue el autor, cuándo se pintó, pudiendo reconocer que se trata del cuadro original pintado por Velázquez o la reinterpretación de Picasso.



Figura 1. *Las meninas* de Velázquez

De todo ello se hablará más en profundidad en los siguientes apartados. Primero, se describirán las líneas de investigación actuales en el campo del análisis de imagen para la detección de objetos y anotación de imágenes a gran escala, mientras que en segundo lugar se mostrará el funcionamiento de la contextualización de las imágenes con el objetivo de disminuir el *semantic gap* enriqueciendo la anotación automática de las imágenes.

2.1. Procesamiento automático de imágenes para anotación del contenido visual

Su objetivo es utilizar tecnologías de visión artificial para detectar qué “elementos” componen la imagen. Como es de suponer, el concepto “elemento” es muy amplio y por tanto, las anotaciones de las imágenes también pueden ser extremadamente extensas, desde la detección de las personas que están presentes hasta la del tipo de árboles existentes o las acciones que se están realizando.

Desde el comienzo de la anotación de imágenes no estaba claro si era posible atravesar el *semantic gap*

Detección de objetos a pequeña escala

Una de las líneas de investigación que tiene más fuerza en el campo de la visión artificial es la categorización de objetos en imágenes, es decir, identificar qué clases de objetos existen en una imagen (persona, coche o edificio).

Esta línea de investigación no es nueva y los primeros trabajos se remontan a los años 70 (Fischler; Elschlager, 1973), con resultados modestos. Hoy en día, la evolución del campo de detección y categorización de objetos está aumentando de forma exponencial. En los últimos trabajos el objetivo es crear y entrenar un algoritmo específico para detectar un tipo de objeto, existiendo varias aproximaciones que se pueden ver en diversos artículos de referencia (Dicarlo, 2009).

Para lograr avanzar en este campo, numerosas comunidades de investigadores están creando bases de datos de imágenes comunes y públicas sobre las que cualquiera pueda realizar pruebas, de tal forma que sea posible establecer



Figura 2. *Las meninas* de Picasso

comparaciones entre los diferentes métodos que se proponen. Un ejemplo es el denominado *Visual object classification challenge* (Everingham et al., 2010). Esta competición se celebra una vez al año y reúne decenas de algoritmos que tratan de obtener la mejor tasa de aciertos sobre un conjunto definido de 20 objetos, entre los que se encuentran por ejemplo “persona”, “botella”, “avión” o “planta”; y en la cual se puede ver el avance del estado de la técnica. En la tabla 1 se comparan los algoritmos ganadores de la competición entre los años 2007-2011. Se muestra para cada año, la mejor puntuación obtenida por el algoritmo ganador, la peor y la puntuación media de los 20 objetos. Además, muestra qué algoritmo ha sido el ganador, así como los objetos que han conseguido la mejor y la peor puntuación.

Año	Mejor objeto detectado	Mejor puntuación (%)	Peor objeto detectado	Peor puntuación (%)	Puntuación media (%)
2007	persona	85,9	botella	33,1	59,4
2008	persona	86,9	planta	29,2	54,9
2009	avión	88,0	planta	36,6	66,5
2010	avión	93,0	planta	48,6	73,8
2011	avión	95,0	planta	56,5	78,6

Tabla 1. Algoritmos ganadores de la competición entre 2007-2011

Desde 2007 hasta 2011 se ha producido una evolución en las tasas de acierto, de tal forma que todas (incluida la menor) han aumentado con el paso de los años.

El salto cuantitativo ha sido muy grande, por lo que cada vez los sistemas de detección de objetos son más perfectos en tareas como detectar un objeto concreto en un determinado tipo de imagen.

Uno de los sistemas de contextualización que más se está trabajando es la relación de una imagen con sus etiquetas asociadas en la página web

Detección de objetos a gran escala

A pesar del avance en la detección de objetos, estos trabajos se están centrando en el análisis de imágenes a pequeña escala (se trabaja con unos pocos cientos de imágenes en escenas más o menos controladas). Cuando se pasa a una escala del tamaño de internet, donde las imágenes pueden contener todo tipo de objetos en todo tipo de situaciones, se deben tener en cuenta otros aspectos. Este ámbito comienza a tener una gran relevancia y se están generando bases de datos comunes para la comparativa, como puede ser *Saipr-TC 12* (Escalante et al., 2010), con unas 20.000 imágenes, *MIR Flickr* (Huiskes; Lew, 2008), con 1 millón de imágenes en su edición de 2010, o las competiciones *The imagenet large scale visual recognition challenge* o *ImageCLEF*.

Existen dos grandes aproximaciones para solventar el problema de la anotación automática: modelizar todo el universo de elementos que existen o reducir el mundo a la hora

de realizar la anotación centrándose en un conjunto limitado de imágenes.

Una aproximación muy empleada, debido a su similitud con los algoritmos clásicos mencionados anteriormente, es modelizar de forma individual todo el universo de objetos a anotar (Hao et al. 2009). Otra que está teniendo muy buenos resultados es la de combinar información estadística de todos los tipos de conceptos a anotar en una imagen y realizar una modelización conjunta (Weston; Bengio; Usunier, 2011). En este caso sigue siendo necesario modelizar todo el universo, pero la combinación de información relativa a relaciones semántica y visual entre conceptos posibilita mejorar la anotación.

Para lograr una mayor riqueza de anotación, destaca una línea de investigación que será muy relevante en el futuro, la “contextualización de la imagen”

Existe una segunda aproximación: reducir el problema de anotación antes de modelizar los objetos. La ventaja clara de este tipo de sistemas es que no es necesario modelizar todos y cada uno de los elementos existentes en el mundo. La propuesta, cuyo máximo exponente se encuentra en el artículo de Makadia, Pavlovic y Kumar (2010), se basa en que a la hora de anotar una imagen, ésta se compare visualmente con un conjunto de entrenamiento ya anotado. Una vez se conozcan las imágenes más similares, se analizan sus anotaciones y se trata de inferir, en base a ellas, las anotaciones de la imagen que deseamos anotar. Esta aproximación es mucho más sencilla que otras propuestas anteriores, pero según sus autores es la que mejores resultados obtiene, con una tasa de aciertos del 35% mientras que con otros métodos de modelización de objetos más complejos se obtienen valores inferiores, como el 32% (Llorente; Manmatha; Rüger, 2010). Debido a su simplicidad, han comenzado a aparecer evoluciones de este método que permiten alcanzar una tasa de acierto hasta del 52% (Verbeek et al., 2010).

2.2. Contextualización de las imágenes en la web

A pesar de las mejoras en el procesamiento de imágenes, hay más información adicional que puede extraerse conociendo el entorno web que rodea a una imagen y que se denomina “contexto de la imagen”. Este contexto web es capaz de mejorar la anotación automática de imágenes gracias a que permite eliminar la ambigüedad de los conceptos a anotar. En este apartado se mostrarán dos de las tecnologías más relevantes de contextualización web, la primera de ellas se basa en las propias “etiquetas” relacionadas con las imágenes, mientras que la segunda trata de un nuevo modelo de contextualización fundamentado en las redes sociales.

Contextualización en base a etiquetas

Uno de los sistemas de contextualización que más se está trabajando en la actualidad es la relación de una imagen con sus etiquetas asociadas en su página del sitio web. Un ejemplo claro es *Flickr*, donde cada usuario puede subir sus imágenes y etiquetarlas en función de sus intereses personales,

y donde gracias a su importancia están surgiendo numerosas iniciativas (**Ulges; Worrying; Breuel, 2011**). Una de las más innovadoras es el hecho de utilizar la estructura del contenido de la web que proporcionan los usuarios. Por ejemplo, la comunidad *Flickr* organiza sus fotos en base a los grupos. Existen 200.000 grupos que han sido definidos y relacionados con todo tipo de temas como “fotos de fiestas” o “fotografía natural”. La utilización de los grupos se centra en la fase de aprendizaje del sistema, donde para cada grupo se genera un modelo. Durante la anotación, la información del grupo se asume como proporcionada por el usuario, es decir utilizan los grupos como fuente adicional de información para eliminar la ambigüedad (**Ulges; Worrying; Breuel, 2011**).

Otros autores consideran que los resultados obtenidos de *Flickr* son muy “ruidosos” es decir, se obtienen muchas imágenes que tienen poco o nada que ver con la imagen anotada. Este fenómeno se puede observar en la figura 3, donde se ilustran los resultados de la búsqueda con la palabra “perro”: entre los primeros 36 resultados obtenemos 4 que no corresponden (11%).

Por ello, cuando se utiliza *Flickr* para construir una base de datos de entrenamiento para la anotación de imágenes, es necesario hacer un filtrado previo, aplicando la máxima de que una imagen es relevante para una cierta etiqueta cuando la etiqueta describa el contenido de una o más regiones de la imagen (**Tang et al., 2011**).

Contextualización en base a redes sociales

Otras iniciativas utilizan las redes sociales (**Elhai; Karlsen; Akselsen, 2009**) para adquirir los metadatos contextuales sociales. Las redes sociales son una de las plataformas más utilizadas por las comunidades online para compartir texto, imágenes y vídeos. Se basan en perfiles de usuarios que ofrecen una descripción de cada miembro. Además de las imágenes subidas por un usuario, su perfil contiene comentarios y opiniones positivas/negativas sobre las mismas.

Un trabajo de análisis del contexto más complejo es el llevado a cabo por **Elhai, Karlsen y Younas (2011)**. En él se propone un sistema que genera de forma semi-automática anotaciones de imágenes sobre la ontología *OntoCAIM*, considerando el contexto de una red social y haciendo uso de las anotaciones manuales de imágenes proporcionadas por el usuario más activo. Para ello, infieren a este usuario mediante *social network analysis* y se toman sus anotaciones como base para las imágenes objetivo.

La ontología *OntoCAIM* reutiliza las ontologías *FOAF* (representa el perfil del usuario de la red social), *SIOC* (representa comunidades online), *EXIF* (representa metadatos básicos de fotografías) y *WordNet* (ayuda en la desambiguación del lenguaje natural), y por lo tanto describe la red social y las imágenes. Como se ha comentado, este sistema gira entor-

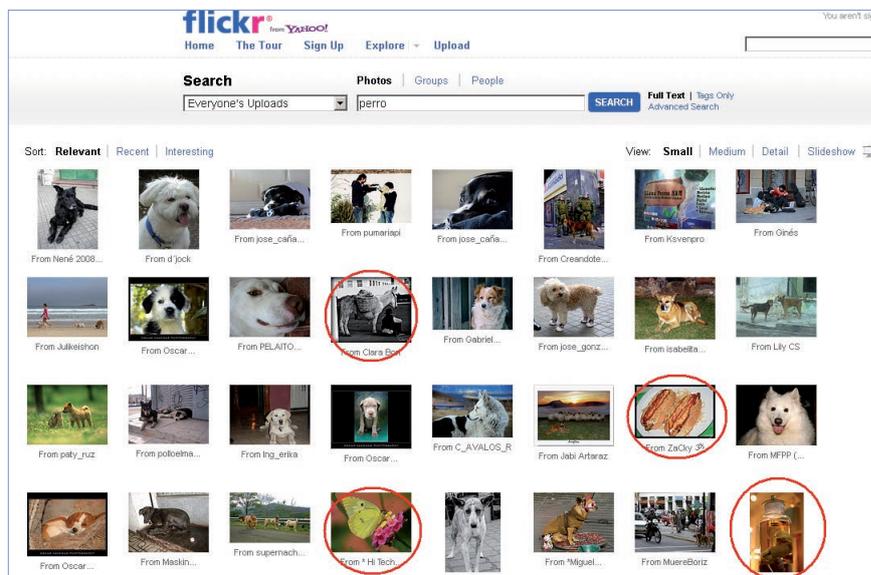


Figura 3. Resultados en *Flickr* para búsqueda con la palabra “perro”

no al contexto social de una imagen que estaría formado entre otros por la geo-referencia, fecha y hora y la granularidad de las relaciones entre actores de la red social:

- para obtener el actor central o correcto utilizan técnicas de SNA (*social network analysis*);
- la fecha y hora es utilizada porque se parte de la hipótesis de que fotografías que hayan sido sacadas en un corto espacio de tiempo, seguramente sean muy similares porque la cámara no puede haber estado en localizaciones muy distantes en el umbral de 5 minutos que plantean inicialmente;
- la geo-referencia se utiliza para agrupar fotografías tomadas dentro de un radio concreto, siempre y cuando contengan en sus metadatos las coordenadas GPS.

Otra opción de utilización de las redes sociales como ayuda a la anotación es el reconocimiento facial para la anotación.

“ Otra opción de utilización de las redes sociales como ayuda a la anotación es el reconocimiento facial ”

En el trabajo de **Stone, Zickler y Darrel (2008)** se propone mejorar los algoritmos de reconocimiento utilizando las redes sociales. Así, una imagen facial puede no ser reconocida por diversos motivos (iluminación, maquillaje). En este caso, se utilizan imágenes parecidas anotadas de forma manual en las redes sociales y se analizan las personas que aparecen en la misma imagen, identificando grupos de amigos. De esta forma, la tasa de anotaciones automáticas de personas aumenta considerablemente gracias al uso del contexto social.

3. Sistemas de búsqueda de imágenes actuales

Se ha visto cómo es posible anotar de forma automática el contenido de una imagen. El objetivo final de esta anotación es permitir a los motores de búsqueda la recuperación de imágenes de forma inteligente. Existen numerosos intentos por crear buenos buscadores de imágenes basados en su

Buscador de imágenes	Referencia web
TinEye	http://www.tineye.com
Cydral	http://www.cydral.com
Quintura (y versión Kids)	http://www.quintura.com
Google similar images	http://images.google.com
Ithaki	http://www.ithaki.net
LTU	http://www.ltutech.com
Pixlogic	http://www.pixlogic.com

Tabla 2. Buscadores de imágenes

contenido. Algunos de los más avanzados se pueden encontrar en la tabla 2.

Hay que destacar el buscador *Google images* en su versión de búsqueda de imágenes similares (opción a la derecha de una imagen hallada). Tras realizar numerosas pruebas, se ha comprobado que es capaz de realizar búsquedas de muy distintos tipos dependiendo de la imagen subida. Por ejemplo, si se le indica la imagen frontal de un coche, puede reconocer la marca y modelo, así como su color. En cambio, si se le muestra el mismo coche en una vista lateral, sólo intuye que es un vehículo y devuelve imágenes genéricas de vehículos. Esto demuestra que los buscadores comienzan a anotar los contenidos para su búsqueda, pero el resultado no es del todo correcto, ya que existen casos en los que se buscan imágenes de un determinado elemento y aparecen imágenes de otros que no esperamos. Este fenómeno está relacionado con la experiencia de uso, que no es satisfactoria porque el usuario espera otra información.

Por ello, como reto futuro en el campo de la recuperación inteligente de imágenes se puede destacar la necesidad de avanzar en la elaboración de herramientas de medición de la satisfacción de la experiencia de usuario, considerando sus expectativas y estado emocional. El resultado de estas mediciones debería influir en el proceso de búsqueda y presentación de los resultados. Así mismo, la anotación de imágenes posee un factor de medición subjetivo, es decir, esperamos que la anotación sea tan buena como aquella que realizaría un ser humano. Por ello, esta subjetividad influye de forma determinante en la recuperación de información: ésta será mejor cuanto mejor sea la anotación. Pero si además se une a la evolución de la capacidad de medición de la calidad de la experiencia -aprendiendo qué es lo que cada usuario espera en cada ocasión- la recuperación inteligente de información daría un paso de gigante.

Existen numerosos intentos por crear buenos buscadores de imágenes basados en su contenido

4. Conclusiones

La búsqueda de imágenes es uno de los retos del futuro de internet. En la actualidad se están proponiendo modelos de anotación automática de imágenes en base a su contenido.

Se ha querido destacar que aunque el campo del análisis de imagen está aún en evolución, va alcanzando unas tasas de

anotación efectiva muy altas. En un futuro las anotaciones según el contenido serán mucho más eficientes.

Por otra parte, se ha descrito cómo las imágenes en la Web están rodeadas de información de contexto, útil para ayudar en el proceso de anotación de las imágenes. Las redes sociales, los sitios para compartir fotografías, etc., son fuentes de dicho contexto y sirven, por un lado para construir nuevas y mejores bases de entrenamiento, y por otro para mejorar la precisión, problemas clave en todo proceso de anotación de imágenes.

Como reto futuro es necesario avanzar en la medición de la satisfacción de la experiencia de usuario, considerando sus expectativas y estado emocional

5. Agradecimientos

Este trabajo ha sido elaborado dentro del proyecto *Buscamedia* de programa *Cenit-E*, y ha sido subvencionado parcialmente por el CDTI (*Centro para el Desarrollo Tecnológico e Industrial*), dependiente del *Ministerio de Ciencia e Innovación*. Los autores también agradecen el trabajo a todos los socios del proyecto.

<http://www.cenitbuscamedia.es>

6. Bibliografía

Csurka, Gabriella; Dance, Christopher; Fan, Lixin; Willamowski, Jutta; Bray, Cédric. "Visual categorization with bags of keypoints". En: *Workshop on statistical learning in computer vision, ECCV, 2004*.

<http://www.cs.cmu.edu/~efros/courses/AP06/Papers/csurka-eccv-04.pdf>

Deselaers, Thomas; Ferrari, Vittorio. "Visual and semantic similarity in imageNet". En: *IEEE Computer vision and pattern recognition (CVPR) conf., 2011*.

http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/es//pubs/archive/37065.pdf

Dicarlo, James. "A strategy for understanding how the brain accomplishes object recognition". En: Dickinson, Sven J.; Leonardis, Ales; Schiele, Bernt; Tarr, Michael J. *Object categorization*. Cambridge University Press, 2009. ISBN 9780511635465

<http://dx.doi.org/10.1017/CBO9780511635465>

Elai, Najeeb; Karlsen, Randi; Akselsen, Sigmund. "A context centric approach for semantic image annotation and retrieval". *Future computing, service computation, cognitive, adaptive, content, patterns, 2009. Computation world'09*. Computation World: IEEE, pp. 665-668.

<http://dx.doi.org/10.1109/ComputationWorld.2009.30>

Elahi, Najeeb; Karlsen, Randi; Younas, Wagas. "Image annotation by leveraging the social context". En: *ACM Icuimc 2011, The 5th Intl conf on ubiquitous information management and communication, 2011*. Seoul, South Korea, 21-23 Febr. ISBN 978 1450305716

<http://dx.doi.org/10.1145/1968613.1968662>

- Escalante, Hugo-Jair; Hernández, Carlos A.; González, Jesús A.; López-López, Aurelio; Montes, Manuel; Morales, Eduardo; Sucar, L. Enrique; Villaseñor, Luis; Grubinger, Michael.** "The segmented and annotated IAPR TC-12 benchmark". *Computer vision and image understanding*, 2010, n. 4, pp. 419-428.
<http://ccc.inaoep.mx/~emorales/Papers/2010/hugo.pdf>
<http://dx.doi.org/10.1016/j.cviu.2009.03.008>
- Everingham, Mark; Van-Gool, Luc; Williams, Christopher K.; Winn, John; Zisserman, Andrew.** "The Pascal visual object classes (VOC) challenge". *Intl journal of computer vision*, 2010, v. 88, n. 2, pp. 303-338.
http://research.microsoft.com/pubs/102944/PascalVOC_IJCV2009.pdf
<http://dx.doi.org/10.1007/s11263-009-0275->
- Facebook. *Facebook Tweet*, 2011.
<http://twitter.com/#!/facebook/status/22372857292005376>
- Fischler, Martin A.; Elschlager, Robert A.** "The representation and matching of pictorial structures". *IEEE transactions on computers*, 1973, v. 22, n. 1, pp. 67-92.
<http://dx.doi.org/10.1109/T-C.1973.223602>
- Gao, Shenghua; Chia, Liang-Tien; Cheng, Xiangang.** "Web image concept annotation with better understanding of tags and visual features". *Journal of visual communication and image representation*, 2010, v. 21, n. 8, pp. 806-814.
<http://202.114.89.42/resource/pdf/5546.pdf>
<http://dx.doi.org/10.1016/j.jvcir.2010.08.005>
- García-Serrano, Ana.** "UNED-UV experiments using multimodal information approaches". *Image CLEF*, 2011, Amsterdam.
- Hao, Su; Sun, Min; Fei-Fei, Li; Savarese, Silvio.** "Learning a dense multi-view representation for detection, viewpoint classification and synthesis of object categories". En: *IEEE 12th Intl conf on computer vision*, 2009, pp. 213-220.
http://www.eecs.umich.edu/vision/papers/SuSunLiSavarese_ICCV2009.pdf
- Huiskes, Mark J.; Lew, Michael S.** "The MIR Flickr retrieval evaluation". En: *Proceedings of the 1st ACM Intl conf on multimedia information retrieval (MIR)*, 2008, pp. 39-43. New York: ACM.
<http://press.liacs.nl/mirflickr/mirflickr.pdf>
<http://doi.acm.org/10.1145/1460096.1460104>
- Lee, Yong-Jae; Grauman, Kristen.** "Object-graphs for context-aware category discovery". En: *IEEE Conf on computer vision and pattern recognition (CVPR)*, 2010, pp. 1-8.
<ftp://ftp.cs.utexas.edu/pub/AI-Lab/tech-reports/AI-14.pdf>
- Llorente, Ainhoa; Manmatha, Raghavan; Rüger, Stefan M.** "Image retrieval using Markov random fields and global image features". *Proc. of the ACM Intl conf on image and video retrieval*, 2010, pp. 243-250.
<http://oro.open.ac.uk/23503/1/p243-llorente.pdf>
<http://doi.acm.org/10.1145/1816041.1816078>
- Makadia, Ameesh; Pavlovic, Vladimir; Kumar, Sanjiv.** "Baselines for image annotation". *Intl journal of computer vision*, 2010, v. 90, pp. 88-105.
http://www.sanjivk.com/ImageAnnotation_IJCV10.pdf
<http://doi.acm.org/10.1007/s11263-010-0338-6>
- Serre, Thomas; Wolf, Lior; Bileschi, Stanley; Riesenhuber, Maximilian; Poggio, Tomaso.** "Robust object recognition with cortex-like mechanisms". *IEEE transactions on pattern analysis and machine intelligence*, 2007, v. 29, n. 3, pp. 411-426.
<http://cbcl.mit.edu/publications/ps/serre-wolf-poggio-PAMI-07.pdf>
<http://doi.acm.org/10.1109/TPAMI.2007.56>
- Stone, Zak; Zickler, Todd; Darrel, Trevor.** "Autotagging Facebook: social network context improves photo annotation", *IEEE Computer Society Conf on computer vision and pattern recognition workshops*, 2008.
http://www.eecs.harvard.edu/~zickler/papers/Autotag_IVW2008.pdf
<http://dx.doi.org/10.1109/CVPRW.2008.4562956>
- Tang, Jinhui; Yang, Shuicheng; Chua, Tat-Seng; Jain, Ramesh.** "Label-specific training set construction from web resource for image annotation". *CoRR*, 2011.
<http://arxiv.org/pdf/1107.2859v1>
- Torresani, Lorenzo; Szummer, Martin; Fitzgibbon, Andrew.** "Efficient object category recognition using classemes". En: *Proc of the 11th European conf on computer vision: Part I*, 2010, pp. 776-789. ISBN 978 3642155482
<http://dl.acm.org/citation.cfm?id=1886063.1886122>
- Tousch, Anne-Marie; Herbin, Stéphane; Audibert, Jean-Yves.** "Semantic hierarchies for image annotation: a survey". *Pattern recognition*, 2011, v. 45, pp. 333-345.
<http://certis.enpc.fr/~audibert/Mes%20articles/PR11.pdf>
<http://dx.doi.org/10.1016/j.patcog.2011.05.01>
- Ulges, Adrian; Worring, Marcel; Breuel, Thomas.** "Learning visual contexts for image annotation from Flickr groups". *IEEE transactions on multimedia*, 2011, v. 13.
http://www.science.uva.nl/research/publications/2011/UlgesITM2011/tmm_final.pdf
<http://dx.doi.org/10.1109/TMM.2010.2101051>
- Verbeek, Jakob; Guillaumin, Matthieu; Mensink, Thomas; Schmid, Cordelia.** "Image annotation with TagProp on the MIR Flickr set". En: *11th ACM Multimedia information retrieval*, 2010, pp. 537-546.
<http://hal.inria.fr/docs/00/60/63/92/PDF/verbeek10mir.pdf>
<http://dx.doi.org/10.1145/1743384.174347>
- Weston, Jason; Bengio, Samy; Usunier, Nicolas.** "Wsabie: scaling up to large vocabulary image annotation". En: *Proc of the Intl joint conf on artificial intelligence, IJCAI*, 2011, pp. 2764-2770.
<http://www.thespermwhale.com/jaseweston/papers/wsabie-ijcai.pdf>
- Youtube Inc. "Thanks, YouTube community, for two big gifts on our sixth birthday!" *Broadcasting ourselves. The official Youtube blog*, 25 mayo 2011.
<http://youtube-global.blogspot.com/2011/05/thanks-youtube-community-for-two-big.html>



FUTURO DE LAS REVISTAS CIENTÍFICAS DE COMUNICACIÓN EN ESPAÑA



Antonio J. Baladrón-Pazos y Beatriz Correyero-Ruiz



Antonio J. Baladrón-Pazos, doctor en publicidad y relaciones públicas y profesor contratado doctor de esta especialidad en el *Departamento de Ciencias de la Comunicación I* de la *Univ. Rey Juan Carlos (URJC)* de Madrid, es licenciado en periodismo y en publicidad y relaciones públicas. Ha sido docente en la *Univ. Pontificia de Salamanca*, la *Univ. Católica San Antonio* y la *Univ. de Málaga*. Ha participado en actividades académicas y como profesor invitado en diferentes universidades extranjeras: *Univ. de Coimbra* (Portugal), *Univ. degli Studi di Sassari* (Italia), *Wroclaw Univ.* (Polonia), *Univ. de Antioquia* (Colombia), *Pontificia Univ. Católica de São Paulo* (Brasil) y *Univ. Autónoma de Yucatán* (México). Participa en varios proyectos de I+D financiados por el *Ministerio de Ciencia e Innovación* y por la *Comunidad de Madrid*.

Universidad Rey Juan Carlos
Edif. Departamental I, desp. 035
Camino del molino, s/n. 28943 Fuenlabrada (Madrid), España
antonio.baladron@urjc.es



Beatriz Correyero-Ruiz es doctora en periodismo por la *Universidad Complutense de Madrid*. Es profesora contratada doctora en la *Universidad Católica San Antonio de Murcia (UCAM)*, donde ejerce como subdirectora del *Departamento de Periodismo*. Ha impartido las asignaturas de información en radio, periodismo digital y producción periodística. Es además, asesora de la revista *Sphera publica* y miembro del grupo de investigación en *Comunicación y Menores* de la *UCAM*.

Universidad Católica San Antonio
Facultad de Comunicación
Campus de los Jerónimos, s/n. 30107 Guadalupe (Murcia), España
bcorreyero@pdi.ucam.edu

Resumen

Se presentan los principales resultados de un estudio cualitativo llevado a cabo entre noviembre de 2010 y febrero de 2011 sobre el futuro de las revistas científicas de comunicación en España. La técnica de investigación utilizada fue el método *Delphi* y se contó con la participación de una muestra significativa de los directores de las principales revistas académicas de comunicación. Las conclusiones revelan las tendencias de futuro más relevantes para estas publicaciones, que en los últimos años han experimentado una gran expansión, tanto en número como en influencia, debido al auge de la investigación comunicológica en España.

Palabras clave

Revistas científicas, Investigación en comunicación, España, Futuro, Tendencias

Title: Future of communication journals in Spain

Abstract

The results of a qualitative investigation carried out between November 2010 and February 2011, on the future of the scientific journals specializing in media and communication in Spain are summarized. The research methodology used was the Delphi method based on the answers provided by the editors of the most important journals in this field. The main purpose was to provide a descriptive analysis of the current situation, along with a prospective analysis to attempt to foresee the future development of these publications.

Keywords

Scientific journals, Media and communication, Research, Communication research, Future, Spain, Trends.

Baladrón-Pazos, Antonio J.; Correyero-Ruiz, Beatriz. "Futuro de las revistas científicas de comunicación en España". *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 34-42.

<http://dx.doi.org/10.3145/epi.2012.ene.05>

Artículo recibido el 12-04-11
Aceptación definitiva: 22-12-11

1. Introducción

El objeto de estudio de este artículo son las revistas científicas de comunicación españolas (en adelante RCCE), que viven una situación de expansión debido al auge de la investigación en comunicación en España. Ese auge comenzó en los años ochenta y fue en los noventa cuando alcanzó su “desarrollo o maduración” (Martínez-Nicolás, 2009). Muestra de la expansión de las RCCE es que entre 2007 y 2008 el número de artículos publicados creció casi un 37% (Fernández-Quijada, 2010, p. 558) y ese número sigue incrementándose año tras año, ya que en 2007 eran 289 los artículos publicados, en 2008 supusieron un total de 395 y en 2010 fueron ya 495 (Fernández-Quijada, 2011, pp. 3-4). Paralelamente se está produciendo una creciente internacionalización de la investigación en comunicación, incrementándose cada vez más el número de artículos de autores españoles en las revistas de comunicación del *Social sciences citation index (SSCI)*; entre 2005 y 2009 la producción española se incrementó en un 255% respecto al período 2000-2004, convirtiéndose España en la cuarta nación europea más productiva en ese área (Masip, 2011, pp. 12-13).

La comunicación se suma, aunque con retraso, al *boom* de la comunicación científica de los años setenta, a menudo sustentando los nuevos proyectos en el libre acceso porque puede ayudar a potenciar el impacto (Abadal; Rius, 2005, p. 8); cuando menos, incrementa la visibilidad, con la influencia que esto tiene en ocasiones sobre el impacto. Es ahora cuando la relación entre publicaciones e investigación se refuerza con la que Perceval y Fornieles-Alcaraz califican como otra “novedosa” relación (2008, p. 218): la de las publicaciones con las acreditaciones, que actúa sobre la demanda de las RCCE. En este contexto, dado que “la valoración de la calidad exclusivamente a través del impacto es incorrecta” (Aleixandre-Benavent; Valderrama-Zurián; González-Alcaide, 2007, p. 10), se utilizan también otros indicadores alternativos, si bien los índices de impacto están siendo la referencia primaria para los comunicólogos interesados en publicar sus trabajos en las revistas científicas del área y por ende un factor determinante en la demanda de unas cabeceras en detrimento de otras.

Por todo ello, cada vez son más los trabajos sobre este tema, algunos ya citados. Pionero fue el de Giménez-Toledo y Alcaín-Partearroyo (2006), que aportó una primigenia jerarquización de las revistas de periodismo basándose en indicadores de calidad. También de Giménez-Toledo (2011) cabe destacar un reciente estudio basado en una encuesta a expertos en comunicación acerca de la valoración que hacen de las RCCE. Asimismo, hay que referir los estudios bibliométricos de Fernández-Quijada (2010, 2011), que trazan un perfil de las RCCE, Masip (2011), Castillo y Carreón (2010) o Martínez-Nicolás y Saperas-Lapiedra (2011), entre otros. A ellos se suman otros ensayísticos, como los descriptivos de Jones (2003, 2005), fundamentales para un mejor conocimiento de las RCCE. Otros trabajos ensayísticos están más cargados de opinión o son abiertamente críticos (Sabés; Perceval, 2009). Finalmente, existen estudios sobre cabeceras específicas como los de Colle (2009) o López-Ornelas (2010), o bien sobre aspectos particulares de las RCCE, por ejemplo ligados a su naturaleza digital (Martí, 2004; De-

Pablos, 2010). En general, estos trabajos suelen abordar el presente o el pasado de las RCCE y en menor medida el futuro. Igualmente sucede en los estudios sobre el tema realizados en otros países, entre los que no existen precedentes de nuestra investigación en cuanto a los objetivos planteados y la metodología utilizada. Se han publicado trabajos sobre diferentes aspectos de estas revistas científicas como los de Lauf (2005), Feeley (2008) o Espanha y Lima (2011) e incluso en ocasiones se ha utilizado una metodología similar basada en la consulta a expertos (Stumpf, 2008), pero en general están centrados en objetos de estudio más particulares.

La comunicación se suma, con retraso, al *boom* de la comunicación científica, a menudo sustentando los nuevos proyectos editoriales en políticas de libre acceso

2. Objetivos y metodología

El objetivo de nuestra investigación era llevar a cabo de forma científica una prospectiva de las RCCE para los próximos diez años y por ello optamos por el método *Delphi*, adecuado para prever evoluciones futuras de situaciones cambiantes como la de las publicaciones científicas de comunicación. Es una técnica de previsión subjetiva basada en un proceso iterativo y sistemático en el que un panel de expertos es interrogado sucesivamente con el fin de obtener un consenso de opinión o una opinión grupal fidedigna.

Para seleccionar los expertos se siguieron estos criterios: nivel de conocimientos, grado de motivación, capacidad predictiva y que fuesen especialistas en el tema y estuviesen activamente implicados en él. Se optó por los directivos de las principales RCCE, primer paso que se debía dar para abordar el objeto de estudio, si bien en otras fases actuales del proyecto se está contando con otros grupos de especialistas. Se contactó con las revistas de comunicación presentes en *In-Recs* (en 2010 eran 21) por ser uno de los principales índices de calidad relativa de las revistas de ciencias sociales. Aunque los índices internacionales tienen mayor relevancia, al estudiarse las cabeceras editadas en España *In-Recs* era una opción más idónea, además de porque es una referencia inexcusable para las agencias de evaluación y para los investigadores en comunicación. La iniciativa tuvo una excelente acogida; los directores que confirmaron su participación inicial fueron: Ana Azurmendi (*Comunicación y sociedad*), Felicísimo Valbuena (*CIC*), Francesc Canosa (*Trípodos*), Francisco García (*Icono14*), Humberto Martínez-Fresneda (*Comunicación y hombre*), José-Antonio Ortega (*Etic@net*), J. Ignacio Aguaded (*Comunicar*), José-Manuel De-Pablos (*Latina*), Juan Rey (*Questiones publicitarias*), Luis Núñez (*Doxa*), Begoña Zabildea (*Zer*), María-del-Mar Rodríguez (*Sphera pública*), María-Jesús Casals (*Estudios sobre el mensaje periodístico*), Miguel Vázquez (*I/C*) y Ramón Reig (*Ámbitos*). Concluyeron el proceso, en febrero de 2011, trece expertos y el índice de abandono fue del 13,3%; en los trabajos publicados se suele situar por encima del 20% (Lan-

leta, 2002, p. 122), de modo que no implicó distorsión de los resultados finales.

La principal misión de las revistas de comunicación es ayudar a la acreditación de los investigadores, lo que da cuenta del alcance del llamado efecto *Aneca*

Se realizaron dos rondas de consultas. En la primera el cuestionario estuvo conformado por preguntas abiertas (tabla 1), dando libertad de respuesta a los expertos. A partir de los datos obtenidos y como corresponde a la ortodoxia del método *Delphi* se extrajeron los ítems de respuesta del segundo cuestionario, compuesto por preguntas cerradas fruto de la agregación de las opiniones de los expertos y sobre las que se solicitaron valuaciones en una escala 0-5. No se hizo una tercera ronda por el riesgo de abandono de los expertos y porque científicamente no es necesaria (Landeta, 2002, p. 102). Por último, se sometieron las respuestas recibidas a tratamiento estadístico; al ser valuaciones, se calculó la mediana como medida de tendencia central de los valores de cada ítem de respuesta y éstos se ordenaron de mayor a menor consenso: de mediana superior a inferior y a igual mediana, de menor a mayor desviación típica. De acuerdo con la filosofía del método *Delphi*, no se ofrecen resultados desagregados por grupos de expertos dado que el propósito era obtener una opinión grupal fidedigna; en todo caso, los niveles de desviación típica en cada ítem dan cuenta del grado de consenso o disenso entre los expertos respecto a cada uno de esos ítems.

3. Resultados

En la tabla 2 se recoge la opinión agregada de los expertos sobre la misión que cumplirán las RCCE. Consideran que su papel primordial será ayudar a la acreditación de los investigadores, dato sin duda sorprendente que da cuenta del

Cuestionario Delphi-1	
1.	En su opinión ¿cuál es el papel o misión que las RCCE desempeñarán en los próximos diez años en el contexto del <i>Espacio Europeo de Educación Superior</i> ? Señale al menos 3 aspectos.
2.	Enumere los 5 principales cambios que en los próximos diez años afrontarán las RCCE en su organización interna.
3.	¿Cuáles son los principales servicios que ofrecerán en los próximos diez años las ediciones electrónicas de las RCCE? Indique los 5 que considere más importantes.
4.	Señale 5 de las principales medidas que en los próximos diez años adoptarán las RCCE para promover su calidad científica.
5.	Enumere 5 de las principales medidas que en los próximos diez años adoptarán las RCCE para promover la consulta y citación de las mismas entre la comunidad científica.
6.	Indique cuáles serán los 5 principales criterios para medir el impacto de las RCCE que en los próximos diez años utilizarán las entidades encargadas de esa medición.
7.	¿Cree que en los próximos diez años surgirán más revistas especializadas en comunicación en España? Enumere al menos 3 razones para justificar su opinión.

Tabla 1

alcance del llamado efecto *Aneca*. Difundir la investigación española en comunicación y promover su internacionalización son otras cuestiones muy valoradas. Con puntuación algo menor los participantes señalan el papel de las RCCE para mejorar la investigación o incrementar el conocimiento en España de las investigaciones comunicológicas que se hacen en otros países. Ser entornos de interacción para investigadores o fuentes de documentación para estudiantes son otras misiones consideradas importantes, si bien los grados de disenso son más altos. En el resto de ítems las desviaciones típicas también son elevadas, evidenciándose la disparidad de opiniones.

Sobre los cambios que afrontarán las RCCE en su organización interna (tabla 3), destaca por su alto consenso la implementación de políticas de acceso libre; todos los expertos apuntan a ello, al igual que dan suma importancia a la sustitución del papel por lo digital –aspecto relacionado– y a la internacionalización de los consejos editoriales. Con similar puntuación, aunque menor consenso, señalan otros aspectos vinculados (reducción de gastos de producción, presencia en redes sociales, mejora de la infraestructura para soportar buenas ediciones electrónicas), así como la necesidad de dedicar más esfuerzos a la inclusión en índices de impacto. En segundo plano, otros cambios son: publicación multilingüe, inclusión de contenidos multimedia, gestión online, universalización de la evaluación por pares o incremento en la información a autores.

Como se preveía que la digitalización era una tendencia en las RCCE, los expertos fueron preguntados por los servicios que ofrecerán las ediciones electrónicas de estas publicaciones. Destacan la búsqueda avanzada en los contenidos y el acceso a los textos completos o la imbricación en la comunicación 2.0. Otras cuestiones sobre las que el consenso

Misión en el EEES				
Puesto	Ítem	Mediana	Media	Desv.
1	Ayudar a la acreditación de los investigadores	5	4,54	0,50
2	Difundir la investigación y el conocimiento en comunicación	5	4,46	0,63
3	Internacionalizar la investigación española en comunicación	5	4,31	0,91
4	Incrementar el conocimiento de la investigación comunicológica extranjera	4	3,69	0,72
5	Contribuir a la innovación y mejora de la investigación en comunicación	4	3,92	0,83
6	Ser entorno de interacción para investigadores	4	4,00	1,04
7	Ser fuentes de documentación para estudiantes universitarios	4	3,69	1,14
8	Ampliar la cultura editorial científica entre los investigadores	4	3,77	1,37
9	Contribuir a la innovación y mejora de la docencia universitaria	3	3,31	1,07
10	Incentivar la competitividad entre grupos de investigación	3	3,54	1,15
11	Intensificar las relaciones entre universidades y empresas	3	2,54	1,45
12	Ser foro de debate sobre el EEES	3	2,46	1,55

Tabla 2

alcanzado es importante son: servicios de acceso a bases de datos, repositorios externos o revistas científicas, versiones en inglés y mayor información sobre sí mismas y su impacto entre los investigadores. En la tabla 4 se detallan otros servicios de interés, como los relativos a la referida gestión y evaluación online.

En cuanto a las medidas que adoptarán las RCCE para promover su calidad científica (tabla 5) sobresale, por la puntuación y el consenso obtenido, lo referente a la colabora-

Cambios en organización interna				
Puesto	Ítem	Mediana	Media	Desv.
1	Acceso abierto y a texto completo en internet	5	4,85	0,36
2	Ampliación e internacionalización de los consejos editoriales	5	4,46	0,63
2	Progresiva sustitución del papel por ediciones electrónicas	5	4,54	0,63
3	Reducción de gastos de producción	5	4,38	0,74
4	Presencia en redes sociales	5	4,54	0,84
5	Mejora de la infraestructura técnica para soportar buenas ediciones electrónicas	5	4,38	0,92
6	Mayor esfuerzo para la inclusión en índices de impacto	5	4,23	1,05
7	Publicación en varios idiomas, fundamentalmente inglés	4	4,31	0,82
8	Inclusión de contenidos multimedia	4	4,08	0,83
9	Realización online de los procesos de gestión y edición	4	4,23	0,89
9	Universalización del sistema de evaluación externa, ciega y por pares	4	4,23	0,89
10	Más transparencia e información a los autores en la evaluación	4	4,00	0,96
11	Mejoras en la promoción y difusión	4	3,85	1,03
11	Optimización de los equipos de evaluadores	4	3,85	1,03
12	Continuidad o actualización constante frente a periodicidad fija	4	4,00	1,04
13	Mejoras en el diseño	4	3,77	1,05
14	Mayor exigencia de calidad de los artículos	4	3,69	1,14
15	Búsqueda de nuevas fuentes de financiación	4	3,62	1,15
16	Mayor profesionalización de los equipos humanos, principalmente en nuevas tecnologías o índices de impacto	4	3,69	1,20
17	Mayor garantía del anonimato en las evaluaciones	4	3,46	1,39
18	Consolidación de acuerdos entre revistas	4	3,69	1,49
19	Renuncia al anonimato de los evaluadores	3	3,38	1,27
20	Mayor preocupación por los derechos de autor	3	2,92	1,33

Tabla 3

ción de las revistas en los sistemas de medición del impacto mediante proyectos conjuntos y al mayor cumplimiento de los requisitos para su indexación. Si lo segundo tiene más que ver con la implementación por parte de cada revista de las medidas para cumplir con los requisitos que los sistemas de medición del impacto exigen, lo primero hace referencia al establecimiento de cauces de colaboración entre las distintas cabeceras encaminados a coordinar colectivamente sus políticas de gestión editorial armonizándolas con las contempladas por esos sistemas de medición del impacto; de hecho, ya existen iniciativas conjuntas de las que forman parte las principales revistas y con las que se pretende una mayor fuerza colectiva a la hora de conseguir un mayor peso de las RCCE.

Por otra parte, el impulso a las ediciones electrónicas y al acceso abierto son, una vez más, otras dos medidas fundamentales para los expertos, así como el cumplimiento escrupuloso de los procedimientos de evaluación. Muy importantes, aunque con menor consenso, son: la utilización de criterios más objetivos y claros por parte de las propias revistas en los procesos de evaluación de los artículos, la potenciación de investigación innovadora y aplicada, la lucha contra lo que consideran prácticas corruptas de algunas cabeceras para favorecer sus índices de impacto, prácticas que algunos de los expertos participantes ejemplifican con

Servicios de ediciones electrónicas				
Puesto	Ítem	Mediana	Media	Desv.
1	Búsqueda avanzada en los contenidos	5	4,31	0,91
2	Acceso a texto completo	5	4,46	1,08
3	Acceso a redes sociales	5	4,23	1,19
4	Acceso a bases de datos o repositorios externos	4	4,23	0,70
5	Acceso a otras revistas científicas	4	4,23	0,80
6	Versiones en varios idiomas, fundamentalmente inglés	4	4,23	0,89
7	Exhaustiva información sobre la revista (impacto, normas, etc.)	4	4,15	0,95
8	Consulta en línea del estado de los originales sometidos a evaluación	4	3,94	1,00
9	Foros para la interacción entre investigadores	4	3,77	1,05
10	Envío y autoedición online de los artículos	4	4,08	1,07
11	Incorporación de nuevos formatos en los artículos	4	3,77	1,12
12	Mayor interacción con directivos o evaluadores	4	3,85	1,23
13	Informaciones útiles para investigadores (congresos, <i>call for papers</i> , links de interés, etc.)	4	3,77	1,25
13	Continuidad y actualización constante frente a periodicidad fija	4	3,77	1,25
14	Información personalizada a cada lector	4	3,46	1,39
15	Más números monográficos	4	3,23	1,53
16	Más secciones temáticas	4	3,23	1,58

Tabla 4

Medidas para promover la calidad científica				
Puesto	Ítem	Mediana	Media	Desv.
1	Impulsar las ediciones electrónicas	5	4,62	0,62
2	Cumplir con puntualidad escrupulosa los procedimientos	5	4,46	0,63
2	Cumplir los requisitos para estar indexadas en índices de referencia	5	4,54	0,63
3	Colaborar en los sistemas de medición del impacto	5	4,38	0,74
4	Mejorar la accesibilidad a los contenidos en internet	5	4,31	0,91
5	Utilizar criterios de evaluación de artículos más objetivos, claros y públicos	5	3,92	1,33
6	Potenciar la investigación innovadora aplicada y de interés social	5	3,92	1,49
7	Luchar contra prácticas corruptas de las propias cabeceras para entrar y mantenerse en índices de impacto	5	4,00	1,62
8	Fomentar los servicios de búsqueda avanzada en los contenidos	4	4,15	0,86
9	Abrir espacios de debate para los investigadores	4	4,08	0,92
9	Afrontar la profesionalización de los procesos de producción (redacción, diseño) dejando los contenidos a los comunicólogos	4	4,08	0,92
10	Renovar la infraestructura técnica para soportar buenas ediciones electrónicas	4	4,15	0,95
11	Editar en inglés	4	3,69	0,99
12	Incrementar el número de revistas	4	3,69	1,07
12	Internacionalizar los <i>call for papers</i> y los colaboradores	4	3,69	1,07
12	Unificar las normas editoriales y de evaluación entre las distintas revistas	4	3,69	1,07
13	Exigir el cumplimiento absoluto de requisitos formales en las colaboraciones	4	3,85	1,17
13	Incluir documentación audiovisual en los artículos	4	3,85	1,17
14	Mejorar la transparencia e interactividad con los autores	4	4,00	1,18
15	Exigir que los artículos se basen en metodología científica	4	4,00	1,24
15	Mejorar el diseño y edición	4	4,00	1,24
16	Incrementar el nivel académico de los consejos editoriales	4	3,69	1,26
17	Dar mayor importancia a las reseñas de libros de calidad	4	3,69	1,32
18	Utilizar el sistema DOI	4	3,85	1,35
19	Valorar la actualización e internacionalización de las bibliografías	4	3,77	1,42
20	Extremar las exigencias en la evaluación	4	3,85	1,46
20	Universalizar la evaluación externa, ciega y por pares	4	3,85	1,46

Tabla 5

la promoción excesiva de la autocitación. Este último aspecto evidencia los recelos de algunos directivos consultados respecto al comportamiento de otras revistas distintas a la suya.

Otros aspectos ya señalados también son considerados relevantes para promover la calidad: servicios de búsqueda avanzada, espacios para la interacción entre los investigadores, profesionalización de los procesos productivos, mejora de las ediciones electrónicas o edición en inglés.

Medidas para promover citación				
Puesto	Ítem	Mediana	Media	Desv.
1	Optimizar la edición electrónica	5	4,62	0,62
2	Colaborar en los sistemas de medición del impacto	5	4,38	0,84
3	Mejorar el posicionamiento en buscadores	5	4,31	0,91
4	Intensificar la promoción	5	4,54	0,93
4	Incrementar el uso de comunicación 2.0	5	4,54	0,93
5	Luchar contra prácticas corruptas de las propias cabeceras para entrar y mantenerse en índices de impacto	5	4,46	1,39
6	Incorporarse a sistemas de indexación internacionales solventes	5	4,15	1,51
7	Editar en varios idiomas, fundamentalmente inglés	4	4,08	0,83
8	Enviar avisos de citas localizadas a los autores	4	3,92	1,07
8	Incorporarse a directorios del tipo <i>DOAJ</i>	4	4,08	1,07
9	Enviar información personalizada a los investigadores	4	3,54	1,15
10	Organizar y participar en encuentros sobre comunicación	4	3,69	1,2
11	Garantizar el rigor y la exigencia en la selección de artículos	4	4,08	1,21
11	Incorporar profesionales especializados en índices de calidad	4	3,92	1,21
12	Fomentar la innovación en la investigación en comunicación	4	4,00	1,24
13	Utilizar prescriptores o recomendadores de contenido	4	3,69	1,32
14	Disponer de metadatos del tipo <i>Dublin Core</i> y exponerlos a los lectores	4	3,62	1,33
15	Mejorar el nivel investigador de los autores	4	3,77	1,37
16	Promover asociaciones entre cabeceras	4	3,69	1,43
17	Incluir recomendaciones a artículos relacionados	4	3,38	1,44
18	Aumentar la periodicidad	4	3,31	1,64
19	Promover la citación de la revista entre los colaboradores	4	2,92	1,73
20	Intensificar la distribución física y los intercambios	3	2,77	2,04

Tabla 6

Dada la relación entre calidad e impacto, cuando los expertos fueron interrogados sobre medidas que adoptarán para promover la citación, parte de las respuestas más valoradas siguen la misma línea que las recogidas en la tabla 5: mejorar las ediciones electrónicas, colaborar en los sistemas de medición del impacto evitando prácticas corruptas por parte de las propias cabeceras, incorporarse a bases de datos internacionales, editar en inglés, etc. Pero también valoran otras como mejorar el posicionamiento web, dedicar más esfuerzos a la promoción o incorporarse a plataformas de comunicación 2.0; en definitiva, potenciar las políticas de marketing. Otras medidas a las que los participantes en el *Delphi* conceden cierta relevancia son las relativas al envío de información a los investigadores sobre los artículos publicados y las citas localizadas.

Respecto a los criterios que utilizarán las entidades encargadas de la medición del impacto de las RCCE llama la atención que la respuesta mejor valorada sea la referida a los índices de lectura o difusión de las revistas, aunque el elevado disenso permite relativizar este dato. Otros de los criterios apuntados tienen más que ver con los referentes

actuales: número y diversidad de citas, transparencia en las evaluaciones y evaluación por pares, nivel académico de los consejos editoriales, cumplimiento de criterios de calidad regulados, así como edición en inglés.

Finalmente se interrogó a los expertos sobre el nacimiento de nuevas cabeceras en los próximos diez años y la respuesta afirmativa fue unánime. Según indican, sucederá sobre todo para dar respuesta a la necesidad de publicar que tiene la creciente comunidad de comunicólogos. También afirman que influirán los bajos costes de las ediciones electrónicas o su utilidad para dar visibilidad a grupos de investigación emergentes. En la tabla 8 se recogen otras razones, muy vinculadas al auge de la investigación en comunicación que se vive en España.

Entre los cambios que afrontarán en los próximos diez años las RCCE, destacan la implementación de políticas de libre acceso o la sustitución del papel por lo digital

Criterios para medir impacto				
Puesto	Ítem	Mediana	Media	Desv.
1	Índices de lectura o difusión: consultas, descargas online, etc.	5	4,23	1,31
2	Variedad de las citas a la revista (en otras áreas, etc.)	4	4,08	0,83
3	Nivel de transparencia en los procesos de evaluación	4	4,15	0,86
4	Existencia de evaluadores externos, ciegos y por pares	4	4,23	0,89
5	Nivel investigador del consejo editorial	4	4,08	0,92
5	Número de citas en otras revistas	4	3,92	0,92
6	Presencia en la web 2.0: blogs, redes sociales	4	4,15	0,95
6	Rigor y cumplimiento de criterios de calidad regulados	4	4,15	0,95
7	Edición en varios idiomas, fundamentalmente inglés	4	3,77	0,97
8	Credibilidad e influencia entre la comunidad científica	4	4,08	1,00
8	Regularidad en la publicación	4	4,08	1,00
9	Número de citas en actas de congresos de excelencia	4	3,85	1,03
9	Número de citas en publicaciones de otros países o idiomas	4	3,85	1,03
10	Apertura editorial	4	4,00	1,04
11	Presencia en actos académicos: congresos, etc.	4	3,69	1,14
12	Nivel investigador de los autores	4	4,08	1,21
13	Fuentes mencionadas por los autores	4	3,77	1,31
14	Impacto social de las investigaciones publicadas	4	3,85	1,35
14	Originalidad de las contribuciones	4	3,85	1,35
15	Número de artículos publicados	3	3,31	1,20

Tabla 7

Conclusiones

De las opiniones de los expertos consultados se desprende cierta falta de madurez y profesionalización de las RCCE. De hecho, algunas de las tendencias de futuro señaladas son práctica habitual en las revistas científicas de otras áreas. Es el caso de la universalización del sistema de evaluación ciega por pares y de una mayor internacionalización de los consejos editoriales, además de ser necesarios en general mayores esfuerzos para la inclusión en índices de impacto. Puede explicarse por la propia juventud e inmadurez de la disciplina, caracterizada por una "insuficiencia metodológica generalizada" (Martínez-Nicolás y Saperas-Lapiedra, 2011, p. 124), y por la urgencia que los procesos de evaluación de

Razones para nuevas cabeceras				
Puesto	Ítem	Mediana	Media	Desv.
1	Para dar respuesta a la necesidad de publicar de los investigadores pendientes de acreditación	5	4,46	0,75
2	Por la disminución de costes de la edición electrónica	5	4,15	1,17
3	Porque son útiles para dar visibilidad a grupos de investigación	5	4,23	1,25
4	Por el incremento de centros e investigadores especializados en comunicación	4	3,85	1,03
5	Porque hay cabida para revistas especializadas en áreas de la comunicación poco representadas	4	3,92	1,27
6	Por la mayor valoración de la investigación comunicológica entre los estamentos científicos	4	3,77	1,48
7	Por la importancia de la comunicación en la sociedad	3	3,62	1,27
8	Por el incremento en el número de lectores	3	3,08	1,38

Tabla 8

la actividad investigadora imprimen a la evolución reciente de estas publicaciones.

Algunas cabeceras se han encontrado con un importante desconocimiento respecto al conjunto de medidas y dinámicas que adoptar para mejorar su impacto, por lo que sería recomendable un mayor asesoramiento de documentalistas y profesionales de la información. Si bien es incuestionable que la gestión de los editores para mantener niveles exigentes de calidad científica de los contenidos de sus revistas es la principal vía que puede servir para consolidar un mayor impacto de las mismas, lo cierto es que en los últimos años, por ese insuficiente conocimiento sobre el tema, muchas RCCE han tardado en incorporar a su gestión las necesarias medidas y requisitos para tener más posibilidades de un mayor impacto. Es una situación que, de hecho, se ha empezado a corregir en algunos casos en los que los editores han buscado la asesoría de expertos en el campo, por supuesto paralelamente a una adecuada gestión en el nivel de calidad científica de los contenidos.

El futuro pasará por optimizar los procesos de evaluación de los artículos, ya que las respuestas de los expertos evidencian que los procedimientos no son siempre suficientemente rigurosos. No en vano una de las medidas que proponen para promover la calidad científica es la utilización de criterios de evaluación de los artículos más objetivos, públicos y claros por parte de las propias revistas; si bien es cierto que esto es una realidad cada vez más constatable, dicha opinión podría estar relacionada con esas carencias que, según indican, persisten en algunas publicaciones en lo referente a la evaluación anónima y transparente ante los autores.

Hay cierta falta de madurez de las RCCE, ya que algunas de las tendencias de futuro señaladas ya son práctica habitual en las revistas científicas de otras áreas

También urge ampliar las posibilidades que ofrecen las ediciones electrónicas. Salvo en casos excepcionales, los expertos consideran que los servicios que ofrecen (búsqueda avanzada, información sobre la propia revista, gestión online, etc.) son muy deficientes, cuando no inexistentes, quedando al descubierto el retraso tecnológico que sufren debido a los problemas de financiación o a la falta de personal especializado en nuevas tecnologías. El futuro implicará un mayor impulso a la digitalización favorecida por el consiguiente ahorro en los costes de edición, factor importante para muchas cabeceras. Además, según los especialistas, esa digitalización irá imbricada en políticas de libre acceso. La juventud de muchas RCCE y las características propias de este mercado editorial justifican que a corto plazo el acceso abierto se imponga para permitir una mayor visibilidad e intentar con ello potenciar un mayor impacto de las revistas. A diferencia de lo que ocurre en otros mercados en los que es práctica más habitual el acceso restringido y un mayor sometimiento a los intereses lucrativos de las empresas editoras, casi todas las RCCE son editadas por instituciones públicas.

Por otra parte, cada vez será más necesaria la edición en inglés; así se desprende de la opinión de los encuestados en varias preguntas del cuestionario, aunque hay que aclarar que a lo que la práctica totalidad de ellos se refiere es a la edición bilingüe en inglés y español, precisamente para abarcar una comunidad científica más amplia sin menoscabo de la relevancia que cualquiera de estos dos idiomas pueda tener como vehículo de comunicación científica en distintas áreas geográficas. Cabe señalar, en todo caso, una cierta confusión respecto a la relación entre la calidad científica de las revistas y la edición en inglés. De la lectura literal de los datos obtenidos en la investigación se podría concluir que vinculan estos aspectos como si el idioma utilizado condicionase la calidad, cuando de una lectura más global de los cuestionarios se podría deducir que relacionan la utilización del inglés con la mayor visibilidad de las cabeceras y por tanto con mayores posibilidades para acreditar su impacto o calidad; de hecho, la edición en inglés también es subrayada cuando los expertos son preguntados sobre las medidas que adoptarán las RCCE para incrementar su impacto. Además, al referirse a las medidas para promover la calidad científica también se señalan como destacados otros factores vinculados a la visibilidad antes que a la calidad en sí misma, lo cual se podría explicar por esa equiparación que hacen entre visibilidad y calidad científica, aunque en el cuestionario no se exprese como tal. Prueba de ello es que hay especialistas que aportan algunas respuestas coincidentes en las dos preguntas, lo que ha contribuido a esta distorsión, mientras que otros han separado claramente ambos aspectos, por lo que podría ser necesario en próximas fases del estudio implementar medidas para evitar esa distorsión. En todo caso, se exponen objetivamente esos datos porque se han querido redactar las opiniones agregadas recogidas en el segundo cuestionario del *Delphi* manteniendo el sentido de las respuestas dadas al primero y sin perder información real, para cumplir así con las exigencias de esta técnica de investigación.

El futuro de las revistas científicas de comunicación pasa por una mayor internacionalización y mayores esfuerzos para la inclusión en índices de impacto

En la investigación realizada se dejan entrever algunas críticas hacia el comportamiento de las RCCE con relación a los sistemas de evaluación bibliométrica de la calidad científica de las revistas. Así, en varias ocasiones se habla literalmente de la necesidad de superar prácticas corruptas de las propias cabeceras para favorecer sus índices de impacto, refiriéndose a la autocitación y promoción de las citas entre colaboradores. Aunque en las tablas se recogió la opinión agregada, analizando las opiniones individuales se comprueba la disparidad de pareceres al respecto, como queda patente en las desviaciones típicas referidas a estos ítems de respuesta. Existe un grupo minoritario de expertos que critica esos sistemas de evaluación bibliométrica de las revistas, pero sobre todo en lo referente a las posiciones adoptadas al respecto por algunas cabeceras, mientras que otros defienden esas posiciones como válidas. En muchas ocasiones

estas críticas hacia el propio sector editorial son achacadas a razones externas, principalmente a imposiciones de las entidades encargadas de esa evaluación, si bien los propios encuestados son también autocríticos con muchas carencias que arrastran las RCCE, algunas de las cuales se han indicado anteriormente. En cualquier caso y en contraste con algunas de esas opiniones de los participantes en el *Delphi*, conviene subrayar el importante papel que, a pesar de las limitaciones que puedan tener, cumplen los sistemas de valoración de la investigación científica a partir de indicadores bibliométricos para garantizar un proceso lo menos arbitrario posible en la evaluación y acreditación de la producción científica de los investigadores y para contribuir a una mayor excelencia de la comunidad científica.

Pese a las carencias señaladas en párrafos anteriores, las RCCE tendrán un papel cada vez más importante en la difusión de la investigación en comunicación si bien para los expertos consultados su principal utilidad práctica será la de ayudar a los procesos de acreditación. El objetivo de esta investigación era recoger la opinión agregada de los directores de las principales revistas y de ahí lo significativo de este dato, ya que sin duda resulta sorprendente que consideren que sus cabeceras son herramienta para la acreditación de los investigadores antes que para promover la difusión del conocimiento científico; sobre todo por lo que ello implica en la definición de los objetivos, estrategias y tácticas que marcarán en los próximos años para el desarrollo de esas cabeceras y la misión que le asignarán a las mismas. El estudio *Delphi* ha permitido constatar esa opinión, pero sin duda sería necesario profundizar en éste y otros datos con una investigación cualitativa complementaria.

También es llamativo que los expertos defiendan la creación de nuevas RCCE para dar visibilidad a grupos de investigación, cuando, lejos de procedimientos endogámicos, las revistas científicas debieran ser vehículos abiertos y competitivos, nunca cerrados a los intereses de equipos de investigación más o menos emergentes.

Por último, cabe destacar la unanimidad de los expertos respecto a la necesidad de que surjan nuevas cabeceras en los próximos diez años. La principal razón es el crecimiento exponencial de la comunidad científica de investigadores en comunicación con la implantación de estos estudios en medio centenar de centros universitarios. Ahora bien, mientras la mayoría de esos investigadores pugnan por publicar en las principales revistas del área, otras revistas tienen problemas para recibir originales de calidad. Por ello, hay que valorar la tendencia contraria. La reconfiguración del mercado de RCCE tendrá que ir paralela a un incremento aún mayor en la calidad global de las publicaciones, siendo necesario, como se concluye en un reciente estudio (López-Rabadán; Vicente-Mariño, 2011, p. 13), “una fase de apuntamiento, en la que las revistas con mayor bagaje deben ejercer una función de referencia”. En la misma línea, Giménez-Toledo (2011, p. 8) aboga por una “agrupación de revistas generalistas” que permitiría concentrar esfuerzos. La mayor madurez del conjunto de las RCCE contribuiría sin duda a la consolidación de algunas de las publicaciones emergentes surgidas en los últimos años, así como al prestigio de otras cabeceras más veteranas que

empiezan ya a ser conscientes de la importancia de que su calidad sea acreditada mediante procedimientos de evaluación bibliométrica.

Bibliografía

- Abadal, Ernest; Rius, Lluís.** “Revistas científicas digitales: características e indicadores”. *Revista de universidad y sociedad del conocimiento*, 2006, v. 3, n. 1, pp. 6-20.
http://www.uoc.edu/rusc/3/1/dt/esp/abadal_rius.pdf
- Aleixandre-Benavent, Rafael; Valderrama-Zurián, Juan-Carlos; González-Alcaide, Gregorio.** “El factor de impacto de las revistas científicas: limitaciones e indicadores alternativos”. *El profesional de la información*, 2007, v. 16, n. 1, pp. 4-11.
<http://eprints.rclis.org/handle/10760/9489>
- Castillo, Antonio; Carretón, María-del-Carmen.** “Investigación en comunicación. Estudio bibliométrico de las revistas de comunicación en España”. *Comunicación y sociedad*, 2010, v. 23, n. 2, pp. 289-327.
http://www.unav.es/fcom/comunicacionysociedad/es/articulo.php?art_id=372
- Colle, Raymond.** “La temática de *Revista latina de comunicación social*, 1998-2008”. *Revista latina de comunicación social*, 2009, n. 64, pp. 71-85.
http://www.ull.es/publicaciones/latina/09/art/07_806_13_revista/Raymond_Colle.html
- De-Pablos, José-Manuel.** “Análisis de las revistas españolas de comunicación tras la actualización del índice de impacto de 2009”. En: *II Congreso Intl Latina de comunicación social*, 2010. ISBN: 978 8493842802
http://www.revistalatinacs.org/10SLCS/actas_2010/217_De_Pablos.html
- Espanha, Rita; Lima, Tiago.** “Open access and multilingual approach to communication journals. The case and the editor’s perspective of *Observatorio journal* and the importance of open science for the knowledge society”. *Online journal of communication and media technologies*, 2011, v. 1, n. 4, pp. 97-120.
<http://ojcmt.net/articles/144.pdf>
- Feeley, Thomas-Hugh.** “A bibliometric analysis of communication journals from 2002 to 2005”. *Human communication research*, 2008, v. 34, n. 3, pp. 505-520.
http://www.iimahd.ernet.in/library/PDFs/NICMAN/A_biblio_metric_analysis_of_communication_journals.pdf
- Fernández-Quijada, David.** “El perfil de las revistas españolas de comunicación (2007-2008)”. *Revista española de documentación científica*, 2010, v. 33, n. 4, pp. 553-581.
<http://redc.revistas.csic.es/index.php/redc/article/download/671/746>
<http://dx.doi.org/10.3989/redc.2010.4.756>
- Fernández-Quijada, David.** “De los investigadores a las redes. Una aproximación tipológica a la autoría en las revistas españolas de comunicación”. En: *Simposio AE-IC Investigar la comunicación en España*, 2011.
<http://eprints.rclis.org/handle/10760/15566>
- Giménez-Toledo, Elea; Alcaín-Partearroyo, Dolores.** “Estu-

dio de las revistas españolas de periodismo". *Comunicación y sociedad*, 2006, v. 19, n. 2, pp. 107-131.

http://www.unav.es/fcom/comunicacionysociedad/es/articulo.php?art_id=62

Giménez-Toledo, Elea. "La opinión de los expertos sobre las revistas españolas de comunicación y otros indicadores de calidad". En: *Simposio AE-IC Investigar la comunicación en España*, 2011.

<http://www.revistacomunicar.com/pdf/2011-04-Elea-Gimenez.pdf>

Jones, Daniel. "Les revistes teòriques sobre comunicació audiovisual a lberoamèrica". *Quaderns del CAC*, 2003, n. 17, pp. 57-66.

http://www.cac.cat/pfw_files/cma/recerca/quaderns_cac/Q17jones.pdf

Jones, Daniel. "Las revistas de comunicación en España". *Telos*, 2005, n. 64.

<http://www.sociedadinformacion.fundacion.telefonica.com/teelos/articulorevista.asp?idarticulo=1&rev=64.htm>

Landeta, Jon. *El método Delphi*. Barcelona: Ariel, 1999. ISBN 8434428369

Lauf, Edmund. "National diversity of major international journals in the field of communication". *Journal of communication*, 2005, v. 55, n. 1, pp. 139-151.

http://newmedia.cityu.edu.hk/enjhzhu/ruc/readings/Lauf_2005_NationalDiversity.pdf

López-Ornelas, Maricela. "Estudio cuantitativo de los procesos de comunicación de *Revista latina de comunicación social*, 1998-2009". *Revista latina de comunicación social*, 2010, n. 65, pp. 538-552.

http://www.revistalatinacs.org/10/art3/917_Mexico/39_Maricela.html

<http://dx.doi.org/10.4185/RLCS-65-2010-917-538-552>

López-Rabadán, Pablo; Vicente-Mariño, Miguel. "Métodos y técnicas de investigación dominantes en las revistas científicas españolas sobre comunicación (2000-2009)". En: *Simposio AE-IC Investigar la comunicación en España*, 2011.

<http://www.revistacomunicar.com/pdf/2011-04-Lopez-Vicente.pdf>

Martí, Daniel. "Las revistas académicas ibéricas y latinoamericanas de comunicación en internet en el contexto tecnológico actual". *Razón y palabra*, 2004, n. 41.

<http://www.razonypalabra.org.mx/antecedentes/n41/dmarti.html>

Martínez-Nicolás, Manuel. "La investigación sobre comunicación en España. Evolución histórica y retos actuales". *Revista latina de comunicación social*, 2009, n. 64, pp. 1-14.

http://www.ull.es/publicaciones/latina/09/art/01_800_01_investigacion/Manuel_Martinez_Nicolas.html

Martínez-Nicolás, Manuel; Saperas-Lapiedra Enric. "La investigación sobre comunicación en España (1998-2007). Análisis de los artículos publicados en revistas científicas". *Revista latina de comunicación social*, 2011, n. 66, pp. 101-129.

http://www.revistalatinacs.org/11/art/926_Vicalvaro/05_Nicolas.html

Masip, Pere. "Los efectos del efecto *Aneca*: análisis de la producción española en comunicación en el SSC". En: *Simposio AE-IC Investigar la comunicación en España*, 2011.

<http://www.revistacomunicar.com/pdf/2011-04-Masip.pdf>

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

Perceval, José-María; Fornieles-Alcaraz, Javier. "Confucio contra Sócrates: la perversa relación entre la investigación y la acreditación". *Anàlisi*, 2008, n. 36, pp. 213-224.

<http://www.raco.cat/index.php/analisi/article/viewFile/94533/119748>

Sabés, Fernando; Perceval, José-María. "Retos (y peligros) de las revistas científicas de comunicación en la era digital". En: *I Congreso Intl Latina de comunicación social*, 2009, pp. 1-16. ISBN 978 8499410012

http://www.revistalatinacs.org/09/Sociedad/actas/27_sabes.pdf

Stumpf, Ida. "Peer review in communication journals: viewpoint of editors, authors and referees". *Perspectivas em ciência da informação*, 2008, v. 13, n. 1, pp. 18-32.

Anota en tu agenda:

2ª Conferencia sobre calidad de revistas de ciencias sociales y humanidades (CRECS 2012)

Presentación del

Anuario ThinkEPI, vol. 6

Valencia, 10 de mayo de 2012

<http://www.thinkepi.net/crecs2012>





WORLD SHARES OF PUBLICATIONS OF THE USA, EU-27, AND CHINA COMPARED AND PREDICTED USING THE NEW WEB OF SCIENCE INTERFACE VERSUS SCOPUS



Loet Leydesdorff



Loet Leydesdorff received a Ph.D. in the *Faculty of Social Sciences* (sociology) in 1984; a M.A. in philosophy in 1977, a M.Sc. in biochemistry (1969) and B.Sc. in chemistry (1969). Since 2010 he is professor of *Communication and Innovation in the Dynamics of Science and Technology* at *University of Amsterdam*. Since 2007 he is a visiting professor at the *Institute of Scientific and Technical Information of China (ISTIC)* in Beijing, and honorary fellow of *SPRU (Science and Technology Policy Research of the University of Sussex)*. He has published extensively in fields of systems theory, social network analysis, scientometrics, and the sociology of innovation.

Amsterdam School of Communications Research (ASCoR), University of Amsterdam
Kloveniersburgwal, 48. 1012 CX Amsterdam, The Netherlands

loet@leydesdorff.net

<http://www.leydesdorff.net>

Abstract

The new interface of the *Web of Science* (Thomson Reuters) enables users to retrieve sets larger than 100,000 documents in a single search. This makes it possible to compare publication trends for China, the USA, EU-27, and smaller countries with the data in the *Scopus* (Elsevier) database. China no longer grew exponentially during the 2000s, but linearly. Contrary to previous predictions on the basis of exponential growth, the cross-over of the lines for China and the USA is postponed to the next decade (after 2020) according to this data. These long extrapolations, however, should be used only as indicators and not as predictions. Uncertainty in trends can be specified using the coefficient of determination of the regression (R^2) and confidence intervals. Along with the dynamics in the publication trends, one also has to take into account the dynamics of the databases used for the measurement.

Keywords

World share of publications, EU-27, China, USA, Cross-over, Measurement, *Scopus*, *Science citation index*, *SCIE*, *Thomson-Reuters*.

Título: Comparación y predicción de las cuotas mundiales de publicación de los EUA, EU-27 y China usando la nueva interfaz *Web of science* versus *Scopus*

Resumen

La nueva interfaz de la *Web of science* (de *Thomson Reuters*) permite ahora recuperar conjuntos de más de 100.000 documentos en una sola búsqueda. Esto hace posible comparar las tendencias de publicación de los EUA, UE-27 y China —y evidentemente también de los países más pequeños—, cosa que ya se podía hacer con los datos de la base de datos *Scopus* de *Elsevier*. China ya no creció de manera exponencial durante la década de 2000, sino de forma lineal. Contrariamente a las predicciones anteriores basadas en un crecimiento exponencial, el cruce de las líneas de China y de EUA se ha pospuesto a la próxima década (después de 2020). Sin embargo estas extrapolaciones a largo plazo sólo deben utilizarse como indicadores y no como predicciones. La incertidumbre en las tendencias se puede especificar utilizando el coeficiente de determinación de la regresión (R^2) e intervalos de confianza. Junto a la dinámica de las propias tendencias de publicación científica, también hay que tener en cuenta la dinámica de las bases de datos utilizadas para la medición.

Palabras clave

Cuota mundial de publicación, EUA, UE-27, China, Punto de cruce, Medición, *Scopus*, *Science Citation Index*, *SCIE*, *Thomson-Reuters*.

Leydesdorff, Loet. "World shares of publications of the USA, EU-27, and China compared and predicted using the new *Web of Science* interface versus *Scopus*". *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 43-49.
<http://dx.doi.org/10.3145/epi.2012.ene.06>

Nota: Este artículo puede leerse traducido al español en: <http://elprofesionaldelainformacion.com/leydesdorff.html>

Artículo recibido el 20-12-11

Aceptación definitiva: 10-01-11

1. Introduction

On March 28, 2011, the *BBC-online* had a headline that the *Royal Society* –the UK’s national science academy– had issued a report warning that “China (was) ‘to overtake US on science’ in two years” based on Elsevier’s *Scopus* data (Clarke *et al.*, 2011; Plume, 2011; see figure 1).

In the weeks thereafter, this news led to discussions on the email list of the *US National Science Foundation’s* “Science of science policy” (*scisip@listserv.nsf.gov*) about the quality of the prediction based on *Scopus* data. More recently, in July 2011, Thomson Reuters launched Version 5 of the *Web of science* (*WoS*) which allows the user –as in *Scopus*– to search directly for countries’ shares of contributions, whereas in the previous version one had to overcome the limits of a recall of more than 100,000 publications in each search (Arencibia-Jorge *et al.*, 2009).

Both *Scopus* and the *Science Citation Index* now allow for direct access to large numbers of items in the retrieval. In this article, the new *WoS*-version of the *Science Citation Index-Expanded* (*SCIE*) is first used to show the long-term trends of a few leading nations in science and also some smaller ones. The 10-year trendlines for the USA, China, and the EU-27 can be compared using confidence intervals (at the 95% level) for the prediction. These results are compared with those of the *Royal Society* and the latter will be reproduced using the online version of *Scopus*, but including data for 2009 and 2010. However, the *Elsevier/Royal Society* team used *Scopus* including the social sciences and humanities, while these were not included using *SCIE* for the measurement. After correction for this, the decline of both the EU-27 and the US since 2004 disappears using *Scopus* data. The significant differences between using the two databases and the different assumptions for the measurement raise questions about the reliability of the prediction.

2. Theoretical relevance

The measurement of national publication outputs has been a methodological issue on the research agenda of scientometrics from the very beginning of the *Science Citation Index*. Both Narin (1976) and Small & Garfield (1985) conceptualized this database as a matrix organized along the two dimensions of journals versus countries. The “decline of British Science” in the 1980s (under the Thatcher government), for example, spurred a debate about whether such a decline could perhaps be a scientometric artifact of parameter choices (Anderson *et al.*, 1988; Braun *et al.*, 1989 and 1991; Leydesdorff, 1988 and 1991; Martin, 1991).

At the time, the main database used for the *Science and Engineering Indicators* of the *US National Science Board* (since 1982)² was based on two assumptions made by the contracting firm (Narin’s *Computer Horizons Inc.*):

- 1 internationally coauthored articles were attributed proportionally to the contributing nations (this is also called “fractional counting”), and
- 2 a fixed journal set was extracted from the *Science Citation Index* for the purpose of longitudinal comparisons (Narin, 1986).

Leydesdorff (1988) argued that both these assumptions

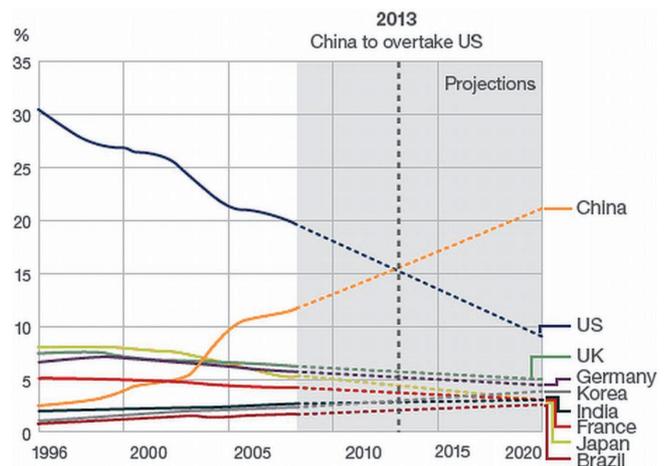


Figure 1. Linear extrapolation of future publication trends based on *Scopus* (1996-2008). Source: Clarke *et al.*, 2011, figure 1.6, at p. 43.¹

had an effect on the measurement of national outputs: the ongoing internationalization of coauthorship patterns decreased the national output *ceteris paribus*, and authors in advanced nations such as the UK can be expected to publish above average in new journals associated with newly developing fields of science.

The issue led to a debate and eventually a special issue of *Scientometrics* in 1991 (Braun *et al.*, 1991). Braun *et al.* (1989) distinguished 28 possible parameter choices. The sensitivity of the measurement for relatively minor decisions at the methodological level questions the role of policy advice based on these trendlines for both nations and units at lower levels of aggregation (Leydesdorff, 1989; 1991). How reliable are these data for comparisons between years? One would expect random fluctuations to be averaged out at a high level of aggregation, and thus uncertainty to be reduced. Nowadays, one can additionally ask whether the two major databases (*Scopus* and *WoS*) can provide us with similar results. What may be sources of misspecification and therefore potential misrepresentations in the policy arena? (Leydesdorff, 2008).

The issue of China as a leading nation in science is particularly salient to the science-policy debate today. How much of the spectacular increase of China’s world share of publications during the 1990s and 2000s can be attributed to internationalization, to the detriment of national publication outlets? (Wagner, 2011). Zhou & Leydesdorff (2006) conjectured that, different from the linear growth witnessed in the case of internationalization (and Anglification) of national research outputs (e.g., Scandinavia and the Netherlands during the 1980s; Italy and Spain during the 1990s), a reservoir of Chinese scientists who hitherto had no access to other than national journals was tapped and now competing for access to the international literature.

China has also a large number of national journals. Zhou (2009) estimated that China had 9,468 journals in 2006, including 4,758 in science and technology and 2,339 in the social sciences (Jiang, 2007; Ren, 2007; cf. Ren; Rousseau, 2002). China has also two citation indices in science and technology (Wu, 2004) and two more in the social sciences (Zhou *et al.*, 2010). The number of journals covered by these databases has increased during the last two decades. Thus, it seems that the growth in international presence

adds to China's national publication system (Jin & Leydesdorff, 2006).

Is there a justified concern about "the West losing ground" in the sciences (Leydesdorff; Wagner, 2009a; Shelton, 2010; Wagner; Wong, 2011)? Reflexively, the bibliometric analyst can ask how reliably one can provide policy advice in these matters (Leydesdorff, 2008)? How can the bibliometric analysis be improved (Rafols *et al.*, 2011)?

3. Methods and materials

All searches were performed between September 23 and 25, 2011 (unless specified otherwise), using the web interfaces of *Scopus* and *WoS-v.5*, respectively:

<http://www.scopus.com>

<http://apps.webofknowledge.com>

Searches were limited to the so-called citable items: articles, proceedings papers, and reviews. Using these databases, internationally co-authored papers are attributed to contributing nations as whole numbers (so-called "integer counting"; cf. Andersen *et al.*, 1988; National Science Board, 2010). For the European Union-27, a search string with the names of all member states was composed with a Boolean OR. In the *WoS*, one additionally has to use "England OR Scotland OR Wales OR Northern Ireland" for the UK.

In the *WoS*, the years were delimited in terms of tape-years, that is, from January 1 to December 31 of each year, respectively. In *Scopus*, the corresponding search string for the USA in 2010, for example, can be formulated as follows: "AFFILCOUNTRY(United States) AND (DOCTYPE(ar) OR DOCTYPE(re) OR DOCTYPE(cp)) AND PUBYEAR is 2010". This search provides us with a replication of the report of the team of the *Royal Society* and *Elsevier/Scopus* (Moed *et al.*, 2011). However, the data from this search in *Scopus* includes also the social sciences and the humanities while this database enables us to exclude these domains by adding to the searches "AND NOT (SUBJAREA(Arts) OR SUBJAREA(Soci) OR SUBJAREA(Econ) OR SUBJAREA(Psyc) OR SUBJAREA(Deci) OR SUBJAREA(Busi))".³

The data gathering is otherwise straightforward. I distinguish additionally the group of 12 countries that joined the EU in May 2004 because these results may help to explain some of the differences between the USA and the EU-27 during the 2000s (Leydesdorff, 2000). The analysis is confined to the years 2000-2010. For the extrapolation, using SPSS v.18 enables users to draw the confidence intervals in the graphs.⁴ The other figures are drawn from the database in Excel.

4. Results

4.1 *WoS* data

Contrary to previous analyses that included also the 1990s (e.g., Jin; Rousseau, 2004; Moed, 2002), the focus on the last 10 years shows that the growth of China's percentage share of publications has been increasing linearly during this period ($R^2 > 0.99$). The exponential growth of China in these terms during the second half of the 1990s and the early 2000s was spectacular. Using *WoS* data, figure 2 shows an extrapolation of the linear regression lines for China, the USA, and the EU-27. The decline of the EU-27 and the USA

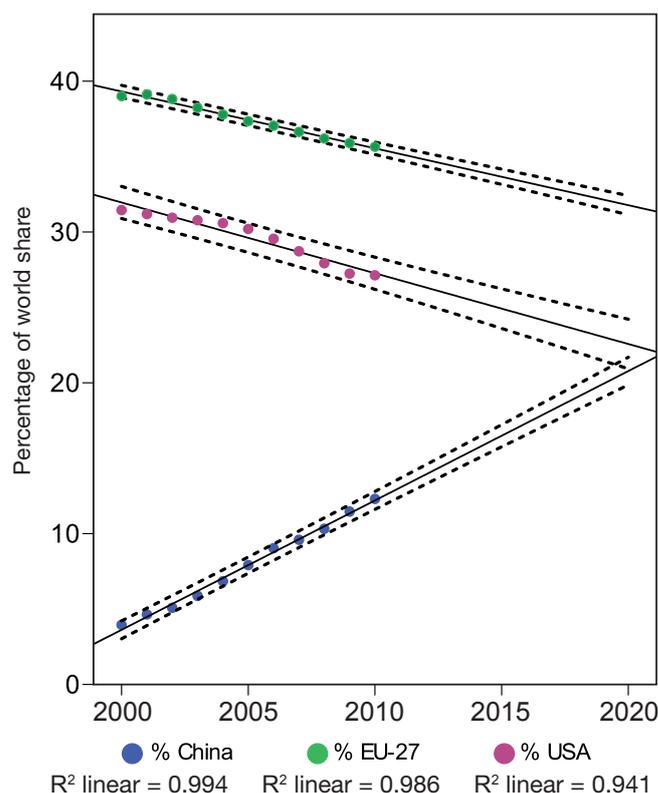


Figure 2. Percentages of world share of publications in SCIE (articles, proceedings papers, and reviews) for the USA, EU-27, and China. Source: *Web of science*; confidence levels indicated at the 95% level

in terms of percentages of world share is partly a function of the increase of other countries (although the percentages do not have to add up to 100% given that international collaborations are counted for all contributing nations; cf. Anderson *et al.*, 1988).

As against earlier predictions (e.g., Shelton; Foland, 2009; Leydesdorff; Wagner, 2009b) that found exponential growth for China (during the 1990s), the linear growth in this projection postpones the cross-over between the USA and China until well into the next decade. This graph predicates an even later date than a previous prediction based on using *WoS-v.4* data (Shelton; Leydesdorff, in press). As said, the construction of datapoints was hitherto less straightforward and perhaps less reliable (Arencibia-Jorge *et al.*, 2009).

Figure 3 extends the analysis to some middle-sized and smaller economies. At the top of the figure, one can see that the middle-sized countries (UK, Germany, Japan, and France; cf. Daraio; Moed, 2011) are in decline at approximately the same rate as the USA, but Japan has a steeper decline rate in the share of publications. China surpassed the UK (in this database) in 2005, and Germany and Japan in 2006.

In the lower half of figure 3, one can see that Korea has been growing similarly to China, but this curve is not linear (Leydesdorff; Zhou, 2005; Park *et al.*, 2005; Park; Leydesdorff, 2010). The curve for Korea happens to be an excellent match for a third-order polynomial ($r^2 > 0.98$) indicating a slowing down of growth in the middle years of the decade under study. Over this whole period, the 12 new accession countries to the EU increased their shares of publication (cf. Leydesdorff; Wagner, 2009), but this growth potential seems to approach saturation during the last three years. Smaller

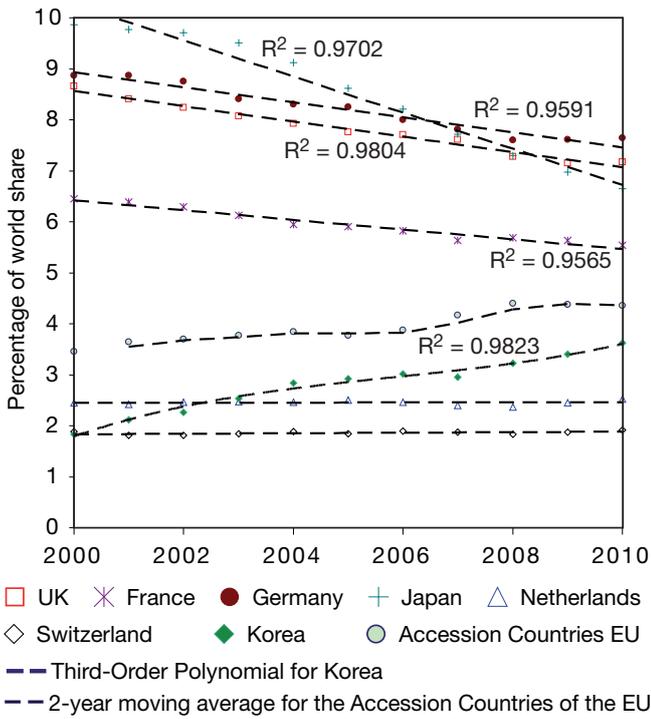


Figure 3. Percentages of world share of publications in *SCIE* (articles, proceedings, and reviews) for some middle-sized and smaller nations

European countries such as Switzerland and the Netherlands have been able to maintain their percentage shares during this decade; at 1.86 (± 0.01)% and 2.45 (± 0.01)%, respectively. This precision provides further confidence in these data.

4.2. Scopus data

Using *Scopus* data, one obtains a very different perspective on the shares of publications of the US, China, and EU-27 (figure 4). The data for China again fit best with a linear regression line ($R^2 > 0.97$), but the lines for the EU-27 and the USA are shaped differently. The two or three most recent years show an upward trend that cannot be found using the *WoS* data. Using *Scopus*, however, the years 2010 and 2011 already fall within the 95%-confidence interval for the prediction that China might take over the first position from the USA. Thus, this effect is even stronger than the one reported previously by Clarke *et al.* (2011) and Plume (2011). However, the quality of the correlation with the linear regression is so low for the USA and the EU-27 that one can be hesitant to draw these regression lines. The confidence lines show the uncertainty.

5. Social sciences and humanities in the Scopus data

National performance using *WoS* data is often measured using the *Science Citation Index-Expanded* (6,650 journals) or the *Science Citation Index* (CD-Rom version; appr. 3,700 journals; *National Science Board*, 2010). However, the study of the *Royal Society* and *Elsevier* was based on the entire *Scopus* database including also the social sciences and humanities, while these fields are separately indexed in the *WoS*.

The social sciences and humanities can be excluded in *Scopus* by using the appropriate subject areas in the search string as

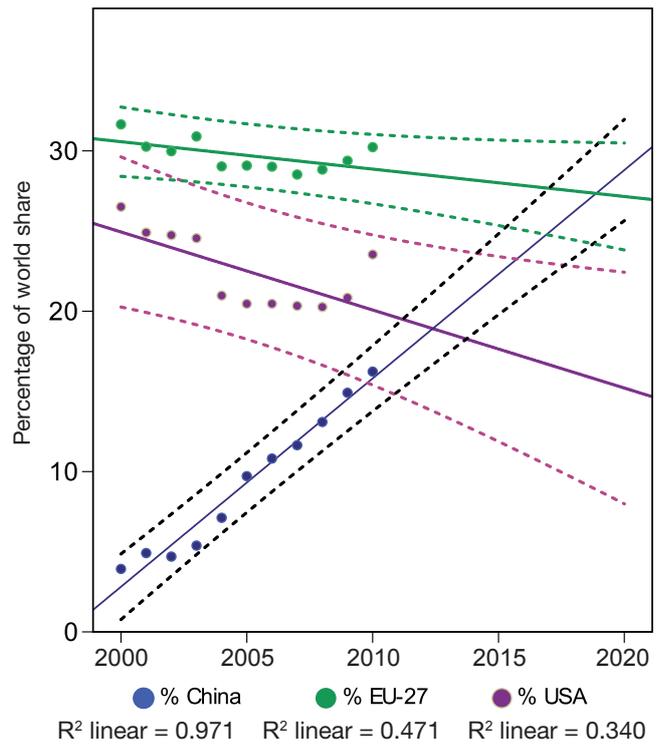


Figure 4. Percentages of world share of publications (articles, proceedings, and reviews) for the USA, EU-27, and China. Source: *Scopus*; confidence levels indicated at the 95% level

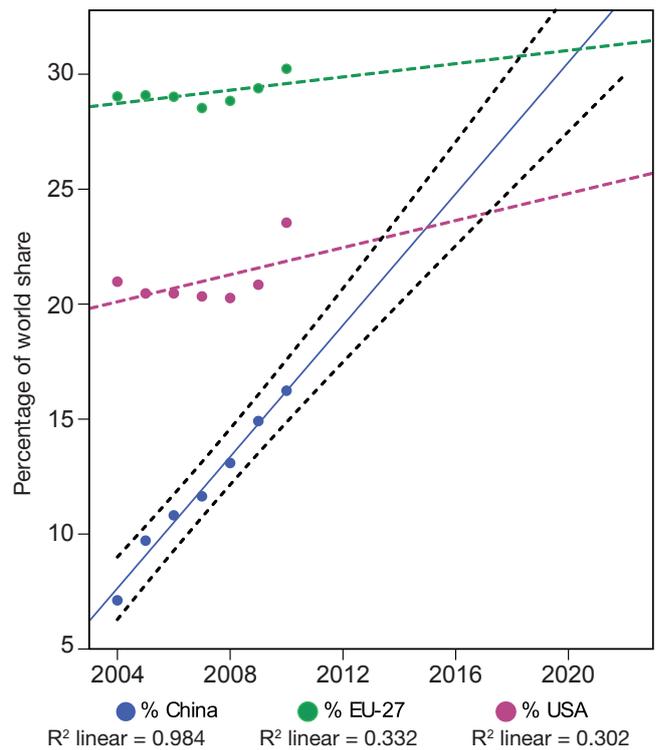


Figure 5. Percentages of world share of publications (articles, proceedings, and reviews) for the USA, EU-27, and China, after correction for the social sciences and the humanities. Source: *Scopus*, November 29, 2011; confidence levels indicated at the 95% level

specified above (in the methods section). The general pattern (figure 4) does not change by this refinement, but the upward trend in the data for the EU-27 and the USA since 2004 is more pronounced than before, and highlighted in figure 5. The message of the *Royal Society/Elsevier* team would be mistaken on the basis of this extrapolation of *Scopus* data.

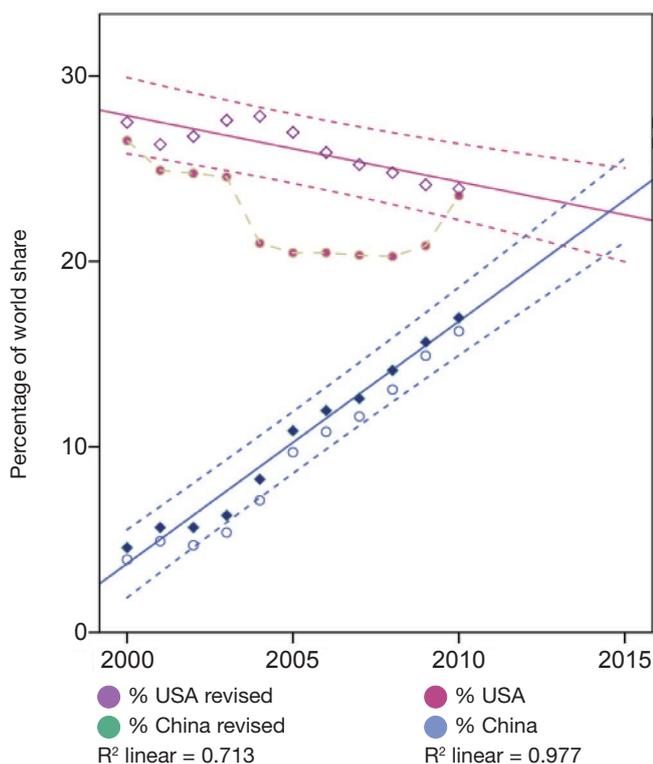


Figure 6. Revised percentages of world share of publications (articles) for the USA (◊) and China (○). Source: the bibliometric version of *Scopus*; confidence levels indicated at the 95% level.

6. A recent revision of the prediction in *Research trends*

In reaction to a preprint version of this paper, the staff of *Scopus* published a reply in *Elsevier's* online journal *Research trends* (Moed et al., 2011) in which one argues that *Elsevier* publishes a version of *Scopus* on the internet, but also maintains a bibliometric version of this database in which the data are subjected to more intensive cleaning processes. As stated: "Especially the results for the USA differ considerably between the two *Scopus* versions. These discrepancies are due to the fact that not all author affiliations contain the name of the country in which the author's institutions are located. This is especially true for US affiliations: many indicate the US State but not the country name."

Figure 6 is based on a reconstruction of this revised data of the *Scopus* team (Leydesdorff, 2011; Moed et al., 2011). The values from figure 4 are penciled in for the comparison. Indeed, the differences are largest for the USA in almost all years. However, even in this corrected data the previous prediction of a cross-over in 2013 is not fully warranted and the fit for the linear regression in the case of USA data remains relatively poor ($R^2 = 0.71$).

7. Discussion

What might cause these differences between the measurements in the respective databases? Let me first stipulate that in both databases I used 2000-2010, whereas the team of the *Royal Society* and *Elsevier* used 1996-2008 for their prediction. When this report appeared in March 2011, I replicated the measurement and found some deviation for

2009 and 2010, but assumed that this could be an artifact because the publication year 2010 was not yet completed by March/April 2011. Publications may arrive with the timestamp of 2010 at a later date in 2011, and practice may vary for publications from different world regions. However, a repeat of the measurement in September did not change these results.

I deliberately used the data since 2000 because *Scopus* data are only reliable since 1996 (Ove Kahler, personal communication, 28 August 2009), and the database was gradually improved in terms of coverage during the initial years. As against the *Web of science*, *Scopus* claims to include more regional journals among the 18,000 journals covered by this database. See at:

<http://www.info.sciverse.com/Scopus/Scopus-in-detail/facts>

The *Web of science* nowadays covers approximately 11,500 journals including approximately 3,000 journals added since 2008 (Testa, 2011). Thomson Reuters first announced this as an expansion of regional coverage in May 2008, possibly in response to competition from *Scopus*. The comparison of figures 2 and 4 above, however, teaches us that the focus in the *Web of science* has remained on Europe and the USA more than in *Scopus*.

Figure 7 shows that the percentage of share of Chinese publications in the *WoS-v.5* is 12.30% in the *Web of science*, while it is 17.24% in *Scopus* (after correction for the social sciences and humanities). Similarly, the USA has 22.54% in *Scopus* data as against 27.13% in the *Web of Science*. The differences are approximately five percentage points on either side, and thus add up to more than 9.5%. For the EU-27, the difference between the two databases is even larger, with 30.12% in *Scopus* and 35.65% in the *WoS*, a difference of 5.53 percentage points; this difference is of the size of the contribution of France.

8. Conclusions

The bibliometric contribution to the policy debate about the ranking of national and institutional science systems, in my opinion, should focus on the specification of uncertainty

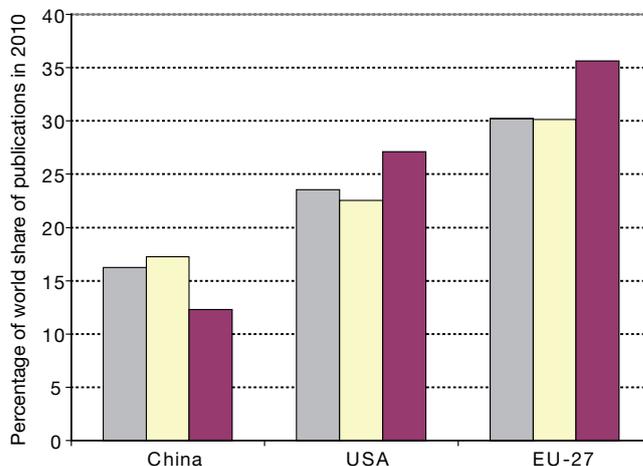


Figure 7. Percentages of world share of publications for China, the USA, and the EU-27 in 2010 in both *Scopus* (with and without correction for the social sciences and humanities) and *WoS-v.5*.

and possible sources of error and potential misinterpretation (cf. Leydesdorff, 2008). In a seminal paper, Martin & Irvine (1983) suggested to rely on “converging partial indicators” for the assessment: results of the bibliometric analysis can be considered as more reliable when the results indicate the same trends or differences in rankings. Using the *Scopus* database, however, one could even make a case for a relative increase of the shares of publications for the US and the EU since 2004 (on the basis of figure 5 above).

In the above comparisons between *WoS* and *Scopus* data, the confidence lines and fit provide an argument to build policy advice preferentially on *WoS* data since the uncertainty is lower. However, the difference in coverage between these two major databases is significant at this high level of aggregation: *Scopus* is more oriented to the Chinese publication system and less to the US and the EU than *SCIE*. As noted, the differences can be in the order of five to ten percentage points. In my opinion, such differences are worrisome and worth noting in the case of policy advice (Clarke *et al.*, 2011; Rafols *et al.*, in press).

Although strong growth remains indicated for the case of China, the USA cannot be expected to continue declining linearly. Whereas the world sum of publication is not a zero-sum game because of the steady increase of international coauthorship relations (Persson *et al.*, 2004; Wagner, 2008), the competition drives in the direction of decreasing marginal returns because all nations are investing in order to improve their share of publications (and citations). In addition to the dynamics of the competition, the above exercise reminds us that the dynamics of the databases also need to be taken into account.

9. Acknowledgement

I am grateful for comments of Lutz Bornmann and Henk Moed.

10. Notes

1. We use the remake of the figure by the BBC:
<http://www.bbc.co.uk/news/science-environment-12885271>
2. The 1988 edition of these indicators was named “Science & engineering indicators”. Before this date the title was “Science indicators”.
3. Whereas it is not possible to search online within the *Scopus* database with the four major categories of journals –Life sciences (>4,300 titles), Health sciences (>6,800 titles, 100% *Medline* coverage), Physical sciences (>7,200 titles) and Social sciences & humanities (>5,300 titles)– *Scopus* offers a concordance table to 27 subject area classifications that can be searched using the function “subjarea()” in the advanced search engine.
http://help.scopus.com/robo/projects/schelp/h_subject_categories.htm
4. The graphs are produced by the subroutine *Chart builder* within *SPSS*. Different regression lines and curve fits can be added to the graphs in the *Chart editor*. Various forms of regression analysis were also performed in *SPSS*; for example, for determining the β coefficient.

11. References

- Anderson, Joe; Collins, Peter M. D.; Irvine, John; Isard, Phoebe A.; Martin, Ben R.; Narin, Francis *et al.* “On-line approaches to measuring national scientific output: A cautionary tale”. *Science and public policy*, 1988, n. 15, pp. 153-161.
- Arencibia-Jorge, Ricardo; Leydesdorff, Loet; Chinchilla-Rodríguez, Zaida; Rousseau, Ronald; Paris, Soren W. “Retrieval of very large numbers of items in the Web of Science: an exercise to develop accurate search strategies”. *El profesional de la información*, 2009, v. 18, n. 5, pp. 529-533.
- Braun, Tibor; Glänzel, Wolfgang; Schubert, András. “Assessing assessments of British science. Some facts and figures to accept or decline”. *Scientometrics*, 1989, v. 15, n. 3, pp. 165-170.
- Braun, Tibor; Glänzel, Wolfgang; Schubert, András. “The bibliometric assessment of UK scientific performance. Some comments on Martin’s reply”. *Scientometrics*, 1991, v. 20, pp. 359-362.
- Clarke, Luke; Plume, Andrew *et al.* *Knowledge, networks and nations: Global scientific collaboration in the 21st century*. London: The Royal Society, 2011.
<http://royalsociety.org/policy/projects/knowledge-networks-nations/report>
- Daraio, Cinzia; Moed, Henk F. “Is Italian science declining?”. *Research policy*, 2011, v. 40, n. 10, pp. 1380-1392.
- Jiang, X. P. 我国现行科技期刊评价状况分析 (*Analysis on China’s present situation in journal evaluation*). 2008 中国科协学术建设发布 (Announcement on the academic research project of China Association for Science and Technology), 2007.
<http://www.cast.org.cn/n435777/n435799/n1105056/n1109014/n1110359/n1110369/40193.html>
- Jin, Bihui; Leydesdorff, Loet. “A comparison of the aggregated journal citation structures in the Chinese Science Citation Database and the Science Citation Index”. In: *Eighth Intl Conf on Science and Technology Indicators*, Leiden, 23-25 September 2004.
- Jin, Bihui; Rousseau, Ronald. “Evaluation of research performance and scientometric indicators in China”. In: H. F. Moed, W. Glänzel & U. Schmoch (Eds.), *Handbook of quantitative science and technology research*, pp. 497-514. Dordrecht, etc.: Kluwer Academic Publishers, 2004.
- Leydesdorff, Loet. “Problems with the ‘measurement’ of national scientific performance”. *Science and public policy*, 1988, v. 15, n. 3, pp. 149-152.
- Leydesdorff, Loet. “On the ‘scientometric decline’ of British science. One additional graph in reply to Ben Martin”. *Scientometrics*, 1991, v. 20, pp. 363-368.
- Leydesdorff, Loet. “Is the European Union becoming a single publication system? *Scientometrics*, 2000, v. 47, n. 2, pp. 265-280.
- Leydesdorff, Loet. “Caveats for the use of citation indicators in research and journal evaluation”. *Journal of the American Society for Information Science and Technology*, 2008, v. 59, n. 2, pp. 278-287.

- Leydesdorff, Loet.** "When can the cross-over between China and the USA be expected using *Scopus* data". *Research trends*, 2011, November, n. 25.
<http://www.researchtrends.com/issue25-november-2011/is-science-in-your-country-declining-or-is-your-country-becoming-a-scientific-super-power-and-how-quickly>
- Leydesdorff, Loet; Wagner, Caroline S.** "Is the United States losing ground in science? A global perspective on the world science system". *Scientometrics*, 2009a, v. 78, n. 1, pp. 23-36.
- Leydesdorff, Loet; Wagner, Caroline S.** "Macro-level indicators of the relations between research funding and research output". *Journal of informetrics*, 2009b, v. 3, n. 4, pp. 353-362.
- Leydesdorff, Loet; Zhou, Ping.** "Are the contributions of China and Korea upsetting the world system of science?". *Scientometrics*, 2005, v. 6, n. 3, pp. 617-630.
- Martin, Ben R.** "The bibliometric assessment of UK scientific performance. A reply to Braun, Glänzel and Schubert". *Scientometrics*, 1991, v. 20, pp. 333-357.
- Martin, Ben R.; Irvine, John.** "Assessing basic research: some partial indicators of scientific progress in radio astronomy". *Research policy*, 1983, n. 12, pp. 61-90.
- Moed, Henk F.** "Measuring China's research performance using the Science Citation Index". *Scientometrics*, 2002, v. 53, n. 3, pp. 281-296.
- Moed, Henk F.; Plume, Andrew; Aisati, Mohamed; Bervkens, Peter.** "Is science in your country declining? Or is your country becoming a super power and when?". *Research trends*, 2011, November, n. 25.
<http://www.researchtrends.com/issue25-november-2011/is-science-in-your-country-declining-or-is-your-country-becoming-a-scientific-super-power-and-how-quickly>
- Narin, Francis.** *Evaluative bibliometrics: the use of publication and citation analysis in the evaluation of scientific activity*. Washington, DC: National Science Foundation, 1976.
- National Science Board. *Science and engineering indicators*. Washington DC: National Science Foundation, 2010.
<http://www.nsf.gov/statistics/seind10>
- Park, Han-Woo; Leydesdorff, Loet.** "Longitudinal trends in networks of university-industry-government relations in South Korea: The role of programmatic incentives". *Research policy*, 2010, v. 39, n. 5, pp. 640-649.
- Park, Han-Woo; Hong, Heung-Deug; Leydesdorff, Loet.** "A comparison of the knowledge-based innovation systems in the economies of South Korea and the Netherlands using triple helix indicators". *Scientometrics*, 2005, v. 65, n. 1, pp. 3-27.
- Persson, Olle; Glänzel, Wolfgang; Danell, Rickard.** "Inflationary bibliometrics values: The role of scientific collaboration and the need for relative indicators in evaluative studies". *Scientometrics*, 2004, v. 60, n. 3, pp. 421-432.
- Plume, Andrew.** "Tipping the balance: The rise of China as a science superpower". *Research trends*, 2011, March, n. 22.
<http://www.researchtrends.com/issue22-march-2011/tipping-the-balance-the-rise-of-china-as-a-science-superpower>
- Rafols, Ismael; Leydesdorff, Loet; O'Hare, Alice; Nightingale, Paul; Stirling, Andy.** (in press). "How journal rankings can suppress interdisciplinary research: A comparison between innovation studies and business & management". *Research policy*.
- Ren, Sheng-Li.** 我国科技期刊的现状及其在学术交流中的作用 ("Status quo of China's sci-tech journals and role in academic exchanges"). *中国科技期刊研究 (Chinese journal of scientific and technical periodical)*, 2007, v. 18, n. 3, p. 357.
- Ren, Sheng-Li.; Rousseau, Ronald.** "International visibility of Chinese scientific journals". *Scientometrics*, 2002, v. 53, n. 3, pp. 389-405.
- Shelton, Robert D.; Foland, Patricia.** "The race for world leadership of science and technology: status and forecasts". *Science focus*, 2010, n. 5, pp. 1-9.
- Shelton, Robert D.; Foland, Patricia.** "The race for world leadership of science and technology: status and forecasts". In: 12th Intl conf of the Intl Society for Scientometrics and Informetrics, Birger Larsen and Jacqueline Larsen (Eds.), vol. 1, pp. 369-380, Rio de Janeiro, Brazil, July 14-17, 2009.
- Shelton, Robert D.; Leydesdorff, Loet** (in press). "Publish or patent: bibliometric evidence for empirical trade-offs in national funding strategies". *Journal of the American Society for Information Science and Technology*.
- Small, Henry; Garfield, Eugene.** "The geography of science: disciplinary and national mappings". *Journal of information science*, 1985, v. 11, n. 4, pp. 147-159.
- Testa, James.** *The globalization of the Web of Science*. Thomson Reuters, 2011.
<http://wokinfo.com/media/pdf/globalwos-essay.pdf>
- Wagner, Caroline S.** *The new invisible college*. Washington, DC: Brookings Press, 2008.
- Wagner, Caroline S.; Wong, Shing K.** "Unseen science: representation of BRICs in global science". In: *2011 Atlanta conf on science and innovation policy*, Atlanta, September 15-17, 2011.
- Wu, Yishan; Pan, Yuntao; Zhang, Yuhua; Ma, Zheng; Pang, Jingan; Guo, Hong; Xu, Bo; Yang, Zhiqing.** "China scientific and technical papers and citations (CSTPC): History, impact and outlook". *Scientometrics*, 2004, v. 60, n. 3, pp. 385-397.
- Zhou, Ping.** *Mapping knowledge production and scholarly communication in China*. Unpublished Ph.D. thesis, University of Amsterdam, Amsterdam, 2009.
- Zhou, Ping; Leydesdorff, Loet.** "The emergence of China as a leading nation in science". *Research policy*, 2006, v. 35, n. 1, pp. 83-104.
- Zhou, Ping; Su, Xinning; Leydesdorff, Loet.** "A comparative study on communication structures of Chinese journals in the social sciences". *Journal of the American Society for Information Science and Technology*, 2010, v. 61, n. 7, pp. 1360-1376.

VALORACIÓN DE EDITORIALES ESPECIALIZADAS EN COMUNICACIÓN, BIBLIOTECONOMÍA Y DOCUMENTACIÓN: ENCUESTA A PROFESORES E INVESTIGADORES

Elea Giménez-Toledo y Carlos Tejada-Artigas



Elea Giménez-Toledo, científica titular del CSIC, es licenciada y doctora en documentación, y responsable del *Grupo de Investigación de Evaluación de Publicaciones Científicas (EPUC)* del *Centro de Ciencias Humanas y Sociales*. Su actividad investigadora se centra en los procesos de evaluación de la actividad científica en ciencias humanas y sociales. Diseña y aplica indicadores de calidad para revistas científicas y editoriales de libros. Es coautora de las plataformas de evaluación de revistas *DICE*, *RESH* y *CIRC* y es responsable de la parte española de *Latindex*.

G. I. Evaluación de Publicaciones Científicas (EPUC)
Instituto de Estudios Documentales sobre Ciencia y Tecnología (Iedcyt)
Centro de Ciencias Humanas y Sociales (CCHS)
Consejo Superior de Investigaciones Científicas (CSIC)
Albasanz, 26-28. 28037 Madrid
elea.gimenez@cchs.csic.es



Carlos Tejada-Artigas es doctor en documentación por la *Universidad Carlos III* de Madrid y licenciado en geografía e historia por la *Universidad Complutense de Madrid (UCM)*. Profesor titular de la *Facultad de Ciencias de la Documentación* de esta última universidad. Miembro de la junta directiva de *Sedic*, coordinador editorial de la revista *El profesional de la información* y coordinador del *Grupo de Trabajo de Perfiles Profesionales* del *Consejo de Cooperación Bibliotecaria* del *Ministerio de Cultura*. Sus principales líneas de investigación son el mercado de trabajo en información y documentación, competencias profesionales y publicaciones científicas.

Facultad de Ciencias de la Documentación
Universidad Complutense de Madrid
Santísima Trinidad 37. 28010 Madrid
cmtejada@pdi.ucm.es

Resumen

Se analizan los procesos de publicación de libros científicos y las editoriales en Biblioteconomía y Documentación y en Comunicación. El método empleado ha sido un cuestionario dirigido a 263 investigadores y profesores españoles de estas áreas, del que se obtuvo una tasa de respuesta del 35,5%. Los resultados permiten estudiar los indicadores de calidad de las editoriales que pueden tener una incidencia directa en la calidad de los libros que se editan, y los hábitos, opiniones y actitudes de los investigadores ante el proceso de selección y publicación de sus libros. Como resultado más pragmático se han identificado las editoriales españolas y extranjeras más prestigiosas para los investigadores de estas áreas.

Palabras clave

Editoriales, Libros, Indicadores de calidad, Evaluación de la investigación, Biblioteconomía y documentación, Comunicación.

Title: Assessment of book publishers in Communication and LIS: survey to researchers and professors

Abstract

Publishers of scholarly books about library and information science and communications are analyzed, as well as their publishing processes. The research method was a questionnaire sent to 263 Spanish researchers and teachers in these areas. The survey response rate was 35.5%. Thanks to these responses, main quality indicators for publishers were identified. Researchers' working habits, opinions and attitudes related to the selection, editing and publishing processes of their books are also examined. Finally, as a more pragmatic result, the most prestigious Spanish and foreign publishers are identified.

Keywords

Publishers, Books, Quality indicators, Research evaluation, Library and information science, LIS, Communication.

Artículo recibido el 25-11-11
Aceptación definitiva: 29-12-11

Giménez-Toledo, Elea; Tejada-Artigas, Carlos. “Valoración de editoriales especializadas en Comunicación, Biblioteconomía y Documentación: encuesta a profesores e investigadores”. *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 50-62.

<http://dx.doi.org/10.3145/epi.2012.ene.07>

0. Introducción

La intensa actividad actual de evaluación de la investigación con el fin de asignar recursos, sexenios, acreditaciones, etc., ha traído consigo la necesidad de especificar cada vez con más claridad y precisión los indicadores de calidad, especialmente los de publicaciones.

Existen diversas fuentes internacionales y nacionales que aportan indicadores de calidad de distinta naturaleza sobre revistas (desde *WoS*, *Scopus*, *ERIH* y *CIRC* hasta *DICE*, *RESH*, *In-Recs*, *MIAR*, etc.), pero para los libros la situación es diferente. Los estudios de calidad de monografías tienen aún poca trayectoria, a pesar de que han sido demandados directa o indirectamente por humanistas y científicos sociales, la comunidad científica que más las valora y utiliza como vehículo de comunicación científica. Algunos precedentes en la evaluación de libros y editoriales (Cullars; 1992; 1998; Moed; Luwel; Nederhof, 2001; Hemlin; Gustafsson, 1996; Glänzel; Schoepflin, 1999) son revisados con detalle en el trabajo de Giménez-Toledo y Román-Román (2009) y pueden resumirse en:

- valoraciones de la calidad de las editoriales por parte de expertos;
- propuestas metodológicas para evaluar monografías;
- proyectos experimentales de evaluación cualitativa de libros¹.

Exceptuando *Google Scholar* y el recién estrenado *Book Citation Index (Thomson Reuters)*, que ofrecen datos de citas a libros, no existen fuentes que proporcionen indicadores de calidad directos o indirectos. Sin embargo parece necesario que existan, a la luz de la precisión que han de tener los actuales sistemas de evaluación y para contribuir a que sean lo más objetivos posible.

En esta línea de trabajo se inscribe este artículo que pretende dar voz a la comunidad científica española de humanistas y científicos sociales en asuntos relacionados con la evaluación de monografías. El trabajo se ha llevado a cabo en el marco de un proyecto más amplio² que ha tenido como objetivo estudiar los indicadores que pueden tener una incidencia directa en la calidad de los libros que se editan, en los hábitos, opiniones y actitudes de los investigadores ante el proceso de selección y publicación de sus libros y, como objetivo más pragmático, identificar las editoriales españolas y extranjeras más prestigiosas para los investigadores. Este proyecto ha sido desarrollado para todas las disciplinas de las humanidades y las ciencias sociales y lo que se presenta en este trabajo es un análisis sectorial en las áreas de Comunicación y Documentación, ya que son las de interés específico para los lectores de esta revista.

El estudio de las editoriales propiamente dichas es objeto de un artículo complementario a éste, publicado en otra revista, que pretende constituirse en una herramienta de con-

sulta para esclarecer cuáles son las editoriales prestigiosas según la comunidad científica.

1. Objetivos

Con el objetivo general de conocer mejor los indicadores de calidad para libros científicos y objetivar en cierto modo el indicador “prestigio de la editorial” –referido por las agencias de evaluación pero poco concretado–, este trabajo se plantea los siguientes objetivos específicos:

- validar con los investigadores unos indicadores que puedan determinar la calidad de las editoriales con publicaciones científicas en ciencias sociales y humanidades;
- conocer las editoriales españolas y extranjeras que tienen más prestigio entre los investigadores;
- identificar los métodos más comunes de selección de originales seguidos por las editoriales a partir de la propia experiencia de los investigadores. Sin la pretensión de describir o analizar los casos particulares, el objetivo es conocer el comportamiento más habitual de las editoriales de carácter científico;
- saber si las editoriales científicas son transparentes en los procesos de selección de manuscritos;
- conocer en qué tipo de editoriales científicas publican los investigadores;
- determinar el proceso de toma de contacto entre investigadores y editoriales para la publicación de un texto;
- identificar los criterios de los investigadores a la hora de seleccionar una editorial para publicar;
- conocer si los investigadores pagan a las editoriales para publicar y/o si estarían dispuestos a ello.

Los estudios de calidad de los libros tienen aún poca trayectoria, a pesar de que han sido demandados por los humanistas y científicos sociales

2. Metodología

El método empleado ha sido la encuesta. El cuestionario utilizado contó con un total de catorce preguntas, divididas en tres apartados:

- perfil del investigador;
- valoración de la calidad de una editorial con publicaciones científicas;
- valoración del proceso de edición de una editorial con publicaciones científicas.

Antes de difundirla entre los destinatarios, fue probada por varios investigadores y expertos en la materia para tener la certeza de que las preguntas estaban claramente planteadas, así como para detectar faltas de funcionalidad en el diseño.

El cuestionario se realizó en html y se alojó en un servidor del *Centro de Ciencias Humanas y Sociales* del CSIC.

Al enviar el cuestionario, los datos alimentaron mediante php una base de datos MySQL y una vez concluido el período de respuesta se exportaron al programa estadístico SPSS 17.

La encuesta se difundió entre investigadores y profesores, solicitando su cooperación por correo electrónico. Se les animó a participar mediante dos mensajes, el 15 de junio y el 15 de julio de 2010. La encuesta estuvo abierta a la participación durante el período del 15 de junio al 22 de agosto de 2010. Con un universo de 11.647 profesores e investigadores españoles de universidad y del CSIC que cuentan al menos con un sexenio de investigación, se obtuvieron 3.045 respuestas, lo que supone una tasa de respuesta del 27%.

Para las disciplinas de Biblioteconomía y Documentación y de Comunicación (que engloba las áreas de Periodismo, Comunicación audiovisual y Publicidad), la encuesta fue enviada a 263 académicos y respondieron 94, lo que supone una tasa de respuesta del 35,5%, un porcentaje elevado para este tipo de estudios.

En el caso de Biblioteconomía y Documentación, la encuesta fue enviada a 92 académicos y respondieron 41, el 44,57%. Hay que destacar que fue el tercer colectivo que más participó, tras Filología eslava y Didáctica de la expresión corporal, que alcanzaron las tasas de respuesta más altas. Por su parte, de Comunicación contestaron 53 de 171, el 31% (27 de Comunicación Audiovisual y Publicidad de 86, y 26 de Periodismo de 85).

A continuación se analizan los resultados obtenidos por los diferentes apartados, según los objetivos propuestos.

3. Resultados

3.1. Perfil del investigador

Organismo en el que trabaja

El 97% de las 91 personas que contestaron a esta pregunta provienen de la universidad, mientras que el 2,2% proceden del CSIC. Las instituciones que más participaron fueron la *Universidad Complutense de Madrid* (18 académicos), *Autònoma de Barcelona* (12) y *Carlos III de Madrid* (10). En la tabla 1 se presentan todas las respuestas obtenidas.

3.2. Valoración de la calidad de una editorial con publicaciones científicas

Características que avalan la calidad de una editorial

En el cuestionario se presentaba una serie de indicadores agrupados para que los investigadores señalaran cuáles les parecían importantes para medir la calidad de una editorial. Analizar esta información puede servir para conocer el grado de aceptación que tienen en la comunidad científica los indicadores que actualmente emplean las agencias de evaluación.

Como muestra la tabla 2, los seis indicadores más importantes a juicio de los investigadores son:

- trayectoria continuada de la editorial, indicado por 80 investigadores (el 85,1% de los que respondieron);

- sistema de evaluación mediante revisores externos (77 investigadores, el 81,9%);
- presencia en bibliotecas especializadas nacionales y extranjeras (73, el 77,7%);
- presencia en bases de datos internacionales de las monografías publicadas por la editorial (70, el 74,5%);
- editorial con colecciones especializadas (67, el 71,3%);
- presencia en librerías especializadas nacionales y extranjeras (67, el 71,3%).

Entre los indicadores más importantes de calidad editorial están la trayectoria continuada y el sistema de evaluación mediante revisores externos

En algunos indicadores hubo diferencia entre las tasas de respuesta de los académicos de Comunicación y de Biblioteconomía y Documentación. Por ejemplo estos últimos señalaron más porcentualmente los siguientes indicadores:

- especialización de la editorial (70,7% Biblioteconomía frente a 37,7% Comunicación);
- coordinación de la obra (46,3% frente a 24,5%);

Organismo	Acrónimo	Frecuencia	Porcentaje
<i>Universitat Autònoma de Barcelona</i>	UAB	12	13,2
<i>Universidad de Alcalá de Henares</i>	UAH	2	2,2
<i>Universitat de Barcelona</i>	UB	3	3,3
<i>Universidad Carlos III de Madrid</i>	UC3M	10	11,0
<i>Universidad Cardenal Herrera CEU</i>	UCH-CEU	1	1,1
<i>Universidad Complutense de Madrid</i>	UCM	18	19,8
<i>Universidad de Granada</i>	UGR	3	3,3
<i>Universidad Jaime I</i>	UJI	1	1,1
<i>Universidad de La Laguna</i>	ULL	2	2,2
<i>Universidad de Murcia</i>	UM	5	5,5
<i>Universidad de Málaga</i>	UMA	3	3,3
<i>Universidad de Extremadura</i>	UNEX	3	3,3
<i>Universidad de León</i>	UNI-LEON	1	1,1
<i>Universidad de Zaragoza</i>	UNIZAR	3	3,3
<i>Universitat Pompeu Fabra</i>	UPF	3	3,3
<i>Universidad del País Vasco</i>	UPV/EHU	4	4,4
<i>Universidad Rey Juan Carlos</i>	URJC	2	2,2
<i>Universitat Rovira i Virgili</i>	URV	3	3,3
<i>Universidad de Sevilla</i>	US	4	4,4
<i>Universidad de Salamanca</i>	USAL	2	2,2
<i>Universidad de Valencia</i>	UV	3	3,3
<i>Universidad de Valladolid</i>	UVA	1	1,1
<i>Consejo Superior de Investigaciones Científicas</i>	CSIC	2	2,2
Total		91	100,0

Tabla 1. Número de respuestas por organismos

- evaluación positiva de la editorial por parte de las agencias (61% frente a 39,6%);
- marketing editorial (24,4% frente a 11,3%);
- estructura adecuada (75,6% frente a 47,2%).

En cambio y sorprendentemente, la presencia de la editorial en bibliotecas especializadas nacionales y extranjeras fue señalada más por los investigadores de Comunicación que por los de Biblioteconomía (84,9% frente a 68,3%).

Conscientes de que la relación de indicadores propuesta podía resultar incompleta, se reservó un espacio para que los encuestados pudieran sugerir otros o matizar el significado de algunos.

Uno de los indicadores que se añaden, y que es mencionado por varios investigadores, es el número de citas que recibe cada obra. No se incorporó en el momento de enviar la encuesta por la falta de fuentes que podían aportar estos valores. No obstante, *Google Scholar* calcula las citas recibidas por libros y *Book Citation Index* también lo hace, aunque

con un universo más limitado. Las reticencias de los humanistas y de algunos científicos sociales hacia los indicadores de impacto también influyeron a la hora de excluir el indicador de la relación.

Al mismo tiempo que se mencionan las citas recibidas por las monografías como indicador de calidad, algunos comentarios se refieren a la cantidad de personas que leen y usan la obra. Al margen de las citas, es difícil conocer estos datos de “consumo” de las obras. La tirada, las ventas de la obra y el número de ediciones y traducciones, las consultas y préstamos en bibliotecas, así como las citas serían los datos más cercanos y más precisos sobre consumo.

Otro parámetro es la diversidad en el tipo de obras que publica una editorial; es decir, que acoja tanto obras individuales como colectivas, de creación original o procedentes de congresos y textos de distinta naturaleza.

También se alude a la profesionalización de la edición, esto es, al hecho de que la editorial cuente con responsables es-

Indicador	Total		Comunicación		Biblioteconomía y documentación	
	Frecuencia	%	Frecuencia	%	Frecuencia	%
Sistema de evaluación de originales mediante revisores						
Comités de lectura especializados	54	57,4	31	58,5	23	56,1
Revisores externos	77	81,9	42	79,2	35	85,4
Información sobre el proceso de selección de originales	52	55,3	29	54,7	23	56,1
Especialización temática						
Editorial especializada	49	52,1	20	37,7	29	70,7
Editorial con colecciones especializadas	67	71,3	39	73,6	28	68,3
Cantidad de publicaciones sobre la materia	19	20,2	11	20,8	8	19,5
Prestigio						
Trayectoria continuada de la editorial	80	85,1	45	84,9	35	85,4
Reseñas positivas en las mejores revistas del área	55	58,5	28	52,8	27	65,9
Número de títulos y tirada de cada uno de ellos	18	19,1	12	22,6	6	14,6
Editor de relevancia en la colección o en la editorial	55	58,5	26	49,1	29	70,7
Coordinador de la obra	32	34,0	13	24,5	19	46,3
Que las agencias de evaluación (<i>Aneca, Cneai</i> , etc.) valoren positivamente a esa editorial	46	48,9	21	39,6	25	61,0
Sistema de difusión y distribución						
Presencia en librerías especializadas nacionales y extranjeras	67	71,3	41	77,4	26	63,4
Presencia en bibliotecas especializadas nacionales y extranjeras	73	77,7	45	84,9	28	68,3
Presencia en bases de datos internacionales de las monografías publicadas por la editorial	70	74,5	38	71,7	32	78,0
Traducciones en distintos idiomas de las monografías publicadas por la editorial	40	42,6	21	39,6	19	46,3
Marketing editorial	16	17,0	6	11,3	10	24,4
Proyección internacional de la editorial	54	57,4	29	54,7	25	61,0
Que la editorial coedite	10	10,6	5	9,4	5	12,2
Calidad de las ediciones						
Presentación formal	51	54,3	29	54,7	22	53,7
Estructura adecuada	56	59,6	25	47,2	31	75,6
Índices de autores, temas, onomásticos, etc.	54	57,4	30	56,6	24	58,5
Formato de las monografías (que también se ofrezcan como libro electrónico o <i>ebook</i>)	41	43,6	20	37,7	21	51,2

Tabla 2. Indicadores de calidad de las editoriales

pecializados en la dirección editorial y en los procesos técnicos, a que haya un control editorial serio y riguroso, no guiado por intereses comerciales, y a que exista un esquema de trabajo bien definido y asumido en cada una de sus fases por especialistas.

La correcta identificación de todos los datos que constituyen la referencia bibliográfica, que son indispensables para catálogos de bibliotecas y bases de datos bibliográficas, es señalada como otra característica importante. Sin duda, se trata de un aspecto que valoran los especialistas de estas áreas (Ciencias de la información) y que probablemente no haya sido destacado en otras disciplinas.

Algún investigador también mencionó la calidad de la obra establecida por los propios expertos del área, es decir, una evaluación cualitativa similar a la llevada a cabo por la agencia catalana de evaluación, en colaboración con la UAB, a modo de proyecto piloto (*AQU Catalunya*, 2010).

Otros aspectos señalados son:

- la selección de títulos por su interés científico y no tanto por sus posibilidades de éxito comercial;
- la permanencia de la obra a lo largo de los años para que no desaparezcan del mercado títulos interesantes;
- que las editoriales acojan obras de autores jóvenes y no sólo de los consagrados;
- que la evaluación no sea anónima;
- que publique manuales de referencia relevantes;
- que sus obras sean útiles tanto para la investigación como para la docencia;
- que editen a los científicos e intelectuales más reconocidos internacionalmente;
- que la editorial tenga proyección internacional; y
- que no esté influenciada por los valores o prioridades del grupo de comunicación o académico al que pertenezca, es decir, que sea imparcial y plural. Este “indicador” es sin duda deseable, pero algo utópico y, sobre todo, difícil de valorar de manera objetiva.

El prestigio de una editorial es una cualidad a la que se refieren las agencias de evaluación a pesar de ser un concepto subjetivo

Valoración de las características generales que avalan la calidad de una editorial

Se pedía asimismo que puntuaran de 0 al 5 los indicadores de calidad de una editorial tratados en la pregunta anterior. Las medias obtenidas por los distintos indicadores son parecidas (gráfico 1) y no presentan desviaciones típicas significativas³, siendo el sistema de evaluación de originales mediante revisores el que logra una mayor valoración (4). En el otro extremo está el formato de libro electrónico que es la característica que menos puntuación ha tenido (2,62). Las otras valoraciones por orden de importancia han sido: especialización temática de la editorial (3,95); prestigio (3,93); sistema de difusión y distribución (3,64); y calidad de las ediciones (3,52).

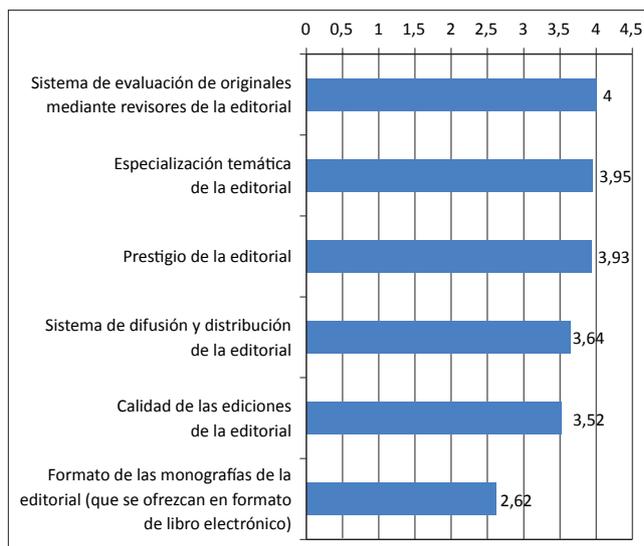


Gráfico 1. Valoración de las características de calidad de una editorial (medias)

Razones de elección de una editorial para publicar

Se incluía una pregunta en la que se pedía a los investigadores que señalaran las razones que les habían llevado a elegir una determinada editorial para publicar. La pregunta era multirespuesta, ya que podían elegir varias opciones de una lista cerrada. El procesamiento de los datos obtenidos señala las siguientes razones por orden de importancia:

- prestigio de la editorial (69% de los que contestaron);
- especialización temática de la editorial o de la colección (65%);
- buen sistema de difusión y distribución en librerías y bibliotecas (46,8%);
- presencia en bases de datos internacionales de los libros publicados (31%);
- buena calidad de la edición (29,8%);
- facilidad en la aceptación del texto ya que se trata de una editorial de la institución en la que trabaja el investigador (12,8%);
- facilidad en la aceptación del texto por tratarse de una editorial en expansión (3,2%).

Si se comparan los resultados de las dos áreas, se observa que en Comunicación se valora más que en Documentación el prestigio de la editorial (77% frente al 59%) y el buen sistema de difusión y distribución en librerías y bibliotecas de los libros publicados (62% frente a 27%), tal como ya se observaba en la pregunta anterior. En el gráfico 2 se muestra la representación de todos los porcentajes obtenidos.

Editoriales españolas más prestigiosas

La evaluación de las editoriales es una posible aproximación a la calidad de los títulos que publica cada una. Aunque éstas pueden exigir comportamientos distintos a cada colección —algunas más científicas y rigurosas, otras más divulgativas y/o comerciales, etc.— lo cierto es que un sello editorial suele estar asociado a un nivel de calidad determinado. Sin duda alguna, el análisis cualitativo de cada obra publicada es el que mejor puede juzgar el trabajo de un investigador. Sin embargo, el volumen de obras publicadas, así como los variados procesos de evaluación que existen hacen impo-

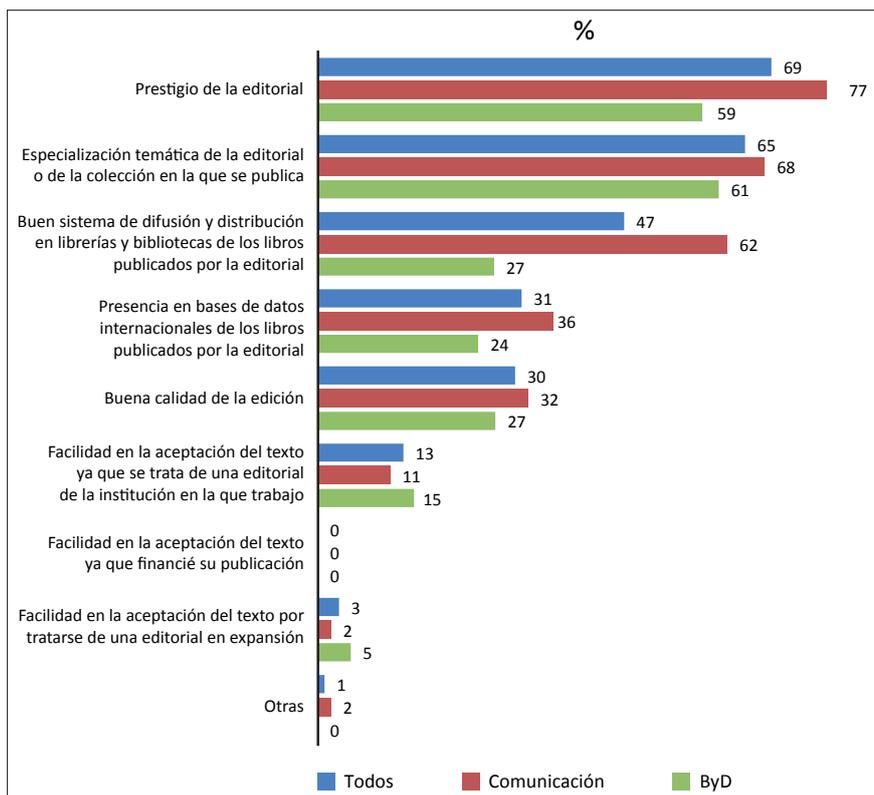


Gráfico 2. Razones de elección de la editorial

sible abordar un análisis de esas características. Por eso, la evaluación de las editoriales representa un ejercicio viable, que permitirá a los gestores de política científica contar con una herramienta orientativa en sus procesos de evaluación.

Conocer las editoriales “prestigiosas” ayuda en los procesos de evaluación

El prestigio de una editorial es factor al que se refieren habitualmente las agencias de evaluación de la actividad científica como indicador de calidad de los libros publicados (Cneai, 2009; Aneca, ANEP, etc.). Sin embargo, el prestigio es en sí mismo un concepto ciertamente subjetivo. Un investigador puede considerar que una editorial es prestigiosa porque haya leído buenos libros publicados por ella, porque le ha publicado los suyos, porque sus ediciones son cuidadas y atractivas, porque sus obras hayan pasado por una evaluación rigurosa en esa editorial o porque la mayor parte de la comunidad científica opine que esa editorial es prestigiosa.

Conocer qué editoriales son descritas como “prestigiosas” por los expertos de un área puede resultar una vía interesante para contar con listados de editoriales de referencia que ayuden o guíen en los procesos de evaluación. Resulta necesario disponer de alguna herramienta en ese sentido, desde el momento en que apenas hay indicadores de calidad directos para evaluar la producción científica a través de los libros y sabiendo que es un elemento clave en los procesos de evaluación.

Así pues, en la encuesta se pedía a los participantes que señalaran, por orden de importancia, las tres editoriales españolas y las tres extranjeras que, a su juicio, tienen más

prestigio en su área. Al ser una pregunta abierta se tuvieron que normalizar todas las editoriales mencionadas.

Los editoriales señalados y los votos que recibieron se muestran en las tablas 3 y 4.

Editoriales extranjeras más prestigiosas

En el caso de las editoriales extranjeras, el planteamiento de la pregunta fue similar a la anterior. Estos resultados indican una mayor dispersión que en el caso de las editoriales españolas, ya que los investigadores han señalado 32 editoriales extranjeras distintas (frente a las 17 españolas) en Biblioteconomía y Documentación, y 43 en Comunicación (frente a las 29 españolas) (tablas 5 y 6).

3.3. Valoración del proceso de edición de una editorial con publicaciones científicas

Investigadores que han publicado una monografía científica en una editorial

Como resulta obvio, interesaba que opinaran y valoraran el proceso de edición en las editoriales aquellos investigadores que tuvieran libros publicados y, por tanto, hubieran pasado por un proceso de selección y de edición. Por esta razón, la primera pregunta que se les hacía era si alguna vez habían publicado libros. Como era de esperar, una mayoría

Editorial	Nº total de votos	Nº votos en 1ª posición	Nº votos en 2ª posición	Nº votos en 3ª posición
Trea	27	18	7	2
Síntesis	19	4	11	4
Fundación Germán Sánchez Ruijérez	6	2	2	2
Pirámide	5	1	1	3
Ra-Ma	2	1	1	
Cátedra	2	1	1	
Arco Libros	3		1	2
CSIC	2	1		1
Ollero y Ramos	1	1		
Anaya	1	1		
Santillana	1	1		
Díaz de Santos	2		1	1
Alianza	2		1	1
Troa	1		1	
Eunsa	1			1
Paidós	1			1
Pearson Educación	1			1
Total	77	31	27	19

Tabla 3. Editoriales españolas con más prestigio en Biblioteconomía y Documentación

de los académicos (93,4%) respondió afirmativamente a la pregunta.

Tipo de editorial en la que han publicado los investigadores

Los investigadores que contestaron afirmativamente fueron preguntados sobre el tipo de editorial en la que habían publicado sus libros. Puesto que usualmente los investigadores tienen más de un libro publicado, la pregunta es multirespuesta y los datos que se ofrecen son frecuencias de publicación en los distintos tipos de editoriales.

Predomina la publicación en:

- editoriales científicas universitarias españolas (55,3% de los investigadores que han respondido a la encuesta);
- editorial científica comercial española (53,2%);
- editorial comercial española con colecciones científicas (41,5%);

Editorial	Nº total de votos	Nº votos en 1ª posición	Nº votos en 2ª posición	Nº votos en 3ª posición
Paidós	30	14	8	8
Cátedra	22	15	1	6
Gedisa	12	1	8	3
Ariel	9	4	3	2
Síntesis	9		7	2
Tecnos	8	2	2	4
Bosch	4	3	1	
Alianza	4	2	2	
Gustavo Gili	3	1		2
Biblioteca Nueva	3		2	1
Fragua	3		1	2
Eunsa	2		1	1
Rialp	2		1	1
Tirant lo Blanch	2		1	1
Universidad del País Vasco	1	1		
Iberoamericana/Vervuert	1		1	
Trotta	1		1	
Universitat de València	1		1	
Caja España (Aprender a mirar)	1		1	
Delta	1		1	
Aldea Global	1		1	
Filmoteca Española	1			1
Pearson Educación	1			1
Pirámide	1			1
Anthropos	1			1
Universitat Oberta de Catalunya	1			1
Castalia	1			1
Crítica	1			1
Diputación de Sevilla	1			1
Total	128	43	44	41

Tabla 4. Editoriales españolas con más prestigio en Comunicación

- editorial científica de organismo de investigación español (18,1%);
- editorial comercial extranjera con colecciones científicas (17%);
- editorial científica universitaria extranjera (14,9%);
- editorial científica comercial extranjera (14,9%);
- autoedición del texto (6,4%);
- editorial científica de organismo de investigación extranjero (3,2%).

Editorial	Nº total de votos	Nº votos en 1ª posición	Nº votos en 2ª posición	Nº votos en 3ª posición
Elsevier	14	6	5	3
Springer	10	7	1	2
McGraw Hill	10	3	4	3
American Library Association	9	3	4	2
De Gruyter	6	2	1	3
Aslib	2	1	1	
Libraries Unlimited	3	1		2
Pearson Education. Addison-Wesley	3		2	1
Universidad Nacional Autónoma de México. CUIB	2		2	
Facet	2		2	
IGI Global	1	1		
John Wiley and sons	1	1		
Asist	1	1		
Oxford University Press	1	1		
Reichenberger	1	1		
Vid. 5	1	1		
ACM	1	1		
Chandos Publishing	1	1		
Haworth Press	1		1	
International Council on Archives	1		1	
Pearson Education	1		1	
Blackwell	1		1	
Scare Crow Prees, Metuchen N. J.	1		1	
Taylor and Francis	1		1	
Academic Press	1		1	
La Documentation Francaise	1			1
Auger	1			1
Pergamon	1			1
Sage	1			1
Unesco	1			1
Unlimited Books	1			1
Alfagrama	1			1
Total	83	31	29	23

Tabla 5. Editoriales extranjeras con más prestigio en Biblioteconomía y Documentación

Editorial	Nº total de votos	Nº votos en 1ª posición	Nº votos en 2ª posición	Nº votos en 3ª posición
Sage	20	12	6	2
Routledge	16	8	6	2
McGraw-Hill	6	1	5	
Macmillan	4	2	1	1
Oxford University Press	6		1	5
Blackwell	3	1	2	
Taylor and Francis	3	2		1
Cambridge University Press	5		2	3
Amorrortu	3	1	1	1
Harmattan	3	1		2
Harvard University Press	2	1	1	
Intellect	4		1	3
Pearson Education	3	1		2
Hampton Press	2		2	
Focal Press	2	1		1
Fondo de Cultura Económica	2	1		1
Hachette	1	1		
Iowa University Press	2		1	1
John Libbey Media	1	1		
Klincksieck	1	1		
Manchester University Press	2		1	1
MIT Press	1	1		
Mondadori	2		1	1
Pearson	1	1		
Praeger	1	1		
Princeton University Press	1	1		
California University Press	2		1	1
Facultade Social de Bahia	1	1		
Ashgate	1		1	
Nomos	1		1	
Trillas	2			2
Columbia University Press	1		1	
Gallimard	1		1	
Éditions du Nouveau Monde	1		1	
La Crujja	1			1
La Découverte	1			1
Polity	1			1
Presses Universitaires de France	1			1
Unesco	1			1
University of Illinois	1			1
De Boeck	1			1
Elsevier	1			1
Free Press	1			1
Total	117	40	37	39

Tabla 6. Editoriales extranjeras con más prestigio en Comunicación

Estos valores se refieren al conjunto de las dos áreas.

Al comparar los datos obtenidos entre investigadores de Biblioteconomía y Documentación y de Comunicación, se aprecia que estos últimos publican más en el extranjero. Sólo 4 investigadores en Biblioteconomía señalaron haber publicado en editoriales extranjeras, un dato indicativo del tipo de investigación más bien local que se realiza en el área o de la escasa proyección internacional que se le está dando a investigaciones que sí que podrían tener interés fuera. Así el 24,5% de los investigadores en Comunicación han publicado en editoriales científicas universitarias extranjeras frente a tan solo al 2,4% de los de Biblioteconomía. Lo mismo ocurre con las publicaciones en editoriales científicas de organismos de investigación extranjeros (5,7% frente al 0%); la editorial científica comercial extranjera (20,8% frente a 7,3%); y la editorial comercial extranjera con colecciones científicas (30,2% frente al 0%).

Los investigadores de Comunicación publican más en el extranjero que los de Biblioteconomía y Documentación

Otra diferencia acusada entre las áreas se da en las editoriales de organismos de investigación españoles y extranjeros. En el primer caso, destaca Biblioteconomía y Documentación sobre Comunicación (con una ventaja de 14 puntos) por una clara razón: no hay apenas centros de investigación no vinculados a universidades en el área de la Comunicación, algo que sí ocurre con la Biblioteconomía pues, al menos dos institutos del CSIC se dedican a la Documentación científica.

El gráfico 3 muestra la diferencia de frecuencias para cada una de las dos disciplinas.

Proceso de toma de contacto con la editorial para la publicación del texto

En cuanto al proceso que siguieron los investigadores para que la editorial les publicara el texto, las opciones más señaladas por orden de importancia fueron:

- a) los investigadores se pusieron en contacto con la editorial en un 55,3% de los casos;
- b) la editorial se puso en contacto con los investigadores en un 48 % de las ocasiones;
- c) en un 51% fue el coordinador de una obra colectiva quien se puso en contacto con ellos para la redacción de un capítulo.

Otra forma, aunque ya de importancia menor, es la existencia de un concurso de carácter científico cuyo premio era la publicación del texto (2,1%).

Los datos obtenidos muestran que en Biblioteconomía y Documentación es más frecuente que sean los editores quienes se pongan en contacto con los autores, mientras que en Comunicación es más habitual que sea el autor quien tome la iniciativa.

Los distintos porcentajes de respuesta pueden observarse en el gráfico 4.

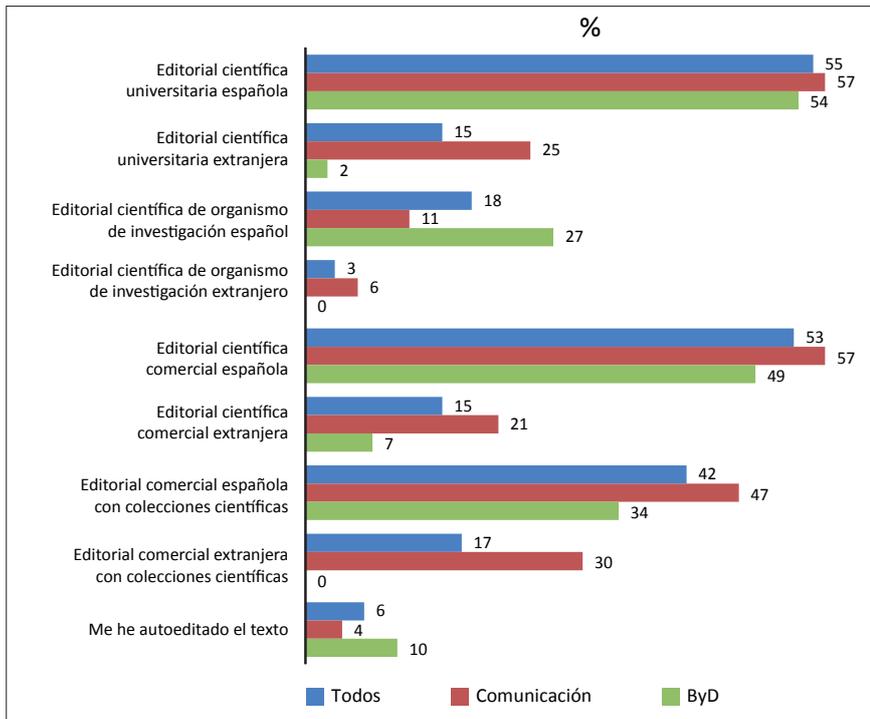


Gráfico 3. Tipo de editorial en la que han publicado

Claridad de la información que ofrecen las editoriales en relación con el envío de textos para su publicación

Uno de los aspectos que más diferencia a las editoriales de monografías frente a las revistas científicas es que no suele ser muy clara la información sobre la propuesta de textos para su publicación. Así se desprende de los resultados de esta encuesta, que incluía una pregunta sobre este asunto. Así, el 52% de los investigadores que contestaron a esta pregunta señalaron que las editoriales sí que ofrecen una información clara para el envío de textos, pero también un 25% indicó que no y un 22% no tiene una opinión definida al respecto (gráfico 5).

Tanto en las respuestas positivas como en las negativas, parece que el colectivo de investigadores en Comunicación es más crítico en relación a la información proporcionada por las editoriales. El grupo de editoriales sobre el que pueden opinar es mayor y también lo es la masa crítica de investigadores. Por otra parte, esta comunidad científica se ha mantenido bastante crítica en relación a los asuntos de publicación y evaluación científica.

Los comentarios abiertos que los académicos realizaron en torno a esta pregunta resultaron especialmente expresivos. La opinión más extendida es que las editoriales son poco transparentes e incluso “opacas” en cuanto a la información que ofrecen para enviar un manuscrito. En algún caso se menciona:

- este comportamiento es propio de editoriales españolas, pero no de las estadounidenses;
- con la falta de información tratan de frenar el aluvión de originales que hay, debido a la presión por publicar;

- los autores deben ser insistentes en la petición de información a las editoriales hasta que la logran obtener;
- los autores desconocen, en términos generales, si su obra puede resultar interesante o no para una editorial, precisamente por esa falta de información.

Claridad de la información que ofrecen las editoriales en relación al proceso de aceptación de los textos.

Complementando la pregunta anterior se pedía a los académicos que señalaran si les parecía claro el proceso de aceptación de textos. Predominan las respuestas afirmativas (39%), aunque contestaron negativamente un 32% y, por otra parte, un 29% no tienen una clara información al respecto. Estos dos últimos porcentajes son lo suficientemente altos como para que cobre un mayor interés el análisis de las preguntas abiertas. En términos generales, se observa que hay un cierto desconcierto

con la información que ofrecen las editoriales y se reitera la idea de que el autor desconoce el proceso de evaluación que sigue una editorial. Por una parte, está claro que las editoriales no siguen un protocolo informativo hacia sus autores parecido al que se sigue en las revistas científicas, donde el autor suele estar informado en cada momento de la fase y la decisión que afecta a su original. Entre los comentarios, se repite la idea de que entre el envío del texto completo y la recepción de alguna información sobre su interés para la editorial, media una falta de comunicación entre autor y editor. Incluso se llega a afirmar que el autor ha de realizar una continuada labor de intento de contacto con la editorial para lograr información sobre el manuscrito.

Un aspecto que diferencia a las editoriales de monografías de las de revistas científicas es que no suele ser muy clara la información sobre la propuesta de textos para su publicación

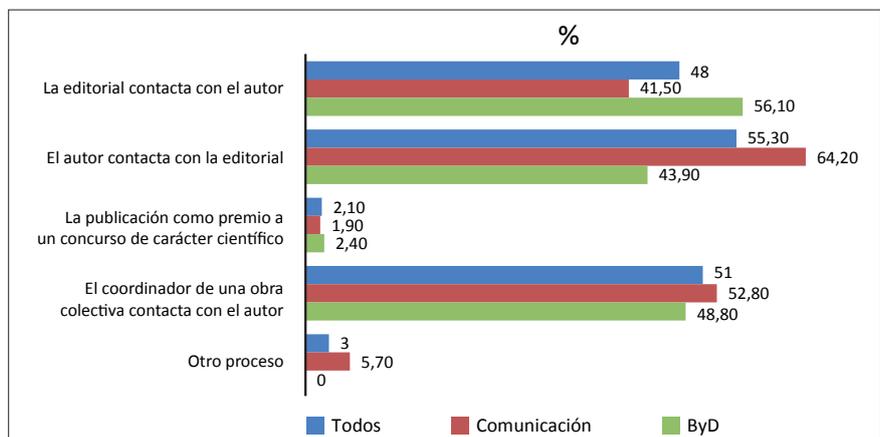


Gráfico 4. Forma de contacto entre la editorial y el autor

Por otra parte, se mencionan algunas cuestiones realmente delicadas y también interesantes que convendría analizar, como por ejemplo, que las decisiones finales de publicación por parte de una editorial no sólo están relacionadas con la calidad de la obra sino también con su vinculación a asignaturas troncales que garanticen, de algún modo, la viabilidad económica de la edición. También se alude a que los grupos de presión impiden el acceso a la publicación a algunos autores, a la fase de “letargo” en que entran los originales a la espera de una evaluación e incluso a la creencia de que las editoriales publican todo lo que se les envía.

En cuanto a la comparativa por áreas, que se muestra en el gráfico 6, nuevamente se detecta una mayor actitud crítica de la comunidad especializada en Comunicación que puede tener que ver con el movimiento crítico que se ha dado en la

profesión en torno a los sistemas de evaluación, o bien con el propio comportamiento, en conjunto, de las editoriales de esa área, que las haga distintas a las del área de Biblioteconomía y Documentación.

Proceso de revisión científica por parte de la editorial

Para conocer el proceso de revisión científica que llevan a cabo las editoriales, en primer lugar se preguntaba a los investigadores si se habían sometido a estos procesos. El 38% señaló que sí; en algunas ocasiones un 32,5%; un 19% que lo desconoce; y un 10% responde negativamente.

Forma del proceso de revisión científica por parte de la editorial

Con posibilidad de múltiple respuesta, el 51,5% de los investigadores señalaron el comité de lectura de la propia editorial como el órgano encargado de la evaluación de la obra.

Con porcentajes ya menores aparecen el resto de opciones: evaluación mediante un informe de un evaluador externo (28,7%); revisión por pares anónimos (26,6%); y revisión por pares no anónima (2,1%). Además, dos investigadores señalaron que la revisión fue realizada por el coordinador de la colección o del libro entero (gráfico 7).

Comunicación de los motivos de un rechazo editorial

La encuesta incluyó una pregunta sobre si las editoriales habían dado razones a los autores para justificar el rechazo de un manuscrito. Contestaron 59 personas, de las cuales 32 indicaron que no procedía, es decir, que sus obras no habían sido rechazadas. 10 indicaron que siempre se les había proporcionado esa información y 9 que no la habían recibido. También cabe señalar que 3 investigadores indicaron que sí se les proporcionó la información, pero no con suficientes detalles.

Pago a la editorial para la publicación

En una pregunta directa se pedía a los investigadores que señalaran si alguna vez habían pagado a una editorial para la edición y publicación de un texto científico suyo. 6 investigadores señalaron que sí que lo habían hecho, lo que representa el 8,3% de los que contestaron a la pregunta. Las 6 respuestas procedieron del área de Comunicación (gráfico 8).

Intención de pagar a la editorial para la publicación

Complementando la pregunta anterior, se les pedía que indicaran si estarían dispuestos a pagar a una editorial para la edición de un texto científico suyo.

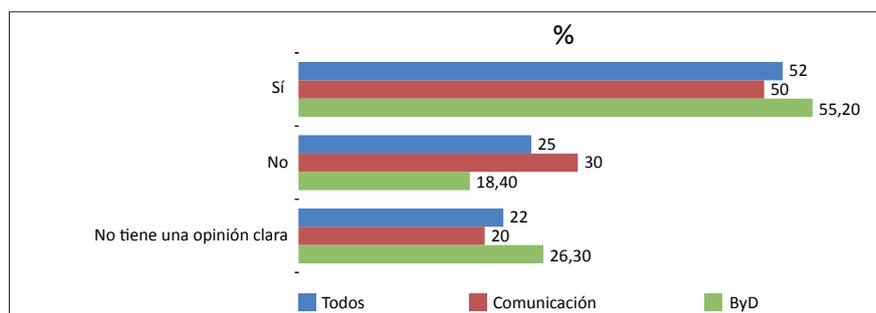


Gráfico 5. Opinión sobre la claridad de las editoriales en el proceso de envío de textos

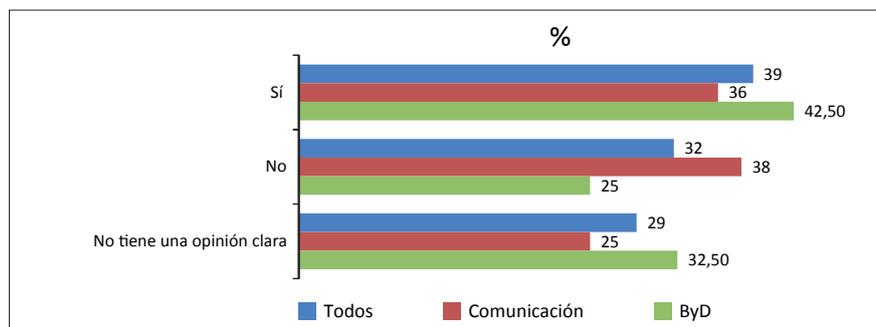


Gráfico 6. Opinión sobre la claridad de las editoriales en el proceso de revisión de textos

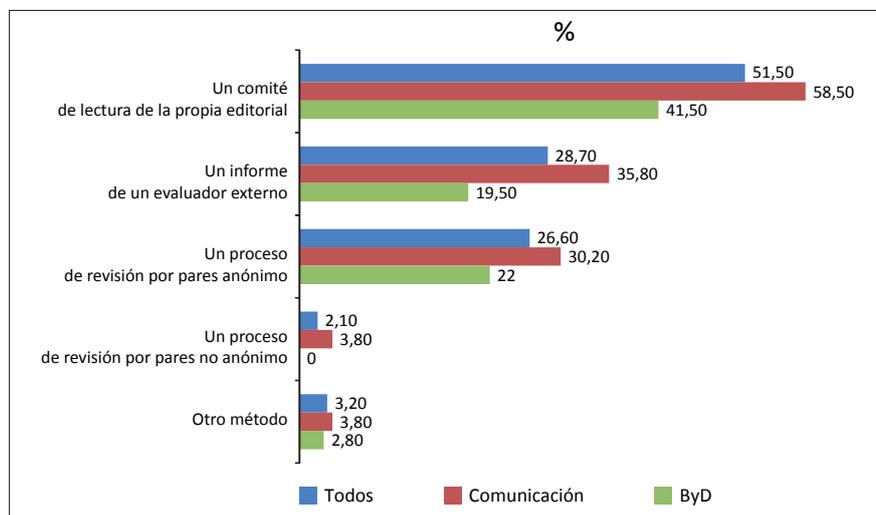


Gráfico 7. Proceso de revisión científica

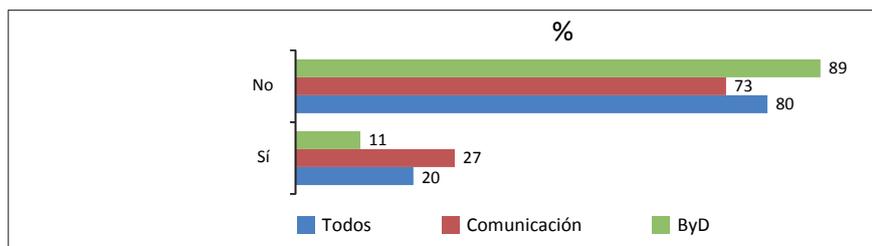


Gráfico 8. Investigadores que pagarían a una editorial para publicar

También el porcentaje mayoritario de respuesta ha sido el “no” (80%), aunque hubo 17 casos que señalaron que sí que podrían pagar (un 20%). Por áreas, los investigadores de Comunicación se muestran más abiertos a la idea del pago para publicar.

4. Discusión y conclusiones

Son muchos los aspectos que intervienen en el concepto de “calidad” de una editorial. Sin embargo, parece que son sólo unos pocos los realmente determinantes. Al poner en relación los resultados de los dos primeros apartados de este trabajo (valoración de las características de calidad de una editorial y razones para elegir una editorial para publicar) se observa bastante coincidencia en los resultados. Básicamente, se considera muy relevante la especialización temática, el prestigio y el sistema de difusión y distribución editorial. Resulta llamativo que la evaluación por expertos se considere el aspecto fundamental (y el que presenta más consenso en la comunidad científica) para conocer la calidad de una editorial y, sin embargo, no aparece mencionado como un criterio selectivo. Cabe imaginar que todos los investigadores presuponen que esa evaluación del original se llevará a cabo de forma rigurosa pero, en cualquier caso, destaca que no se señale como una condición para seleccionar una editorial.

Las editoriales deberían ser más transparentes en cuanto a sus procesos de selección, habida cuenta de que desempeñan su actividad en el ámbito científico

En cuanto al “prestigio editorial” que también aparece destacado, existen problemas para definirlo y objetivarlo, pero al mismo tiempo es un aspecto clave en la evaluación de la actividad científica. Es relevante conocer qué editoriales son consideradas prestigiosas por la comunidad científica, tratando así de facilitar una aproximación sobre esa difusa variable de calidad.

Atendiendo a la relación de editoriales españolas más prestigiosas que se obtiene de esta encuesta destaca que, tanto en Documentación como en Comunicación, las editoriales más “votadas” son las más especializadas de cada disciplina, las que publican más títulos sobre ambas disciplinas. Probablemente no todos los títulos de esas editoriales tengan la misma calidad o no todas hayan pasado por el mismo proceso de selección, pero la especialización hace que la editorial sea referente para el investigador. En Documentación destacan las editoriales *Trea*, *Síntesis* y *Fundación Germán*

Sánchez Ruipérez, mientras que en Comunicación lo hacen *Paidós*, *Cátedra*, *Gedisa* y *Ariel*.

Cabe añadir que se citan editoriales que, aunque no mantengan su actividad editorial o hayan cerrado, los investigadores reconocen como relevantes en el área por las obras editadas en el pasado (como la *Fundación Germán*

Sánchez Ruipérez). También se citan editoriales que no son especializadas en el área (como *Pirámide* y *Ra-Ma*) pero que publican obras de disciplinas conexas para el área como la Informática.

Entre el resto de editoriales mencionadas predominan las de carácter multidisciplinar, aunque de vez en cuando aparece alguna más especializada como *Arco libros* y *Eunsa* para Documentación o *Fragua* y la *Filmoteca española* para Comunicación. Esta tendencia se invierte en el caso de las editoriales extranjeras ya que las que aparecen como más relevantes son reconocidas en el resto de disciplinas, grandes editoriales multidisciplinarios o grandes grupos editoriales. Varían, eso sí, de una disciplina a otra. En Documentación destacan *Elsevier*, *Springer* y *McGraw Hill*, mientras que en Comunicación son *Sage*, *Routledge* y *McGraw Hill* las que aparecen mejor posicionadas.

Se detecta una mayor dispersión en las editoriales extranjeras citadas, quizá porque la oferta editorial extranjera y especializada sea mayor. Será interesante comprobar si esto es así en todas las disciplinas.

Otro de los aspectos analizados en este trabajo se refiere al tipo de editoriales elegidas para publicar. Los resultados muestran una predominante pauta de publicación: se publica más en editoriales españolas que en las extranjeras. Es decir, a partir de las personas que contestaron la encuesta, se puede afirmar que la internacionalización, medida por la publicación en editoriales extranjeras, aún no es una realidad. Sin embargo, es muy probable que este comportamiento vaya cambiando, tal y como ha sucedido con la publicación de artículos en el campo de la Comunicación (Masip, 2011; Fernández-Quijada, 2011). Sí es cierto que se detectan mayores porcentajes de publicación en editoriales extranjeras en el caso de la Comunicación que en el de la Documentación, lo que tiene que ver —probablemente— con una mayor madurez de la Comunicación como disciplina, una mayor trayectoria en investigación y una mayor masa crítica.

Aunque la participación podría haber sido más alta, los resultados son sin duda orientativos. Una próxima edición de la encuesta —prevista para 2012— permitirá ratificar o corregir los resultados. Esta primera fase del proyecto permite conocer el núcleo de editoriales reconocidas por los investigadores. En cualquier caso, si estos resultados se emplearan de forma orientativa —sólo podría hacerse así— en algún proceso de evaluación, sería necesario acompañarlos de la opinión de algunos expertos del área que pudiera valorar aquellas editoriales que no hayan aparecido en el estudio y que puedan resultar relevantes.

Respecto a la forma de contacto entre autor y editor para iniciar el proceso de publicación de una obra, se identifica una

práctica predominante distinta en cada caso. En Documentación es más frecuente que la editorial contacte con el autor, y ello puede tener que ver con la menor trayectoria de las editoriales especializadas en el área. La Documentación como disciplina universitaria y de investigación es joven y eso tiene su reflejo en un menor número de editoriales especializadas. Puede explicarse, por tanto, que en estos primeros años de andadura, las editoriales hayan tenido que buscar a sus autores. Nuevamente, y como señal de una mayor madurez, los hábitos en Comunicación son diferentes y el mayor porcentaje de respuestas (64%) corresponde a la fórmula en la que la editorial se pone en contacto con el autor.

Predominan las críticas de los investigadores hacia la falta de información y transparencia de las editoriales sobre política editorial, selección y envío de manuscritos, etc.

Una de las conclusiones más destacadas de este trabajo se refiere a la percepción que tienen los autores de las editoriales: predominan las críticas hacia la falta de información y transparencia sobre la política editorial, la selección y envío de manuscritos, etc. Esto, sin duda, debe ser conocido por las editoriales que, actualmente, lo quieran o no, desempeñan su labor en un entorno muy competitivo (el académico) y deben adecuarse a las mismas exigencias de calidad que se dan en las revistas científicas. Al fin y al cabo, el núcleo de su actividad es la investigación llevada a cabo y escrita por los investigadores. Sin las obras no hay editoriales. Y parece que la comunidad científica demanda más claridad y rigor en los procedimientos.

Los porcentajes de “desconocimiento” de las prácticas editoriales son abultados y señal inequívoca de que las editoriales deberían ser más transparentes en cuanto a sus procesos de selección, habida cuenta de que desempeñan su actividad en el ámbito científico, en el que la evaluación por expertos debe ser un aspecto clave. De hecho, tal y como apuntan los resultados de este estudio, la evaluación por expertos es un elemento determinante de la calidad de una editorial. El hecho de que las editoriales adopten una política de transparencia en este asunto no anularía su autonomía como empresas, sino más bien ayudaría a mejorar su marca, más vinculada en ese caso a calidad editorial.

Los porcentajes registrados en la pregunta sobre procesos de revisión científica a los que se han sometido los investigadores no son buena señal de lo extendidos (o poco extendidos) que están los procesos de revisión. Sin duda, convendría que las editoriales no sólo ofrecieran información más clara y completa sobre los procesos de evaluación sino, sobre todo, que los aplicaran y fuera evidente para autores y para lectores.

Otra de las conclusiones es que la revisión externa de los originales no es la más extendida, sino que es más habitual la realizada por los comités de lectura de la propia editorial. Nuevamente surge aquí la inevitable comparación con las revistas científicas, a las que se exige, cada vez más, que la

evaluación sea externa a la entidad editora y a los consejos de redacción. Sería necesario reflexionar sobre la autonomía de los procedimientos de selección en editoriales –mayoritariamente privadas– frente a los cada vez más reglamentados procedimientos de selección en revistas científicas –mayoritariamente editadas por instituciones públicas–. Dicho de otro modo, cabe plantearse si deben prevalecer las condiciones de la publicación científica (revisada por expertos) o si deben hacerse distinciones en función de los objetivos de las editoriales.

Por otra parte, aunque las respuestas sobre la práctica del pago por publicar son muy limitadas y de ellas se deduciría que no es habitual, lo cierto es que entre los comentarios surge de forma recurrente la idea de que muchos investigadores pagan por publicar precisamente por la presión que existe sobre ellos de contar con publicaciones. También se alude a que en la solicitud de proyectos ya se puede contemplar alguna partida para publicar pero ciertamente esto abre un debate sobre la financiación de la investigación y la publicación, sobre los intereses que se crean y sobre la posible preponderancia del dinero sobre el criterio de excelencia científica.

La comunidad científica de humanistas y, en menor medida, científicos sociales viene reclamando desde hace tiempo la consideración o mejor valoración de las monografías en los procesos de evaluación de la actividad científica. Los comentarios vertidos en esta encuesta son una muestra más de ello –además de reflejar la creciente presión por publicar–, y los resultados de este estudio representan un primer paso en la concreción de esas valoraciones.

La revisión externa de los originales no es la más extendida, sino que es más habitual la realizada por los comités de lectura de la propia editorial

5. Notas

1. *Agència per a la Qualitat del Sistema Universitari de Catalunya (AQU Catalunya)*. <http://www.aqu.cat>
2. *Categorización de publicaciones científicas en Ciencias Humanas y Sociales*. 2004410E607. Financiado por el CSIC.
3. La mayor desviación típica (1,3) se ha dado en la opción de respuesta sobre el formato electrónico.

6. Agradecimientos

Los autores agradecen su participación a los investigadores y profesores que respondieron la encuesta, así como el trabajo técnico de **Sonia Jiménez**, y de la *Unidad de Análisis Estadístico* y de la *Unidad TIC del Centro de Ciencias Humanas y Sociales (CCHS)* del CSIC.

7. Bibliografía

Aneca. *Programa Academia. Principios y orientaciones para la aplicación de los criterios de evaluación*. Madrid: Aneca, 2008.

http://www.aneca.es/var/media/557171/academia_ppiosorientaciones_100616.pdf

Anep. *Criterios de calidad en la investigación en humanidades*. Madrid: Anep/Fecyt, 2007.

AQU Catalunya. *L'avaluació de la recerca en humanitats i ciències socials*. Barcelona: AQU Catalunya, 2010.
http://www.aqu.cat/publicacions/quaderns_qualitat/qqhum_ccsoc.html

BOE. *Resolución de 18 nov. 2009, de la Presidencia de la Comisión Nacional Evaluadora de la Actividad Investigadora, por la que se establecen los criterios específicos en cada uno de los campos de evaluación*, 2009, pp. 102491-102505.
<http://www.boe.es/boe/dias/2009/12/01/pdfs/BOE-A-2009-19218.pdf>

Cullars, John M. "Citation characteristics of monographs in the fine arts". *Library quarterly*, 1992, v. 62, n. 3, pp. 325-342.

Cullars, John M. "Citation characteristics of English-language monographs in philosophy". *Library & information science research*, 1998, v. 20, n. 1, pp. 41-68.

Fernández-Quijada, David. "De los investigadores a las redes: una aproximación tipológica a la autoría en las revistas españolas de comunicación". En: *Simposio AEIC 2011. Investigar la comunicación en España: proyectos, metodologías y difusión de resultados*.
<http://eprints.rclis.org/handle/10760/15566>

Giménez-Toledo, Elea; Román-Román, Adelaida. "As-

essment of humanities and social sciences monographs through their publishers: a review and a study towards a model of evaluation". *Research evaluation*, 2009, v. 18, n. 3, pp. 201-213.
<http://dx.doi.org/10.3152/095820209X471986>

Glänzel, Wolfgang; Schoepflin, Urs. "A bibliometric study of reference literature in the sciences and the social sciences". *Information processing and management*, 1999, v. 35, pp. 31-44.

Hemlin, Sven; Gustafsson, M. "Research production in the arts and humanities: a questionnaire study of factors influencing research performance". *Scientometrics*, 1996, v. 37, n. 3, 417-432.

Masip, Pere. "Los efectos del efecto *Aneca*: análisis de la producción española en Comunicación en el *Social Science Citation Index* (1999-2009)". En: *Simposio AEIC 2011. Investigar la comunicación en España: proyectos, metodologías y difusión de resultados*.
<http://www.revistacomunicar.com/pdf/2011-04-Masip.pdf>

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

Moed, Henk F.; Luwel, Marc; Nederhof, Anton J. "Towards research performance in the humanities". *Library trends*, 2001, v. 50, n. 3, pp. 498-520.
http://findarticles.com/p/articles/mi_m1387/is_3_50/ai_88582627

¿QUÉ MEJOR QUE
DOMINAR LOS PROCESOS
DE LA INFORMACIÓN
EN LA SOCIEDAD DE
LA INFORMACIÓN?

ESTUDIOS DE CIENCIAS DE LA INFORMACIÓN Y DE LA DOCUMENTACIÓN

GRADOS

INFORMACIÓN Y DOCUMENTACIÓN
COMUNICACIÓN

POSGRADOS

GESTIÓN DE LA INFORMACIÓN Y EL CONOCIMIENTO
ANÁLISIS Y MEDICIÓN DE LA INFORMACIÓN UOC-ACCESO
ARQUITECTURA DE LA INFORMACIÓN Y REDES SOCIALES
EL PROFESIONAL DE LA INFORMACIÓN DE SALUD EN LA SOCIEDAD
2.0 UOC-ANIS
PERIODISMO MULTIMEDIA UOC-ACN
PERIODISMO DIGITAL UOC-EL PERIÓDICO-LAVINIA
INNOVACIÓN EN CREACIÓN DE CONTENIDOS AUDIOVISUALES
UOC-DIGITALENT
COMUNICACIÓN CORPORATIVA E INSTITUCIONAL
ENTRETENIMIENTO UOC-TV3
PUBLICIDAD EN INTERNET UOC-DOUBLE YOU
APRENDIZAJE Y TRABAJO COLABORATIVO EN RED

Acceso Abierto



Consulta todos los programas de los Estudios de Ciencias de la
Información y Comunicación en www.uoc.edu





USO Y CREDIBILIDAD DE FUENTES PERIODÍSTICAS 2.0 EN PORTUGAL Y ESPAÑA



João Canavilhas y Begoña Ivars-Nicolás



João Canavilhas, licenciado en comunicação social por la *Universidade da Beira Interior, UBI* (Portugal) y doctor por la *Universidad de Salamanca* (España), es profesor en la *UBI* donde imparte clases de webperiodismo, periodismo radiofónico e infografía multimedia. Además es director del Mestrado em Jornalismo, vicepresidente del *Centro de Investigação Labcom-UBI* y director del diario online *URBI*, el primer periódico universitario portugués en la Web.

*Universidade da Beira Interior
Faculdade de Artes e Letras
Av. Marquês d'Ávila e Bolama
6200 001 Covilhã, Portugal
jc@ubi.pt*



Begoña Ivars-Nicolás es profesora de periodismo en la *Universidad Miguel Hernández (UMH)* de Elche y de comunicación audiovisual en el *Centro de Estudios Ciudad de la Luz* adscrito a esta misma universidad. Doctora en antropología social, centra sus investigaciones en los medios de comunicación en la Red y en la repercusión de las TIC en la docencia. Coedita la revista digital de comunicación audiovisual *Croma* y codirige las ediciones del *CICAP Congreso Internacional de Comunicación Audiovisual y Publicidad* de la *UMH*.

*Universidad Miguel Hernández
Av. de la Universidad, s/n
03202 Elche (Alicante), España
bivars@umh.es*

Resumen

Internet ha cambiado las rutinas de producción en el periodismo. La búsqueda de información ha sido una de las áreas más afectadas, con la aparición de fuentes 2.0 como blogs y redes sociales. A pesar del aparente valor añadido de estas nuevas fuentes de información, ¿qué utilización práctica les dan los periodistas portugueses y españoles y qué credibilidad les atribuyen? Se intenta responder a estas dos cuestiones, definiendo las aplicaciones que pueden considerarse fuentes 2.0 y en qué medida los periodistas las estiman creíbles y fiables. Los resultados muestran que los periodistas de ambos países siguen utilizando internet diariamente para buscar material para su trabajo, documentarse, recibir y enviar información y verificarla, pero aún confían poco en las fuentes 2.0.

Palabras clave

Internet, Periodismo, Fuentes de noticias, Fuentes de información, Credibilidad, Web, Web 2.0.

Title: Use and credibility of news sources 2.0 in Portugal and Spain

Abstract

Internet has changed news making. The search for information has been one of the most affected areas, with the emergence of news sources 2.0, such as blogs and social networks. Despite the apparent value of these new sources of information, what practical use and credibility do Portuguese and Spanish journalists give to these sources? In this paper we define what kind of tools can be considered "2.0 sources" and the ways in which these journalists consider them credible and reliable. The results show that the journalists from both countries are using the internet daily to find material for their work, get factual information, receive and send information and verify it, but still have little confidence in the 2.0 sources.

Keywords

Internet, Journalism, News sources, Information sources, Credibility, Web, Web 2.0.

Canavilhas, João; Ivars-Nicolás, Begoña. "Uso y credibilidad de fuentes periodísticas 2.0 en Portugal y España". *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 63-69.

<http://dx.doi.org/10.3145/epi.2012.ene.08>

Introducción

En el origen de toda noticia existe una fuente de información. Pero ¿de qué hablamos cuando nos referimos a fuentes de información? De acuerdo con **Herbert Gans**, en el contexto del periodismo, fuentes son “*los actores que los periodistas observan o entrevistan, incluidos los entrevistados que hablan por los medios o que son citados en artículos de revista, y también los que sólo suministran información de antecedentes o sugieren temas y narraciones*” (2004, p. 80).

Los actores observados o entrevistados a los que se refiere **Gans** asumen diferentes posiciones en la taxonomía de fuentes según sea su implicación con el hecho objeto de la noticia. Si tomamos como punto de partida la clasificación más generalista, las fuentes pueden ser oficiales o no oficiales, y es en este campo donde internet ha introducido cambios significativos, con nuevas formas de buscar información, localizar expertos y contactar con fuentes no oficiales (**Pavlik**, 2001).

En este trabajo hacemos un recorrido por los cambios en las rutinas periodísticas de los profesionales ibéricos, sobre todo en la fase de recoger temas e informaciones para noticias. Seguidamente definimos el término fuentes 2.0, integrándolo en el concepto grupal de web 2.0. Para terminar el marco teórico, se discute la credibilidad de las fuentes de información en la Web y los criterios que emplean los usuarios para seleccionar la información más creíble. En el estudio empírico se presentan los resultados de una encuesta a periodistas de España y Portugal sobre las fuentes 2.0 que utilizan y el grado de credibilidad que les reconocen.

1. Internet en las rutinas de los periodistas: Portugal y España

Desde los años 90 del siglo pasado la práctica profesional del periodismo ha experimentado notables cambios como consecuencia del uso de internet en las redacciones. El aspecto más visible ha sido el nacimiento de periódicos online, pero la Red también ha transformado las rutinas de trabajo. La búsqueda de material para reportajes, la recogida de informaciones para complementar las noticias y el contacto con las fuentes han sufrido importantes alteraciones. La integración de aplicaciones online en el trabajo diario ha acelerado el proceso de producción de noticias.

1.1. Portugal

La primera investigación sobre internet y periodistas realizada en Portugal (**Bastos**, 1998) reveló que éstos lo utilizaban para buscar información (100%), obtener datos más actualizados (93,5%), leer y enviar correo electrónico (87,5%) e identificar expertos para aclarar dudas relacionadas con el material de sus noticias (62,5%).

Seis años después, **Canavilhas** (2004) repitió la encuesta obteniendo cifras similares en búsqueda de información (97,5%), obtención de datos actualizados (83,9%) y utilización del correo electrónico (92,6%), pero distinta en la identificación de expertos (30,9%). El cambio podría estar justificado porque en este estudio existía la opción alternativa “Contactar fuentes” que recibió el 53,5% de las preferen-

cias. Las aplicaciones más utilizadas eran la Web (97,5%), el correo electrónico (93,8%), los foros (18,5%), los grupos de noticias (16%) y los chats (9,9%). Se preguntaba igualmente por la percepción que tenían los periodistas de la fiabilidad de la información de internet: el 32% la consideró fiable, el 37% se mostraba sin opinión y el 21% no la consideraba creíble.

En 2006 **Obercom** realizó un estudio más detallado en el que participaron 341 periodistas de 14 medios de comunicación portugueses. El 99,4% de los encuestados valoró internet como “importante”¹ para el ejercicio de sus funciones profesionales. En general, los más jóvenes atribuyeron más valor al papel de este medio en sus rutinas diarias. El 99,7% lo utilizaba en su profesión, casi todos con un uso diario (99,4%), y el 87% manejaba diferentes aplicaciones online más de cinco veces al día.

Con referencia a las aplicaciones de la Red, el 99,1% consideró los sitios web “importantes” para su trabajo y el correo electrónico presentó valores muy similares (98,8%). En el análisis de este medio como fuente, el 96,7% consideró las búsquedas online “importantes”. Subrayando este hecho, el 85,2% respondió que la utilizaba diariamente. El uso de grupos de noticias como fuente fue poco valorada: solamente el 22% la consideró “importante”, una cifra inferior a los 30,5% de las listas de correo y los 31,4% de los blogs.

En 2010 **Obercom** ha llevado a cabo otro estudio con el propósito de “obtener una percepción de los valores, prácticas y actitudes que caracterizan la profesión” (2010, pág. 7). En un capítulo dedicado a las fuentes, internet destaca una vez más como primera posibilidad de recogida de información (51,9%), superando en más del doble a la siguiente fuente, la prensa (24,6%). Los periodistas de prensa son los que más optan por internet (67,1%), seguidos por los de radio (52,2%) y los de televisión (40,6%). La encuesta igualmente ha indagado sobre el tipo de aplicaciones que utilizan los encuestados: de las opciones propuestas, los lectores de *feeds* son la fuente más utilizada (30,2%), seguido de los *UGC* (*user-generated content*) (29,2%), los blogs (25,5%), el *crowdsourcing* (22,6%) y el *podcast* (18,9%). Hay que señalar que varias de estas categorías se solapan.

1.2. España

Entre los trabajos que abordan el uso de internet en las redacciones españolas o su credibilidad, se encuentra el realizado por **AccesoGroup** (2002) que encuestó a 616 periodistas. Los datos obtenidos mostraron que los profesionales la empleaban como fuente de información (97%): el 49% usaba la Red para todas sus noticias y el 48% sólo en la mitad de sus trabajos. El 96% le atribuyó mucha credibilidad; el 60% la consideró una fuente de información y el 40% una fuente de documentación: entre los dos grupos, los primeros dieron mayor credibilidad a internet.

El *Col·legi de Periodistes de Catalunya* estudia desde 1992 el perfil del periodista mediante encuestas. En un estudio de 2002, **Masip** concluyó que el 85% de los profesionales catalanes utilizaban internet diariamente: correo electrónico (75%), cibermedios, buscadores y webs institucionales. Sin embargo, el 95% de los periodistas consideraba la Red como una fuente de información carente de confianza de-

bido a la dificultad en verificar la autoría y a la falta de credibilidad, prefiriendo sistemas tradicionales como el cara a cara, el teléfono y la prensa. En 2005, **Masip** reforzó su estudio cuantitativo de 2002 con la observación de campo de tres redacciones catalanas obteniendo resultados similares que además fortalecieron el uso de la Red. Los periodistas catalanes encontraron ventajoso usar internet para documentarse sobre noticias ya previstas, de largo recorrido, sustituyendo los centros de documentación de los medios y fomentando el “periodismo de datos” para complementar y enriquecer las noticias.

En 2009 se llevó a cabo otra investigación en el País Vasco y Navarra (**Díaz-Noci et al.**, 2010) y los datos obtenidos poco distaban de los recogidos en Catalunya. En cuanto al origen de las noticias, los profesionales vascos y navarros no otorgaron fiabilidad a internet. Opinaron que abundaba la información de mala calidad (38,7%), era difícil establecer la credibilidad de un sitio web (25,8%), así como su autoría (14,5%).

En 2011 se ha publicado la cuarta edición del *Estudio sobre periodismo digital 2011* llevado a cabo por *Oriella PR Network*. Esta investigación internacional, realizada a 475 periodistas de 15 países, incluido España, ha analizado por primera vez el uso de los medios sociales para buscar y comprobar información en las redacciones. Los resultados han concluido que en España un tercio utiliza redes sociales como *Twitter* (31%) y *Facebook* (28%) para buscar y contrastar noticias. También acceden a blogs tanto de autores conocidos (31%) como desconocidos (15%). No obstante, los profesionales españoles consideran estas fuentes de información no fiables.

Por fuentes 2.0 entendemos los suministradores de información que lo hacen por libre y espontánea voluntad del autor, sea un individuo o un grupo, utilizando herramientas colaborativas

2. Fuentes 2.0

Por web 2.0 se entiende la Web colaborativa, es decir, donde los lectores son más que consumidores pasivos, contribuyendo con su conocimiento al enriquecimiento de la misma. En este concepto, que tiene una vertiente más tecnológica y otra más social, la Web se aleja del patrón inicial de plataforma de distribución de información y se sitúa en el eje del paradigma descentralizado donde la participación de los interesados es el combustible del medio.

En el campo de las fuentes de información periodísticas ocurre lo mismo, siendo posible dividir las fuentes en dos grupos: las propias de la primera web, de naturaleza más informativa, con predominio de las oficiales, y las fuentes de la web 2.0, basadas en la contribución de la inteligencia colectiva y, en ese sentido, son predominantemente no oficiales.

En su concepto original, inteligencia colectiva es la forma en que “las organizaciones sociales (...) se agrupan para com-

partir y colaborar, para encontrar una ventaja individual y colectiva mayor que si cada participante hubiese permanecido solo” (**Noubel**, 2006, p. 5). Esta inteligencia colectiva inicial está condicionada por dos factores: el número de participantes y el espacio, ya que su eficacia depende de que el grupo tenga un número adecuado de participantes y que estos estén físicamente cercanos. Las dos limitaciones han sido superadas con internet, gracias a que las aplicaciones informáticas “acercan” virtualmente a los lectores, posibilitando el trabajo comunitario, independientemente del número de participantes. En esta nueva realidad, **Noubel** propone que inteligencia colectiva sea “la capacidad de un grupo de personas para colaborar en orden a decidir sobre su propio futuro y alcanzarlo en un contexto complejo” (2006, p. 16). Esta capacidad de colaboración (vertiente social de la web 2.0) es potenciada por las nuevas aplicaciones informáticas (vertiente tecnológica) y da lugar a nuevas fuentes de conocimiento online que, sobre todo en el caso de las grupales, son lo que llamamos fuentes de información 2.0.

Manuel Pinto (1999) hace un recorrido por las diferentes taxonomías de fuentes e identifica siete tipos, en función de su naturaleza, origen, duración, área geográfica, implicación en los hechos, actitud de cara al periodista, identificación y estrategia de actuación. En el actual ecosistema mediático donde internet juega un papel central, proponemos un nuevo tipo basado en el rol del profesional en la manera cómo utiliza en su trabajo las nuevas aplicaciones online. Esta taxonomía nos parece acertada porque se ubica en una variable propia de los medios online, la interacción, sea por iniciativa del profesional o del lector.

- Fuentes *push* (el usuario es “empujado” por la información): se acercan a las propiedades de las fuentes tradicionales. El rol del periodista es más pasivo porque se limita a recibir informaciones de organizaciones, grupos o personas. Esa información puede ser recibida mediante una inscripción previa en un grupo/*feed*, o porque las relaciones públicas de esa organización añadieron su contacto al grupo. En este caso hablamos de aplicaciones como el correo electrónico, los grupos de usuarios, o los lectores de *feeds*.
- Fuentes *pull* (el usuario tira de la información): el profesional toma la iniciativa de buscar la información porque sabe lo que pretende y realiza búsquedas por temas o palabras clave. Las aplicaciones típicas de este tipo de investigación son motores de búsqueda, webs, foros, chats, wikis, blogs o redes sociales.

Aunque las redes sociales y los blogs son las fuentes 2.0 más utilizadas en las rutinas periodísticas, los profesionales consideran más fiables los repositorios y los wikis

La taxonomía propuesta facilita la separación de las fuentes en dos grupos: las que son simplemente online y las que además pueden ser consideradas 2.0. ¿Cuáles son las particularidades o cualidades que marcan la diferencia?

Por fuentes 2.0 entendemos los suministradores de información que lo hacen por libre y espontánea voluntad del autor, sea un individuo o un grupo, utilizando las herramientas colaborativas.

- a) Blogs y microblogs: nacidos como diarios online, se han transformado en fuentes de opinión alternativas muy utilizadas por los periodistas. Su sencillez de edición ha hecho posible que meros ciudadanos y expertos sin cabida en los medios tradicionales hayan creado una vasta oferta de opinión alternativa muy especializada en determinadas disciplinas. Al utilizar abundante hipertexto en los *posts*, los blogueros crean una red compleja entre blogosfera y mediasfera, montando un sistema informativo contextualizado que materializa un ejemplo de inteligencia colectiva.
- b) Redes sociales: aunque en una primera fase no se hayan asumido como fuentes de información por tratarse de simples aplicaciones personales, desde su utilización por Barack Obama en las elecciones presidenciales norteamericanas de 2008 y a raíz del enorme crecimiento de *Facebook*, se han transformado en un novedoso palco para las figuras públicas. En este caso, más que la opinión anónima o de expertos, las redes sociales funcionan como una forma de acceso directo y alternativo a las fuentes.
- c) Foros o listas de correo: no son una novedad, pero el nacimiento de aplicaciones *user friendly* ha aumentado su éxito. El posterior nacimiento de las redes sociales, con la posibilidad de constituir grupos, y los wikis, han reducido su impacto. Aun así, hay millones de comunidades muy especializadas donde es posible recoger información fuera de los circuitos principales.
- d) Chats: son una forma de contacto interpersonal o grupal, por audio o vídeo, que utiliza internet como canal.
- e) Wikis: están hechos con un programa que posibilita que varias personas editen un mismo documento. Han dado lugar a crear enciclopedias colaborativas, siendo *Wikipedia* la más conocida. En sus primeros años, la edición de textos era muy accesible, por lo que se han creado muchas páginas falsas o con informaciones incorrectas, lo que manchó su credibilidad como fuente de información. En la actualidad, hay grupos de voluntarios (wikipedistas) que controlan las actualizaciones.
- f) Repositorios multimedia: en este grupo de fuentes se integran las bases de datos de fotos, sonidos y vídeos. *Flickr*, *Soundcloud* o *Youtube* se encuentran entre los más utilizados por los medios de comunicación. Su éxito está conectado con la popularización de las grabadoras de vídeo y, más recientemente, de los móviles con capacidad de grabar sonidos, vídeos y fotos.

Los motores de búsqueda no son una fuente directa de información, sino un intermediario que facilita encontrarla.

3. Fuentes y credibilidad en el contexto digital

En 2002, *Online News Association (ONA)* llevó a cabo un estudio sobre noticias online y credibilidad. Un dato sobresaliente es que el 91,5% de los encuestados consideró más creíbles los sitios web de organizaciones informativas tradicionales. Esta orientación se confirmó en otros estudios (*Cassidy*, 2007) donde la credibilidad percibida por los lec-

tores siempre asocia el medio online a su versión tradicional. Pero ¿qué ocurre cuando no hay un referencial perceptible, es decir, cuando el usuario que busca información no conoce el actor que está detrás del medio?

Metzger (2007) hace una revisión bibliográfica en autores de diferentes áreas e identifica cinco criterios que los usuarios emplean para evaluar la credibilidad de una información en internet:

- exactitud: ausencia de errores y posibilidad de confirmar la información en otro soporte;
- autoridad: información sobre el autor, afiliaciones y reconocimiento de la comunidad;
- objetividad: identificación de los propósitos del sitio o utilización de un lenguaje persuasivo;
- actualización periódica: ritmo de oferta de informaciones;
- cobertura/alcance: profundidad y amplitud de la información ofrecida.

Si la exactitud y la objetividad son criterios conectados con la credibilidad, independientemente de que la información original sea digital o tradicional, los criterios de autoridad, actualización y cobertura nos conducen a fenómenos grupales de colaboración. En ese sentido, la credibilidad de las fuentes online tendría una fuerte conexión con las características de las herramientas 2.0 (vertiente tecnológica), que son el motor de la inteligencia colectiva (vertiente social) enunciada por **Noubel** (2006). Esto es, la información irrumpe desde un hecho casi insignificante o una información incompleta, y por la fuerza de la colaboración existente entre los expertos online y de las posibilidades de los instrumentos, crece hasta tener un valor informativo para los periodistas.

El criterio de autoridad de la fuente se fundamenta en su forma de legitimación, es decir, en el proceso que permite su reconocimiento en la comunidad online. Por ejemplo, los blogs se legitiman entre pares: el ranking de *Technorati* es una forma de legitimación interna que cruza el número de visitas, los enlaces dirigidos a los blogs y el prestigio de quien publica esos enlaces. En el caso de *Twitter*, importa más el contenido compartido que los atributos de los usuarios (**Asur et al.**, 2011), lo que reduce el peso del actor y subraya el del instrumento comunitario.

“ Para evaluar la credibilidad de una información en internet se emplean cinco criterios: exactitud, autoridad, objetividad, actualización y profundidad ”

En el caso de la actualización permanente hay un vínculo con las características del medio (internet), pero también con la percepción de que los temas actuales y más atractivos suscitan una mayor participación. En este caso, el mejor ejemplo son los foros o listas, espacios frecuentados por especialistas en un determinado tema que mantienen una permanente vigilancia sobre la calidad de la discusión.

Por último, la cobertura nos lleva una vez más a espacios de inteligencia colectiva. Es el caso de los wikis, donde la infor-

mación no tiene el ritmo de actualización de los foros, pero presenta semejante nivel de profundidad y de vigilancia.

Para otros autores (O'Keefe, 2002; David; Glore, 2010), el diseño y la usabilidad son igualmente factores que influyen en la credibilidad de la información online. En ese sentido las herramientas de la web 2.0, diseñadas para una fácil utilización, son una referencia de usabilidad y, en consecuencia, la información que se difunda en ellas tiene potencial para generar mayor percepción de credibilidad.

4. Metodología y procedimiento

Este estudio intenta conocer el uso por parte de los periodistas portugueses y españoles de fuentes 2.0 y el grado de credibilidad que les reconocen. Se elaboró un cuestionario online realizado entre el 27 de julio y el 17 de agosto de 2011 a periodistas de diferentes medios de comunicación social de Portugal y España. La distribución del enlace al cuestionario se hizo de dos formas: por correo electrónico a los periodistas solicitando que rellenaran en documento online, o contactando con el director del medio para que difundiera el enlace desde la red interna del mismo.

<http://www.in2web.es/fuentes20>

Se optó por enviar el enlace a medios de comunicación con fuerte presencia en internet por considerar que serían los que más utilizan las fuentes 2.0.

En Portugal se envió a periodistas de los más importantes:

- diarios: *Correio da manhã, Jornal de notícias, Diário de notícias, Público e I;*
- semanarios: *Expresso y Sol;*
- radios informativas: *Antena 1 y TSF;*
- televisiones: *RTP, SIC y TVI.*

En España se escogieron los medios de comunicación nacional más visitados según datos del ranking *Alexa* en la fecha de realización del estudio:

- televisiones: *Telecinco, RTVE, Cuatro y Antena3;*
- radios nacionales: *Cadena SER, Los 40 principales, Intereconomía, Cope, Onda cero, Cadena 100, Punto radio, Europa FM, y Catalunya radio;*
- prensa escrita nacional: *El país, El mundo, 20 minutos, ABC, La vanguardia, Xornal Galicia y La razón;*
- prensa digital: *Libertad digital, Minuto digital y Estrella digital.*

En total se han obtenido 202 cuestionarios válidos, 58 de Portugal y 144 de España. No seapuró la tasa de retorno porque el estudio coincidió con el período de vacaciones y en muchos casos las redacciones funcionaban con menos personal del habitual.

La mitad de los periodistas ibéricos encuestados realiza su labor en prensa impresa, alrededor de una cuarta parte lo hace en televisión (23,7%) y radio (23,2%), y apenas un 20% en cibermedios. Por país, los periodistas portugueses más representados son los de prensa (46,6%), seguidos de los de radio (27,6%), televisión (22,4%) y cibermedios (17,2%). En cuanto a los españoles, el 56,9% de los encuestados trabaja en prensa impresa, el 25% en televisión, el 19,4% en internet y el 18,8% en radio. Un factor interesante es que el 8,6%

de los profesionales lusos y el 18% de los españoles afirman trabajar en dos medios a la vez.

El cuestionario constaba de cinco bloques de preguntas, pero para este artículo se han utilizado solamente tres:

- el primero dirigido a conocer datos personales del profesional y sus rutinas con respecto al uso de la Red en su labor diaria, con cuestiones cerradas de opción múltiple;
- el segundo grupo buscaba medir la percepción de credibilidad que los periodistas depositan en las informaciones procedentes de las llamadas fuentes 2.0. Para evaluarlo, se utilizó la *escala de Likert* desde 1 (nada creíble) a 5 (muy creíble);
- el tercer bloque de cuestiones buscaba saber qué tipo de fuentes 2.0 había estado en el origen de noticias publicadas por los encuestados.

5. Resultados

En 2011 la Red se utiliza diariamente en las tareas periodísticas en cualquiera de los cuatro medios. Más del 75% de los informadores la usa con la finalidad de buscar material para su trabajo, documentarse y verificar la información, contactar fuentes y recibir y enviar información.

Pero ¿cuáles son las principales fuentes o aplicaciones que utilizan en internet para obtener la información?

Los datos obtenidos revelan que las más empleadas en las redacciones portuguesas son los contactos personales (75,9%), notas de prensa (*press releases*) (72,4%), otros medios de comunicación (72,4%), motores de búsqueda (70,7%), y artículos científicos (67,2%). Las redes sociales son usadas por 50% de los periodistas portugueses, los blogs por 31%, los repositorios por 25,9% y los wikis por 24,1%. Los foros (6,9%) y los chats (5,2%) cierran las preferencias.

En España, los periodistas recurren igualmente a contactos personales (87,5%), motores de búsqueda (62,5%), redes sociales (62,5%) y otros medios de comunicación (56,3%). Las fuentes que reúnen menos de la mitad en preferencias son wikis (43,8%), artículos científicos (31,3%), repositorios y blogs (25% cada uno) y, en último lugar, chats y foros (12,5% cada uno). Es notable la diferencia de uso de algunas fuentes entre ambos países como sucede con los artículos científicos (Portugal: 67,2%; España: 31,3%) y los wikis (Portugal: 24,1%; España: 43,8%).

Con relación a las fuentes de internet, los resultados totales indican que algo más de la mitad de los periodistas utiliza los motores de búsqueda (67,3%) y las redes sociales (56,3%). Menos populares son los wikis (34%), los blogs (28%), los repositorios (25,5%), los chats (8,9%) y los foros (9,7%).

La parte siguiente del cuestionario valora el nivel de confianza de estas fuentes. Se destaca que alrededor del 20% de los periodistas ibéricos no sabe qué responder a estas cuestiones. Son los encuestados que utilizan con menos frecuencia la Red como medio de información, documentación y comunicación en su trabajo.

En Portugal ninguna fuente ha obtenido valores de credibilidad por encima del 50%: más de la mitad opina que los chats (65,5%), los foros (58,6%), las redes sociales (53,4%) y los

	Portugal %	España %
Contactos personales	75,9	87,5
Otros medios de comunicación	72,4	56,3
Motores de búsqueda	70,7	62,5
Artículos científicos	67,2	31,3
Redes sociales	50,0	62,5
Blogs	31,0	25,0
Repositorios	25,9	25,0
Wikis	24,1	43,8
Foros	6,9	12,5
Chats	5,2	12,5

Tabla 1. Fuentes más empleadas

blogs (51,7%) son fuentes “poco creíbles” o “nada creíbles”. Respecto a la credibilidad de los wikis y los repositorios, el 34,5% los valora fiables, y estas son las únicas fuentes 2.0 con más evaluaciones positivas que negativas.

Entre los españoles hay una fuente 2.0 que destaca por ser considerada “creíble” o “muy creíble” por más de la mitad de los encuestados: son los repositorios (56,3%). Las redes sociales (43,8%) y los wikis (37,5%), también presentan más opiniones positivas que negativas. Los chats (75%), los foros (43,8%) y los blogs (37,5%) son considerados fuentes “poco creíbles” o “nada creíbles”.

Se señala una vez más alguna discrepancia entre periodistas portugueses y españoles, en el sentido de que éstos últimos confían más en las fuentes 2.0. Aunque con valoraciones diferentes, repositorios y wikis están entre las tres fuentes más creíbles para ambos. En el caso español se destaca el valor de credibilidad obtenido por las redes sociales (43,8%) comparado con los portugueses (25,9%).

Si analizamos los resultados conjuntos de periodistas portugueses y españoles, los repositorios (46,2%), los wikis (36%), las redes sociales (34,8%) y los blogs (29,4%) son los instrumentos calificados como “creíbles” o “muy creíbles”. No obstante, en ningún caso estas aplicaciones gozan de la confianza de más del 50% de los encuestados de ambos países.

Con este planteamiento de las fuentes 2.0 como origen de noticias, se preguntó a los periodistas si habían publicado alguna noticia basada en informaciones recogidas en estas plataformas. Las redes sociales son la fuente que más noticias originó (68%), con valores muy semejantes entre españoles (68,8%) y portugueses (67,2%). Los blogs se posicionan en segundo lugar (49,5%), siendo los portugueses (55,2%) los que han recurrido más a ellos que los españoles (43,8%). En el caso portugués se trata de una consecuencia de la ligación histórica entre periodistas y blogueros. Los foros ocupan la tercera posición en el origen de noticias (27,7%), pero en este caso los españoles (31,3%) presentan resultados más altos que los portugueses (24,1%). Wikis (24,5%), repositorios (20,9%) y chats (10,5%), presentan resultados equilibrados entre los dos países.

Resulta interesante señalar que las redes sociales son consideradas “creíbles” o “muy creíbles” por solamente el 34,8% de los periodistas ibéricos, pero el 68% publicó noticias con

origen en informaciones recogidas en esta fuente. La discrepancia es más destacada en el caso de los periodistas portugueses: solamente 25,9% las considera fiables pero son la fuente 2.0 que más noticias originó (67,2%). Los blogs, segunda fuente entre las que más noticias originaron en los dos países, también muestra esta discrepancia entre credibilidad percibida (29,4%) y utilización en noticias (49,5%).

6. Conclusiones

La intención de este trabajo es definir las fuentes 2.0 y, esencialmente, conocer el uso que los periodistas hacen de las mismas y el grado de fiabilidad que depositan en ellas.

Se confirma que los periodistas ibéricos, independientemente del medio de comunicación en el que trabajan, utilizan la Red diariamente para buscar material para su trabajo, documentarse, contactar con fuentes, recibir o enviar información y verificarla. Si nos centramos en las fuentes 2.0, los resultados obtenidos muestran que las más utilizadas son las redes sociales seguidas de lejos por los wikis, blogs y repositorios.

Mediante estas aplicaciones abiertas a la contribución colectiva de cualquier individuo, internet facilita el acceso a contenidos, ideas, etc., que de otro modo serían imposibles de alcanzar. No obstante, en ambos países los profesionales prefieren informarse y documentarse mediante fuentes reconocidas en internet, menos participativas y colaborativas, como son otros medios de comunicación o contactos personales, que les infunden más credibilidad que las fuentes 2.0. La excepción son las redes sociales, pero hay que destacar que ésta es la fuente que presenta más semejanzas con un contacto personal, al tratarse de una plataforma casi siempre de utilización personal donde los usuarios se presentan con su fotografía.

Aunque las redes sociales y los blogs sean las fuentes 2.0 más utilizadas en las rutinas periodísticas, los profesionales otorgan más fiabilidad a los repositorios y los wikis. Esta conclusión confirma una conexión entre el concepto de inteligencia colectiva, motor de wikis y repositorios, y la percepción de credibilidad de los periodistas con relación a estas fuentes de información.

Se constata una mejora gradual con respecto a estudios anteriores en cuanto a la credibilidad de las informaciones encontradas en redes sociales o blogs

La insuficiente confianza depositada en algunas fuentes 2.0 sigue siendo el principal inconveniente en las redacciones: las dudas en cuanto a la exactitud y objetividad pueden originarse por la dificultad que tiene el periodista para identificar la autoridad, la actualización periódica y la cobertura, criterios más visibles en los wikis y repositorios, por ejemplo. Esta hipótesis podría explicar que este tipo de fuentes sean consideradas más fiables.

Aun así, se constata un aumento gradual de la credibilidad de informaciones encontradas en las redes sociales o los

blogs con respecto a estudios anteriores. Este progreso puede ser debido a una mejor preparación de los periodistas en habilidades relacionadas con el uso de la Red, que les capacita para recuperar, buscar, identificar, seleccionar y verificar la información en este medio esquivando obstáculos como pueden ser informaciones anónimas, de mala calidad, incorrectas o falsas. Pero también puede estar relacionado con el hecho de que los periodistas sean conscientes de que las nuevas aplicaciones informáticas facilitan el funcionamiento de los distintos tipos de comunidades online y la vigilancia permanente que sus participantes ejercen para asegurar la fiabilidad de la información publicada.

7. Nota

1. Se utiliza la palabra “importante” para las cifras relativas a las respuestas “importante” y “muy importante”.

8. Bibliografía

Asur, Sitaram; Huberman, Bernardo; Szabo, Gabor; Wang, Chunyan. *Trends in social media: persistence and decay*, 2011.

http://www.hpl.hp.com/research/scl/papers/trends/trends_web.pdf

AccesoGroup. *Estudio sobre el uso de la Red en los medios de comunicación*. Madrid: Deloitte & Touche, 2002.

<http://www.pymesonline.com/formacion/index.php?action=file&id=84>

Bastos, Helder. *Jornalismo electrónico*. Coimbra: Minerva, 1998.

Canavilhas, João. “Os jornalistas portugueses e a internet”. En: *V Congreso Ibero-Americano de Jornalistas na Internet*. Salvador da Bahia, 2004.

<http://www.bocc.ubi.pt/pag/canavilhas-joao-jornalistas-portugueses-internet.pdf>

Cassidy, William P. “Online news credibility: an examination of the perceptions of newspaper journalists”. *Journal of computer-mediated communication*, 2007, v. 12, n. 2, article 7. <http://jcmc.indiana.edu/vol12/issue2/cassidy.html>

David, Alicia; Glore, Peyton. “The impact of design and aesthetics on usability, credibility, and learning in an online environment”. *Online journal of distance learning administration*, 2010, v. 13, n. 4.

http://www.westga.edu/~distance/ojdla/winter134/david_glore134.html

Díaz-Noci, Javier; Meso, Koldobika; Larrondo-Urreta, Aina; Salaverría-Aliaga, Ramón; Sadaba-Chalezquer, María-Rosario. “Presencia y uso de internet en las redacciones de los diarios regionales vascos y navarros”. *Mediatika*, 2010, n. 12, pp. 301-319.

<http://www.euskonews.com/0383zkbk/gaia38301es.html>

Gans, Herbert. *Deciding what's news: a study of CBS evening news, NBC nightly news, Newsweek and Time* (25th anniversary edition). Evanston: Northwestern University Press, 2004. ISBN 978 0810122376

Metzger, Miriam J. “Making sense of credibility on the Web: models for evaluating online information and recommendations for future research”. *Journal of American Society for Information Science and Technology*, 2007, v. 58, n. 13, pp. 2078-2091.

http://www.ischools.org/conference08/pc/WC17_iconf08.pdf

<http://dx.doi.org/10.1002/asi.20672>

Masip-Masip, Pere. “Presencia y uso de internet en las redacciones catalanas”. *Zer*, 2002, n. 14.

<http://www.ehu.es/zer/zer14/presenciayuso14.htm>

Masip-Masip, Pere. “Rutinas periodísticas e internet en la información diaria”. En: *Noves competències professionals. III Congrés intl comunicació i realitat*, 2005, pp. 561-576.

<http://robertoigarza.files.wordpress.com/2008/11/conrutinas-periodisticas-e-internet-masip-2006.pdf>

Noubel, Jean-François. *Inteligencia colectiva, la revolución invisible*, 2006.

<http://solocreatividad.files.wordpress.com/2011/01/intelco.pdf>

Obercom. *O jornalismo hoje: uma análise de 14 redacções de TV, rádio e jornais*. Research report, 2006.

<http://www.obercom.pt/client/?newsId=29&fileName=rr1.pdf>

Obercom. *Desafios do jornalismo 2010*. Lisboa: CIES.

http://www.obercom.pt/client/?newsId=428&fileName=desafios_do_jornalismo.pdf

O'Keefe, Daniel J. *Persuasion: theory and research*. Thousand Oaks: Sage, 2002. ISBN 978 0761925392

Online News Association. *Digital journalism credibility study*, 2001.

http://banners.noticiasdot.com/termometro/boletines/docs/marcom/prensa/ona/2002/ona_credibilitystudy2001report.pdf

Oriella PR Network. *The state of journalism in 2011*, 2011.

<http://orielladigitaljournalism.com/files/assets/downloads/publication.pdf>

Pavlik, John. *Journalism and new media*. New York: Columbia University Press, 2001. ISBN 978 0231114837

Pinto, Manuel. “Fontes jornalísticas: contributos para o mapeamento do campo”. *Comunicação e sociedade*, 2000, v. 14, n. 1-2, pp. 277-294.

http://repositorium.sdum.uminho.pt/bitstream/1822/5512/1/CS_vol2_mpinto_p277-294.pdf

LA PRECARIEDAD TE HACE DÓCIL. PROBLEMAS QUE AFECTAN A LA PROFESIÓN PERIODÍSTICA

Mònica Figueras-Maz, Marcel Mauri-Ríos, Salvador Alsius-Clavera y Francesc Salgado-De-Dios



Mònica Figueras-Maz es doctora en periodismo por la *Universitat Pompeu Fabra* de Barcelona (UPF) y licenciada en ciencias de la información por la misma universidad. Trabaja desde 1998 como profesora en el *Departamento de Comunicación* de la UPF. Su línea de investigación se basa en los aspectos deontológicos de la comunicación, especialmente en relación con la juventud, el género y la inmigración.

Universitat Pompeu Fabra
Campus de la Comunicació - Poblenou
Roc Boronat, 138. 08018 Barcelona
monica.figueras@upf.edu

Marcel Mauri-Ríos es doctor en periodismo por la *Universitat Pompeu Fabra* de Barcelona (UPF). Sus principales líneas de investigación son la historia del periodismo y la deontología periodística. Imparte docencia en el *Departamento de Comunicación* de la UPF y en los estudios de comunicación audiovisual de la *Universitat Oberta de Catalunya* (UOC). Es miembro del *Grup de Recerca en Periodisme*. Ha realizado estancias de investigación predoctorales en la *Université Paris II* y postdoctoral en la *Columbia University* de Nueva York.

Universitat Pompeu Fabra
Campus de la Comunicació - Poblenou
Roc Boronat, 138. 08018 Barcelona
marcel.mauri@upf.edu

Salvador Alsius-Clavera es profesor titular de periodismo audiovisual en la *Universitat Pompeu Fabra* (UPF) y actualmente dirige la carrera de periodismo en dicha universidad. Es doctor en periodismo, licenciado en ciencias económicas y en ciencias de la información. Pertenece al *Grupo de Investigación en Periodismo* (GRP) y es el investigador principal del equipo que trabaja en ética de la información. A lo largo de su carrera ha alternado o compaginado una dedicación intensa a la actividad periodística (en prensa, radio y muy especialmente en televisión) con la docencia, la investigación y la gestión universitaria. Fue decano del *Colegio de Periodistas de Catalunya* y actualmente es miembro del *Consejo de la Información* de Catalunya.

Universitat Pompeu Fabra
Campus de la Comunicació - Poblenou
Roc Boronat, 138. 08018 Barcelona
salvador.alsius@upf.edu

Francesc Salgado-De-Dios es profesor de la facultad de Comunicación de la *Universitat Pompeu Fabra* de Barcelona (UPF), en la que se doctoró en 2009 con una tesis sobre Manuel Vázquez Montalbán. Imparte cursos de redacción e historia del periodismo. Investiga la ética periodística y la historia de la prensa contemporánea. Licenciado en filosofía y ciencias de la educación por la *Universitat de Barcelona* (UB) y en periodismo por la *Universitat Pompeu Fabra* (UPF).

Universitat Pompeu Fabra
Campus de la Comunicació - Poblenou
Roc Boronat, 138. 08018 Barcelona
francesc.salgado@upf.edu

Resumen

Se muestran los problemas que son percibidos como los más importantes por los periodistas de cuatro comunidades autónomas españolas: Catalunya, Madrid, País Vasco y Andalucía. De acuerdo con los resultados obtenidos en un trabajo de campo que mezcla metodología cuantitativa y cualitativa, la precariedad y la inseguridad laboral son los problemas que más preocupan a los periodistas de todas las comunidades consultadas. Las variables, edad, sexo e ideología son las que ofrecen más diferencias en las respuestas.

Artículo recibido el 26-07-11

Aceptación definitiva: 16-11-11

Palabras clave

Ética periodística, Problemas de la profesión, Periodistas, Comunicación, Precariedad laboral, Intrusismo, Intereses económicos y políticos, España.

Title: Insecurity makes you docile. Perception of problems facing journalists

Abstract

The problems that are perceived as the most important by journalists from four Spanish regions: Catalonia, Madrid, Basque Country and Andalusia are reported. Fieldwork using both quantitative and qualitative methods found that precarious and uncertain job security are the issues of most concern to journalists in all the regions surveyed. Of the variables analyzed, age, sex and ideology accounted for the greatest differences in responses.

Keywords

Media ethics, Professional issues, Journalists, Professional precarity, Intrusion, Economic and political interests, Spain.

Figueras-Maz, Mònica; Mauri-Ríos, Marcel; Alsius-Clavera, Salvador; Salgado-De-Dios, Francesc. "La precariedad te hace dócil. Problemas que afectan a la profesión periodística". *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 70-75.

<http://dx.doi.org/10.3145/epi.2012.ene.09>

Introducción

La crisis económica global ha afectado de forma muy clara a los medios de comunicación, y esta circunstancia, sumada a la imparable revolución tecnológica –comparable a la que cinco siglos antes se produjo con la llegada de la imprenta–, ha traído profundos cambios a la industria periodística (Díaz-Noci, 2010, p. 561). Varios trabajos han demostrado cómo desde 2009 las empresas periodísticas se han visto inmersas en una espiral de pérdidas económicas, concentración de medios y cierre de proyectos (Almiron, 2009), en un momento en que el trabajo de las redacciones tradicionales está siendo desafiado y cuestionado en buena parte por una nueva era de ciudadanos-productores de noticias (Compton; Benedetti, 2010).



El principal objetivo de este artículo es mostrar los problemas de la profesión periodística a partir de un estudio empírico que analiza en profundidad la percepción de los propios periodistas de cuatro comunidades autónomas de España: Cataluña, País Vasco,

Madrid y Andalucía¹. Esta investigación tiene como principal antecedente el *Libro blanco de la profesión periodística en Cataluña* (2006), aunque sólo se hiciera en esta comunidad y con otra finalidad. También cabría señalar el *Informe anual de la profesión periodística* de la *Asociación de la Prensa de Madrid* (APM) y el *Libro blanco de la prensa diaria* elaborado por la *Asociación Española de Editores de Diarios* (AEDE),

estudios que contemplan algunos de los problemas profesionales aunque no se ciñen exclusivamente a estos asuntos (Maciá-Barber; Herrera-Damas, 2010).

Fuera de España destacan los estudios de **Noordenstreng y Hannikainen** (1984) sobre la deontología y su aplicación en la profesión periodística en Finlandia y Europa, y de **Claude-Jean Bertrand** (2000, p. 108), pionero en la introducción en el ámbito de la comunicación de los estudios sobre los sistemas de rendición de cuentas de los medios (*media accountability systems*). En los Estados Unidos el estudio de referencia es *The American journalist*, de **David Weaver y Cleveland Wilhoit** (1986), actualizado en varias ocasiones, la última en 2007. Con su investigación estos autores realizan un retrato fidedigno de la evolución de la profesión periodística, sus valores y condiciones laborales en Estados Unidos, que es extrapolable al mundo occidental. Cabe señalar la contribución de **Susanne Fengler y Stephan Russ-Mohl** (2008) acerca del comportamiento del periodista ante los retos y contextos profesionales contemporáneos. Son también muchas las investigaciones recientes sobre los condicionantes políticos (Hallin; Mancini, 2004), los imperativos económicos que afectan a la tarea de los periodistas en su trabajo tradicional (Benson; Hallin, 2007; McManus, 2009) o la convergencia que provoca internet (Casero-Ripollés, 2010). El análisis desde la perspectiva de los propios periodistas lo encontramos en **Hanitzsch y Mellado** (2011), que comparan el grado de presiones e influencias que perciben en el trabajo de periodistas de 18 países, entre ellos España. Lo hacen mediante una encuesta llevada a cabo en un trabajo previo (Hanitzsh *et al.*, 2010) que contempla seis dimensiones (influencias políticas, económicas, organizacionales, profesionales, de procedimiento y de grupos de referencia). Los factores políticos y económicos son claramente los denominadores más importantes en las diferencias nacionales: así, los españoles, junto con los australianos, perciben influencias políticas y económicas, a diferencia del resto de periodistas de países occidentales que no lo sienten así.



Sala de control de una televisión

De todos los problemas planteados, la cuestión que más preocupa a los profesionales es la precariedad y la inseguridad laboral

Metodología aplicada

Para el estudio que presentamos en este artículo se tomó como punto de partida un análisis de los códigos y documentos con normativa deontológica existentes en España, a partir de los trabajos de compilación de textos realizados por **Hugo Aznar** (1999) y **Juan-Carlos Pérez-Fuentes** (2004) que sirvió de marco teórico para la investigación: un análisis cualitativo (entrevistas en profundidad) y cuantitativo (encuestas online) para conocer los criterios éticos interiorizados por los profesionales del periodismo. El trabajo de campo (entrevistas en profundidad y encuesta online) se delegó en empresas expertas en investigación de mercados o, en el caso de Catalunya y Madrid, en los colegios profesionales de sociólogos.

El número de entrevistas en cada comunidad autónoma varió en función del censo de periodistas: en Madrid y Catalunya se efectuaron 30, en el País Vasco 20, y en Andalucía 24. Se llevaron a cabo entre abril y junio de 2008 y los resultados obtenidos fueron utilizados como punto de partida para diseñar la encuesta online compuesta por 43 preguntas. Esta fase del trabajo de campo tuvo lugar a lo largo de tres semanas a finales de 2008 y principios de 2009. El número de encuestas obtenidas fue de 2.012, repartidas de la siguiente manera: Catalunya, 1.198; Madrid, 417; País Vasco, 172; Andalucía, 225.

Para realizar los cuestionarios se eligieron los temas más candentes recogidos por los códigos éticos, libros de estilo, estatutos de redacción y recomendaciones específicas de instituciones autorreguladoras: defensores de los lectores, consejos de la información, etc. También se recogieron los problemas estructurales y coyunturales que tiene el sistema periodístico español, en la medida que forma parte del modelo mediterráneo o pluralista polarizado —en el que la influencia política y el partidismo de la prensa es muy importante— según la definición que hacen **Daniel Hallin** y **Paolo**

Mancini (2004, pp. 89). Precisamente, en otro estudio de **Hallin** y **Papathanassopoulos** (2002) se concluye que existe un alto nivel de percepción de influencia política en los países con fuertes tradiciones de pluralismo polarizado, como el sur de Europa y América Latina.

En este artículo se desgranar los resultados de una de las preguntas de la encuesta online, la que se refería a los que posiblemente son percibidos como principales problemas de la profesión periodística, a la vez que se explotan las respuestas que se obtuvieron de las entrevistas en profundidad.

La pregunta se formuló de la siguiente forma: *¿Cuál es su percepción sobre los siguientes problemas que supuestamente afectan a la profesión periodística? Y se planteaban ocho ítems:*

- precariedad laboral,
- concentración de medios,
- influencia de los índices de audiencia,
- predominio de los intereses políticos y económicos por encima de los periodísticos,
- falta de ética profesional,
- confusión de tareas de otras profesiones y el intrusismo profesional,
- bajo respeto hacia la independencia de los profesionales, y
- falta de conciencia hacia la responsabilidad social de los medios.

No hay grandes diferencias entre la opinión de hombres y mujeres, salvo en el caso del intrusismo

Precariedad e inseguridad como principales problemas de la profesión

De los problemas planteados, la cuestión que más preocupa a todos los periodistas encuestados es la precariedad y la inseguridad laboral. El segundo aspecto que suscita más inquietud es el predominio de los intereses económicos y políticos sobre los periodísticos. En la fase cualitativa estos dos temas son también recurrentes.

Los profesionales andaluces aparecen especialmente sensibilizados con estos dos problemas y no dudan en considerar la profesión como “una selva donde muchos mercenarios trabajan gratis o por muy poco dinero”. Asimismo, consideran que la inseguridad laboral condiciona la calidad y la libertad de expresión porque “el periodista precario es más dócil”. La novedad en el discurso generalizado está en Catalunya, donde 8 de las 30 personas consultadas encuentran “lógica” y “admisible” y en algún caso hasta positiva la influencia que ejercen los intereses económicos y políticos en el periodismo.

Tras esos dos problemas, la influencia de los índices de audiencia, la falta de ética profesional y la baja conciencia de la responsabilidad social del periodista son consideradas preocupaciones bastante o muy importantes en todas las co-

munidades analizadas (más de 8 de cada 10 encuestados). La excepción son Andalucía y el País Vasco. En la comunidad andaluza hay una preocupación ligeramente inferior por los índices de audiencia (76,8%); y los periodistas vascos se muestran menos afectados por la baja conciencia de la responsabilidad social del periodista, con un 75% de respuestas positivas. Sin embargo, la concentración de los medios preocupa especialmente en el País Vasco y los entrevistados de esta comunidad reconocen que no tienen otra opción que plegarse a las condiciones impuestas por las empresas.

El intrusismo profesional y la confusión con las tareas propias de otras profesiones, si bien son temas considerados importantes para más de la mitad de los comunicadores, representan una inquietud menor. En ambos casos, los más preocupados son los vascos y los menos, los madrileños. En un término medio se sitúa la preocupación por la falta de respeto a la autonomía de los profesionales y, nuevamente, los vascos (90,1%) son los que tienen una percepción superior del problema y los madrileños, menor (76,6%). Las entrevistas en profundidad en el País Vasco refuerzan esta impresión: “en este sector ha entrado cualquiera que tuviese unas mínimas dotes comunicativas”.

En la tabla 1 pueden apreciarse las diferencias y similitudes en las respuestas dadas por los profesionales consultados.

Diferencias de percepción según variables

Mientras que en Andalucía y el País Vasco no se detectan diferencias significativas según variables independientes, en Catalunya y Madrid las respuestas varían en función de la edad, sexo, ingresos e ideología de los encuestados. En el caso catalán, condiciona el medio de trabajo y en Madrid lo hacen el nivel de estudios, el tamaño de la empresa, la categoría profesional y el área temática.

Los de mayor edad y los de izquierdas, los más preocupados

Respecto a la edad, se produce una relación directamente proporcional con la importancia otorgada a los problemas, de manera que cuanto mayor es la persona más importancia concede a los factores planteados. Los mayores de 50 años manifiestan mayor inquietud por la falta de respeto a la autonomía de los profesionales, y por la concentración de medios. Esto se invierte en el caso de la precariedad labo-

ral y el intrusismo profesional: son los menores de 35 años los más sensibilizados con este problema. Por ejemplo, el 73,3% de los profesionales catalanes menores de 35 años lo considera muy importante, mientras que sólo preocupa al 58,6% de los de otras edades. La franja de menor edad de este territorio también es la que más se preocupa por los índices de audiencia.

La ideología aparece así mismo como una variable que refleja diferencias: los comunicadores que se declaran de izquierdas son los que muestran mayor grado de inquietud respecto a la práctica totalidad de los problemas planteados. La excepción a esta tendencia se produce en Catalunya y respecto al problema de la falta de ética profesional: los profesionales catalanes de centro y derecha conciben el problema de forma ligeramente superior (37,6% frente al 31,6% de izquierda).

Los profesionales de izquierda muestran mayor grado de inquietud respecto a los problemas planteados

Ellas, más sensibilizadas con el intrusismo que ellos

No hay grandes diferencias entre la opinión de los hombres y mujeres consultados, salvo en el caso del intrusismo. Las periodistas están más sensibilizadas por esta cuestión que los hombres. En el caso catalán, esto se constata con un 71,8% y un 51,4% de respuestas positivas respectivamente. En Madrid, 7 de cada 10 mujeres lo consideran un problema bastante o muy importante frente a casi 5 de cada 10 hombres. También en esta comunidad, el predominio de los intereses económicos y políticos sobre los periodísticos preocupa más a los hombres que a las mujeres.

El intrusismo preocupa más a los licenciados en periodismo y a los trabajadores de empresas públicas

Por nivel de estudios, el único supuesto que presenta diferencias notables es el del intrusismo profesional. Esto se debe a que los licenciados en periodismo –en un entorno creciente de periodismo ejercido por no licenciados o ciudadanos-productores– perciben el problema de forma más im-

	Madrid	Catalunya	Andalucía	País Vasco
Precariedad laboral	93,7	95,0	97,8	96,5
Concentración de medios	70,5	74,8	75,9	90,1
Índices de audiencia	83,6	81,2	76,8	89,5
Intereses económicos y políticos	90,0	92,9	95,1	93,1
Falta de ética profesional	84,9	82,1	87,4	83,7
Confusión de tareas de otras profesiones	52,9	57,0	60,6	77,9
Intrusismo profesional	54,4	60,1	69,1	79,7
Falta de respeto a la autonomía de los profesionales	76,6	79,8	81,8	90,1
Baja conciencia de la responsabilidad social del periodista	81,0	81,0	86,7	75,0

Tabla 1. Problemas percibidos como “bastante” o “muy” importantes (%). Las columnas están ordenadas de menor (Madrid) a mayor (País Vasco) intensidad problemática. Los números en rojo marcan los valores máximos de las respuestas

portante que los que no lo son. Los andaluces entrevistados en la fase cualitativa creen que no hay forma de acabar con el intrusismo, al cual se le considera la lacra actual. Algunos dicen que se debe restringir el acceso a los que no son licenciados; en cambio otros no dudan en afirmar que se puede hacer muy buen periodismo sin ser licenciado. También se aprecian algunas diferencias, sobre todo en Catalunya, entre trabajadores de empresas públicas y privadas—más sensibles a esta cuestión los primeros que los segundos— ante la preocupación por el intrusismo.

Los ingresos, la situación y la categoría laboral marcan algunas diferencias

Se observa que a mayores ingresos, menor preocupación por los diferentes problemas. Por ejemplo, los profesionales catalanes que cobran más de 3.000 euros mensuales son los que menos inquietud muestran por la concentración de los medios de comunicación, la confusión con las tareas propias de otras profesiones, la falta de ética profesional, el intrusismo y la falta de autonomía profesional.

La preocupación por la precariedad y la inseguridad laboral también se reduce a medida que aumenta el salario. Los profesionales que se encuentran sin trabajo, autónomos y asalariados conceden mayor importancia a este asunto que los empresarios con trabajadores. Del mismo modo, los redactores o equivalentes conceden más importancia a este problema que los jefes de sección o los directivos, por lo que se puede deducir que a mayor jerarquía menor importancia se concede al problema. En la fase cualitativa se constató que los cargos intermedios, quienes se ven en la posición de transmitir las presiones, parecen haberlas internalizado o asumido, de manera que podrían ser menos conscientes de su gravedad. Por el contrario, los trabajadores que ocupan posiciones inferiores en la jerarquía de sus empresas expresan con mayor frecuencia su incomodidad respecto de la falta de autonomía en el desempeño de su trabajo cotidiano.

“ En el País Vasco no hay tanta preocupación por la conciencia de la responsabilidad social y en cambio, mucha por la falta de autonomía profesional ”

Conclusiones

La situación económica mundial y española, de la que la industria periodística no se escapa, afecta de forma directa a los profesionales de la información. Prueba de ello es que la precariedad e inseguridad laboral son los principales problemas aludidos por los periodistas. Que más de un 90% de los encuestados señalen esta cuestión como problemática demuestra el grado de inestabilidad laboral de los periodistas, y más si se tiene en cuenta que el estudio se hizo en 2008 y 2009, cuando la crisis todavía no había mostrado toda su dureza.

La segunda preocupación más señalada, el predominio de los intereses económicos y políticos sobre los periodísticos,

es compartida por más del 90% de los profesionales encuestados, que ponen de manifiesto la vulnerabilidad de la información ante los poderes políticos y/o financieros.

En general no hay diferencias significativas por sexos pero resulta interesante observar cómo, en el caso de la precariedad, las mujeres se preocupan más que los hombres: ellas tienen mayor sensación de vulnerabilidad en sus puestos de trabajo que ellos. También se sienten más inseguros laboralmente los jóvenes periodistas, un hecho que parece lógico en un momento de crisis de medios y de difícil acceso al mercado laboral. En cambio, cuanto más mayor se es, más preocupación existe por el resto de supuestos de componente más ético.

“ En el País Vasco y en Andalucía, la confusión de tareas profesionales y el intrusismo preocupan bastante más que en Madrid y Catalunya ”

Los resultados son bastante parecidos cuando se comparan los cuatro territorios analizados. En general, los índices de audiencia, la falta de ética profesional y la baja conciencia de la responsabilidad social del periodista aparecen como preocupaciones bastante o muy importantes en todas las comunidades analizadas. Pero los resultados no son tan transversales cuando se trata de problemas laborales concretos. Así, en el País Vasco no hay tanta preocupación por la conciencia de la responsabilidad social y en cambio mucha por la falta de autonomía profesional, quizá debido a la sensación de presión al trabajar en un entorno acosado por el terrorismo. Sería, pues, interesante observar si los resultados pueden cambiar a partir de ahora con el abandono de las armas de ETA. También hay claras diferencias en la preocupación por la concentración de medios, 20 puntos más en el País Vasco que en Madrid, un diferencial que puede encontrar explicación en el censo de profesionales mucho más reducido en el País Vasco, y en el temor a ser absorbidos por grupos empresariales con sede fuera de esta comunidad y, con ello, la posible pérdida de la idiosincrasia local.

Finalmente, cabe destacar que en el País Vasco y en Andalucía, la confusión de tareas profesionales y el intrusismo preocupan bastante más que en Madrid y Catalunya. Parece que en entornos donde el censo de periodistas es menor, se es más sensible a un periodismo amateur muy vinculado a las nuevas tecnologías de la información: cualquier ciudadano con un smartphone puede compartir información. En futuras encuestas sobre la profesión se podrá ver si esto es sólo una cuestión coyuntural (debido, también, a la crisis) o no.

Nota

1. Esta investigación es resultado del proyecto “Ética y excelencia informativa. La deontología periodística frente a las expectativas de los ciudadanos”, nº de referencia SEJ 2006-05631-C05-01, del Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica 2004-2007 del Ministerio de Educación y Ciencia.

Bibliografía

- Almiron, Núria.** "Grupos privados propietarios de medios de comunicación en España: principales datos estructurales y financieros". *Comunicación y sociedad*, 2009, v. 22, n. 1, pp. 243-273.
http://www.unav.es/fcom/comunicacionsociedad/es/articulo.php?art_id=36
- Aznar, Hugo.** *Ética y periodismo: autorregulación, códigos, estatutos de redacción y otros documentos*. Barcelona: Paidós, 1999. ISBN 8449306531
- Benson, Rodney; Hallin, Daniel.** "How states, markets and globalization shape the news. The French and US national press, 1965-97". *European journal of communication*, 2007, v. 22, n. 1, pp. 27-48.
<http://dx.doi.org/10.1177/0267323107073746>
- Bertrand, Claude-Jean.** *Media ethics & accountability systems*. New Brunswick: Transaction Publishers, 2000. ISBN 1560004207
- Casero-Ripollés, Andreu.** "Prensa en internet: nuevos modelos de negocio en el escenario de la convergencia". *El profesional de la información*, 2010, v. 19, n. 6, pp. 595-601.
<http://eprints.rclis.org/bitstream/10760/15015/1/595-601.pdf>
<http://dx.doi.org/10.3145/epi.2010.nov05>
- Compton, James R.; Benedetti, Paul.** "Labour, new media and the institutional restructuring of journalism". *Journalism studies*, 2010, v. 11, n. 4, pp. 487-499.
<http://dx.doi.org/10.1080/14616701003638350>
- Díaz-Noci, Javier.** "Medios de comunicación en internet: algunas tendencias". *El profesional de la información*, 2010, v.19, n. 6, pp. 561-567.
http://www.elprofesionaldelainformacion.com/contenidos/2010/noviembre/medios_comunicacion.pdf
<http://dx.doi.org/10.3145/epi.2010.nov.01>
- Fengler, Susanne; Russ-Mohl, Stephan.** "Journalists and the information-attention markets: towards an economic theory of journalism". *Journalism*, 2008, v. 9, n. 6, pp. 667-690.
http://www.journalistik-dortmund.de/fileadmin/content/fengler/Journalists_and_Information-Attention-Markets.pdf
<http://dx.doi.org/10.1177/1464884908096240>
- Hallin, Daniel; Mancini, Paolo.** *Comparing media systems: three models of media and politics*. Cambridge: Cambridge University Press, 2004. ISBN 0521543088
- Hallin, Daniel; Papathanassopoulos, Stylianos.** "Political clientelism and the media: Southern Europe and Latin America in comparative perspective". *Media, culture & society*, 2002, v. 24, n. 2, pp. 175-95.
<http://www.portalcomunicacion.com/both/opc/client.pdf>
- Hanitzsch, Thomas; Mellado, Claudia.** "What shapes the news around the world? How journalists in eighteen countries perceive influences on their work". *International journal of press/politics*, 2011, v. 16, n. 3, pp. 404-426.
<http://dx.doi.org/10.1177/1940161211407334>
- Hanitzsch, Thomas; Anikina, Maria; Berganza, Rosa; Cangoz, Incilay; Coman, Mihai; Hamada, Basyouni; Hanusch, Folker et al.** "Modeling perceived influences on journalism: evidence from a cross-national survey of journalists". *Journalism & mass communication quarterly*, 2010, v. 87, n. 1, pp. 7-24.
- Maciá-Barber, Carlos; Herrera-Damas, Susana.** "La deontología periodística: praxis, disfunciones y retos desde la perspectiva de los profesionales de la comunicación en la Comunidad de Madrid (2006-2009)". *Comunicación y sociedad*, 2010, v. 23, n. 1, pp. 77-104.
http://www.unav.es/fcom/comunicacionsociedad/es/articulo.php?art_id=351
- McManus, John H.** "The commercialization of news". En: Wahl-Jorgensen, Karin; Hanitzsch, Thomas. *The handbook of journalism studies*. New York: Routledge, 2009, pp. 218-235. ISBN 0203877683
- Noordenstreng, Kaarle; Hannikainen, Lauri.** *The mass media declaration of Unesco*. Norwood: Ablex, 1984. ISBN 0893910775
- Pérez-Fuentes, Juan-Carlos.** *Ética periodística: principios, códigos deontológicos y normas complementarias*. Bilbao: Servicio Editorial de la Universidad del País Vasco, 2004. ISBN 8483735946
- Real-Rodríguez, Elena.** "La autoregulación. Valoración de los códigos. Conocimiento de los mecanismos de autoregulación". En: Alsius, Salvador; Salgado, Francesc (eds.). *La ética informativa vista por los ciudadanos*. Barcelona: UOC, 2010, pp. 275-291. ISBN 978 8497889353
- Russ-Mohl, Stephan.** "Regulating self-regulation: the neglected case of journalism policies". *Communication and medicine*, 1993, v. 18, n. 2, pp. 151-168.
- Weaver, David; Wilhoit, Cleveland.** *The American journalist*. Indianapolis: Indiana University Press, 1991 (2ª ed.). ISBN 0253363640

Suscripción EPI sólo online

Pensando sobre todo en los posibles suscriptores latinoamericanos, ya no es obligatorio pagar la suscripción impresa de EPI para acceder a la online.

EPI se ofrece a instituciones en suscripción "sólo online" a un precio considerablemente más reducido (99 euros/año), puesto que en esta modalidad no hay que cubrir los gastos de imprenta ni de correo postal.



Te entendemos”

Sistemas de gestión de bibliotecas Open Source
Interfaces interactivas y OPACs
Repositorios OAI
Gestión documental y de archivos
Digitalización
Outsourcing de servicios documentales
Desarrollo de sitios web / multimedia / e-learning
Comunicación y e-marketing de servicios de información



Oficinas Centrales:
C/Garcilaso 15-B
46003 Valencia

Tel.: 96 369 41 23
Fax: 96 369 34 39
info@masmedios.com



EL FUTURO DE LA WEB ANTE LA NEUTRALIDAD DE LA RED: ESTADO DE LA CUESTIÓN EN LA UNIÓN EUROPEA

Cristina Cullell-March



Cristina Cullell-March es doctora en ciencias de la comunicación, licenciada en derecho y periodismo. Ha realizado estancias de investigación en el *Media Management and Transformation Centre* de la *Universidad de Jönköping* (Suecia) y en el *Centre for European Governance* de la *Universidad de Exeter* (Reino Unido). Acaba de publicar el libro *La regulación del espacio radioeléctrico. Los servicios de comunicaciones electrónicas en la Unión Europea*, Bosch, 2011.

Universitat Jaume I
Castelló de la Plana, España
cullell@uji.es

Resumen

El futuro de la Web está condicionado por la neutralidad de la Red. Para que la Web pueda desarrollarse en plenitud todos los contenidos, blogs y plataformas deben ser tratados de la misma forma. Se indica el origen y las motivaciones presentes en el debate de la neutralidad. Se centran las dos posiciones encontradas en este debate: los aperturistas, partidarios de una Red abierta sin restricciones; y los desreguladores que defienden la capacidad de intervenir sobre los contenidos que circulan, privatizándola. Finalmente se describe el estado de la cuestión respecto a la regulación de la neutralidad de la Red en la *Unión Europea*.

Palabras clave

Neutralidad de Red, Web, Tráfico, Libertad de expresión, Censura, Unión Europea.

Title: The future of the web: the challenge of Net neutrality in the European Union

Abstract

The future development of the web largely depends on net neutrality, that is, ensuring that all content, sites and platforms are treated equally. This article assesses the origins and motivations that exist in the debate about network neutrality. Then it focuses on the two distinct views regarding net neutrality: the Openists, supporters of an open net without restrictions, on one side and Deregulationists, naturally suspicious of any government regulation, on the other. Finally, this article describes the net neutrality regulatory framework within the European Union.

Keywords

Net neutrality, Web, Traffic, Freedom of expression, Censorship, European Union.

Cullell-March, Cristina. "El futuro de la Web ante la neutralidad de la Red: estado de la cuestión en la Unión Europea". *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 77-82.

<http://dx.doi.org/10.3145/epi.2012.ene.10>

Introducción

El futuro de la Web se juega en múltiples campos y frentes. No sólo en el contexto tecnológico, de contenidos, de la evolución hacia la web 3.0 o de las redes sociales, sino también en el campo de la regulación. Si internet se mantiene como una plataforma abierta, innovadora y neutral dependerá en gran medida del tipo de regulación que se acabe imponiendo. En Europa el debate sobre la neutralidad de la Red está más vivo que nunca, especialmente ahora que las

autoridades europeas deben decidir si se aplican medidas adicionales para evitar que los proveedores de servicios de internet (PSIs) limiten, dificulten o impidan el acceso a determinados contenidos online. El futuro de la Web está en gran medida condicionado por la neutralidad de la Red.

1. El debate sobre la neutralidad de la Red

A mediados del los 90 se publicó la *Declaración de la independencia del ciberespacio* (Barlow, 1996), en la que se

rechazaba tajantemente cualquier tipo de intervención política o económica sobre este espacio. Esta declaración puso las bases de lo que hoy conocemos como “neutralidad de la Red”. Este concepto fue acuñado por primera vez por el profesor de la *Universidad de Columbia* **Tim Wu** para explicar que internet es únicamente un transportador o *carrier* de contenido online que no distingue entre una web u otra. La idea central de este concepto es que internet en tanto que red pública debe tratar de la misma forma todos los contenidos, webs y plataformas que circulan a su través (**Wu**, 2003). Los proveedores de servicios de internet (PSIs) deben cobrar por el acceso a la Red sin favorecer un contenido por encima de otro (**Serrano-Cobos**, 2011). Internet es por naturaleza abierta y libre, es un canal a través del cual puede circular todo tipo de contenido a nivel global (**Brown**, 2008). La neutralidad es por tanto una característica definitoria y fundacional de la Red (**Cerf**, 2009).

Si internet se mantiene como una plataforma abierta, innovadora y neutral dependerá en gran medida del tipo de regulación que se acabe imponiendo

Sin embargo desde hace unos años se han producido cambios que están cuestionando dicho principio fundacional. Nos referimos a abusos o usos inadecuados que obligan a establecer mecanismos para evitar el spam, el cibercrimen, etc., a la posibilidad que tienen los PSIs para diferenciar y discriminar servicios o, también, a las demandas de más y mejores aplicaciones con el consiguiente incremento de tráfico en la Red (**Cordón-García**, 2011). A pesar de que cada una de estas tendencias no es por ella misma determinante para transformar el diseño de internet, sí favorecen un cambio en una misma dirección que podría motivar una modificación profunda tanto de sus principios básicos como de su diseño (**Blumenthal; Clark**, 2001).

El futuro de la Web está en gran medida condicionado por la neutralidad de la Red

El debate sobre la neutralidad viene motivado fundamentalmente por tres factores:

- 1) El incremento en el tráfico, que se multiplicará por cuatro en los próximos cinco años (**Cisco**, 2011). La causa principal es el consumo privado, que se manifiesta en una mayor demanda de vídeos. Se está avanzando así hacia una progresiva audiovisualización de internet¹ que exige irremediablemente un aumento de capacidad para dar soporte a los nuevos servicios y, a la vez, satisfacer las demandas de unos usuarios cada vez más exigentes.
- 2) La capacidad de los PSIs para interferir y controlar cualquier dato que circula a través de la Red gracias a la evolución tecnológica. Se han elaborado diferentes sistemas para monitorizar o manipular aplicaciones y tráfico, como los cortafuegos -diseñados para bloquear accesos no au-

torizados- o el *deep packet inspection (DPI)*, que permite a los PSIs rastrear y analizar el contenido que circula y determinar si un paquete de información debe ser bloqueado –por ejemplo, si estuviéramos ante conductas delictivas–, retrasado o priorizado aplicando mecanismos de ralentización de tráfico.

Podríamos distinguir varias causas de gestión ralentizada que aplican los PSIs:

- al compartir archivos P2P; fue llevada a cabo por la operadora norteamericana *Comcast*, posteriormente sancionada por la *Comisión Federal de Comunicaciones (FCC)*²;
- por exceso de consumo de ancho de banda o en horas concretas del día, por ejemplo de 5 de la tarde a 3 de la madrugada (**Marcon, et al.**, 2011);

Por otro lado, los PSIs deben hacer importantes inversiones para implementar redes de nueva generación, y a su vez su capacidad para interferir sobre el contenido que circula por ellas incrementa. Ello les permite fijar tarifas de conexión en base al tráfico generado o demandado, y en función de las cuotas pagadas priorizar unos contenidos y ralentizar otros.

3) Procesos de integración vertical, es decir, la posibilidad de los PSIs de comercializar aplicaciones o servicios empaquetados, que pueden generar eventuales riesgos para la libre competencia en el mercado de las comunicaciones electrónicas (**Yoo**, 2007; **Crocioni**, 2011). Un ejemplo del efecto nocivo que pueden tener los procesos de integración vertical para la neutralidad de la Red puede observarse en la fusión en 2001 entre la productora norteamericana *Time Warner* y la operadora de telecomunicaciones *AOL*. Existía el riesgo de que *AOL* en tanto que PSI priorizase en beneficio propio los contenidos producidos por *Time Warner* en detrimento de otros (**Hazlett; Caliskan**, 2008).

La visión aperturista intenta asegurar una red de acceso abierta que no distinga ni priorice un tráfico sobre otro

Internet es uno de los instrumentos más poderosos del siglo XXI y su protección constituye un valor fundamental para la evolución de la democracia y el progreso económico y social (**Castells**, 2009). Prueba de ello son las declaraciones de organizaciones internacionales vinculando directamente el acceso a internet a derechos fundamentales como la libertad de expresión e información (**Unesco**, 2011) o hablando incluso del derecho a internet como derecho fundamental (**ONU**, 2011; **Consejo de Europa**, 2010). Internet ha creado nuevas formas de poder (**Castells**, 2011)³ y libertad (**Benkler**, 2011), y las redes sociales (**Boyd; Ellison**, 2007) juegan un importante papel en la creación de estas nuevas estructuras de poder como demuestran las revoluciones árabes de la primavera de 2011 o el movimiento 15M en España.

2. Intereses en conflicto

Se puede enfocar el debate sobre la neutralidad de la Red desde dos posiciones: una aperturista y otra desreguladora (**Wu**, 2004).

El enfoque aperturista se basa en tres principios básicos que hacen internet diferente a otras plataformas:

- El primero hace referencia a la infraestructura, y parte de la premisa básica de que lo más importante para una red de comunicaciones es la existencia de una infraestructura pública. Su principal valor es indirecto ya que no recae en los propietarios de una red sino en la capacidad de innovación de sus usuarios y los programadores de aplicaciones. A pesar de que hoy en día difícilmente podemos hablar de infraestructuras públicas, sí existe el concepto de servicio universal (**Carlón-Ruiz**, 2006) y la conexión a internet constituye uno de ellos⁴. Estamos ante un servicio mínimo al que todos los ciudadanos tienen derecho y los poderes públicos imponen a los PSIs unas obligaciones o cargas del servicio a un precio asequible.

El enfoque aperturista se basa en 3 principios básicos que hacen internet diferente a otras plataformas: la infraestructura, la neutralidad y el *end-to-end*

- El segundo principio es el de neutralidad, que impide que se restrinja el acceso y uso de servicios y aplicaciones a los usuarios de internet, y para ello postula una separación clara entre Red y contenidos.
- Por último, un principio muy ligado a la neutralidad es el *end-to-end*, uno de los fundacionales de internet en tanto que Red descentralizada favorecedora de la innovación (**Lessig**, 2006; **Wu**, 2003; **Marsden**, 2011). Lo que se plantea desde una visión aperturista es asegurar una red de acceso abierta que no distinga ni priorice un tráfico sobre otro. Internet es un *common* o bien común de la sociedad y nadie puede ejercer control sobre él.

Frente a la posición aperturista se encuentra la desreguladora, cuyo punto de partida es la convergencia de redes, que no es otra que la evolución tecnológica natural a las redes multiservicio (voz, vídeo y datos). Los principios básicos de los desreguladores son la defensa de la propiedad privada de las redes como forma óptima de garantizar su máxima eficiencia. Un acceso abierto como el que proponen los aperturistas acabaría en la denominada “tragedia de los comunes”⁵ en la cual cualquier tipo de comunicación resultaría imposible. Para evitarla, los propietarios deben poder condicionar y restringir el acceso a la Red para que desarrolle su máximo potencial.

El segundo principio constitutivo de la posición desreguladora se refiere a la existencia de incentivos. El desarrollo de las redes requiere grandes inversiones que se acometerán si existen unas expectativas de retorno de la inversión y para ello los PSIs buscan nuevas vías de ingresos, como puede ser una oferta diferenciada de servicios fijando tarifas en base a los contenidos consumidos o consultados. Este argumento es a menudo presentado por los PSIs para justificar prácticas de ralentización de tráfico pero en realidad existe cierta ambigüedad respecto a esta cuestión (**Arlandis**; **Baranes**, 2011) ya que algunos estudios han demostrado que las cuotas de priorización no necesariamente incentivarían inversiones en

infraestructuras por parte de los PSIs (**Cheng**; **Bandyopadhyay**, 2011)⁶.

Finalmente, otro principio característico de esta posición es el de desregulación. Desde su óptica liberal cualquier intervención gubernamental que no sea para proteger el mercado o la propiedad privada es sospechosa en sí misma y el éxito de internet se encuentra precisamente en la no intervención de los gobiernos en su evolución. Son los operadores privados los que deben guiar el desarrollo de las redes de nueva generación y la configuración de internet en éstas. Para los defensores de esta posición, la regulación sería un obstáculo para la inversión y la innovación, y desaprueban rotundamente una regulación que limite sus facultades a favor de la neutralidad. La Red es el centro de todo cambio e innovación tecnológica, y en tanto que propiedad suya, se creen legitimados para actuar sobre los datos que circulan por ella siempre que lo crean conveniente (**Crocioni**, 2011; **Baker**; **Carlton**; **Sider**, 2010; **Yoo**, 2010; **Cave**; **Crocioni**, 2007; **Faulhaber**, 2007; **Sidak**, 2007).

El debate sobre la neutralidad de la Red es complejo y no se limita únicamente a una dimensión de regulación económica favorecedora de la competencia en el mercado sino que también están en juego los derechos fundamentales (**ONU**, 2011; **Unesco**, 2011). Es necesario un enfoque equilibrado que tenga en cuenta aspectos de regulación económica y a la vez fortalezca el derecho fundamental de acceso a internet (**Marsden**, 2010).

Los principios básicos de los desreguladores son la defensa de la propiedad privada de las redes, la existencia de incentivos y la desregulación

3. Neutralidad de la Red en la Unión Europea

A pesar de que fue objeto de discusión en la reforma de las comunicaciones electrónicas de la Unión Europea (UE), finalmente no quedó definida ni codificada de forma expresa en la reforma aprobada a finales de 2009. La única referencia se encuentra en un anexo de la *Directiva 2009/140/CE* que incluye la Declaración de la Comisión sobre la neutralidad de internet en la que manifiesta su compromiso a la hora de “preservar un carácter abierto y neutral de internet”. Con esta Declaración, sin trascendencia jurídica pero de indiscutible valor político, se consagra la neutralidad de internet como un objetivo político y un principio regulador en la UE⁷.

Los tres aspectos que constituirían la neutralidad de la Red en la UE son: capacidad de elección, transparencia y calidad de servicio

Sin embargo, de la reforma se desprenden tres aspectos que constituirían la neutralidad de la Red en la UE: capacidad de elección, transparencia y calidad de servicio. Se debe garan-

tizar a los usuarios capacidad para acceder, distribuir o utilizar todo aquello que libremente escojan sin discriminación alguna sobre su elección. La transparencia implica que los contratos de conexión a internet deben incluir información sobre las limitaciones de acceso y utilización de servicios y aplicaciones de los usuarios. Finalmente, las prácticas restrictivas del tráfico han de ser mínimas a fin de preservar la calidad del servicio y prevenir su degradación.

La falta de concreción de la normativa europea a la hora de fijar qué se considera información relevante y los requisitos mínimos de calidad del servicio ha planteado el debate sobre la posibilidad de promover una regulación complementaria que asegure la neutralidad de la Red en Europa⁸. El ente asesor de la Comisión, el *Body of European Regulators for Electronic Communications (Berec)* –Organismo de Reguladores Europeos de las Comunicaciones Electrónicas (Orece)⁹-, está trabajando en un informe sobre la transparencia y los requisitos de calidad del servicio¹⁰ previsto para principios de 2012. Para el Orece la transparencia informativa es una cuestión fundamental, por ello en el borrador del informe se enumeran algunas de las propiedades que debe tener la información: accesible, comprensible, completa y concisa. En definitiva, que permita al usuario elegir con pleno conocimiento de causa (Trainar, 2011).



<http://erg.eu.int>

4. Conclusiones

El potencial de internet para fortalecer las diferentes dimensiones que adquiere el derecho a la libertad de expresión del siglo XXI depende enormemente de las condiciones de acceso. Las amenazas a la neutralidad representan uno de los últimos intentos de ejercer control sobre las diferentes formas de comunicación y expresión de la sociedad mediante actos de censura, en este caso, de censura privada practicada por los PSIs.

A pesar de que la UE se inclina hacia una protección del principio de neutralidad, su posición todavía no está cerrada. Europa tiene la oportunidad de convertirse en un referente mundial de la regulación de internet si promulga una norma que prevea mecanismos concretos para garantizar la neutralidad de la Red.

5. Notas

1. Según un reciente estudio de *Interactive Advertising Bureau (IAB)* –asociación que representa al sector de la publicidad y la comunicación digital en España– los internautas en España dedican el 25% de su tiempo en internet al consumo de vídeos. Además, más de la mitad de los usuarios (un 53%) lo hacen diariamente (IAB, 2011).

2. El enfrentamiento entre la FCC y Comcast llegó a la Corte de Apelación del Distrito de Columbia en 2010, que sentenció que la FCC no tenía autoridad para regular la forma en que los PSIs gestionan el tráfico en internet. *The Economist: Comcast vs the FCC: raze the mystery house* (10 abril de 2010).

<http://www.economist.com/node/15867976>

3. Resulta interesante enumerar los distintos tipos de poder en la Red que propone Castells (2011): *networking power*, *network power*, *networked power* y *networking-making power*. Cada uno de ellos se organiza y estructura de forma distinta.

4. La Ley 2/2011 de Economía sostenible fija la conexión a internet en banda ancha de 1 Mbyte como servicio universal.

5. Tecnicismo utilizado para designar los casos en que una red de comunicaciones queda saturada y es inapta para cualquier tipo de comunicación.

6. Estos autores se preguntan si los PSIs tienen más incentivos para mejorar las infraestructuras en caso de eliminarse la neutralidad de la Red y llegan, precisamente, a la conclusión contraria (Cheng; Bandyopadhyay, 2008).

7. España todavía no ha adoptado la reforma europea de las comunicaciones electrónicas a su ordenamiento interno; se prevé que lo haga a lo largo de 2012.

8. Mientras que la Comisión Europea considera que el marco regulatorio es suficiente, en noviembre de 2011 el Parlamento Europeo aprobó por unanimidad una resolución en la que llamaba a la Comisión a adoptar medidas complementarias para asegurar la neutralidad en la UE.

<http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+MOTION+B7-2011-0572+0+DOC+XML+V0//ES>

9. Creado en 2009 con el Reglamento (CE) 1211/2009 del Parlamento Europeo y del Consejo, de 25 de noviembre de 2009.

10. Puede consultarse el borrador del informe en:

http://erg.eu.int/doc/berec/bor_11_44.pdf

6. Bibliografía

Arlandis, Antonin; Baranes, Edmond. "Interactions between network operators, content producers and internet intermediaries: empirical implications of net neutrality". *Intereconomics*, 2011, v. 46, n. 2, pp. 98-105.

<http://www.springerlink.com/content/x682l67q21w252p8/>
<http://dx.doi.org/10.1007/s10272-011-0370-y>

Becker, Gary; Carlton, Dennis; Sider, Hal. "Net neutrality and consumer welfare". *Journal of competition law and economics*, 2010, v. 6, n. 3, pp. 497-519.

<http://dx.doi.org/10.1093/joclec/nhq016>

Barlow, John P. *A declaration of the independence of cyberspace*, 1996.

http://w2.eff.org/Censorship/Internet_censorship_bills/barlow_0296.declaration

Benkler, Yochai. "Networks of power, degrees of freedom". *International journal of communication*, 2011, v. 5, pp. 721-775.

<http://ijoc.org/ojs/index.php/ijoc/article/viewFile/1093/551>

Brown, Ian. "Internet filtering – be careful what you ask for". En: Kirka, Süheyla; Hanson, Luett. *Freedom and prejudice: approaches to media and culture*. Istanbul: Bahcesehir University Press, 2008, pp. 74-91. ISBN: 978 9756437810 <http://ssrn.com/abstract=1026597>

Blumenthal, Marjory; Clark, David. "Rethinking the design of the internet: the end-to-end arguments vs. the brave new world". En: Compaine, Benjamin; Greenstein, Shane (eds.). *Communications policy in transition: the internet and beyond*. Cambridge: MIT Press, 2001, pp. 91-139. ISBN: 0262032929 http://www.csd.uoc.gr/~hy558/papers/Rethinking_2001.pdf

Boyd, Danah; Ellison, Nicole. "Social network sites: definition, history and scholarship". *Journal of computer mediated communication*, 2007, v. 13. n. 1, pp. 210-230. <http://jcmc.indiana.edu/vol13/issue1/boyd.ellison.html>

Castells, Manuel. *Comunicación y poder*. Madrid: Alianza. 2009. ISBN 9788420684994

Castells, Manuel. "A network theory of power". *International journal of communication*, 2011, v. 5, pp. 773-787. <http://ijoc.org/ojs/index.php/ijoc/article/viewFile/1136/553>

Carlón-Ruiz, Matilde. "El servicio universal de telecomunicaciones". *Revista de administración pública*, 2006, n. 171, pp. 35-78. <http://www.cepc.es/es/Publicaciones/revistas/revistas.aspx?IDR=1&IDN=622&IDA=26482>

Cave, Martin; Crocioni, Pietro. "Does Europe need a network neutrality rules?". *International journal of communication*, 2007, v. 1, pp. 669-679. <http://ijoc.org/ojs/index.php/ijoc/article/viewFile/157/80>

Consejo de Europa. *Declaración del Comité de Ministros sobre la neutralidad de la Red*, 2010. <https://wcd.coe.int/ViewDoc.jsp?id=1678287>

Cordón-García, José-Antonio. "El final del libro y el principio de la lectura: los libros electrónicos y el fenómeno iPad". *Anuario ThinkEPI*, 2011, v. 5, pp. 132-138. <http://www.thinkepi.net/el-final-del-libro-y-el-principio-de-la-lectura-los-libros-electronicos-y-el-fenomeno-ipad>

Cheng, Hsing; Bandyopadhyay, Shubho; Guo, Hong. "Traffic jams, ISPs and net neutrality", *GigaOm*, 13 Nov. 2011. <http://gigaom.com/broadband/traffic-jams-isps-and-net-neutrality/>

Cheng, H. Kenneth; Bandyopadhyay, Subhajyoti; Guo, Hong. "The debate on net neutrality: a policy perspective". *Information systems research*, 2011, v. 22, n. 1, pp. 1-27. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=959944

Cisco. *Cisco visual networking index: forecast and methodology 2010-2015*, 2011. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html

Cerf, Vint. "The open internet. What is it and why it matters". *Telecommunications journal of Australia*, 2009, v. 59, n. 2, pp. 18.1-18.10. <http://tja.org.au/index.php/tja/article/view/104>

Crocioni, Pietro. "Net neutrality in Europe: desperately seeking a market failure". *Telecommunications policy*, 2011, v. 35, n. 1, pp. 1-11. <http://dx.doi.org/10.1016/j.telpol.2010.12.007>

Faulhaber, Gerald. "Network neutrality: the debate evolves". *International journal of communication*, 2007, v. 1, pp. 680-700. <http://ijoc.org/ojs/index.php/ijoc/article/view/151/85>

Hazlett, Thomas; Caliskan, Anil. "Natural experiments in US broadband regulation". *George Mason law & economics research paper*, n. 08-04, 2008. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1093393

IAB. *Vídeo marketing y publicidad en el vídeo online: aproximación desde la perspectiva del usuario*, 2011. <http://www.iabspain.net/ver.php?mod=noticias&identificador=110>

Lessig, Lawrence. *Code and other laws of the cyberspace*. New York: Basic Books. 2006. ISBN 0465039146

Marcon, Massimiliano; Dischinger, Marcel; Gummadi, Krishna P.; Vahdat, Amin. "The local and global effects of traffic shaping in the internet". En: *Third intl conf on communication systems and networks (Comsnets)*, 2011. ISBN 978 14244895-0 <http://www.mpi-sws.org/~mmarcon/Traffic-Shaping-COMSNETS.pdf>

Marsden, Christopher. "Network neutrality: history, regulation and future". En: Cerrillo-Martínez, Agustí et al. *Net neutrality and other challenges for the future of the internet*. Barcelona: Huygens, 2011, pp. 29-49. ISBN 978846947037-4 <http://openaccess.uoc.edu/webapps/o2/handle/10609/8341>

Marsden, Christopher. *Net neutrality: towards a co-regulatory solution*. Londres: Bloomsbury Publ., 2010. ISBN 978 1849660068

ONU. *Informe del Consejo de Derecho Humanos sobre la promoción y protección del derecho a la libertad de expresión y opinión*, 2011. http://www2.ohchr.org/english/bodies/hrcouncil/docs/17session/A.HRC.17.27_en.pdf

Serrano-Cobos, Jorge. "Informe de situación. Tendencias tecnológicas de 2010". *Anuario ThinkEPI*, 2011, v. 5, pp. 241-248.

Sidak, Gregory. "What is the network neutrality debate really about?" *Intl journal of communication*, 2007, v. 1, pp. 377-388. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1006587

Tramullas, Jesús. "El que tiene los datos, tiene el poder". *Anuario ThinkEPI*, 2011, v. 5, pp. 117-120. <http://www.thinkepi.net/el-que-tiene-los-datos-tiene-el-poder>

Trainar, Nadia. "Berec's approach to net neutrality". *Communication & strategies*, 2011, v. 84, pp. 56-71.

Unesco. *Freedom of connection, freedom of expression*, 2011.
http://portal.unesco.org/ci/en/ev.php-URL_ID=31398&URL_DO=DO_TOPIC&URL_SECTION=201.html

Yoo, Christopher. "Network neutrality or internet innovation?". *Regulation*, 2010, v. 33, pp. 22-29.

<http://www.cato.org/pubs/regulation/regv33n1/regv33n1-6.pdf>

Wu, Tim. "Network neutrality, broadband discrimination". *Journal of telecommunications and high technology law*, 2003, v. 2, pp. 141-177.

http://papers.ssrn.com/sol3/papers.cfm?abstract_id=388863

Wu, Tim. "The broadband debate: a user's guide". *Journal of telecommunications and high technology law*, 2004, v. 3, n. 69, pp. 69-96.

http://papers.ssrn.com/sol3/papers.cfm?abstract_id=557330

Jornada sobre Modelos de negocio con software libre de gestión de contenidos

Fecha: jueves, 29 de marzo de 2012

Horario: 9:00-14:00 h y de 16:00-19:30 h

Lugar: Fundación Carlos de Amberes

Claudio Coello, 99. 28006 Madrid

http://www.fcamberes.org/paginas/es/stc_ins.htm

Organizan:

Baratz Servicios de Teledocumentación SA

Revista *El profesional de la información* (EPI)

Entre otros intervendrán:

Marshall Breeding (consultor)

Patricia Russo (consultora)

Alicia Sellés (MASmedios)

Juan Repiso (Baratz)

Ricardo Éito-Brun (UC3M)

Jesús Casado (Baratz)

Jesús Tramullas (Unizar)

Juan Antonio Pastor (UMurcia)

Tomàs Saorín (UMurcia)...

Entrada libre previa inscripción

Inscripciones:

eventos@baratz.es



baratz

gestionando el conocimiento

El profesional de la

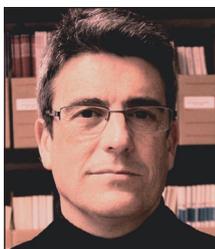
— **información** —



EL LIBRO ELECTRÓNICO: INVARIANZAS Y TRANSFORMACIONES



José-Antonio Cordón-García y Carlos A. Lopes



José-Antonio Cordón-García es profesor titular de la *Universidad de Salamanca*. Miembro del Consejo Ejecutivo del *Instituto de Historia del libro y de la Lectura*, ha publicado numerosos artículos y monografías sobre la industria editorial en general y el libro electrónico en particular. Dirige el grupo de investigación *E-lectra* centrado en el análisis de la edición electrónica y las nuevas formas de creación y comunicación escrita. Es director de la revista *Pliegos de Yuste: revista de ciencia, cultura y pensamiento europeos*, y presidente de la *Asociación Española de Bibliología*. Forma parte del proyecto *Territorio Ebook*, de la *Fundación Germán Sánchez Ruipérez*.

Universidad de Salamanca
Facultad de Traducción y Documentación
Francisco Vitoria, 6-16. 37008 Salamanca, España
<http://www.pliegosdeyuste.eu>
jcordon@usal.es



Carlos A. Lopes es director del *Centro de Documentación del Instituto Universitário de Ciências Psicológicas, Sociais e da Vida (ISPA)*. Es profesor del *ISPA-IU* donde imparte asignaturas relacionadas con la alfabetización informacional (alfin), recursos y comunicación en psicología. Es doctor en biblioteconomía y documentación por la *Universidad de Salamanca* y por la *Universidade de Coimbra*. Su trayectoria investigadora se centra en el estudio de la alfin y la calidad en bibliotecas universitarias. Es autor de diferentes artículos en revistas especializadas. Ha formado parte del comité *SciELO Portugal* y de la *Biblioteca do Conhecimento Online (B-on)*.

*Instituto Univ. de Ciências Psicológicas,
Sociais e da Vida (ISPA)*, Centro de Documentação
Jardim do Tabaco, 34. 1149 041 Lisboa, Portugal
<http://www.ispa.pt>
clopes@ispa.pt

Resumen

Los libros electrónicos han renovado el concepto tradicional de libro, favoreciendo la socialización de la lectura, la participación del usuario y la ruptura de la cadena tradicional de la edición. Las empresas editoriales se han ido adaptando al nuevo contexto con propuestas novedosas, en las que el contenido reviste una importancia cada vez mayor. Se está produciendo una lucha entre los grandes grupos como *Amazon*, *Apple* y *Google* por el control del mercado a medida que éste se consolida. Esas marcas se van internacionalizando y durante 2011 han desembarcado en varios países europeos y latinoamericanos, manteniendo posiciones encontradas.

Palabras clave

Libros electrónicos, Lectura social, *Amazon*, *Apple*, *Google*.

Title: The e-book: invariances and transformations

Abstract

Electronic books have renewed the traditional concept of the book, encouraging the socialization of reading and user participation while breaking the traditional pattern of publishing. The publishing companies have been adapting to the new context with new proposals, in which the content is of increasing importance. There is a struggle between large groups such as *Amazon*, *Apple* and *Google* for the control of the market as it matures. These brands are internationalizing and in 2011 landed in several European and Latin American countries, maintaining their positions as direct competitors.

Keywords

Ebooks, Social reading, *Amazon*, *Apple*, *Google*.

Cordón-García, José-Antonio; Lopes, Carlos A. "El libro electrónico: invarianzas y transformaciones". *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 83-90.

<http://dx.doi.org/10.3145/epi.2012.ene.11>

Artículo recibido el 28-10-11
Aceptación definitiva: 11-11-11

1. El futuro del libro

Desde que apareció la informática doméstica y se vislumbró la posibilidad de que los textos escritos se reprodujeran en cualquier ordenador, el futuro del libro se convirtió en objeto de debate por parte de autores, editores, comunicólogos, semiólogos, filólogos y todos aquellos que, de manera directa o indirecta, tenían algo que ver con la cultura escrita. Desde el principio la cuestión se planteó como una disyuntiva en la que la exclusividad, oposición o prevalencia cifraba el núcleo polémico: impreso vs digital. La discusión ha llegado hasta nuestros días, en una formulación reductiva que ignora que estamos no frente a un escenario de desaparición sino de cambio de naturaleza. No se trata de un mero traspaso de un formato a otro, en el que el papel permanece de alguna manera como indicio o referente en lo digital. El fundamento de la digitalización no radica en la conversión de obras impresas ya editadas a formato electrónico, sino en la creación de nuevas formas digitales, en muchos casos provenientes de actores diferentes de los implicados en la industria tradicional del libro impreso.

Existen ya numerosas editoriales que trabajan sólo con el formato digital y que marcan las nuevas tendencias de lo que se aventura como el libro del futuro. Entre ellas: *Atavist* y *Byliner* en Estados Unidos; *Publie.net*, *Walrus Studio* y *NumerikLivre* en Francia; *24 Symbols*, *Amphibia* y *Musa a las 9* en España. Por no hablar de proyectos de integración de soportes y medios como *Pottermore*, modelo de lo que se ha dado en llamar lectura inmersiva, en la que la confluencia de medios conduce a experiencias de lectura enriquecidas y personalizables. Algunos libros infantiles integran textos con la imagen en movimiento, la interactividad y el sonido, transformándose en experiencias multimedia.

En el universo digital surgen nuevos géneros, desde los más frívolos *fan fictions* –grupos de entusiastas de una obra o un personaje que continúan una saga o un título en particular para compartirlo entre ellos-, a las wikinovelas, diginovelas, obras colaborativas, obras transmedia..., toda una nueva narrativa basada en las nuevas tecnologías.

Lo que es fundamental es saber distinguir las tendencias de fondo de lo que son meros epifenómenos. Decía **Henri-Jean Martin**, en el epílogo de su obra *Historia y poderes de lo escrito* (Trea, 1999), que la misión de nuestra generación es comunicar a nuestros descendientes que el progreso técnico no implica obligatoriamente el rechazo irreflexivo hacia los soportes del pasado.

<http://atavist.net>

<http://byliner.com>

<http://www.publie.net>

<http://www.walrus-books.com>

<http://www.numeriklivres.com>

<http://www.24symbols.com>

<http://www.amphibiaeditorial.com>

<http://www.musaalas9.com>

El libro es un documento imperecedero, sometido a cambios y transforma-



Ejemplo de lectura inmersiva. Fuente: <http://www.prweb.com>

ciones permanentes, como han puesto de manifiesto **Martin** (Trea, 1999), **Chartier y Cavallo** (Taurus, 2011), **Manguel** (Lumen, 2005) o **Furtado** (Trea, 2007, *Booktailors*, 2011). No hay que olvidar que la forma código con que ahora lo identificamos tiene sólo 500 años de antigüedad, nada comparado con la era del manuscrito, por ejemplo. Por lo tanto el libro como tal seguirá existiendo como principal vía de comunicación social para pensamientos, reflexiones y discursos que exijan mayor profundidad y extensión.

El fundamento de la digitalización no radica en la conversión de obras impresas a formato digital, sino en la creación de nuevas formas digitales

La unidad de pensamiento puede ser la frase, el párrafo, la sección o el capítulo de un libro. Pero la forma extensa, una historia, un relato unificado, una discusión completa..., ejerce una fascinación, una extraña atracción sobre nosotros. En la Red es lo mismo: podemos cortar el libro en bits y enlazarlos o tejerlos con otros, pero la organización de primer nivel del libro seguirá centrando la atención.



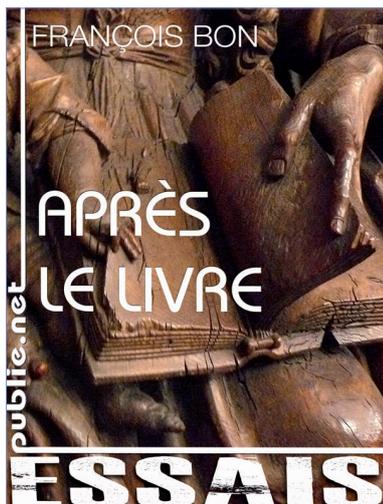
Pottermore es una web para usar mientras se leen libros de Harry Potter, <http://www.pottermore.com/es>

Es cierto que su forma cambiará. Las nuevas tecnologías han abierto su condición hermética convirtiéndolo en un medio en el que la integración multimedia, la participación social y la colaboración son cada vez más importantes. Programas de autopublicación como los de *Amazon*, *Barnes & Noble*, *Apple*, *Smashwords*, etc.; sistemas de lectura social como *aNobii*, *Rethink Books*, *Copia*, *Library Thing*, etc., así lo atestiguan.

También algunas obras tradicionales se van transformando: los diccionarios, por ejemplo, se convierten en servicios de tratamiento lingüístico y terminológico.

El papel no es más que una de las posibles versiones de libro que se pueden desplegar y concurrir simultáneamente en el mercado.

El libro tenderá a convertirse en servicio. De objeto se transformará en flujo, escapando al control de sus productores en beneficio de los que los comunican y favorecen su circulación. Un ejemplo de ello son las obras de la editorial *Publie.net*. Uno de sus títulos, por ejemplo, *Après le livre*, de **François Bon**, se puede leer en pdf, en epub, en mobipocket, en streaming, etc., pero la obra a su vez se va actualizando, de tal manera que, publicada en enero de 2011 a día de hoy (octubre de 2011) ha experimentado ya 9 actualizaciones.



<http://www.publie.net/fr/ebook/9782814504103>

Por otra parte los lectores están cada vez más involucrados en el proceso de las obras mediante sistemas que interactúan con los autores, como *@author* de *Amazon* que permite comentar pasajes del libro con sus creadores, lo mismo que ocurre con el programa *Book country* de *Penguin*, produciéndose una interactividad antes imposible. También se dan versiones personalizadas de las obras, anotaciones y una serie de actividades que antes formaban parte de las prerrogativas exclusivas del autor y el editor. Además, las propias aplicaciones de lectura incluyen ya la posibilidad de socialización de la lectura y personalización de la misma de una manera intuitiva para el usuario. Tal es el caso de *iBooks*, la aplicación de lectura del *iPad*, o de *Stanza* o de *Kobo*. La experiencia de lectura se hace cada vez más colaborativa y abre unos horizontes a la misma, vinculando la experiencia personal a servicios en red como los de *ReadMill* o *Amazon*.

Por lo tanto apreciamos que no sólo cambia la estructura del mercado, sino el mismo concepto de libro, que ha evolucionado en los últimos 40 años de la mano de dos personas reciente y prematuramente fallecidas¹, **Michael Hart**, creador del *Proyecto Gutenberg* en 1971, el primer portal dedicado a la digitalización de obras cuando en Occidente

comenzaba a desarrollarse la informática, y **Steve Jobs**, fundador de *Apple* y creador de ingenios como el *iPod* o, más recientemente el *iPad*.

2. Impacto de los libros electrónicos en España: tendencias actuales y futuras

En España la producción de libros electrónicos ha experimentado un crecimiento progresivo, siendo del 155% en el año 2010 (*Panorámica*, 2011). Frente a los 12.514 ISBNs concedidos en 2009, en 2010 han sido 18.500, producidos por pequeñas y medianas editoriales que aglutinan el 60% del mercado, y con un significativo 8,2% relativo a la autedicción (casi un 30% con respecto al año 2009).

Las cifras de lectura no van a la zaga. Según *Hábitos de Lectura y Compra de Libros* (2011) en 2010 cerca de la mitad de la población de 14 o más años (47,8%) es lectora de formato digital. De ésta, un porcentaje todavía elevado (48%) lee en el ordenador, un 6,6 en el móvil, y un 0,8% en el e-reader, porcentaje que se eleva al 1,1% en la encuesta del 2º cuatrimestre de 2010. El perfil de los lectores digitales es interesante porque ilustra acerca de las tendencias y los desarrollos futuros.

Una evidencia de esta progresión es que en los datos aportados para el primer semestre de 2011, la lectura digital ha pasado de 47,8% al 52,5%. Así pues los lectores en soporte digital han crecido en los primeros 6 meses del año 4,7 puntos respecto al año 2010. En libros electrónicos, 2,1 puntos respecto al 1º cuatrimestre de 2010 y 1,5 puntos respecto a la media del pasado año.

La industria editorial también está cada vez más involucrada en este proceso. Según el último informe *El libro digital en España* (*Federación de Gremios de Editores de España*, 2011) el 75% de los editores ha elaborado planes de digitalización de su catálogo. Sin embargo en España se podría hablar de una cierta dicotomía entre el funcionamiento del mercado y el de la sociedad, en el sentido de que mientras que la recepción por parte de ésta es cada vez más importante, como demuestran las cifras de venta de dispositivos de lectura electrónica, la actitud de los editores evidencia una excesiva prudencia, constatable en el hecho de que la colección de obras electrónicas disponibles es todavía escasa en relación con el nivel del sector, cuarto en importancia en el mundo.

“ En España los usuarios son cada vez más receptivos, pero los editores son excesivamente prudentes ”

Los editores temen la posibilidad de canibalización del libro impreso por el digital, sin entender que lo que se está produciendo es una oportunidad de vender que, de cualquier modo, reviste la contundencia de los movimientos de carácter inexorable. No se trata ya de una opción que se pueda elegir o no, sino de una exigencia del mercado. Y así se ha puesto de manifiesto en los foros profesionales como *Liber*, en el que la edición digital ha sido la protagonista en los

últimos años. O en el último *Ficod* (*Foro internacional de contenidos digitales*) celebrado en Madrid en noviembre de 2011, en el que se presentó el informe del *Observatorio Nacional de las Telecomunicaciones y la Sociedad de la Información* (*Ontsi*, 2011), según el cual en 2010 por primera vez se habían consumido en España más contenidos digitales que impresos. Los informes de *The Cocktail Analysis* (2011) y de *The New Media Consortium* (*The horizon report*, **Johnson et al.**, 2011), entre otros, confirman esta tendencia hacia el consumo de contenidos digitales, en la que las pantallas alcanzan un protagonismo singular, con tablets y smartphones como punta de lanza, y en la que la lectura de libros reviste cada vez más importancia (*ComScore*, 2011).
<http://en.www.mcu.es/libro/MC/ObservatorioLect/Estudios/estudiosOtros/NuevasTec.html>
<http://www.tcanalysis.com>

La progresión es permanente y será cada vez más importante en la medida en que empresas como *24 Symbols* –que propone la distribución de libros electrónicos en la nube con sistemas *freemium* (gratuitos pero con publicidad no invasiva) o *premium* (sin anuncios)- se vayan consolidando. Cada vez hay más propuestas editoriales, y únicamente falta que el mercado madure con unos precios más aquilatados, y que vayan siendo mayoría las jóvenes generaciones de nativos digitales, que constituyen un motor de desarrollo.

Los movimientos de los editores evidencian una implicación cada vez más importante en lo digital que no sólo afecta a las obras de ficción sino que tiene ramificaciones en el libro científico, y en el libro infantil y juvenil. Precisamente en este último sector se han generado importantes desarrollos para tablet principalmente por parte de *Oceanhouse Media*, *Ruckus Media*, *Smashing Ideas*, y *Trilogy Studios*. Valorando la importancia de este emergente sector digital, *Random House* compró recientemente al programador de *Smashing Ideas*. Una muestra de la importancia que va cobrando la constituye el hecho de que en la última feria del libro de Francfort se dedicara un panel específico a la edición infantil con el significativo título de *Children's publishing goes digital*. Pero en el resto de los sectores se multiplican igualmente propuestas novedosas provenientes de actores muy diversos, como puede ser un medio de comunicación como *La vanguardia*, que en noviembre de 2011 lanzó su propia editorial de libros electrónicos, *Ebooks de vanguardia*, con varios títulos ya en cartera y colecciones como *Periodismo de vanguardia*: más de un centenar de reportajes y series de artículos de *La vanguardia* publicados en colaboración con *Amazon* con motivo del lanzamiento de la tienda *Kindle* en España el 1 de diciembre de 2011. La mayor parte de títulos cuestan 0,94 euros.

Ediciones B ha lanzado también su propia propuesta editorial, una colección de libros sin DRM, contraviniendo la

tendencia mayoritaria seguida por el sector editorial de protección de obras mediante este sistema. Se trata de la colección *B de Books* con más de cien títulos ya en el mercado a precios en muchos casos inferiores a los 2 euros.

La editorial *Debate* recurre a la fórmula de lectura en streaming. En 2011, lanzó *Endebate*, una colección sólo digital de textos de 10.000 palabras, disponibles tanto para *iPad* como en formato epub.

“ El poder del libro de ofrecer una historia autónoma, un relato unificado, una discusión completa..., ejerce una extraña atracción sobre nosotros ”

El grupo *Planeta*, por su parte, lanzó igualmente a finales de 2011 *Scyla* y *Zafiro*, dos nuevos sellos editoriales creados para publicar exclusivamente eBooks. *Scyla* está orientada a la ciencia ficción, la fantasía y el terror en tanto que *Zafiro* se ocupa del género romántico
<http://www.oceanhousemedia.com/products>
<http://www.ruckusmediagroup.com>
<http://smashingideas.com>
<http://www.trilogystudios.com>

3. Internacionalización del sector: edición comercial y edición académica

Desde hace muchos años la edición se ha caracterizado por su tendencia a la internacionalización y a la concentración de empresas. Los grandes conglomerados editoriales tienen vocación de monopolio, y esto es un axioma para grupos pre-electrónicos como *Bertelsmann*, *Elsevier*, *Hachette*, *Planeta*, etc.



Web *iBookStore* de Apple, <http://www.apple.com/es/ipad/built-in-apps/ibooks.html>

La edición electrónica en general y los libros electrónicos en particular, no son una excepción, y debido a la globalización dicha tendencia se ha acentuado. Los sistemas de integración vertical como los llevados a cabo por *Apple* o por *Amazon* tienen la misma orientación de acaparar mercado. *Apple* lo ha demostrado con su política de admisiones y rechazos de firmas de la competencia en las apps de *iBookStore*. *Google*, *Amazon* y *Apple* han iniciado una política de expansión territorial que competirá en el ámbito lingüístico, un terreno en el que tienen todas las ventajas de las sinergias generadas desde sus plataformas. Esta expansión se ha materializado en el desembarco de *Amazon* en España en septiembre de 2011, lo mismo que hizo *Apple* que abrió su *iBookStore* en ese mismo mes, con la particularidad de que importantes grupos editoriales como *Planeta*, la empresa multimedia y de edición más importante de España, está incorporando su catálogo a la tienda de *Apple*. Además estarán las obras de *Seix Barral*, *Destino*, *Espasa*, *Crítica*, *Paidós* y *Mondadori*. El fondo de 400.000 obras con que cuenta *iBookStore* se verá incrementado con una mayor disponibilidad de títulos en castellano.

La irrupción más desequilibrante en España puede ser la de *Google eBooks*, aunque se ha retrasado y se anuncia ahora para principios de 2012

Sin embargo, con ser importantes estos movimientos, la irrupción más desequilibrante puede ser la de *Google eBooks*. *Google* cuenta con un catálogo de varios millones de títulos y anuncia unos descuentos en el precio de casi el 80% sobre los libros en papel. Se dará, por tanto, una auténtica pugna por la conquista del mercado. *Google Books* ha acostumbrado a sus usuarios a buscar y encontrar millones de títulos difíciles de localizar en formato papel. Los ha acostumbrado a descubrir citas o referencias en un corpus textual ingente, y ha elevado el libro a la misma condición que las noticias, sitios web y otros servicios propios de internet. Con ello ha ayudado a popularizar un tipo de documento propio de la biblioteca y de la librería, y ha instaurado un hábito muy interesante para el paso siguiente: la creación de *Google eBooks*, transformando su plataforma en vendedora y distribuidora de libros. La facilidad de uso, la variedad de posibilidades y la existencia de un catálogo ingente de títulos aventuran una seria competencia para el resto de empresas del sector. *Google* tiene además la ventaja añadida de la complementariedad con el resto de sus servicios, que permitirá la integración e implementación de *Google eBooks* con sistemas de búsqueda y redes sociales. <http://books.google.com/ebooks>

The screenshot shows the Google eBookstore interface. At the top, there's a search bar with 'sevilla' entered and a dropdown menu showing 'Todos los eBooks en Google'. Below the search bar, there's a navigation bar with 'Página principal' and 'Resultados de búsqueda de "sevilla"'. The main content area displays a grid of book covers with their titles, authors, and review counts. The books listed include 'El Alcázar de Sevilla', 'El Burlador de Sevilla', 'Descripción artística de...', 'La garduña de Sevilla', 'Guía general de Sevilla', 'Anales eclesiásticos y s...', 'Historias y parábolas mo...', 'Noticia histórica del orig...', 'Descripción artística de l...', and 'Sevilla y Cádiz'.

A finales de 2011 en la web *eBookStore* sólo se ofrecían gratuitamente los libros escaneados por *Google* cuyos derechos de autor han caducado, más de 2 millones. <http://books.google.es/ebooks>

No hay que olvidar la derivación latinoamericana del mercado de libros en español, hasta ahora en cierto modo estable, gracias al efecto proteccionista del precio fijo del libro, al que están obligados todos los actores que intervienen en España. Esto dificulta una competencia real, pero previsiblemente la situación irá cambiando en la medida en que las leyes de libre competencia afecten al sector del libro. También afectará la introducción del IVA reducido, pues el actual (del 18% en todos los países europeos) le resta competitividad en un mercado cada vez más globalizado.

De cualquier modo lo que es una realidad es la presencia de *Amazon* en España con una oferta de títulos cada vez más abundante y una política de precios agresiva que se ha hecho extensiva a la venta de su dispositivo estrella, el *Kindle*. A finales de diciembre de 2011, tan sólo tres meses y medio después de su apertura, *Amazon* tenía ya 2,2 millones de usuarios únicos mensuales, la mitad de los que poseía el líder español *El Corte Inglés*. La industria editorial española se está posicionando a través de la creación de sellos y plataformas con propuestas que pueden ser competitivas si consiguen consolidar una oferta atractiva en contenidos, precios y experiencia de usuario.

Amazon no es sólo una empresa, sino que constituye un ecosistema en el que la oferta de títulos tiene su correlato en la existencia de unos dispositivos de lectura con una relación calidad/precio considerablemente buena. El lanzamiento en septiembre de 2011 de la primera tablet de la marca, el *Kindle Fire*, representa otro importante paso en la conquista del entorno global de mercado de los libros electrónicos por parte de esta empresa. <http://www.amazon.es>

En este contexto internacional la edición académica es un nicho de mercado en constante progresión, abocada inevitablemente a su transformación digital. Las universidades que no lo entiendan así quedarán rezagadas y perderán el tren de la modernidad. Ya ocurrió con las revistas científicas y ocurrirá de nuevo con los libros.



Amazon vende casi de todo en sus webs y los libros han pasado a ser un producto secundario. Su principal fuente de ingresos es la venta de alojamiento (hosting) gracias a su potente y fiable estructura informática. <http://amazon.es>

Las iniciativas de edición científica no dejan de proliferar. En octubre de 2010 en España se creó el portal *Unebook*, que agrupa la producción electrónica de todas las editoriales universitarias españolas. <http://www.unebook.es>

En los EUA, *Jstor*, servicio de suscripción de revistas académicas, dio a conocer sus planes para libros electrónicos durante la reunión de enero de 2011 de la *American Library Association (ALA)*. Firmó acuerdos con 5 de las editoriales universitarias más importantes: *Chicago, Minnesota, North Carolina, Princeton y Yale* para poner sus libros online a lo largo de este año 2011. <http://about.jstor.org/books>

Mientras tanto, el proyecto *MUSE*, que al igual que *Jstor* suministra acceso a revistas académicas, ha creado una plataforma denominada *Project MUSE Editions* para distribuir online los libros electrónicos de docenas de editoriales; está previsto que funcione en enero de 2012.

Otro grupo de editores, entre ellos *New York University Press, Rutgers University Press, Temple University Press y University of Pennsylvania Press*, ha elaborado un programa colaborativo de plataforma digital para libros denominado *University Press e-Book Consortium (UPEC)*.

En marzo de 2011 *UPEC* y *Project MUSE Editions* anunciaron su fusión, creando el *University Press Content Consortium (UPCC)*. http://muse.jhu.edu/about/new/ebook_collections.html

Open Access Publishing in European Networks (Oapen) es un proyecto lanzado en octubre de 2010 después de 2 años de elaboración, que tiene la particularidad de distribuir los contenidos en acceso abierto. Los derechos varían en función de las distintas políticas editoriales, aunque la mayoría de las obras están bajo licencia *Creative Commons*. Para la selección de las obras cuenta con un sistema de revisión por pares. <http://www.oapen.org>

La *Association of American University Presses (AAUP)* hizo público en marzo de 2011 un estudio sobre la situación de la digitalización en las diferentes editoriales asociadas. Se trata

de una encuesta para conocer qué estrategias se siguen y en qué medida están adoptando el formato digital para las publicaciones académicas. Todas las editoriales universitarias están llevando a cabo al menos un par de proyectos con el fin de tener criterios para definir su futura estrategia digital.

Las iniciativas en el área del libro científico no se quedan ahí sino que surgen nuevos modelos de negocio para nuevas formas de lectura. Por ejemplo, *CourseSmart*, la plataforma de distribución y comercialización de las principales editoriales de libros de texto en Estados Unidos, ha implementado una versión para acceder a sus contenidos a través del *iPad*. También acaba de actualizar sus versiones para *iPhone* y

iPod touch. Desde el *iPad* ya es posible el acceso a más de 14.000 libros de texto digitales de cinco grandes editoriales especializadas con descuentos de hasta el 60% del precio de las ediciones impresas. Hoy ya son 15 las editoriales que colaboran con este proyecto. Su idea es tratar de mejorar la enseñanza y el aprendizaje, proporcionando a los profesores y estudiantes una mejor exposición y acceso a los materiales del curso mediante el formato digital online, y ahora también mediante el soporte móvil. Las tecnologías digitales permiten la creación de comunidades de aprendizaje presencial y virtual. *Apple* ha creado *Inkling* para la introducción de libros de texto en el *iPad*. <http://www.inkling.com>

El mero acto de entrega de contenido digital no es suficiente. Los editores de hoy necesitan ir más allá con el fin de alcanzar formas innovadoras para mejorar la adquisición de libros electrónicos y para cumplir con las expectativas de los usuarios en cuanto a comodidad, personalización y flexibilidad. Los modelos de alquiler y suscripción de obras académicas caminan en esa línea. Ejemplos son *eCampus.com* y *Cengage* que además del alquiler permiten la compra de capítulos y de paquetes de aprendizaje con materiales complementarios.

<http://www.ecampus.com>
<http://www.cengage.com>

“ La integración multimedia, la participación social y la colaboración son cada vez más importantes en el entorno del libro ”

Proliferan los programas que permiten compartir e intercambiar lecturas científicas y de todo tipo como *Copia*, la plataforma de lectura digital con mayor enfoque social, que firmó un acuerdo con *The Collegiate Retail Alliance (CRA)*, la mayor coalición de tiendas universitarias independientes de Estados Unidos. Esta coalición ya había llegado anteriormente a otros acuerdos con *CaféScribe* y *CourseSmart* para potenciar y satisfacer la demanda de libros electrónicos en las universidades. En marzo de 2011 *Copia* comenzó un pro-

grama piloto con 10 librerías de universidades para ofrecer contenidos digitales con varias funciones: compartir notas, grupos de discusión online... *Copia* tenía previsto implantarse en España en otoño de 2011, incorporando en sus catálogos los libros que comercializan *Librandia*, *Publidisa* y *Bubok*, con una oferta global de 60.000 títulos, que va de los best sellers y las novedades literarias a los libros técnicos.

<http://www.thecopia.com/home/index.html>

<http://www.collegiateretailalliance.com>

<http://www.librandia.com>

<http://www.publidisa.com>

<http://www.bubok.com>

4. Apocalípticos, integrados y asimilados

En junio de 2011, el periodista y filósofo alemán **Florian Rötzer** hizo este comentario:

“Para los jóvenes que han crecido en la era de la información, la privación siquiera por un día del acceso a internet se traduce en síntomas de abstinencia física y psíquica similares a los que tienen los fumadores cuando dejan de fumar. Esto se conoce como trastorno de privación de información (*information deprivation disorder*) y los que lo sufren se sienten excluidos y desconectados del mundo. La función de los espacios públicos y semipúblicos ha comenzado a cambiar. Ya no evolucionamos ni solos ni en grupo, sino en medio de una nube de contactos, amigos e información, a la que a menudo damos más importancia que a nuestro entorno físico inmediato”.

La sociedad va creando su ecosistema según va evolucionando. Cuando se recibe un impacto tan grande como el de internet se precisa un tiempo para que los individuos se adapten al mismo. La asimilación de los cambios depende de los niveles de educación, económicos, etc.

Se habla de los posibles efectos perniciosos de la Red, pero, si se compara, la televisión ha sido mucho más devastadora, pues ha impuesto una actitud pasiva y acrítica. Al menos la Red obliga a la actividad y la participación, y que ésta no sea obligatoriamente presencial representa una gran ventaja sobre todo para aquellos que viven en urbes desconectadas. La Red ha abierto canales de comunicación inimaginables hace unos años. Pensemos en las revueltas de los países árabes, o lo que ocurre en dictaduras como China o Cuba donde gracias a la Red la juventud ha adquirido un grado de concienciación y movilización imposibles de otra manera.

“La Red no tiene los efectos perniciosos de la TV, que ha sido devastadora al propiciar una actitud pasiva y acrítica”

Por lo tanto los vaticinios apocalípticos carecen de sentido, y en la historia se han demostrado siempre exagerados. Los beneficios compensan con creces los inconvenientes que puedan tener. Y el aislamiento y la introversión son síntomas de problemas preexistentes que el uso de la Red pone de manifiesto pero que no han sido generados por ella.

Los libros representan desde siempre el espíritu de la libertad y el progreso, independientemente de los soportes a los que

DER TAGESSPIEGEL Medien

STARTSEITE POLITIK BERLIN WIRTSCHAFT SPORT KULTUR

DIGITAL GAMES UND APPS TEST UND TECHNIK WERBE-BEILAGEN

15.04.2009 00:00 Uhr | Von Barbara Nolte | Kommentare: 0

Im Porträt
Der Eremit

Sie arbeiten als Blogger, Web-Kolumnist, Online-Chef: „Alpha-Journalisten 2.0“. Der Münchner Florian Rötzer ist einer von ihnen. Porträt eines Schnelldenkers

Empfehlen Twittern +1

Florian Rötzer, Chefredakteur von „Telepolis“, ist ein Fremdling unter Deutschlands Online-Journalisten. Ein

Kein Redaktionsschluss, keine Feiertage. Florian Rötzer, 55, ist seit 1996 Chefredakteur des Online-Magazins »Telepolis«. - Foto: pa/dpa

Artículo sobre Florian Rötzer en el periódico Der Tagesspiegel

<http://www.tagesspiegel.de/medien/im-portraet-der-eremit/1496438.html>

estén asociados. Evidentemente el progreso científico e industrial nos hace cada vez más deudores y dependientes del contexto tecnológico y energético. Pero ¿existe una realidad separada del mismo? Hoy en día sería impensable excepto para los integristas de pensamiento radical anclados en propuestas ludistas o misonieístas. Nuestra realidad es tecnológica, imaginativa y en estado de innovación permanente, no sólo en el ámbito del pensamiento o de la cultura en general sino en los aspectos más cotidianos de la vida. La vuelta a la realidad propiciada por una escasez de energía sería una vuelta al paleolítico, no un escenario bucólico o rousseiano de vuelta a la naturaleza, sino apocalíptico como el representado en películas como *Mad Max* y sus derivados.

La realidad es que la composición de átomos va decayendo en beneficio de los bits, por utilizar la conocida expresión de **Negroponte** en *Being digital*.

En septiembre de 2011 se publicó en *Idboox.com* esta crónica futurista, algo catastrófica:

- 2013: Las ventas de libros electrónicos superan todas las ventas de libros, incluso los libros usados. Las ventas de revistas digitales comienzan a igualarse a las de las revistas impresas.
- 2014: Las editoriales subvencionan aparatos lectores de libros digitales. Los editores de periódicos, revistas y libros tratan de crear lectores bloqueados para vender en ellos sólo sus productos. Fallan.
- 2015: Muerte de las librerías locales. Las más pequeñas se convierten en cibercafés. Sobreviven marginalmente las especializadas en libros raros o de coleccionista.
- 2016: Muchas de las revistas de kiosco se publican sólo en digital.
- 2018: La última tienda de *Barnes & Noble* se transforma en cibercafé y en punto de acceso wifi.
- 2019: Hecatombe entre los editores: sólo sobrevive un pequeño número de ellos. Los gigantes como *Random House* y *Penguin* transforman sus filiales más pequeñas en editores exclusivamente digitales.
- 2020: Casi todos los estudiantes de colegios y escuelas secundarias poseen un lector de libros electrónicos. Los libros de texto en papel desaparecen gradualmente.



<http://www.thecopia.com>

- 2023: Los dispositivos de lectura llegan a ser tan delgados como el papel gracias a la evolución del *Epaper*.
- 2025: Termina la transición de los libros impresos a los electrónicos, incluso en los países en vías de desarrollo. En el mejor de los casos el libro se considera una curiosidad y en el peor una molestia. Los libros de colección aún no han desaparecido. Un pequeño número de lectores sigue leyendo libros impresos, pero en general todos los libros están disponibles en versión digital.

Puede parecer una previsión demasiado apocalíptica pero no está exenta de posibilidades, tal y como marcha el mercado en la actualidad.

“ Nuestra realidad es tecnológica, imaginativa y en estado de innovación permanente en el ámbito del pensamiento, de la cultura en general y en los aspectos más cotidianos de la vida ”

5. Nota

1. **Michael Hart** (63) y **Steve Jobs** (56) fallecieron con un mes de diferencia, 7 de septiembre y 7 de octubre de 2011, respectivamente. Habían alumbrado algunas de las ideas más innovadoras del último siglo.

6. Bibliografía

Association of American University Presses (AAUP). *Digital book publishing in the AAUP community, Survey report: Spring 2011*, 11 pp.

<http://www.aaupnet.org/images/stories/data/2011digitalurveyreport.pdf>

Association of American University Presses (AAUP). *Sustaining scholarly publishing: new business models for university presses*, March 2011, 37 pp.

<http://www.aaupnet.org/images/stories/documents/aaupbusinessmodels2011.pdf>

Chartier, Roger; Cavallo, Guglielmo (eds). *Historia de la lectura en el mundo occidental*. Madrid: Taurus, 2011, 566 pp. ISBN 978 84 306 0838 6

The Cocktail Analysis. *Televidente 2.0*, 2011. <http://www.tcanalysis.com/2011/09/21/televidente-20-2011-tablets-television-conectada-y-redes-sociales-enriquecen-el-escenario-de-consumo-de-television/>

ComScore. *Digital omnivores: How tablets, smartphones and connected devices are changing U. S. digital media consumption habits*. ComScore, 2011, 33 pp. <http://www.ipmark.com/pdf/Omnivoros.pdf>

Cordón-García, José-Antonio. *La revolución del libro electrónico*, Colección El profesional de la información. Barcelona: UOC, 104 pp. ISBN 978 84 9788 485 3

Federación de Gremios de Editores de España; Fundación Germán Sánchez Ruipérez. *2ª Encuesta sobre el libro digital en España. Impacto de la digitalización en el catálogo, canales de distribución y de venta y política comercial*. Resultados de la encuesta 2010. Informe elaborado por Neturity. Marzo 2011, 45 pp. Disponible en: Distribuidor de Información del Libro Español en Venta (Dilve.es) <http://www.dilve.es/dilve/getArchivoDocumentacion.do?iddocumento=1541>

Furtado, José-Afonso. *El pixel y el papel: de lo impreso a lo digital: continuidades y transformaciones*. Gijón: Trea, 2007. 109 pp. ISBN 978 84 9704 300 7

Furtado, José-Afonso. *A edição de livros e a gestão estratégica*. Lisboa: Booktailor, 2011. 325 pp. ISBN 978 989 96008 1 2

Hábitos de lectura y compra de libros. Madrid: Federación de Gremios de Editores, 2011.

Idboox. *Ebook: le futur du livre papier, chronique d'une mort annoncée*, 28 septembre 2011. <http://www.idboox.com/ebook/infos-ebooks/ebook-le-futur-du-livre-papier-chronique-d-une-mort-annoncee>

Johnson, Larry; Smith, Rachel S.; Willis, Holly; Levine, Alan; Haywood, Keene. *The 2011 horizon report*. Austin, Texas: The New Media Consortium, 2011. 40 pp. ISBN 978 0 9828290 5 9

Manguel, Alberto. *Una historia de la lectura*. Barcelona: Lumen, 2005. 624 pp. ISBN 84 2641 525 3

Martin, Henri-Jean. *Historia y poderes de lo escrito*. Gijón: Trea, 1999. 526 pp. ISBN 84 95178 40 0

Negroponte, Nicholas. *Being digital*, New York: Alfred A. Knopf, 1995, 243 pp ISBN 0 679 43919 6

Ontsi (Observatorio Nacional de las Telecomunicaciones y la Sociedad de la Información). *La sociedad en red, 2010: informe anual*, edición 2011. <http://www.ontsi.red.es/informes-anuales/articulos/id/5421/informe-anual-2010-edicion-2011.html>

Panorámica de la edición española de libros 2010. Madrid, Ministerio de Cultura, 2011.

Rötzer, Florian. “La réalité augmentée ou la quête d'un nouveau monde”. *Courier international*, 16-06-2011. <http://www.courierinternational.com/article/2011/06/16/la-realite-augmentee-ou-la-quete-d-un-nouveau-monde>



GESTIÓN DE OBJETOS DIGITALES CON CONTENTDM: ESTUDIO DE CASO EN EL CONSORCI DE BIBLIOTEQUES UNIVERSITÀRIES DE CATALUNYA Y EL INSTITUT CARTOGRÀFIC DE CATALUNYA



Rafael Roset, Sandra Reoyo-Tudó y Robert Torre-Marín



Rafael Roset lleva 20 años trabajando en el *Institut Cartogràfic de Catalunya*, en distintos puestos de responsabilidad relacionados con las tecnologías de la información. Desde 2006 es el responsable de la Cartoteca digital.

Institut Cartogràfic de Catalunya
Parc de Montjuïc. 08038 Barcelona
<http://www.icc.cat>
rafael.rosset@icc.cat



Sandra Reoyo-Tudó, licenciada en documentación por la *Universitat de Barcelona*, ha trabajado en el *Catálogo Colectivo de las Universidades de Catalunya (CCUC)* y desde 2004 es responsable de la coordinación y gestión de los repositorios cooperativos *MDC, MDX, RACO, Recercat* y *TDR* del *Consorti de Biblioteques Universitàries de Catalunya (CBUC)*.

Consorti de Biblioteques Universitàries de Catalunya
Gran Capità, 2-4. Edif. Nexus, 4ª pl., desp. 403
08034 Barcelona
<http://www.cbuc.cat>
sreoyo@cbuc.cat



Robert Torre-Marín es licenciado en documentación y diplomado en biblioteconomía y documentación por la *Universitat de Barcelona*. Es responsable del área de producto e implementación en *Doc6*. Ha desarrollado y gestionado varios proyectos de implantación de *ContentDm*.

Doc6. Consultores en recursos de información
Mallorca 272, pl. 3. 08037 Barcelona
<http://www.doc6.es>
rtorremarin@doc6.es

Resumen

La gestión de objetos digitales, ya sean imágenes u otros tipos de ficheros digitales, supone un reto para las organizaciones. Es un proceso estratégico en el que están implicadas varias tareas dentro de la propia institución, desde una posible digitalización del documento físico a una correcta gestión documental pasando por la óptima definición de los requisitos funcionales que debe tener la herramienta que publique y gestione los objetos digitales. Se presenta el caso de dos instituciones, *Consorti de Biblioteques Universitàries de Catalunya* e *Institut Cartogràfic de Catalunya*, que afrontaron este proceso estratégico y escogieron *ContentDm* para la publicación y gestión de sus imágenes digitales.

Palabras clave

ContentDm, Digitalización, Repositorios digitales, Gestión de imágenes, Colecciones digitales, Archivos digitales, WorldCat, OAI-PMH, Metadatos.

Title: *ContentDm's solution for managing digital objects: case study of the Consortium of Academic Libraries of Catalonia and Cartography Institute of Catalonia*

Abstract

Management of digital objects -pictures or other kinds of digital files- is a major challenge for organizations. It is a strategic process that involves several tasks within the institution, from the digitization of a physical document to the right document management process, through proper definition of the functional requirements for the tool used to publish and manage digital objects. A case of two institutions, *Consortium of Academic Libraries of Catalonia* and the *Cartography Institute of Catalonia*, is presented. Both faced this strategic process and choose the *ContentDm* tool for publishing and managing their digital images.

Artículo recibido el 27-05-11
Aceptación definitiva: 25-08-11

Keywords

ContentDm, Digitization, Digital repositories, Images managing, Digital collections, Digital archives, *WorldCat*, OAI-PMH, Metadata

Roset, Rafael; Reoyo-Tudó, Sandra; Torre-Marín, Robert. "Gestión de objetos digitales con *ContentDm*: estudio de caso en el *Consorti de Biblioteques Universitàries de Catalunya* y el *Institut Cartogràfic de Catalunya*". *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 91-97.

<http://dx.doi.org/10.3145/epi.2012.ene.12>

1. Introducción

La gestión de imágenes y otros ficheros digitales es un reto estratégico para cualquier institución. La información digital está creciendo año a año exponencialmente y su preservación va a ser una de las líneas de trabajo más importantes de esta década en el ámbito de la gestión documental.

El problema se acentúa además si tenemos en cuenta que en estos momentos la capacidad de almacenamiento de la información es bastante inferior a la producción digital que genera la propia organización.

Cualquier planificación de gestión de la información digital que queramos desarrollar requiere de una visión estratégica de la organización que permita establecer de forma muy clara los objetivos que se persiguen teniendo en cuenta las herramientas de las que se dispone.

Uno de los pilares básicos sobre los que se sustenta un plan de gestión de imágenes u otros ficheros digitales es el programa de gestión y difusión que se va a utilizar. Habitualmente la decisión sobre el programa a utilizar sería el último paso de la cadena; sin embargo, las posibilidades y utilidades que incorporan programas como *ContentDm* hacen que éstos se puedan situar directamente en el epicentro de cualquier proyecto, partiendo de los requisitos técnicos de la digitalización, de las políticas de difusión o del uso de formatos estándar para preservación y difusión.

En este artículo se expone el caso de dos instituciones referentes en su sector: *Consorti de Biblioteques Universitàries de Catalunya* (en adelante *CBUC*) e *Institut Cartogràfic de Catalunya* (en adelante *ICC*). Ambas organizaciones han necesitado afrontar la misma situación: ¿Cómo gestionar y publicar los fondos de imágenes digitales? ¿Qué programa permite llevar a cabo una correcta gestión de la imagen digital y además publicarla en la Web? Se verá por qué ambas instituciones escogieron *ContentDm*, qué soluciones aportó y qué recomendaciones deben seguir las instituciones que afrontan una situación similar de gestión de sus bancos de imágenes.

2. Antecedentes

El caso *CBUC*

El *Consorti de Biblioteques Universitàries de Catalunya* (*CBUC*) inició su andadura en el ámbito de los repositorios en 1999 con el de *Tesis Doctorales en Red* (*TDR*), bastante antes que éstos recibieran el nombre de repositorios y que se acuñara el término acceso abierto (*open access*). Actualmente funcionan 5 repositorios cooperativos activos que

utilizan 3 programas distintos: *DSpace* para documentos de investigación y materiales docentes (*TDR*, *Recercat* y *MDX*), *OJS* para artículos de revista (*RACO*) y *ContentDm* para imágenes (*Memòria Digital de Catalunya* o *MDC*).

En este apartado se hablará sobre el repositorio *MDC*, coordinado por el *CBUC* e impulsado por las universidades de Catalunya y la *Biblioteca de Catalunya* (*BC*), y que cuenta con la participación de otras instituciones catalanas (entre ellas el *ICC*, que se comenta más adelante). Se puso en funcionamiento a finales de 2006 y actualmente contiene en acceso abierto 325 colecciones digitalizadas de revistas catalanas antiguas, fotografías, mapas, carteles, manuscritos, incunables y un largo etcétera con un total de más de 2 millones de documentos relacionados con Catalunya y procedentes de 18 instituciones diferentes.

<http://www.cbuc.cat/mdc>

Debido a sus posibilidades y utilidades, *ContentDm* puede situarse en el epicentro de cualquier proyecto de gestión y difusión de imágenes en la web

Entre otras se pueden consultar colecciones tan curiosas como las primeras imágenes del gorila albino "Copito de nieve"; o *ARCA* (*Archivo de revistas catalanas antiguas*) que contiene títulos que ya no se publican y que han sido representativas de la cultura y la sociedad catalana, como por ejemplo *Destino* (semanario fundado en Burgos en 1937 por un grupo de catalanes liderados por el editor Josep Vergés).

Este repositorio está abierto a la colaboración de bibliotecas y archivos de Catalunya que deseen difundir sus colecciones digitalizadas especiales con documentos de carácter predominantemente gráfico, relacionados con Catalunya y/o su patrimonio. Se prioriza la incorporación de los documentos más singulares y excepcionales, y los susceptibles de ser consultados por un número más elevado de usuarios. La digitalización de los documentos la llevan a cabo las propias instituciones y la deben realizar teniendo en cuenta los parámetros y criterios indicados en los "Estándares de digitalización: requerimientos mínimos" que el *CBUC* actualiza periódicamente. La incorporación de los documentos se realiza previa consideración de los derechos de propiedad intelectual que puedan afectar a su explotación.

MDC funciona con el software *ContentDm*. Se adquirió en 2005 ante la necesidad de aumentar la visibilidad y consulta del patrimonio documental catalán de las instituciones

Portada de *Memòria Digital de Catalunya*. <http://www.cbuc.cat/mdc>

miembros o colaboradoras del CBUC y por tanto, para dar soporte a la misión de mejorar los servicios bibliotecarios a través de la cooperación que le es propia al CBUC.

MDC contiene en acceso abierto 325 colecciones digitalizadas de 18 instituciones. En total más de 2 millones de documentos

La elección del programa se efectuó tras comprobar que había sido adoptado por la mayoría de grandes bibliotecas americanas con imágenes, para la preservación y difusión de las mismas a través de la Red, y porque demostró ser una buena herramienta para satisfacer las necesidades de las instituciones del CBUC. Actualmente MDC dispone de dos servidores y de dos licencias con objetos ilimitados y se trabaja con la versión 5.4, aunque está previsto migrar a la 6.1.

El caso ICC

La Cartoteca del *Institut Cartogràfic de Catalunya (ICC)* tiene como objetivo velar por la conservación de la documentación cartográfica y geográfica, y ponerla a disposición del público, tal como se recoge en el capítulo IV de la *Ley 16/2005 de la información geográfica y del Institut Cartogràfic*.

En los 20 años transcurridos desde su fundación, la Cartoteca ha reunido un fondo de 300.000 mapas de todo el mundo, 350.000 fotografías aéreas verticales de Catalunya y 50.000 fotografías paisajísticas, así como un archivo de fondos documentales territoria-

les que se estima en 250.000 páginas. Todo ello, juntamente con los 60.000 volúmenes de la biblioteca especializada, constituye un patrimonio cartográfico de gran valor.

El proyecto de digitalización de la Cartoteca de Catalunya se puso en marcha para garantizar la máxima difusión de sus fondos facilitando el acceso remoto a los mismos. Hay que tener en cuenta que una parte de los mismos no son todavía accesibles por internet por tratarse de documentos con derechos de reproducción vigentes que no posee el ICC. El paso de los años permitirá su incorporación progresiva a la Cartoteca digital a medida que la documentación quede libre de derechos.

El ICC inició el proyecto de digitalización de los fondos de la Cartoteca en

2002, en dos fases muy diferenciadas:

1ª fase: estudio y primeras pruebas (2002-2005)

En 2002 se dieron los primeros pasos con la redacción de las primeras propuestas, la definición de los parámetros de digitalización, el establecimiento de los nombres de los archivos digitales, el diseño del flujo de digitalización y la creación del primer esquema de metadatos. Posteriormente se empezó la digitalización de los primeros mapas antiguos aprovechando la estructura productiva del ICC que disponía ya de dos escáneres HP de rollo que aceptaban originales de hasta 90 cm de ancho.

En 2003 se publicaron en internet las primeras imágenes de mapas antiguos: mil mapas a baja resolución y tamaño reducido, y otros 50 mapas en formato *MrSID (multiresolution seamless image database, de LizardTech)* de alta calidad que permitía la visualización con zoom y pan (panorámica).

Colección del naturalista Dr. Jordi Sabater Pi. <http://mdc.cbuc.cat/cdm4/browse.php?CISOROOT=/sabaterpi>



Sala de la Cartoteca, Institut Cartogràfic de Catalunya

Se empezaron a servir a los usuarios de la Cartoteca las primeras impresiones en formato 1:1 y las primeras copias digitales de los materiales.

2ª fase: consolidación (2006 en adelante)

En 2006 se inició una nueva fase del proyecto al dotarse de la Cartoteca con los medios necesarios para llevar a cabo internamente la digitalización. Se adquirió un escáner *Metis* de gran formato que permite digitalizar mapas desde 60x90 cm a 600 ppp hasta 120x180 cm a 300 ppp, y se añadieron un escáner *Creo* fotográfico de tamaño A3 y un escáner plano para documentos hasta A4. También se incorporó personal de perfil técnico para la realización del proyecto.

En estos 20 años desde su fundación, la Cartoteca ha reunido un fondo de más de 300.000 mapas

Paralelamente se inició la evaluación de las distintas opciones de software para la gestión de la colección, desde desarrollos a medida hechos con medios propios dentro del ICC, hasta las diversas de plataformas de código abierto. Y la búsqueda de soluciones se amplió al servidor de contenidos para difundir las colecciones online.

En el mismo 2006 el *CBUC* invitó al *ICC* a participar en la *Memòria Digital de Catalunya*. El *ICC* contribuyó con dos colecciones de documentos: una con 43 mapas de Catalunya, y otra de 106 fotografías aéreas oblicuas. La experiencia resultó muy positiva, y sirvió para finalmente escoger *ContentDm* como el gestor digital de colecciones a utilizar en la Cartoteca Digital, tanto por la experiencia adquirida durante la prueba piloto como por la versatilidad, el soporte y la gran comunidad de usuarios que ofrecía el producto.

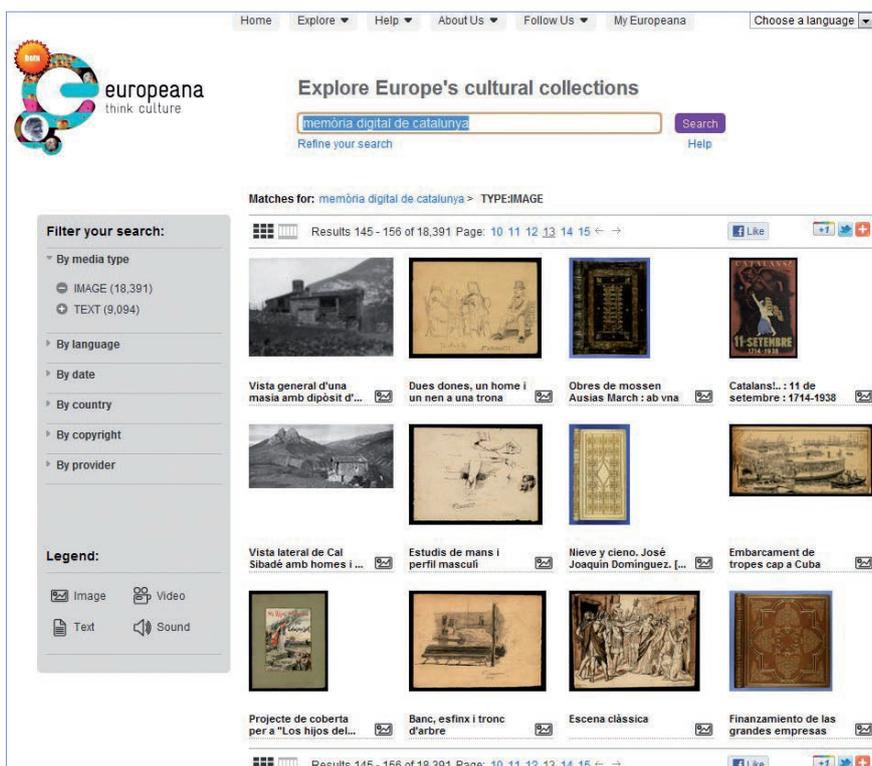
Finalmente la Cartoteca Digital se inauguró el 31 de octubre de 2007, y a día

de hoy recibe una media de 250 visitas diarias de 7 minutos de duración cada una.

3. La experiencia del *CBUC*

ContentDm ha permitido al *CBUC*, con relativamente poco esfuerzo y rapidez por parte de las instituciones participantes, hacer visibles unas colecciones que en la mayoría de los casos por su rareza y/o condición física son de acceso difícil y/o restringido, impulsar la digitalización del patrimonio catalán, permitir su consulta conjunta y facilitar los instrumentos para su preservación. Además, al usar metadatos de descripción *Dublin Core* y el protocolo de interoperabilidad de la *Open Archives Initiative (OAI-PMH)*, también ha hecho posible incrementar aún más la visibilidad de los documentos, al ofrecerse conjuntamente con otros repositorios internacionales como es el caso de *Europeana* (portal europeo de acceso a colecciones culturales que incluye actualmente más de 15 millones de objetos digitales) y *OAIster* (catálogo colectivo internacional integrado en *WorldCat* que incluye más de 23 millones de registros que representan los recursos digitales de más de 1.100 colaboradores).

Lo más positivo de *ContentDm*, teniendo en cuenta el carácter cooperativo del repositorio y las personas e instituciones de diversa índole que participan, es la posibilidad de publicar contenido digital en muy poco tiempo gracias a la sencillez de uso del programa. Además también se destaca la funcionalidad que ofrece para poder personalizar cada colección y poder dar de este modo más visibilidad a las instituciones participantes. Las instituciones personalizan la imagen de la cabecera superior (en la que suelen hacer constar, por ejemplo, el logotipo de la institución y un enlace a su página web, una imagen representativa de la colección y el logotipo y enlace de *MDC*), el menú (en el que se incluye un texto sobre el contenido de la colección) y el pie (en el que se



Imágenes de la *Memòria Digital de Catalunya* presentes en *Europeana*

incluye la dirección de correo electrónico de contacto de la institución). Por otra parte, también es importante la posibilidad de usar plantillas de metadatos para las colecciones que permiten tanto informar los propios de la colección, como los generales del repositorio para identificarnos como grupo, y los obligatorios para participar en *Europeana*.

Lo más positivo de *ContentDm* es la posibilidad de publicar contenido digital en muy poco tiempo gracias a la sencillez de uso del programa

Cabe destacar además que uno de los motivos que en su momento fueron decisivos para adquirir *ContentDm*, fue la posibilidad que ofrece de incorporar conjuntamente con las imágenes las transcripciones del texto obtenido a través del uso de OCR y por lo tanto de facilitar la búsqueda en el texto completo. Esta funcionalidad ha aportado un valor añadido muy destacable a la digitalización de los documentos, sobre todo a aquellos el contenido de los cuales es principalmente textual como son las revistas antiguas de *ARCA*.

Lo más negativo de *ContentDm* en el caso del *CBUC*, dada la gran cantidad de imágenes que incluye, es el límite que tiene actualmente de 300 colecciones (que se ampliará a 350 en la versión 6.1) y que comporta, si se supera este límite, comprar una nueva licencia. Para solucionar esta carencia y debido a que, como se ha comentado anteriormente, *MDC* funciona distribuidamente con dos licencias y dos servidores, se ha hecho necesario el desarrollo de una interfaz intermedia realizada a nivel local que enlaza ambos servidores. En una instalación se hallan las 274 colecciones correspondientes a las revistas del portal *ARCA* y en la otra las 51 colecciones restantes. El aspecto positivo de esta separación es que debido al volumen enorme de datos (sobre todo texto) que contiene *ARCA* cuando un usuario realizaba una búsqueda normalmente recuperaba resultados sólo de este portal y en raras ocasiones de otras colecciones. Con esta interfaz intermedia se realiza la búsqueda y se puede escoger entre visualizar los resultados en colecciones de imágenes, de texto o en revistas de *ARCA*.

Esta limitación de *ContentDm* se minimiza con el hecho de que permite añadir fácilmente funciones locales para mejorarlo y adaptarlo a las necesidades de *MDC*. Como soluciones locales, a parte de la ya comentada de la interfaz intermedia que permite realizar búsquedas en las dos licencias actuales de *ContentDm* y describir tanto las instituciones participantes y sus colecciones, las próximas incorporaciones

de información técnica, también se ha incluido una nube de etiquetas con las palabras más usadas en los títulos de los documentos. Además, para cada documento se ha añadido la opción de pasarlo automáticamente a pdf (sea cual sea su formato de origen) para poder descargarlo y/o imprimirlo con más facilidad.

4. La experiencia del ICC

Al ser *ContentDm* un sistema programado básicamente en php, la facilidad para modificar o añadir funciones es grande, al igual que las posibilidades de modificación del aspecto y del idioma. Ambas características resultaron determinantes en la elección del producto, pues el sitio web de la Cartoteca Digital debía estar en tres idiomas: catalán, español e inglés, y tener un aspecto lo más similar posible al sitio web del *ICC*, para conseguir la máxima integración con éste y respetar la imagen corporativa.

MDC ha incrementado la visibilidad de sus documentos al ofrecerse conjuntamente con otros repositorios internacionales como *Europeana*

Una de las características propias de la Cartoteca Digital que la distingue del resto de instalaciones de *ContentDm* es la opción de descargar las imágenes digitales de los documentos en formato jpg a 300 ppp juntamente con un archivo de

The screenshot shows the 'Cartoteca Digital' website interface. At the top, there is a navigation bar with links for 'Inicio', 'Navegar', 'Búsqueda avanzada', 'Preferencias', 'Mis favoritos', and 'La Cartoteca Digital'. Below this, the main content area is titled 'Colecciones de mapas' and 'Colecciones de imágenes'. Each collection is represented by a small thumbnail image and a text box containing the collection name, a date range, and the number of documents. For example, under 'Colecciones de mapas', there are items like 'Mapas de Catalunya (s. XVII-XX)' with 1011 documents, 'Mapas de España (s. XV-XX)' with 1603 documents, and 'Mapas de Europa (s. XVI-XX)' with 1677 documents. Under 'Colecciones de imágenes', there are items like 'Fotoplanos de ciudades catalanas (1945-1966)' with 80 documents, 'Fotomosaico provincia Barcelona 1:5 000 (1967)' with 919 documents, and 'Fotografías aéreas de Catalunya (Gasper, 1929)' with 268 documents. The footer of the page includes the logo of the Institut Cartogràfic de Catalunya and the text 'Impulsado por CONTENTdm | comuníquese con nosotros'.

Portada de la Cartoteca Digital de Catalunya. <http://cartotecadigital.icc.cat>

acompañamiento con los metadatos básicos de la obra. Para ello se modificó extensivamente el entorno de visualización de *ContentDm* a fin de reutilizar el código y el software que ya estaba implementado en la web del ICC, donde también se ofrece la descarga de productos previo registro gratuito de los usuarios.

A pesar que en versiones anteriores el producto no ofrecía una distinción clara entre las capas de presentación y funcionalidad, era relativamente sencillo modificar el código para añadir funciones no presentes en la instalación base, como el seguimiento a través de *Google analytics*. Pero como contrapartida en cada nueva actualización había que repetir el proceso, tedioso y proclive al error, de aplicar de nuevo todas las modificaciones de aspecto y funcionalidad y fundirlas correctamente con las mejoras y correcciones de la nueva versión.

Otra de las ventajas del producto es su gran base de instalaciones, sobre todo en universidades y bibliotecas, que ha propiciado la creación de una comunidad de usuarios muy activa y participativa que sirve tanto de equipo de soporte alternativo como de desarrollo de pequeñas modificaciones compartidas. Algunas de las modificaciones, como la traducción al catalán, fueron creadas y compartidas por varios usuarios y mantenidas por la comunidad al margen de OCLC, con la ayuda de *Doc6*.

Otra gran ventaja inherente es la facilidad de instalación, la estabilidad y el poco consumo de recursos en el servidor, y la disponibilidad de versiones tanto para entornos *Windows* como *Unix*. En el caso del ICC se optó por instalar *ContentDm* en servidores *RedHat*, que se han mantenido en funcionamiento sin errores ni caídas de servicio durante los 4 años desde que se inauguró la Cartoteca Digital.

La plataforma se escogió para las tareas de difusión: en *ContentDm* se cargan las imágenes en jpg a 300 ppp y los metadatos en *Dublin Core*, que son un subconjunto de los datos almacenados en la herramienta propia de gestión de catálogo desarrollada en el ICC. El flujo de digitalización diseñado en los primeros años del proyecto sigue siendo válido: a las tareas de adquisición, catalogación y digitalización que se llevan a cabo diariamente se añadió la tarea de difusión, que se realiza puntualmente con la publicación de nuevas colecciones o el añadido de nuevos documentos a las colecciones existentes, mediante la exportación de contenidos en el formato apropiado.

“*ContentDm* es un buen programa para exponer y usar los materiales incluidos en el repositorio y técnicamente su funcionamiento es amigable e intuitivo”

Así pues la facilidad de instalación y mantenimiento, junto con las posibilidades de personalización fueron aspectos determinantes en la decisión del ICC de adquirir e instalar el programa en sus propios servidores en lugar de aprovechar el espacio ofrecido por el CBUC. Pero hubo también otra razón de peso: al ofrecer descargas de archivos digitales de

alta calidad los requerimientos de espacio en disco y de ancho de banda eran muy elevados. El volumen actual de las 24 colecciones disponibles supera los 31.000 documentos, entre mapas y fotografías, y ocupa 278 GB de espacio en disco para las copias de descarga.

5. Conclusiones

Una vez detallados ambos casos, podemos destacar las siguientes conclusiones:

En el CBUC se considera que *ContentDm* es un buen programa para exponer y usar los materiales incluidos en el repositorio y técnicamente su funcionamiento es amigable e intuitivo. Lo más difícil son las tareas que se deben realizar y planificar a priori: los proyectos de digitalización que transformen a digitales los documentos en soportes papel u otros; obtener la financiación; priorizar los materiales a digitalizar bien por su interés para la organización, por diferenciar nuestros centros del resto, por una condición física que impida su consulta, o por el valor añadido que pueda darle la digitalización al documento (por ejemplo, transcripciones que permitan búsqueda a texto completo), la utilidad directa para los miembros y usuarios reales o potenciales; comprobar que no han sido ya digitalizados por otras instituciones; asignarles los metadatos de descripción, de digitalización, de preservación; basarse en directrices que normalizan los procesos de digitalización; comprobar los derechos de autor, etc.

En el caso del ICC, se considera que la importancia de la digitalización y de internet en el mundo de la documentación es innegable, pero en el ámbito específico de la cartografía este hecho es aún más evidente. Los hábitos actuales entre los usuarios empujan a las cartotecas públicas hacia nuevas plataformas para servir la información. Un primer paso lo constituye sin duda la digitalización de los documentos cartográficos.

Puesto que la digitalización se acomete para facilitar la consulta de los fondos, disponer de un programa sencillo de recuperación de imágenes por múltiples campos es indispensable, y ofrecer la descarga gratuita de las imágenes de consulta ahorra muchas peticiones a los archiveros y bibliotecarios que se pueden liberar así del trabajo de atención al público.

En las cartotecas, las herramientas de búsqueda y obtención de la información han cambiado de manera notable, y están cambiando mucho más: se debe contemplar la georeferenciación como uno de los elementos clave de los catálogos de cartotecas digitales. Asimismo, se debe ir más allá y posibilitar la introducción en las cartotecas de las IDE (infraestructuras de datos espaciales) de manera que podamos interoperar con todo tipo de catálogos con contenido cartográfico. Por otro lado, la recolección de datos y metadatos siguiendo el estándar *OAI* permite estar presentes en distintos portales y agregadores, multiplicando así la presencia en internet. Y a pesar de que *Worldcat* es un catálogo de obras mayormente en inglés, se puede utilizar la *WorldCat Digital Collection Gateway* para sincronizar los contenidos entre *ContentDm* y *WorldCat* ya que es muy interesante poder sincronizar automáticamente las colecciones digitales

con ese gran catálogo para alcanzar mayor exposición de los contenidos propios.

La apertura de la Cartoteca Digital y la posibilidad de descargar imágenes ha tenido un impacto enorme en el número y la tipología de sus usuarios. La visibilidad se ha incrementado espectacularmente, se ha ampliado la procedencia geográfica de los mismos y sólo ha descendido ligeramente su presencia en las instalaciones de la Cartoteca en la sede del *Institut Cartogràfic* en Barcelona.

La incorporación de funciones sociales en las nuevas versiones de la plataforma añadirá la posibilidad de que los propios usuarios comenten, etiqueten, compartan y califiquen los documentos, lo cual permitirá crear una comunidad alrededor de los mapas antiguos. Como contrapartida, añadirá tareas de moderación y monitorización a realizar por parte del personal de la Cartoteca. Por ello, y también por la actual situación económica y sobre todo la situación legislativa alrededor de los contenidos proporcionados por terceros, habrá que estudiar detenidamente el alcance y las repercusiones de estas herramientas antes de implantarlas en la Cartoteca Digital.

Se debe contemplar la georeferenciación como uno de los elementos clave de los catálogos de las cartotecas digitales

6. Notas

1. "The world's technological capacity to store, communicate, and compute information". *Science express*, 10 feb. 2011. <http://www.sciencemag.org/content/332/6025/60.full>

2. *ContentDm* es un programa desarrollado por el consorcio OCLC que permite la gestión y publicación web de imágenes u otros ficheros digitales. Esta aplicación trabaja con formatos y protocolos estándar, como por ejemplo: *Dublin Core*, OAI-PMH, xml, EAD, Z39.50 en entorno cliente y servidor. <http://www.contentdm.org>
<http://www.oclc.org>

7. Bibliografía

Anglada-Ferrer, Lluís M. "Biblioteca Digital de Cataluña: Continente y contenido". *5ª Jornada sobre la Biblioteca Digital Universitaria*, Los Polvorines (Argentina), 8-9 de nov. de 2007.

<http://hdl.handle.net/2072/4667>

Consorci de Biblioteques Universitàries de Catalunya. *Estándares de digitalización: requerimientos mínimos* (actualizado en octubre de 2010), octubre 2010.

<http://hdl.handle.net/2072/97497>

Ramos, Noelia; Roset, Rafael. "Adaptación de Dublin Core a la Cartoteca digital del Institut Cartogràfic de Catalunya". *Revista catalana de geografia*, dic. 2010, v. 15, n. 41.

<http://www.rcg.cat/articles.php?id=200>

Reoyo-Tudó, Sandra; Sanmartí, Josep. "The Digital Memory of Catalonia, Consortium of Academic Libraries of Catalonia". *D-lib magazine: the magazine of the digital library research*, 2011, Nov./Dec., v. 17, n. 11/12.

<http://www.dlib.org/dlib/november11/11contents.html>

Roset, Rafael; Ramos, Noelia. "Georeferenciación de mapas antiguos con herramientas de código abierto. Un caso práctico". *Mapping interactivo*, marzo 2011, n. 146, pp. 16-30.

IraLIS es el registro de autoridades de nombres de autor del repositorio E-LIS

Voluntades de subir documentos a E-LIS debes iralizarte siguiendo los "Criterios de firma".

<http://eprints.rclis.org>

IraLIS International Registry for Authors: Links to Identify Scientists

Idiomas: Spanish

Menú

Presentación

IraLIS (International Registry of Authors-Links to Identify Scientists) es un sistema de estandarización de las firmas de los autores científicos. Su objetivo es crear un registro de autoridades.

Trabaja en 5 líneas:

1. Concienciar a los autores científicos sobre la importancia de la firma para ser citados correctamente, y para poder recuperar toda la bibliografía a lo largo de su carrera profesional.
2. Redactar criterios de firma normalizada para
 - a) ser indexado correctamente y
 - b) distinguirse de otros autores con nombres

<http://www.iralis.org>

REGISTRO, ALMACENAMIENTO Y ACCESO SIMULTÁNEO DE VÍDEO STREAMING SOBRE PROTOCOLO UDP. EL SISTEMA MRECIPTV DE LA UNIVERSITAT POMPEU FABRA

Frederic Guerrero-Solé e Hibai López-González



Frederic Guerrero-Solé es licenciado en física, postgraduado en *e-commerce* por el idEC y doctor en comunicación pública por la *Universitat Pompeu Fabra (UPF)*, en cuyo departamento de comunicación ejerce actualmente como profesor ayudante. Es miembro del grupo de investigación *Unica* de la misma universidad, en el que ha participado en los proyectos *Televisión y deliberación política* y otros de análisis de campañas electorales. Delegado del director del departamento de las TIC para la investigación. Anteriormente ha sido jefe del departamento de TIC de *EADS Telecom*, y jefe de proyectos en *AGM* e *Ingram Micro*.

Universitat Pompeu Fabra
Roc Boronat, 138. Desp. 53.808
08018 Barcelona
frederic.guerrero@upf.edu



Hibai López-González es licenciado en comunicación audiovisual por la *Universidad del País Vasco*. Trabaja como investigador en la *Universitat Pompeu Fabra (UPF)* de Barcelona en donde imparte docencia en el grado de publicidad y relaciones públicas en el ámbito de los análisis del discurso. Forma parte del grupo de investigación *Unica* de dicha universidad y ha trabajado como analista para el proyecto de investigación *Televisión y deliberación política*.

Universitat Pompeu Fabra
Roc Boronat, 138. Desp. 53.810
08018 Barcelona
hibai.lopez@upf.edu

Resumen

Los investigadores en comunicación se encuentran a menudo con la necesidad de registrar y almacenar gran cantidad de contenido audiovisual. Hasta ahora, en el *Departamento de Comunicación* de la *Universitat Pompeu Fabra* de Barcelona este proceso se realizaba de manera analógica y manual. Tras la instalación de una plataforma de distribución de la señal digital de televisión sobre IP (IPTV) a la red interna, el *Departamento* ha diseñado el sistema *MRecIPTV* que permite su registro, almacenamiento y acceso simultáneo. Esta solución ha transformado la manera de trabajar de los investigadores, simplificando la recogida y el análisis de las muestras. Además, *MRecIPTV* ha permitido la creación de un patrimonio documental para futuras investigaciones.

Palabras clave

Archivo audiovisual, Tecnología, Documentación, Televisión digital, IPTV, Protocolo UDP, Proyectos de investigación.

Title: Recording, storage, and simultaneous access to video streaming over UDP protocol. The MRecIPTV system in the *Universitat Pompeu Fabra*

Abstract

Media researchers frequently need to record and store a huge amount of media content. Until recently, the *Department of Communication* at the *Pompeu Fabra University* in Barcelona handled this problem in an old-fashioned analog and manual way. After migration to iptv, the department designed the *MRecIPTV* software to record, store and access digital media content. This solution has transformed the way researchers work, simplifying data collection and analysis. Moreover, *MRecIPTV* has allowed the creation of a documentary heritage for future research.

Keywords

Media files, Technology, Documentation, Digital television, IPTV, UDP protocol, Research projects.

Guerrero-Solé, Frederic; López-González, Hibai. "Registro, almacenamiento y acceso simultáneo de video streaming sobre protocolo UDP. El sistema *MRecIPTV* de la *Universitat Pompeu Fabra*". *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 98-103.

<http://dx.doi.org/10.3145/epi.2012.ene.13>

Artículo recibido el 22-11-11
Aceptación definitiva: 13-12-11

Introducción

Una de las mayores dificultades que afrontan actualmente los proyectos de comunicación audiovisual que analizan las programaciones televisivas son la grabación, la documentación y el acceso a las muestras audiovisuales para la investigación. Los métodos tradicionales de grabación en formato analógico han quedado obsoletos y, además, dificultan el acceso a las muestras. La televisión sobre IP (*internet protocol TV* o IPTV), que está experimentando un amplio desarrollo desde inicios de siglo (Benoit, 2008; López-Yepes; Sánchez-Jiménez, 2007), permite no sólo el acceso de los usuarios a los contenidos mediante un navegador, sino que posibilita la grabación y catalogación de las muestras televisivas. Aunque en el mercado existe un gran número de softwares para la grabación de contenidos televisivos, en muchos casos tienen precios elevados y no se adecuan a los requisitos de las investigaciones.

El registro analógico de contenido audiovisual es actualmente inviable y se hace necesario migrar a sistemas digitales de mayor versatilidad

Dificultades de registro, difusión y acceso a las muestras audiovisuales

El Departamento de comunicación de la Universitat Pompeu Fabra (UPF) tiene un amplio historial en análisis televisivos. De forma periódica realiza estudios sobre contenidos de la programación y sobre campañas electorales que requieren la grabación multicanal (varios canales a la vez) durante períodos de hasta 4 semanas. Antes de la digitalización de la señal audiovisual y de la implementación de la tecnología IPTV, la recolección de las muestras audiovisuales para los proyectos suponía un desafío para los equipos investigadores. Según el procedimiento estándar, los miembros del equipo se repartían los canales y realizaban las copias de forma manual en equipos de grabación analógicos (primero cinta VHS, posteriormente DVD). Este procedimiento planteaba diversos problemas:

- los investigadores debían programar sus equipos o bien realizar las copias manualmente;
- cada investigador sólo podía realizar una copia, por lo que en los casos en los que era necesario grabar ocho canales, otros tantos investigadores debían realizar el proceso;
- el gran número de errores que se producía en las grabaciones dificultaba poder disponer de las muestras íntegras;
- la duplicación de los contenidos era compleja. En algunos casos se digitalizaban, un proceso que según Spence (2002) es imprescindible ante la inminente obsolescencia de la tecnología analógica, pero que resultaba extremadamente largo en el caso de muestras voluminosas;
- finalmente, estaba el problema del espacio físico que ocupaban las cajas de cintas VHS y los DVD, y su correcta conservación.

Aunque en los últimos períodos de recepción de la señal analógica se dispuso de un software de grabación (*AVerMe-*

dia) que almacenaba los contenidos en formato digital, el hecho de que fuera monocanal y funcionara en un número reducido de equipos hizo que continuaran existiendo las dificultades mencionadas.

MRecIPTV, la solución a medida

La necesidad de una sistematización de los procesos de registro de las muestras surgió a raíz de la aceptación del proyecto de investigación financiado por el Ministerio de Ciencia e Innovación: *Televisión y deliberación política*. El proyecto analizaba los programas de *infotainment* y los noticiarios de ocho canales de televisión españoles durante dos semanas del año 2010. En total la muestra era de aproximadamente 1.000 horas de televisión. Se descartó el proceso de grabación que se había utilizado hasta entonces, por lo que en consecuencia era necesario implementar un nuevo método de obtención de las muestras y la elaboración o adquisición de un sistema informático que realizara las grabaciones.

A partir de 2010 la recogida de muestras en la UPF dejó de ser analógica y pasó a sistematizarse utilizando la nueva red de transmisión de la señal digital de televisión sobre IP. La nueva infraestructura *IP multicast* –envío simultáneo de información a un grupo de receptores utilizando la estrategia más eficiente para entregar los mensajes sólo una vez, creando copias cuando las conexiones a los destinos se dividen (Bernier et al., 2007)–, permite el acceso a la televisión en directo desde cualquier ordenador conectado a la red del campus. Según Simpson (2008, p. 254), las ventajas del *multicast* frente al *unicast* son:

- reduce el ancho de banda requerido;
- las fuentes del vídeo son sencillas;
- permite vídeo de altísima calidad;
- escalabilidad.

La arquitectura del servicio está compuesta por una cabecera IPTV emisora de los canales, y por la red que transmite la señal de televisión a los ordenadores mediante protocolo *UDP (user datagram protocol)* (Postel, 1980), un protocolo extendido en los servicios multimedia en tiempo real en internet (Zheng; Boyce, 2001) y con resultados satisfactorios en el caso de *media streaming* (Treutterer et al., 2010). Para visualizar el contenido cada ordenador debe disponer del programa *VLC Media player* y de un navegador con el que escoger entre los diferentes canales disponibles mediante enlaces m3u.

El investigador es autónomo en la programación de las copias de *media stream*, y puede realizar análisis con mayor flexibilidad gracias a la posibilidad de copias personales y acceso remoto vía VPN

A pesar de la evidente mejora en la recepción de los contenidos, la cuestión de la grabación, la documentación y el acceso a las muestras plantea muchos problemas. Se descartó la compra de software por su inviabilidad económica (la so-

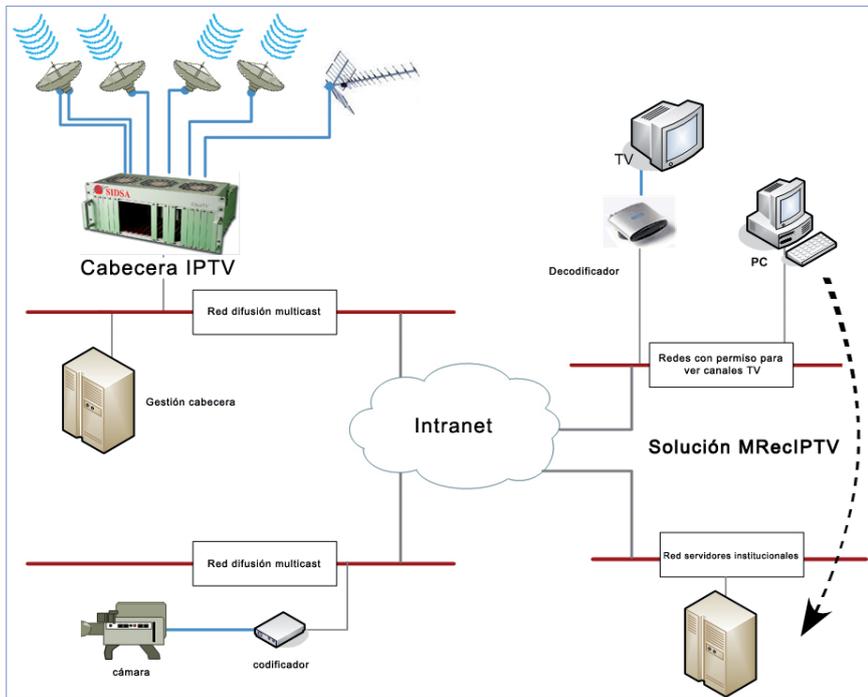


Figura 1. Arquitectura del servicio IPTV

lución *Scriptor multirecorder*, por ejemplo, ascendía a unos 50.000€, y por no responder a las necesidades concretas de las investigaciones. Se procedió, pues, a la elaboración de un sistema a medida y de bajo coste que satisficiera estas necesidades.

Experimentando el sistema MRecIPTV

La solución, a la que se ha llamado *MRecIPTV*, está basada en ejecutables que lanzan la aplicación *VLC Media player* en los ordenadores de los usuarios y crean archivos en formato Mpeg-2. Estos ejecutables están programados con un código sencillo que cualquier usuario puede modificar². Hay un ejecutable para cada uno de los canales de la muestra (*MRecIPTV_TV3* y *MRecIPTV_TVE*, por ejemplo, para los casos de los canales TV3 y TVE), y otro ejecutable que detiene la grabación (*MRecKillIPTV*).

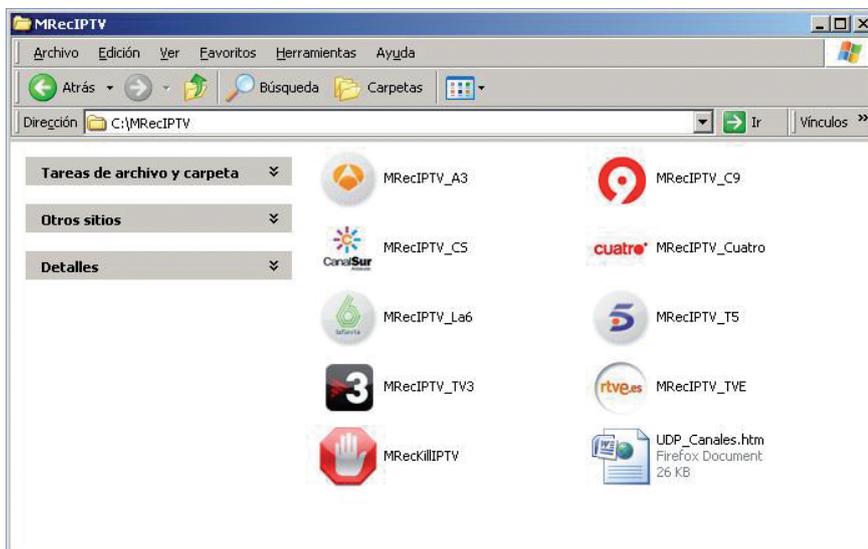


Figura 2. Archivos de ejecución de MRecIPTV

Este código permite designar:

- los archivos (con la estructura estándar *nombre-canal_fecha-registro_hora-inicio.mpg*);
- ubicación;
- dirección *UDP* que designa el canal que se desea registrar;
- codecs y calidad de vídeo;
- codecs y calidad de audio.

En el caso de que los investigadores quieran programar registros de canales que no están en el sistema *MRecIPTV*, tienen acceso a las direcciones *UDP* de todos los canales disponibles a través de una página de la intranet institucional (figura 2).

El modo de proceder de los investigadores debe ser el siguiente:

- identificar los canales que formarán parte de la muestra;
- comprobar que dichos canales son accesibles desde la Red;
- identificar los intervalos de programación a grabar;
- determinar la ubicación de los archivos resultantes;
- determinar los parámetros de grabación (codecs y calidad);
- programar las grabaciones.

“ Las muestras audiovisuales pasan a formar parte del archivo histórico del Departamento y podrán servir en futuras investigaciones comparativas ”

Transformaciones de las rutinas de investigación

Si la digitalización afecta a la cadena de valor de las industrias audiovisuales (**Agirreazaldegui-Berriozabal, 2007; Fernández-Quijada, 2008**), de la misma manera la recepción de los contenidos audiovisuales en formato digital afecta a todo el proceso de investigación audiovisual. La digitalización de la señal de televisión y el almacenamiento de los contenidos de *media stream* han transformado las rutinas de los investigadores en el análisis de los contenidos audiovisuales. Se ha pasado de las copias analógicas, de complicado acceso, catalogación y tareas de duplicado, a archivos digitales a los que pueden acceder todos los investigadores. Entre las ventajas que ofrece la solución *MRecIPTV* están:

- *Grabación multicanal simultánea*: Cada ordenador puede grabar simultáneamente un número ilimitado de

canales, aunque para conservar la calidad del vídeo y evitar archivos corruptos es preferible que no supere los 4 canales simultáneos. Es aconsejable que los archivos tampoco superen los 4 gigabytes, de lo contrario, VLC Media player puede interpretarlos como archivos corruptos y no los visualiza. La grabación también puede realizarse desde un servidor, por lo que el número de canales que pueden grabarse simultáneamente es mucho mayor.

– *Registro manual y automático:* MRecIPTV puede utilizarse de forma manual o programada. La ejecución manual permite que cualquier investigador que tenga instalada la aplicación pueda grabar contenidos. Esta opción requiere que el usuario sea el que lleve a cabo la grabación de los canales, y que sea él mismo quien decida en qué unidad quiere grabar los contenidos y cuándo quiere detener el registro.

Para la grabación automática puede utilizarse *Tareas programadas* de Windows (o cualquier otro software de programación de tareas disponible gratuitamente en internet) y reproducir una parrilla de programación televisiva o intervalos específicos por cada canal. Por ejemplo, si se desea grabar el informativo de TVE, se programará MRecIPTV_TVE (ejecutable con la UDP de *Televisión española*) a las 21 horas y el programa *MKIIIPTV* a las 22:15. Lo mismo puede hacerse con el resto de los programas del mismo canal, y con todos los programas del resto de los 60 canales de televisión accesibles mediante la red IP.

– *Economía de recursos:* El sistema MRecIPTV no requiere servidores de alta capacidad o equipos dedicados. Cualquier ordenador personal puede servir como equipo de grabación. El investigador puede trabajar con su ordenador simultáneamente al proceso de registro, comprobar su estado y visionar el contenido en tiempo real.

– *Deslocalización:* Varios ordenadores pueden estar grabando los mismos canales simultáneamente. La deslocalización de las grabaciones y las operaciones de respaldo permiten garantizar que si por algún motivo uno de los equipos no realiza la grabación, se dispone de otras grabaciones para asegurar la muestra.

– *Catalogación:* MRecIPTV puede crear los directorios y subdirectorios en los que se grabará y asignar a los archivos un nombre compuesto por el nombre del canal grabado, la fecha y la hora de inicio de la grabación. La denominación de los archivos es estándar, por lo que pueden ser catalogados y referenciados unívocamente en las bases de datos de los análisis.

– *Conservación y duplicación:* Los registros audiovisuales pueden grabarse en los discos duros de los usuarios o en unidades de red a las que el usuario esté conectado, e incluso realizar ambas acciones creando un ejecutable MRecIPTV_local y otro MRecIPTV_red. Las copias en las unidades de red son accesibles a todos los investigadores del proyecto. No hay ningún inconveniente en que todos los investigadores tengan acceso simultáneo al mismo archivo gracias al ancho de banda de la red interna de la Universidad.

– *Movilidad:* Los investigadores pueden hacer copias personales en dispositivos móviles, lo que les permite la realización de los análisis desde cualquier ubicación y ordenador, con la única condición de que el equipo tenga instalado el

software VLC Media player u otro que reproduzca archivos audiovisuales con extensión mpg. Además, los usuarios autenticados pueden acceder remotamente a los archivos utilizando la red privada virtual (VPN) de la UPF.

– *Creación de archivo digital:* Los archivos digitales son conservados y pasan a formar parte del patrimonio documental y de la memoria audiovisual del *Departamento de Comunicación*. Son y serán accesibles en cualquier momento y podrán ser utilizados para estudios comparativos más adelante. Esto significa un salto cualitativo respecto a las investigaciones anteriores, en las que resultaba muy costoso encontrar los contenidos de años anteriores de todos los canales de la muestra.

– *Edición:* Los contenidos son de altísima calidad y fácilmente editables y convertibles a cualquier estándar de vídeo. Los investigadores pueden seleccionar los fragmentos a analizar y adaptar el formato y la calidad del vídeo a las necesidades de su investigación.

– *Usabilidad:* MRecIPTV es un sistema que cualquier usuario con unos conocimientos informáticos mínimos puede utilizar.

Sin embargo, la solución también plantea problemas por depender de la recepción de la señal audiovisual desde la cabecera IP de la red. En algunas ocasiones el mal funcionamiento de dicha cabecera o de la estructura de recepción y transporte de la señal ha ocasionado la pérdida de una parte de la muestra, lo que ha obligado a rediseñar los intervalos de grabación.

“ El coste de la solución es prácticamente nulo y permite ahorrar en los presupuestos de los proyectos de investigación ”

Acceso a través del navegador

En el caso del proyecto de investigación de la *UPF Televisión y deliberación política*, las grabaciones han sido editadas y, posteriormente, se han creado páginas web para el acceso a los fragmentos. En el diseño de la *interface* se ha apostado por la sencillez y la usabilidad. Ésta permite navegar de un canal a otro y visualizar fácilmente los contenidos editados. Además, puede agregarse texto o palabras clave para que los analistas identifiquen las temáticas tratadas en los fragmentos y la búsqueda de contenidos resulte más sencilla.

Conclusiones

La recepción de la señal digital de televisión mediante la red IP (IPTV) ha supuesto una transformación radical en la forma de trabajar de los investigadores del *Departamento de Comunicación* de la UPF. El diseño del sistema de grabación MRecIPTV se llevó a cabo en 2010 y ha sido utilizado en dos proyectos de investigación —el ya mencionado *Televisión y deliberación política*, y *Elecciones 2010*, en el que se analizan los informativos de seis canales de televisión españoles a lo largo de una semana—. Asimismo, se ha utilizado en una tesis doctoral del *Departamento*, en la que se grabaron los noticiarios del primer canal de televisión ruso durante 1 mes, y toda la programación del canal durante 3 semanas más, y en

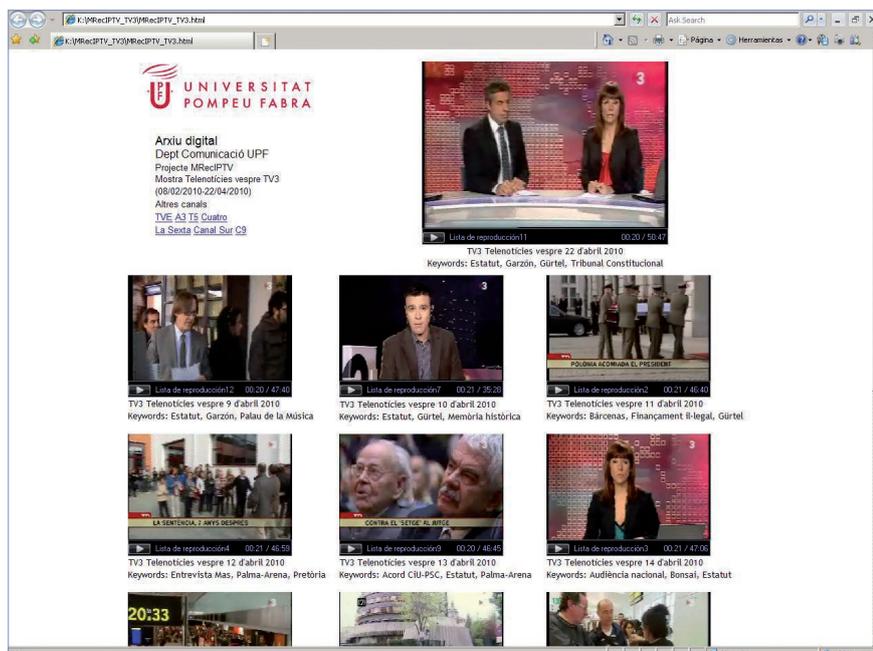


Figura 3. Prototipo de página de acceso a las muestras de proyectos

otra tesis en fase de elaboración en la que se ha grabado la autopromoción de 4 canales de televisión en horario *prime time*.

“MReCIPTV representa un avance en las soluciones de programación de las grabaciones de muestras audiovisuales, un reto que se han propuesto otras universidades a nivel europeo y mundial”

MReCIPTV representa un avance en las soluciones de programación de las grabaciones de muestras audiovisuales, un reto que se han propuesto otras universidades a nivel europeo y mundial (Balaouras; Sartzetakis, 2009). Aunque MReCIPTV es una solución propiedad del Departamento de Comunicación de la UPF, puede ser utilizada gratuitamente por cualquier otra institución, organización o empresa que disponga de una plataforma IPTV y lo solicite a los autores de este artículo.

“Entre los futuros desafíos del Departamento está la integración de las muestras en plataformas de gestión de contenidos Drupal”

Entre los desafíos futuros que afronta el Departamento está la grabación de todos los contenidos audiovisuales recibidos a través de IPTV durante los 365 días del año, y la creación de un fondo audiovisual que pueda difundirse entre todos los grupos de investigación y entre la comunidad universitaria en su conjunto. Junto a esto, se plantea integrar la solución en un sistema de gestión de contenido Drupal que per-

mita a cualquiera poner a disposición del resto del grupo investigador todas las grabaciones de *media stream* que haya realizado y crear una web 2.0 de contenido audiovisual para su análisis académico.

Notas

1. Este artículo es resultado del proyecto *Televisión y deliberación política. La construcción del espacio público a través de los géneros televisivos de la realidad en España (Tedepto)*, financiado por el Ministerio de Ciencia e Innovación (2008-2011), con referencia CSO2008-02589.

2. Un ejemplo de código puede encontrarse en: <http://www.vlc.cat>

Bibliografía

Agirreazaldegi-Berriozabal, Teresa. “Claves y retos de la documentación digital en televisión”. *El profesional de la información*, 2007, sept.-oct., v. 16, n. 5, pp. 433-442. <http://eprints.rclis.org/handle/10760/11133> <http://dx.doi.org/10.3145/epi.2007.sep.05>

Balaouras, Pantelis; Sartzetakis, Stelios. “An IPTV service for academic and research communities”. En: *Contel 2009: Proceedings of the 10th International Conference on Telecommunications*, 2009, pp. 451-456. ISBN 978 9531841306

Benoit, Hervé. *Digital television*. Focal Press, 2008. ISBN 978 0240520810

Bernier, José L.; Gómez-Domínguez, David; Fernández-Rodríguez, Gerardo; Barchéin, Mario J.; Pérez, Gregorio. “CaCoCu: un portal web para la difusión de contenidos multimedia culturales de las universidades públicas andaluzas”. *El profesional de la información*, 2007, mayo-junio, v. 16, n. 3, pp. 216-224. <http://eprints.rclis.org/handle/10760/9433> <http://dx.doi.org/10.3145/epi.2007.may.06>

Bonet, Montse; Fernández-Quijada, David. “El reto de la digitalización del archivo sonoro en los servicios públicos de radiodifusión. El caso de Catalunya Ràdio”. *El profesional de la información*, 2006, sept.-oct., v. 15, n. 5, pp. 390-396. <http://www.elprofesionaldelainformacion.com/contenidos/2006/septiembre/8.pdf>

Fernández-Quijada, David. “Aplicación de arquitecturas peer-to-peer a la distribución de archivos audiovisuales”. *El profesional de la información*, 2008, sept.-oct., v. 17, n. 5, pp. 526-531. <http://www.elprofesionaldelainformacion.com/contenidos/2008/septiembre/06.pdf> <http://dx.doi.org/10.3145/epi.2008.sep.06>

López-Yepes, Alfonso; Sánchez-Jiménez, Rodrigo. “e-Do-culInfo/TVDoc: portal de documentación informativa-canal

temático IPTV". *El profesional de la información*, 2007, sept.-oct., v. 16, n. 5, pp. 456-467.

<http://www.ucm.es/info/tvdoc/documentos/tvdocEPI.pdf>

<http://dx.doi.org/10.3145/epi.2007.sep.08>

Postel, Jon. "User datagram protocol". *Internet engineering task force*, 1980.

<http://tools.ietf.org/html/rfc768>

Simpson, Wes. *Video over ip: iptv, internet video, H.264, P2P, web TV, and streaming: a complete guide to understanding the technology*. 2nd ed. Focal Press, 2008. ISBN 978 0240810843

Spence, John. "Dams & digitization preparedness". *IASA journal*, 2002, n. 20.

Treutler, Wolfgang; Neu, Gregor; Raupp, Gerhard; Zehetbauer, Thomas; Zsche, Dieter; Lüddecke, Klaus; Cole, Richard. "Real-time signal communication between diagnostic and control in Asdex upgrade". *Fusion engineering and design*, 2010, n. 85, pp. 466.

<http://www-fusion-magnetique.cea.fr/tmiaea2009/website/data/articles/000096.pdf>

Zheng, Haitao; Boyce, Jill. "An improved UDP protocol for video transmission over internet-to-wireless networks". *IEEE transactions on multimedia*, 2001, v. 3, n. 3, pp. 356-365.

LIBROS DE BOLSILLO "EL PROFESIONAL DE LA INFORMACIÓN"

Colección de libros de información y documentación

Dirigida por Javier Guallar y Tomàs Baiget

Títulos publicados:

Tecnologías de la web semántica
Juan-Antonio Pastor

La revolución del libro electrónico
José-Antonio Cordón

Sistemas de información en la empresa
Josep Cobarsí-Morales

Información en el móvil
Natalia Arroyo

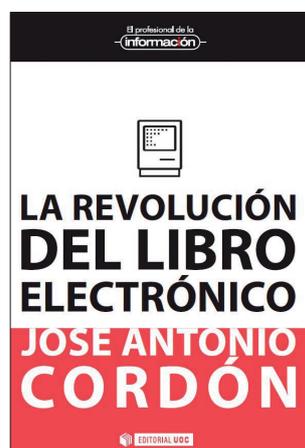
Acceso abierto a la ciencia
Ernest Abadal

En imprenta:

Fuentes de información médica
Pablo Medina-Aguerreberre

Gestión de contenidos
Ricardo Eíto-Brun

Wikipedia de la A a la W
Tomás Saorín



A la venta en librerías y en:

<http://www.editorialuoc.cat> - <http://www.amazon.es>

Un proyecto conjunto de



y



X convocatoria

Gestión de Contenidos 2012

Contenidos Gestionados como valor de Negocio

8 Febrero de 2012 | Hotel Palace | Madrid

22 Febrero de 2012 | Hotel NH Constanza | Barcelona

*“La información digital seguirá experimentando en 2012 fuertes tasas de crecimiento”.
“La tecnología impulsa el cambio y los cambios exigen una transformación en la manera de gestionar la información. Para el año 2020, las proyecciones indican que el 90% de todos los nuevos contenidos, serán no estructurados ”.*

Fuente IDC

IDC ENTREVISTARÁ A:



Pilar Marqués Martínez

Directora de Reputación Corporativa,
Marca y Comunicación Digital
REPSOL

“Primero fue en España, ahora es a nivel europeo: Repsol es la compañía mejor valorada por su transparencia en Internet”

CON LA EXPERIENCIA DE:

IDC
CORREOS
EZ SYSTEMS
FRATERNIDAD MUPRESA
FEDIT
J&A GARRIGUES
MICROSOFT
REPSOL
RTVA
TECNOCOM
OPENTEXT

GOLD SPONSOR:



SILVER SPONSORS:







ASOCIACIONES COLABORADORAS:





SEDIC

ASOCIACIÓN ESPAÑOLA
DE DOCUMENTACIÓN
E INFORMACIÓN

MEDIA SPONSORS:











SOFTWARE DOCUMENTAL



CARROT2: BÚSQUEDA Y VISUALIZACIÓN DE LA INFORMACIÓN



Audilio Gonzales-Aguilar y María Ramírez-Posada



Audilio Gonzales-Aguilar es doctor en derecho y nuevas tecnologías por la *Universidad de Montpellier I*. Es profesor titular de la *Universidad Paul Valéry Montpellier 3* en Montpellier, Francia, donde es miembro del equipo de investigación *Praxiling UMR 5267 CNRS – Université Paul-Valéry– Montpellier 3*. Director del master profesional en gestión de la información y de la documentación. Sus áreas de investigación son el análisis de redes sociales aplicadas a contenidos documentales, hipertextualidad del documento digital y visualización y representación de la información

Universidad Paul Valéry Montpellier 3

Departamento de Documentación

Route de Mende, 34199 Montpellier Cedex 5, Francia

<http://www.univ-montp3.fr>

audilio.gonzales@univ-montp3.fr



María Ramírez-Posada es bióloga marina, candidata a maestría en desarrollo sostenible y medio ambiente (*Universidad de Manizales*, Colombia) y candidata a maestría en gestión de la información y del conocimiento (*Universidad Paul Valéry Montpellier 3*, Francia). Cuenta con experiencia como investigadora en el área del desarrollo sostenible empresarial y visualización de la sustentabilidad y como docente a nivel pre gradual y pos gradual.

Grupo de Investigación Eco-ambiental

<http://201.234.78.173:8080/gruplac/jsp/visualiza/visualizagr.jsp?nro=0000000005056>

webdocumenta@gmail.com

Resumen

Obtener información en cortos períodos de tiempo puede marcar la diferencia en caso de una investigación o de un negocio, por lo cual es importante contar con herramientas que faciliten su búsqueda. Se analiza una técnica de visualización conocida como motor de agrupamiento de información: *Carrot2*, que separa los documentos –tanto encontrados en la Web como por grupos o clústeres, utilizando algoritmos de agrupamiento que, mediante la consulta de diversas fuentes de datos, los procesa y muestra la visualización de la información obtenida. Este software de código abierto reduce significativamente el esfuerzo de la recuperación de la información y su análisis, organizándola por grupos temáticos.

Palabras clave

Visualización de la información, Recuperación de información, Búsqueda web, *Clustering*, Árbol de resultados.

Title: *Carrot2*: Information search and visualization

Abstract

Obtaining information in short periods of time can make the difference for research or business; therefore it is important to have tools that facilitate the search for information. A visualization technique called information-clustering engine is reviewed: *Carrot2* separates the documents found on the Web into groups or clusters, using clustering algorithms that consult various sources of data and then process and display the information obtained. This open-source software significantly reduces the effort involved in information retrieval (IR) and analysis.

Keywords

Information visualization, Information retrieval, Web search, Clustering, Results tree.

Gonzales-Aguilar, Audilio; Ramírez-Posada, María. “*Carrot2*: búsqueda y visualización de la información”. *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 105-112.

<http://dx.doi.org/10.3145/epi.2012.ene.14>

Artículo recibido el 11-01-12

Aceptación definitiva: 18-01-12

Introducción

Con el avance de internet y la existencia de nuevas tecnologías y servicios, el acceso a la información ha tenido un crecimiento sin precedentes. Al buscar información, lo que normalmente se hace es explorar, describir y organizar manualmente los resultados recuperados¹. Pero hoy en día se pueden usar tecnologías eficientes y flexibles, capaces de combinar la recuperación y la posterior organización de la información.

Los motores de búsqueda muestran una lista de resultados por orden de relevancia, un ranking, y el usuario tiene que examinarla de forma descendente, hasta hallar la información solicitada. No hay forma de determinar exactamente qué es relevante para el usuario, ya que las consultas suelen ser muy breves y su interpretación es confusa en ausencia de un contexto. Por ello, a pesar de que los motores de búsqueda son buenos para ciertas tareas, pueden ser menos efectivos para satisfacer determinadas demandas amplias o, al contrario, dar respuesta a preguntas definidas.

El acceso a la información ha tenido un crecimiento sin precedentes con el uso de nuevas tecnologías y servicios

Un enfoque diferente es el de la agrupación de los resultados en los denominados clústeres. El programa *Carrot2* realiza el *clustering*² (categorización o agrupamiento) de los documentos hallados basándose en similitudes entre ellos, sin un conocimiento a priori de sus características (**Goldenberg, 2007**), lo cual permite mejorar la precisión. El usuario

plantea un tema general y posteriormente puede pasar a analizar los temas más específicos creados de forma dinámica a partir de los resultados de la consulta.

<http://project.carrot2.org>

Carrot2 facilita ampliar o modificar la estructura de la búsqueda, comprender mejor tema, y favorece la exploración sistemática de los resultados.

La agrupación de documentos en clústeres o grupos permite mejorar la precisión de la recuperación de información

Características

Carrot2 es un software de recuperación, *clustering* y visualización de documentos y contenidos web. Contiene una colección de algoritmos³ de agrupación que facilitan la exploración del contexto temático de los documentos recuperados a través de los motores de búsqueda o de una colección de textos en un ordenador o en un servidor⁴.

Instalación

Carrot2 puede ser utilizado online como metamotor o buscador de contenidos web a partir de cualquier navegador. Es posible integrar directamente los plug-in, o extensiones, para *Firefox* e *Internet Explorer*.

<http://search.carrot2.org/stable/search>

<http://project.carrot2.org/download-search-plugins.html>

En su versión de escritorio local se puede descargar para los sistemas *Windows*, *Mac OS* y *Linux*. Asimismo, presen-

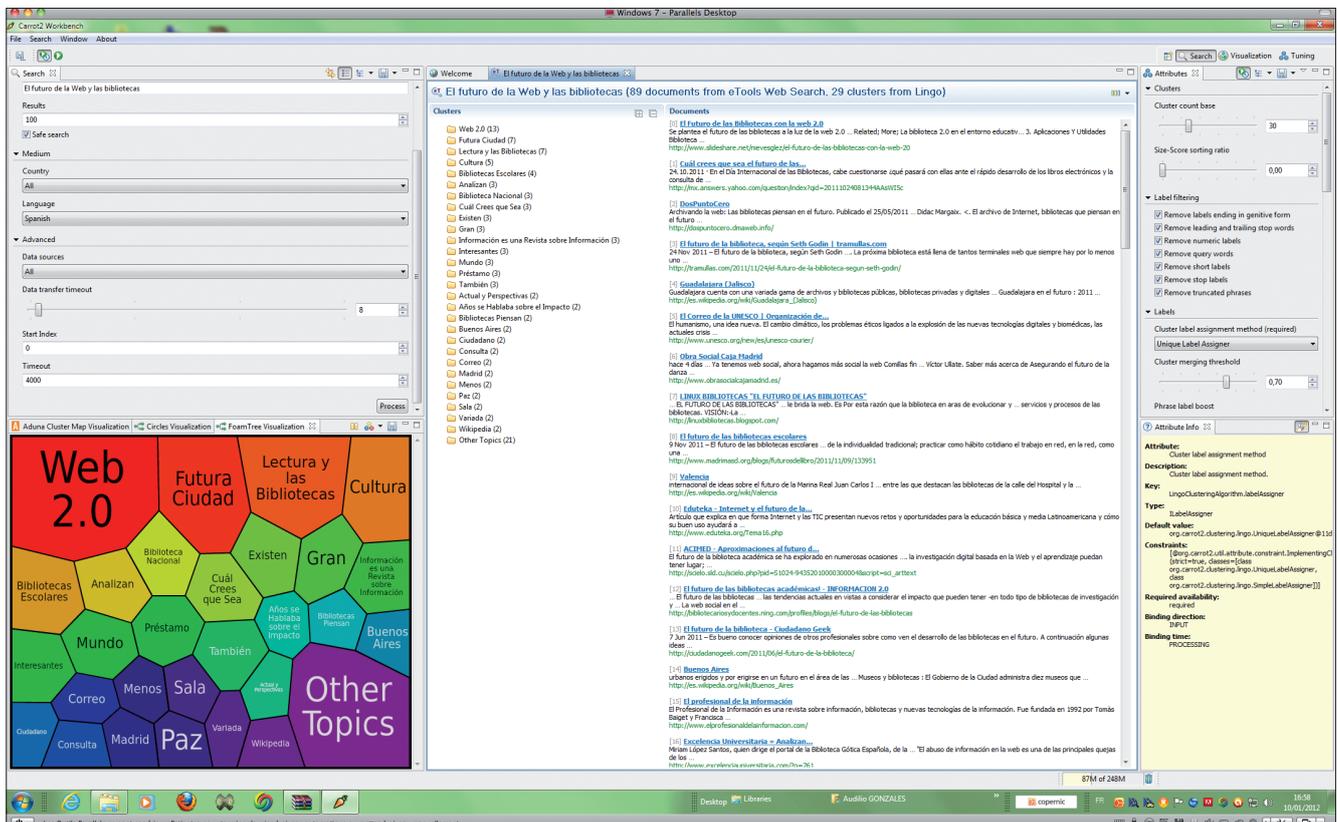


Figura 1. *Carrot2* crea clústeres de etiquetas de un conjunto de documentos.

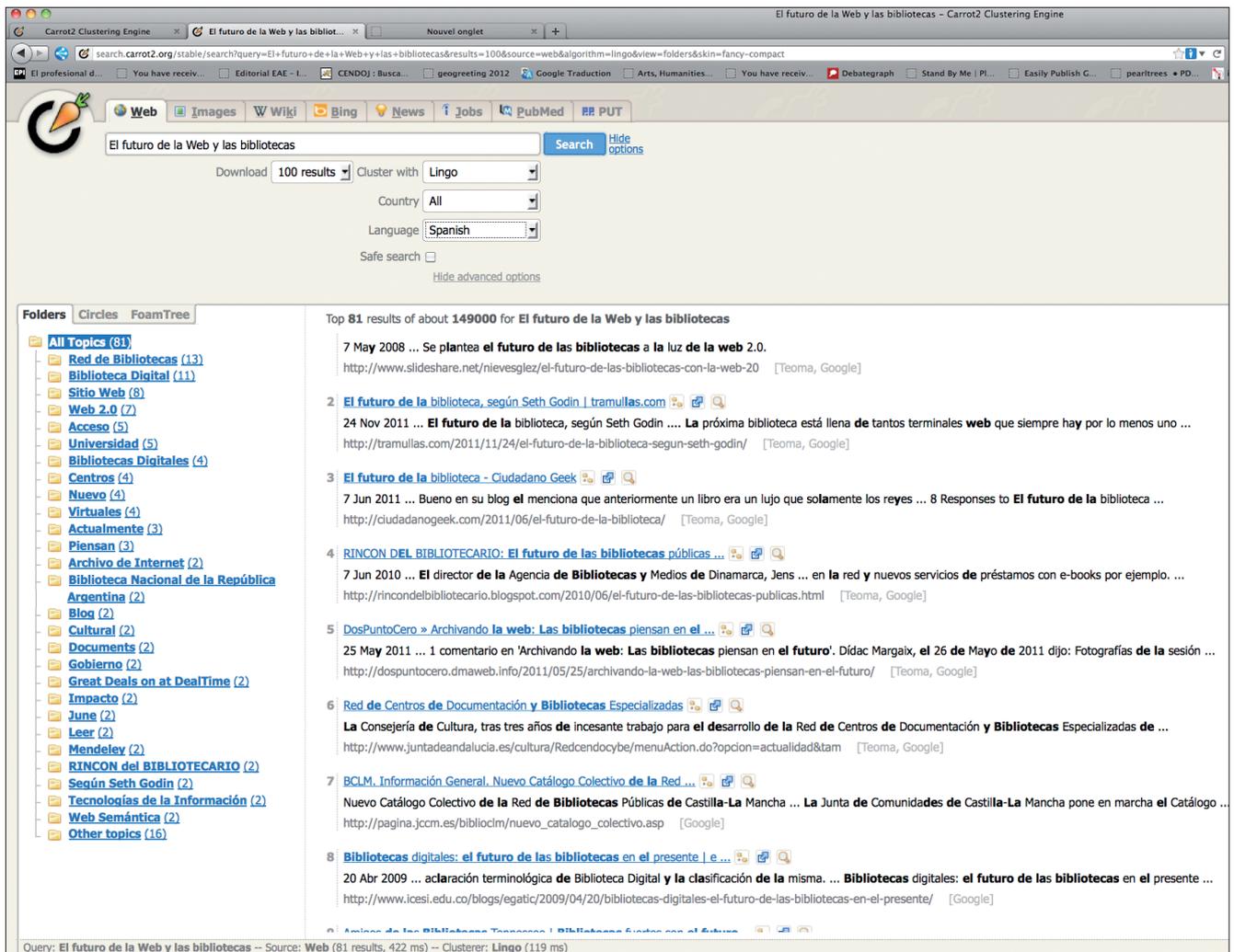


Figura2. Interfaz de Carrot2 en línea

ta otras posibilidades de descarga y utilización: en formato java, integración con PHP, C# y además se puede instalar como una instancia de búsqueda en un sitio web.

<http://project.carrot2.org/download-workbench-win32.html>

<http://project.carrot2.org/download-workbench-macos-cocoa-32bit.html>

<http://project.carrot2.org/download-workbench-linux.html>

<http://project.carrot2.org/download.html>

Carrot2 es un software de recuperación, clustering y visualización de documentos que se puede usar tanto online en su web como en nuestro ordenador

En Windows está disponible para win32 y para win64 (sistemas basados en una arquitectura de 32 ó 63 bits). Una vez descargado, se descomprime y se ejecuta el archivo *carrot2-workbench.exe*. El programa no requiere ninguna instalación adicional. Además, si el ordenador personal tiene instalado *Google Desktop*, Carrot2 recupera toda la información indexada con este medio, e igualmente puede emplearse para la recuperación de documentación dentro de un equipo cliente.

<http://googledesktop.blogspot.com>

Interfaz de Carrot2

Carrot2 tiene una interfaz que corresponde a tres fases del proceso de la información: entrada, filtrado y salida de la información. En la figura 3 se pueden observar los componentes de la aplicación de escritorio y en la figura 4 los de la interfaz en línea.

Opciones de búsqueda

Comprende:

- Fuentes de información, referentes a los motores de búsqueda como *Google*, *Yahoo*, *Wiki*, *PubMed*, etc.
- Algoritmos para la creación de grupos o clústeres como *Lingo* (que se emplea por defecto), *K-mean*, *STC*. También puede agrupar por dominio (.net, .com, .org...).

Páginas de resultados

Muestra un listado de todos los clústeres identificados y de los documentos de cada clúster por orden de relevancia.

Formas de visualización gráfica

Carrot2 presenta tres tipos de representación gráfica, como se observa en la figura 6:

- Esquema relacional *Aduna*: visualiza las relaciones entre los clústeres. Desarrollado por <http://www.aduna-software.com>

- Diagrama circular.
- Mapa de áreas, utilizando colores cálidos para las más pertinentes y colores fríos para las menos en la parte inferior. En la base del mapa aparece el grupo desconectado "Other topics".

Configuración de atributos

En la versión de escritorio *Carrot2* integra una ventana de atributos en la que se pueden configurar y editar las dimensiones de agrupamientos. Esta ventana se encuentra a la derecha de la pantalla del programa.

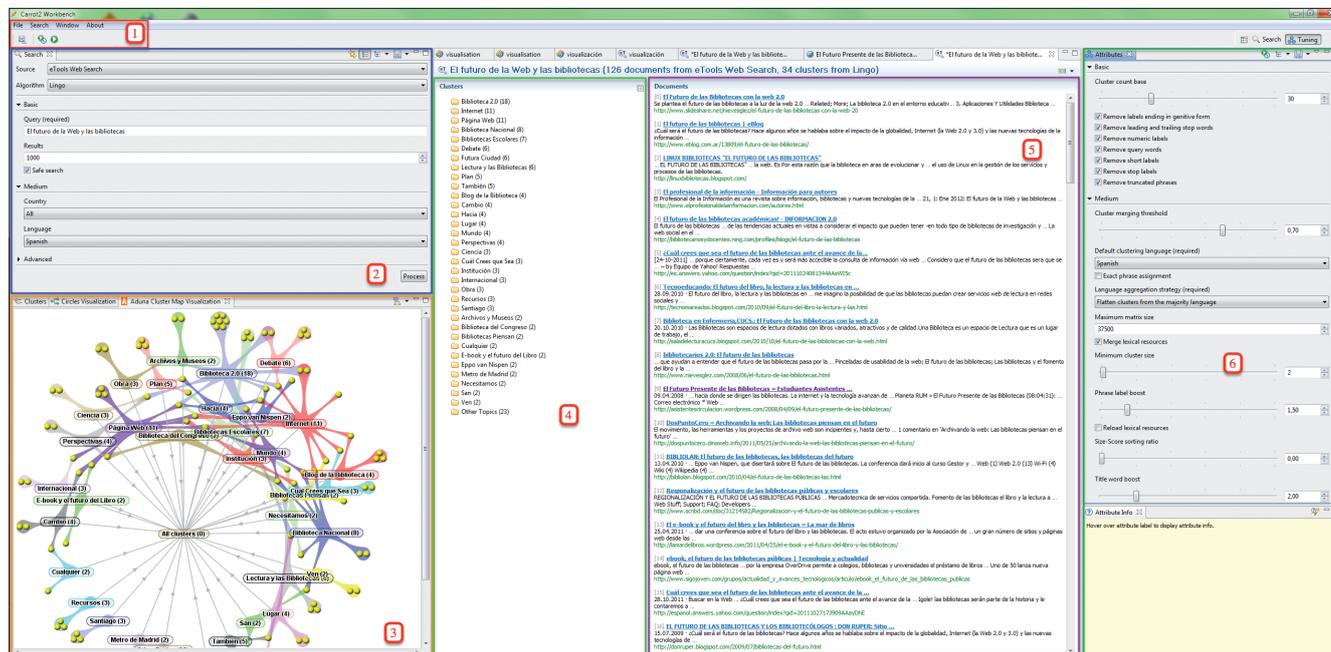


Figura3. Ventanas de la interfaz *Carrot2* en la aplicación escritorio.
 1. Menú del programa, 2. Opciones de búsqueda, 3. Visualización, 4. Ventana de clústeres, 5. Resultados, 6. Edición y atributos de la consulta.

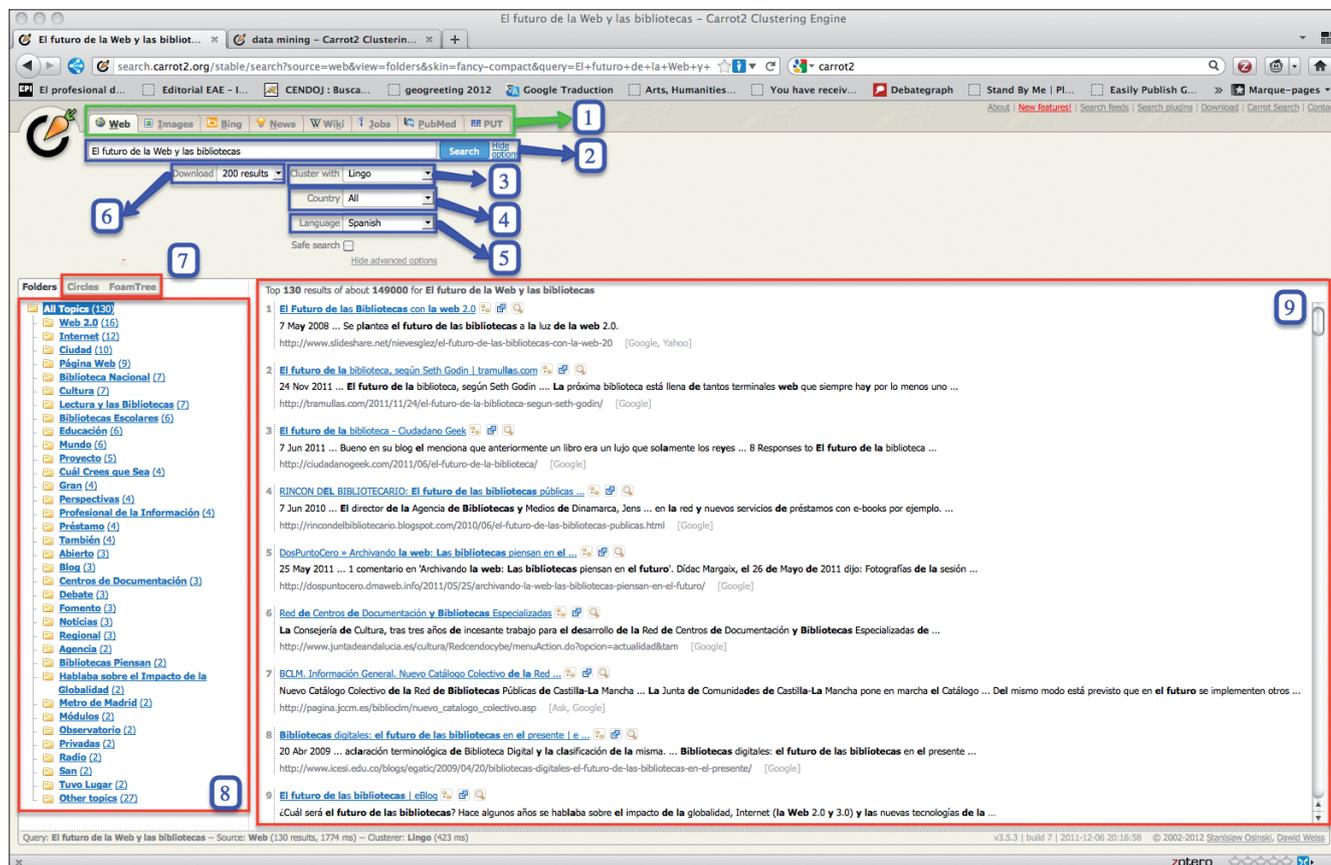


Figura 4. Elementos que componen la interfaz de *Carrot2* en línea.
 1. Opciones de búsqueda, 2. Campo de búsqueda, 3. Ventana de clústeres, 4. País, 5. Idioma, 6. Numero de resultados solicitados, 7. Visualización (diagrama circular, mapa de superficie por grupos), 8. Ventana de clústeres, 9. Resultados.

Conclusión

Es de destacar la flexibilidad de esta aplicación, que permite hacer una *radiografía* de los resultados de una búsqueda de información documental. Cuando se señala que haga una búsqueda en la Web en general (opción *eTools Web Search*) la hace simultáneamente en *Ask*, *Bing*, *Entireweb*, *Teoma*, *Yahoo*, *Google* y *Wikipedia* y forma un conjunto con los primeros resultados de cada buscador. También se pueden especificar determinados espacios de búsqueda como imágenes, wikis, noticias, arte, *PubMed*, *Google Desktop* (sobre los documentos de nuestra computadora), etc. Estaría bien que en el futuro añadieran *Google Scholar*.

Flexibilidad, pertinencia y excelentes opciones gráficas son características básicas de *Carrot2*

Al usar el programa, la lista de los primeros 100 resultados (la cantidad puede variarse) se representa con un árbol clasificándolos a partir de las palabras que figuran en los documentos. Como fallo del sistema –y de todos los buscadores agrupadores– hay que indicar que algunos de los conjuntos se basan en palabras vacías (por ejemplo, en la figura 6 vemos “basados”, “existen”, “habló”, “también”, “cualquier”...), lo cual disminuye mucho su eficacia.

Clasificación jerárquica de los resultados, con una interfaz fácil de trabajar

Sin embargo, a la falta de precisión de las búsquedas, como ocurre con todos los sistemas de recuperación de información, hay que encontrarle las ventajas de la serendipidad: hallar cosas interesantes sin buscarlas específicamente, y aquí con una representación gráfica, al ser los conceptos más evidentes, el descubrimiento es más fácil.

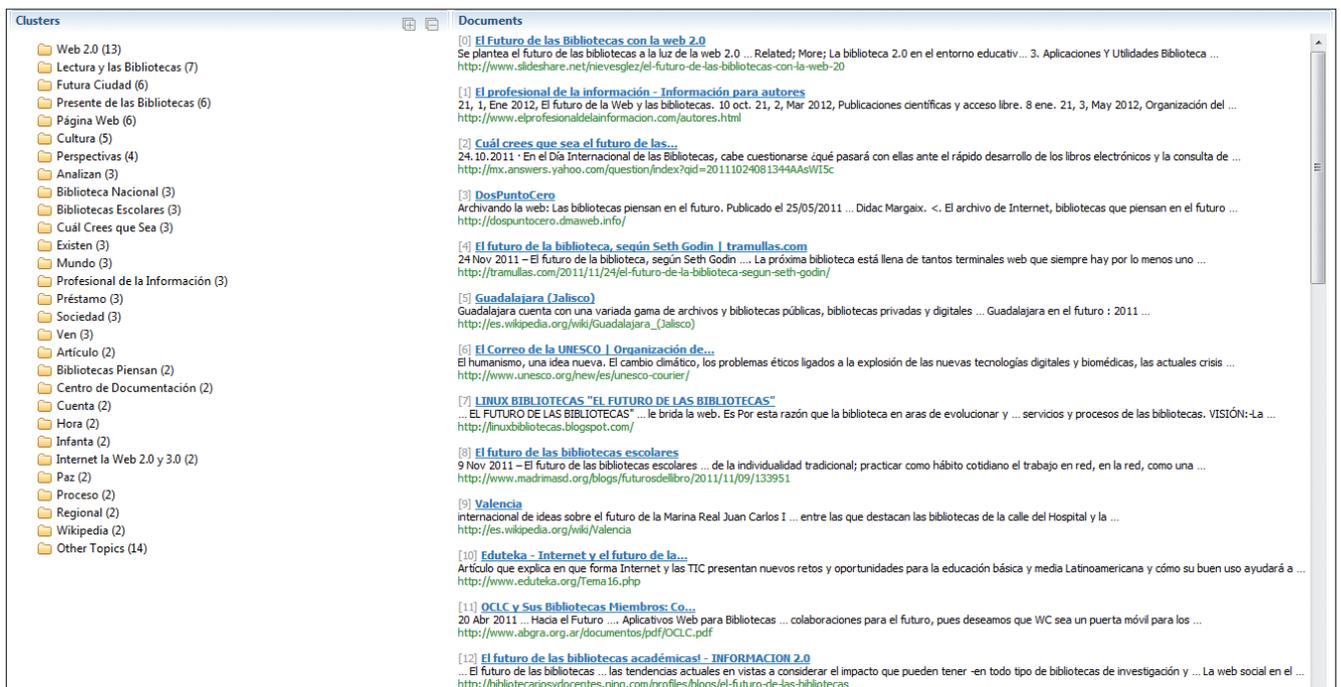


Figura 5. Ventana de resultados (clústeres y documentos)

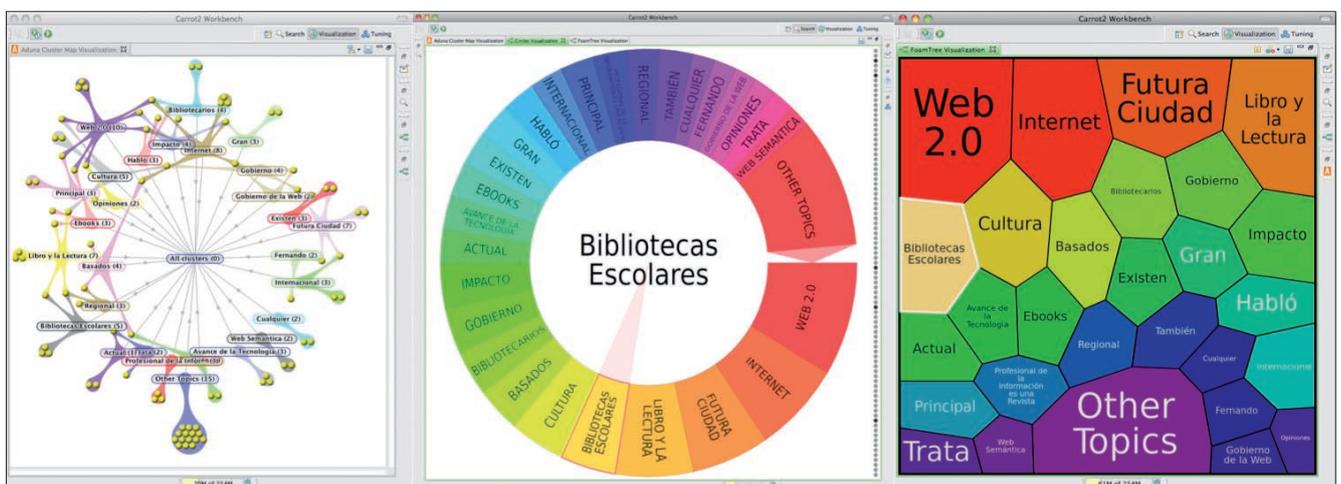


Figura 6. Formas de visualización gráfica de *Carrot2*

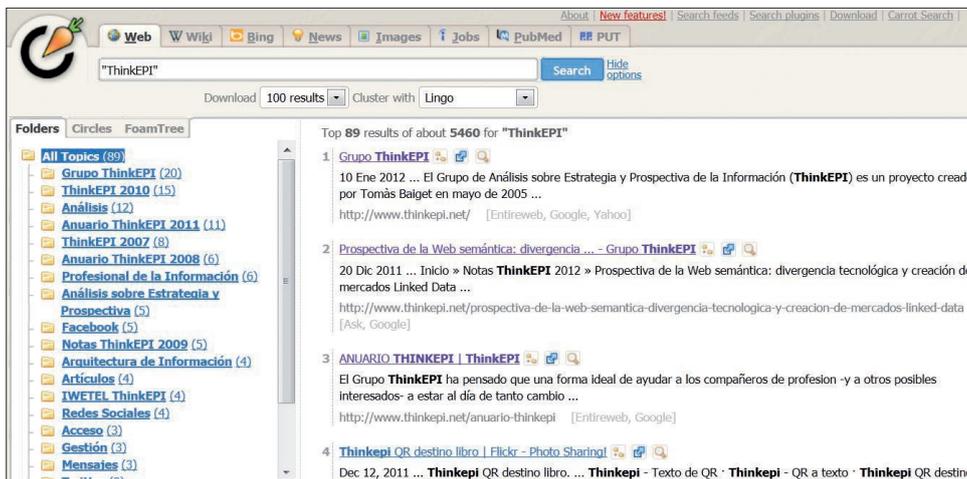


Figura 7. Carrot2 da más precisión añadiendo comillas. Es conveniente ponerlas incluso con palabras sueltas, pues el sistema las descompone. En este ejemplo buscaría también “think” y “epi”

Al hacer clic en cualquiera de los grupos o clúster se despliega una lista de los resultados correspondientes.

La visualización de la información en Carrot2 permite la clasificación de los resultados de búsqueda, la optimización de los mismos y la extracción de palabras clave.

Notas

1. **Dürsteler, Juan C.** Visualización de la información: proceso de interiorización del conocimiento mediante la percepción de información.

<http://www.infovis.net/printMag.php?num=100&lang=1>

2. El término *clustering* o agrupación hace referencia a aquellos sistemas de recuperación que emplean algoritmos de agrupación de contenidos, y cumplen la tarea de particionar un espacio de información no etiquetada en grupos, clases o clústeres.

3. Carrot2 organiza los documentos y contenidos web utilizando algoritmos de agrupación. Lingo es uno de ellos, que logra bastante calidad en etiquetar los grupos basándose en frases recurrentes y crea un árbol de resultados.

4. Los ejemplos típicos para demostrar esta técnica muestran el contexto de las consultas amplias y ambiguas de sinonimia conceptual como “apache” (helicóptero, tribu indígena o software) o “salsa” (danza o comida).

Bibliografía

Anastasiu, David C.; Gao, Byron J.; Buttler, David. “ClusteringWiki: a framework for personalized clustering of search results”. En: *Sigir’11 Procs of the 34th intl ACM Sigir conf on research and development in information retrieval*, 2011. <http://dmlab.cs.txstate.edu/ClusteringWiki/pdf/cw.pdf>

Carpineto, Claudio; Osinski, Stanislaw; Romano, Giovanni; Weiss, Dawid. “A survey of web clustering engines”. *ACM computing surveys*, 2009, July, v. 41, n. 3, pp. 17-38. <http://search.fub.it/claudio/pdf/CSUR09.pdf> <http://dx.doi.org/10.1145/1541880.1541884>

Cigarrán-Recuero, Juan-Manuel. *Organización de resultados de búsqueda mediante análisis formal de conceptos.*

Tesis doctoral. Universidad Nacional de Educación a Distancia, 2008.

<http://e-spacio.uned.es/fez/view.php?id=tesisuned:IngInf-Jcigarran>

Goldenberg, Daniel. *Categorización automática de documentos con mapas auto-organizados de Kohonen.* Tesis de magister. ITBA, 2007.

<http://www.itba.edu.ar/archivos/secciones/goldenberg-tesisdemagister.pdf>

Laufert, Thomas. “Visualization of the Carrot 2 system”. Baltimore: University of Mary

land Baltimore County.

<http://www.csee.umbc.edu/conference/src/conferences/src1/www/papers/laufert.pdf>

Osiński, Stanisław. “Improving quality of search results clustering with approximate matrix factorisations”. En: *Lecture notes in computer science, Procs of the 28th European conf on IR research*. 2006, v. 3936, pp. 167-178.

http://dx.doi.org/10.1007/11735106_16

Osiński, Stanisław; Weiss, Dawid. “Carrot2: design of a flexible and efficient Web information retrieval framework”. En: *Lecture notes in computer science. Procs of the Third intl Atlantic web intelligence conf (AWIC 2005)*, Łódź, Poland, 2005, v. 3528, pp. 439-444.

<http://www.cs.put.poznan.pl/dweiss/site/publications/download/dweiss-carrot2-poster-awic2005.pdf>

http://dx.doi.org/10.1007/11495772_68

Osinski, Stanislaw; Weiss, Dawid. *Clustering search results with Carrot2.* Polonia: Poznan Supercomputing and Networking Center; Institute of Computing Science, Poznan University of Technology, 2007.

<http://project.carrot2.org/publications/carrot2-dresden-2007.pdf>

Senthil-Kumar, R. Subhashini; Senthil-Kumar, V. Jawahar. “The anatomy of web search result clustering and search engines”. *Indian journal of computer science and engineering*, 2010, v. 1, n. 4, pp. 392-401.

Stefanowski, Jerzy; Weiss, Dawid. “Carrot2 and language properties in Web search results clustering”. *Lecture notes in artificial intelligence: advances in Web intelligence. Procs of the First intl Atlantic web intelligence conf*, Madrid, Spain, 2003, n. 2663, pp. 240-249.

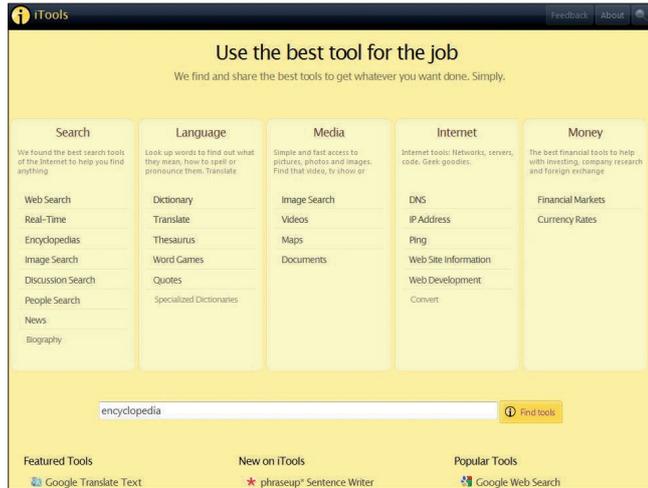
http://dx.doi.org/10.1007/3-540-44831-4_25

Weiss, Dawid; Osinski, Stanislaw. “Carrot2 clustering framework”. *Procs of the First intl conf on open source systems*, 2009, pp. 298-299.

Weiss, Dawid. *Descriptive clustering as a method for exploring text collections.* Tesis doctoral. Institute of Computing Science, Poznan University of Technology, 2006.

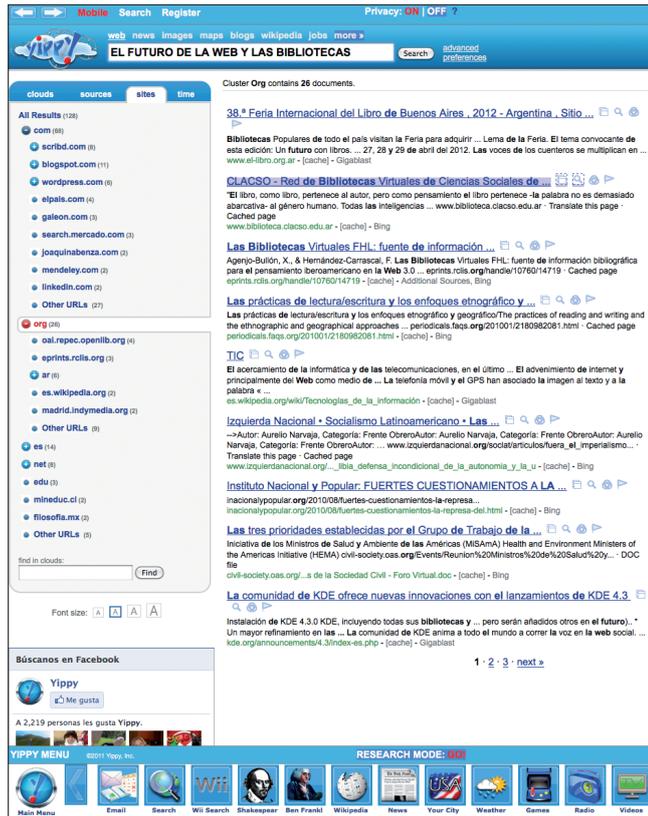
<http://www.cs.put.poznan.pl/dweiss/site/publications/download/dweiss-phd-thesis.pdf>

iTools
http://itools.com



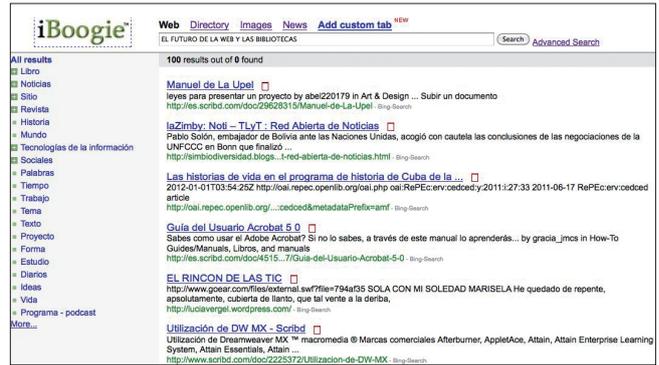
Es un "buscador de herramientas de la Web". En el buscador inicial se pregunta el tipo de información o servicio que se desea (en inglés): enciclopedias, traductores, vídeos, idiomas, finanzas..., y el sistema ofrece diferentes opciones donde buscar. Además es como un portal con multitud de recursos agrupados por tipos, que se abren y se usan haciendo clic en ellos sin más.

Yippy
http://www.yippy.com



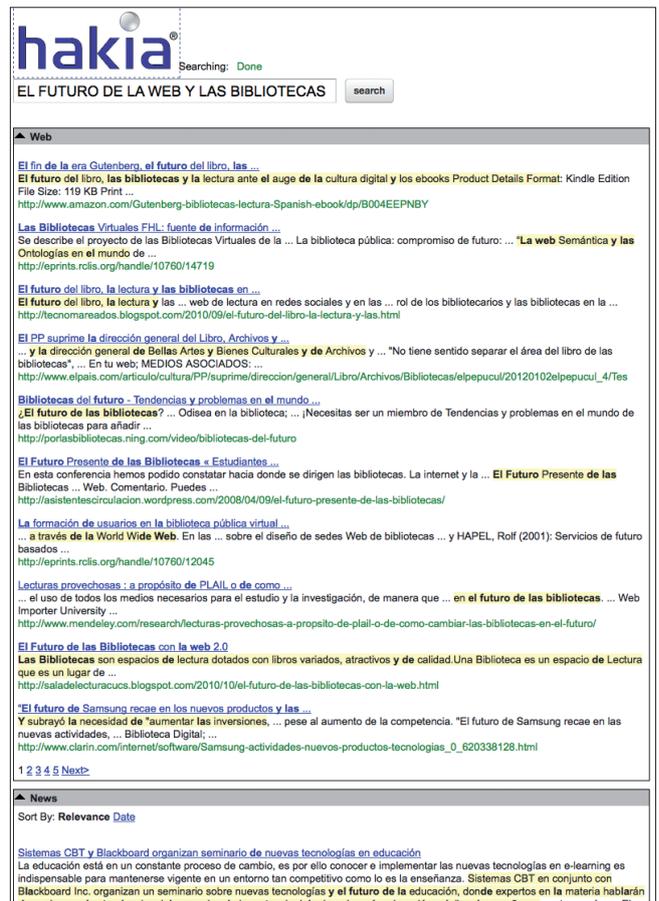
Metabusador bastante estándar, como los ya descritos anteriormente. Tiene una pestaña llamada *Yippy labs*, con aplicaciones concretas como un buscador adaptado para ser usado desde la *Wii* de *Nintendo*, nube de etiquetas, buscador sobre Shakespeare, y buscador sobre Benjamin Franklin. Tiene un filtro para contenidos no apropiados para niños.

iBoogie
http://www.iboogie.com



Las etiquetas de los grupos que forma aparecen a la izquierda. Su característica distintiva es *Custom tab*, que permite añadir webs concretas pre-seleccionadas de Blogs, Alimentos, Juegos, Gobierno de los EUA, Editorial IDG, Israel, Bibliotecas, Medicina, Buscadores, Noticias, etc.

Hakia
http://www.hakia.com



Uno de los mejores buscadores de esta lista. Organiza los resultados por diferentes tipos de fuentes de información, incluida la "credibilidad", que contiene los sitios recomendados. Se presenta como *buscador semántico* sobre medicina (ha realizado <http://newpubmed.com> para *PubMed*), finanzas para inversores (*MoodTrade*), industria aeronáutica (*AeroHakia*). Dice no sólo "leer" las palabras, sino que las "interpreta".

INDICADORES



SISTEMA DE INDICADORES CULTURALES LOCAL (SICLO)



Vicente Coll-Serrano, Salvador Carrasco-Arroyo, Olga Blasco-Blasco y Luis Vila-Lladosa



Vicente Coll-Serrano es profesor del *Departamento de Economía Aplicada* de la *Universidad de Valencia* y miembro del grupo de investigación *MC2 (Métodos Cuantitativos para la Medición de la Cultura)*. Como investigador ha participado en diversos proyectos y contratos de transferencia vinculados con el diseño e implementación de sistemas de indicadores culturales para distintos organismos y administraciones públicas.

*Facultad de Economía
Departamento de Economía Aplicada
Avda. dels Tarongers, s/n. 46021 Valencia
vicente.coll@uv.es*



Salvador Carrasco-Arroyo es profesor del departamento de *Economía Aplicada* de la *Universidad de Valencia* y director del grupo de investigación *MC2 (Métodos Cuantitativos para la Medición de la Cultura)*. Ha publicado artículos tanto en revistas nacionales como internacionales sobre técnicas de análisis de datos en estudios de fenómenos culturales, educativos. Como investigador principal ha dirigido proyectos sobre diseño de instrumentos de medición, análisis y evaluación de políticas de cultura y desarrollo, y sistemas de información para el seguimiento y evaluación de la acción cultural de entidades públicas.

*Facultad de Economía
Departamento de Economía Aplicada
Avda. dels Tarongers, s/n. 46021 Valencia
salvador.carrasco@uv.es*



Olga Blasco-Blasco es directora y profesora del *Departamento de Economía Aplicada* de la *Universidad de Valencia*. Como miembro del grupo *MC2 (Métodos Cuantitativos para la Medición de la Cultura)* ha participado como investigadora en el diseño e implementación de sistemas de indicadores para la evaluación de políticas culturales locales. Participa en el proyecto *Evaluación de Resultados de Aprendizaje* coordinado por el *Bologna Experts Team (BET)* - España.

*Facultad de Economía
Departamento de Economía Aplicada
Avda. dels Tarongers, s/n. 46021 Valencia
olga.blasco@uv.es*



Luis Vila-Lladosa es profesor del *Departamento de Economía Aplicada* de la *Universidad de Valencia*, donde imparte cursos de economía de la educación. Es investigador colaborador del *Ceges (Estudios en Gestión de la Educación Superior)* y miembro del grupo de investigación *MC2 (Métodos Cuantitativos para la Medición de la Cultura)*. Su actividad investigadora y de transferencia se encuentra enfocada a la evaluación econométrica de los efectos socio-económicos de las inversiones educativas tanto desde perspectivas macro como microeconómicas.

*Facultad de Economía
Departamento de Economía Aplicada
Avda. dels Tarongers, s/n. 46021 Valencia
luis.vila@uv.es*

Resumen

Se ofrece una visión sintética del proceso que ha dado lugar al *Sistema de Indicadores Culturales Local (Siclo)*, una aplicación orientada a mejorar el conocimiento sobre el estado de la acción cultural en el ámbito de la administración local. Pretende proporcionar información de apoyo que facilite la evaluación de las políticas culturales locales. Se estructura en cinco áreas temáticas y tres módulos. Actualmente sólo se encuentra operativo el módulo básico que está integrado por 46 indicadores.

Palabras clave

Sistemas de información cultural, Sistema de indicadores, Indicadores culturales, Política cultural, Administración local, Gobierno local, Ayuntamientos.

Artículo recibido el 21-10-11
Aceptación definitiva: 08-12-11

English title: A local cultural indicators system: *Siclo*

Abstract

A synthetic view of the process that led to the Local Cultural Indicators System (*Siclo* in its Spanish acronym) is presented. *Siclo* is an application designed to improve knowledge about the status of cultural actions at the local administration level. This tool seeks to provide supporting information to facilitate the evaluation of local cultural policies. It is structured into five thematic areas and three modules. The basic module, comprising a total of 46 indicators, is the only one which is operative at the moment.

Keywords

Cultural information systems, Indicators systems, Cultural indicators, Cultural policy, Local governments, City councils, Town halls.

Coll-Serrano, Vicente; Carrasco-Arroyo, Salvador; Blasco-Blasco, Olga; Vila-Lladosa, Luis. "Sistema de indicadores culturales local: *Siclo*". *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 113-117.
<http://dx.doi.org/10.3145/epi.2012.ene.15>

1. Introducción

El concepto de cultura, así como sus implicaciones para el desarrollo económico, social, territorial, urbano y comunitario, ha sido ampliamente debatido en distintos foros. En este documento se delimita la reflexión al ámbito local debido al importante papel que juegan los gobiernos locales en la potenciación del sector cultural como mecanismo de desarrollo. El análisis de la información permitirá apreciar la realidad en un determinado escenario así como investigar las estructuras de las relaciones causales entre cultura y desarrollo. La disponibilidad de información, de datos, es imprescindible para comprender la realidad e intentar mejorarla; la información se revela como el recurso (intangible) más valioso para orientar los cambios.

Sin embargo, el sector de la cultura adolece, en general, de información confiable, homogénea y comparable, lo que dificulta su análisis (Carrasco, 2006). La disponibilidad y el acceso a información suele estar asociado con el grado de desarrollo económico, político y social de un país; aunque cada vez con más frecuencia son las administraciones públicas quienes toman la iniciativa de buscar y acceder a la información cultural y promueven el diseño e implantación de indicadores culturales. Más concretamente, por proximidad a las inquietudes ciudadanas, los gobiernos locales se convierten en actores principales que necesitan estructurar y sistematizar la información cultural a través de un sistema de información confiable y útil que permita satisfacer, por un lado, la necesidad de someter a análisis rigurosos la gestión de las políticas públicas locales y, por otro lado, evaluar la adecuación de la planificación a través del análisis de las relaciones entre los objetivos planificados y los logros alcanzados. Los objetivos últimos del sistema de información son (i) apoyar el proceso de adopción de decisiones para fortalecer la política pública cultural mediante el incremento de la eficacia y la eficiencia en el uso de los recursos y (ii) reforzar los mecanismos de rendición de cuentas de cara a los ciudadanos. Lograr estos objetivos supone mejorar la realidad cultural, que debe ampliar las oportunidades económicas de los individuos y grupos sociales, elevando la renta de las personas y reduciendo la pobreza, a la vez que se mejora la educación, la salud, la conservación

del medioambiente, la accesibilidad y participación cultural, así como todas las relaciones simbólicas que se establecen en un escenario complejo y extraordinariamente cambiante como es el urbano (Rausell, 2008).

Establecida la importancia de disponer de un sistema de información cultural a nivel local, en este artículo se describen las principales líneas de trabajo que se han realizado en esta dirección y los resultados alcanzados: el *Sistema de indicadores culturales local (Siclo)*.

2. Antecedentes

Para diseñar un sistema de información cultural local es necesario tomar como marco de referencia la *Agenda 21 de la Cultura (Ag21)* de las *Ciudades y Gobiernos Locales Unidos (UCLG, 2004)*. Éste es el primer documento con espíritu global que apuesta por establecer las bases de un compromiso entre las ciudades y los gobiernos locales para el desarrollo cultural. Según Pascual (2005), es el documento internacional que propone con mayor énfasis un marco normativo para la acción pública en el ámbito de la cultura. Entre sus principales rasgos destacan su valor universal frente a los conflictos que conlleva la globalización cultural, la vocación de transversalidad y la centralidad de las políticas culturales. La *Ag21 de la Cultura* realiza una serie de recomendaciones a los Estados en materias tan decisivas como los derechos humanos, el medio ambiente y el territorio, la inclusión social y la promoción económica, recalcando el soporte de la cultura como cuarto pilar; pero también recomienda a los gobiernos locales la creación de sistemas de indicadores culturales a partir de los cuales pueda efectuarse el seguimiento y evaluación de las políticas culturales, de sus objetivos y de sus logros.

En base a este marco, en 2005 la *Comisión de Cultura* de la *FEMP* en colaboración con el *Ministerio de Cultura* inició un proyecto para diseñar un sistema de indicadores basado en la necesidad de (i) orientar las políticas culturales locales hacia las tesis planteadas en la *Ag21* e (ii) implantar sistemas de información que permitan realizar una planificación estratégica de la política cultural de las administraciones públicas.

El resultado de este trabajo se materializó en la *Guía para la evaluación de las políticas culturales locales (FEMP, 2009)*.

Del sistema de indicadores que a nivel teórico se propone en este documento cabe destacar que (i) se centra en la búsqueda de indicadores sencillos, confiables, útiles y conseguibles a nivel local que permitan implantar un sistema de seguimiento y evaluación, y (ii) no contempla todos los principios y recomendaciones que contiene la *Ag21 de la Cultura* sino sólo las áreas temáticas más relevantes desde el punto de vista de las políticas culturales locales, a saber:

- Cultura y desarrollo;
- Transversalidad de las políticas culturales;
- Cultura y participación ciudadana;
- Accesibilidad de la oferta cultural; y
- Papel de las iniciativas relacionadas con la memoria y la innovación en la construcción de la identidad local.

Bajo el auspicio de la *FEMP* y con el apoyo del *Ministerio de Cultura*, a principios de 2010 el grupo de investigación *Métodos Cuantitativos para la Medición de la Cultura (MC2)* de la *Universidad de Valencia* se responsabilizó del diseño de una aplicación tutelada que permitiese la implementación efectiva del sistema de indicadores propuesto a nivel teórico en la guía de evaluación (*FEMP, 2009*). Esta aplicación recibe el nombre de *Siclo (Sistema de indicadores culturales local)* y actualmente se encuentra en funcionamiento en: <http://www.mcmc-uv.es/Siclo>
<http://www.mc2-uv.es/Siclo>

En el siguiente apartado se comentan las principales acciones que se han llevado a cabo hasta hacer operativo *Siclo*.

3. Sistema de indicadores culturales local

Siguiendo la filosofía establecida en su antecesora, la guía de evaluación, el *Siclo* se configura en las siguientes áreas temáticas:

Área A/B. La cultura como factor de desarrollo económico, social y urbano/territorial

La cultura como factor de desarrollo, en el sentido más amplio del término, es la cuestión central que genéricamente originó y articuló el sistema de indicadores propuesto y hacia la que se orientan las restantes cuestiones. Como se ha dicho, el valor social de la actividad cultural responde a su función como elemento de promoción del desarrollo social y económico de las sociedades y, especialmente en su vertiente local, también del desarrollo territorial. El *Siclo* aspira a proporcionar información sobre la trayectoria pasada, situación presente y evolución previsible de la cultura. Al mismo tiempo, se considera relevante definir y cuantificar, en su caso, cuáles son los rendimientos, en términos de desarrollo del tejido cultural local, que se derivan de la inversión en cultura, puesto que dichos rendimientos garantizarán la sostenibilidad futura del propio sistema cultural local.

Área C. Transversalidad de la cultura

La percepción de que la cultura influye en otras políticas locales no relacionadas directamente con la política cultural se ha afianzado en el entorno local debido, fundamentalmente, a la implantación de fórmulas de desarrollo local

sostenible, a la importancia estratégica del cambio cultural, y a la consideración de la cultura como instrumento para la proyección exterior de los territorios, instituciones y organizaciones. El *Siclo* pretende extraer y recopilar información sobre la transversalidad de la cultura desde diversas perspectivas y en distintos niveles. Se aspira a cuantificar la presencia de la cultura en el discurso global de la entidad local, en sus políticas estructurales (social, económica y urbanística) y en ciertas áreas afines (mujer, juventud, participación ciudadana, deportes, turismo) considerando simultáneamente los niveles de transversalidad conceptual, de gestión y de organización.

Área D. Acceso a la cultura

Siclo aportará información sobre el volumen y naturaleza de las intervenciones encaminadas a favorecer la equidad en el acceso a la oferta cultural pública por parte de todos los ciudadanos y, particularmente, sobre el grado de efectividad de dichas intervenciones. Se propone evaluar el alcance de las políticas en términos del grado de desactivación o reducción de las barreras y factores disuasorios que puedan dificultar el acceso de la ciudadanía a la oferta cultural local, sean éstas de índole pecuniaria, de información, de socialización, psicológicas, físicas, cronológicas, de confort o de ausencia de predisposición al hecho cultural.

Área E. Cultura y participación ciudadana

Este ámbito está estrechamente relacionado con el anterior, en el sentido de que las condiciones de accesibilidad delimitan, a su vez, las posibilidades de participación de los ciudadanos en dos dimensiones principales: como usuarios de servicios culturales locales y como configuradores de la oferta cultural local. *Siclo* aspira a proporcionar orientación e información sobre el grado de aplicación y, en especial, sobre la efectividad de las acciones de los gobiernos locales dirigidas a incorporar el mayor número posible de personas a la vida cultural local en las dos vertientes anteriormente enunciadas. Para ello se pretende evaluar la situación actual y la evolución futura tanto de los niveles de utilización de los servicios culturales y de asistencia a las actividades culturales como, especialmente, de la implicación y participación activa de los ciudadanos, bien individual o bien colectivamente, en la definición, configuración y gestión de la actividad cultural local.

Área F. Memoria, innovación e identidad

Este ámbito temático considera, en un contexto crecientemente globalizado, las relaciones entre la memoria colectiva local y la creación contemporánea en el proceso de construcción dinámica de la identidad local. Concretamente, el *Siclo* obtiene y analiza información sobre la situación del patrimonio cultural, tanto en sus aspectos tangibles como intangibles, entendido como la concreción del proceso de recuperación y proyección de la memoria colectiva. De forma análoga, *Siclo* también se propone recabar información para el análisis de la vitalidad de la creación cultural endógena y su difusión, puesto que se trata de un elemento clave que interviene en la construcción de la identidad local.

Atendiendo a esta estructura, el diseño y construcción del *Siclo* arrancó con una propuesta de selección de indicado-

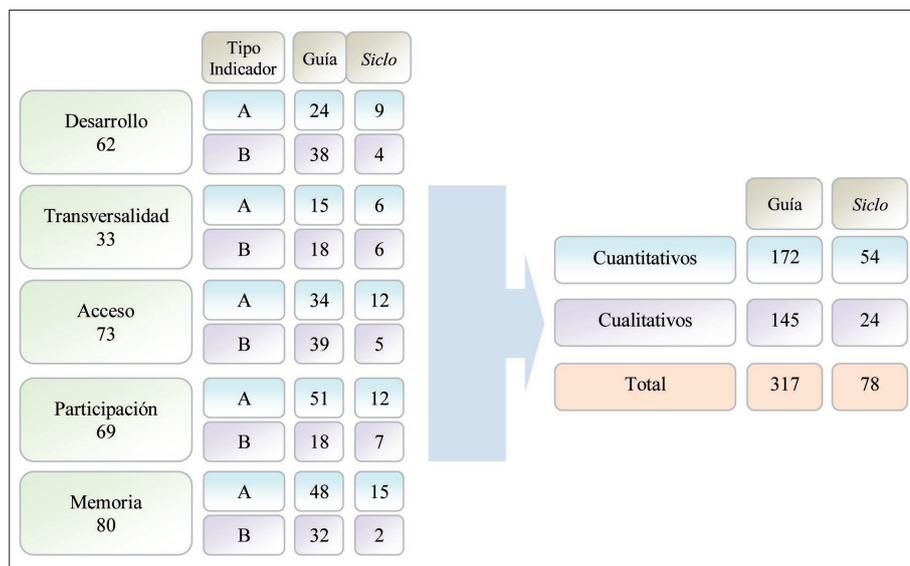


Figura 1. Distribución del número de indicadores. Comparativa guía de evaluación y Siculo. A: Indicadores cuantitativos, B: Indicadores cualitativos

res, puesto que el sistema originalmente estaba formado por un total de 317, lo que resultaba inviable a nivel práctico. Esta selección –basada en el análisis y consenso de un grupo de expertos formado por gestores culturales, académicos y técnicos de cultura tanto de la FEMP como de varios ayuntamientos– condujo a que se retuviese un total de 78 indicadores. La figura 1 muestra la distribución del número de indicadores considerados atendiendo al área de interés y su naturaleza (cuantitativos y cualitativos).

Reducida la dimensión del sistema, se realizó una prueba piloto (pre-test) con el objetivo de contrastar la calidad de los cuestionarios confeccionados para obtener la información requerida, la calidad de los indicadores propuestos y recibir retroalimentación de los responsables de cultura de los gobiernos locales. Para ello se confeccionó un cuestionario telemático que fue cumplimentado por los usuarios potenciales del sistema. En esta fase fueron invitados a participar gobiernos integrantes de la *Comisión de Cultura* de la FEMP en función de su tipo. Se pretendía que los participantes proporcionasen el mayor volumen posible de información para mejorar los contenidos del instrumento y su operatividad. Los gobiernos que finalmente participaron en la aplicación inicial contribuyeron aportando no sólo los datos de que disponían sobre la cultura y la política cultural en la esfera local sino también numerosos comentarios y sugerencias relativos tanto a los cuestionarios como a los mecanismos de cumplimentación y administración de la información resultante. Esta retroalimentación resultó muy valiosa de cara a la calibración final del sistema de indicadores.

Tras la prueba piloto se realizó un taller de trabajo en la FEMP para analizar los resultados obtenidos. Entre las principales conclusiones a las que se llegó cabe destacar:

- que seguía siendo necesario reducir la dimensión del sistema de indicadores;
- y, sobre todo, que dada la gran heterogeneidad de los gobiernos locales en España era necesario modularlo en función de su tipo.

Así, el Siculo se moduló en tres niveles: Básico, Intermedio y Avanzado. El módulo Básico, núcleo central del sistema, se

encuentra integrado por un número limitado de indicadores básicos aplicables a todos los gobiernos locales. De cara al futuro será posible delimitar módulos adicionales (Intermedio y Avanzado) constituidos por grupos de indicadores más específicos que sean de aplicación sólo a determinados municipios en función de determinadas características, por ejemplo de su tamaño.

Actualmente se encuentra operativo el módulo básico del Siculo, que está constituido por un total de 18 indicadores cualitativos y 28 cuantitativos. Debido a restricciones de espacio no es posible abordar con detalle todas las áreas temáticas y todas las cuestiones que en ellas se

plantearon a los responsables de las políticas culturales de los gobiernos locales para recabar información. Por esta razón, a modo de ilustración se considera en este artículo sólo el área de Cultura y participación ciudadana (Participación).

El módulo básico de Siculo en el área de Participación está constituido por 10 indicadores: 3 cualitativos (codificados como EQ), que se evalúan en una escala de 1 a 5, y 7 indicadores cuantitativos (codificados como EN). En esta área se pretende obtener información sobre el uso que los ciudadanos hacen de la oferta cultural promovida por los ayuntamientos así como sobre la actuación de éstos en el fomento de una mayor diversidad de agentes activos (configuradores y gestores de oferta cultural) en la vida cultural local.

Para conocer la demanda efectiva de la ciudadanía con respecto a la oferta cultural promovida por el gobierno local, se pregunta:

- el número de equipamientos culturales, o si existe algún inventario de ellos, distinguiendo los de titularidad o gestión municipal (EQ1a) del resto (EQ1b);
- si existe un sistema integral de recogida de información de la oferta global de actividades y servicios culturales (EQ2);
- si hay canales para que los ciudadanos puedan participar y valorar el diseño de la política cultural (EQ6).

En el aspecto cuantitativo se requiere información, por el lado de la oferta, de:

- butacas/plazas en teatros, auditorios, etc., que son de titularidad, o son gestionados, por el ayuntamiento (EN1);
- metros cuadrados dedicados a actividades de formación cultural y prácticas artísticas (EN2); y
- puestos de lectura/consulta en bibliotecas (EN3).

Y por el lado de la demanda se solicita:

- número de asistentes a las actividades (distinguiendo entre conciertos de música en vivo, representaciones teatrales, danza, etc.), tanto directamente programadas por el gobierno local (EN4a) como de las que reciben su apoyo (EN4b);

- distribución porcentual de asistentes a los actos escénicos programados o apoyados por la entidad local en 1 año por segmento de público (Niños, Jóvenes, Adultos, Personas mayores) y sexo (EN5);
- usuarios del servicio de bibliotecas en 1 año según sexo (EN6); y
- visitantes a museos de titularidad o gestión del ayuntamiento (EN7).

A partir de los datos suministrados voluntariamente por los municipios se procede a la generación de las variables que componen la base de datos general del sistema y la propia de cada usuario. La base de datos general proporciona diversas variables construidas a partir de las respuestas emitidas por los gobiernos locales y constituye el núcleo informativo para el análisis estadístico de los datos y para el proceso de síntesis de la información en un conjunto de indicadores multidimensionales. Los resultados que se derivan del tratamiento de los datos contenidos en el *Siclo* son facilitados a los usuarios a través de una herramienta diseñada especialmente para ello y que se conoce como *Báculo* (*Barómetro cultural local*).
<http://www.mc2-uv.es/Baculo>

4. Consideraciones finales

Las ciudades son los principales espacios de consumo cultural y son los gobiernos locales quienes planifican la política cultural, efectuando el mayor gasto en cultura de la administración pública. Su acción tiene que estar basada en información y análisis y no cabe la toma de decisiones por intuición u olfato político. El *Sistema de indicadores culturales local (Siclo)* es una aplicación que mejora el conocimiento de la acción cultural en el ámbito de la administración local, lo que facilita la evaluación de las correspondientes políticas. *Siclo* es el resultado de implementar en la práctica el sistema de indicadores conceptualizado en la *Guía para la evaluación de las políticas culturales locales (FEMP, 2009)*.

De cara al futuro inmediato los principales objetivos que se plantean son incrementar consistentemente el número de ayuntamiento usuarios que alimenten con información el *Siclo* y mantener el sistema activo en años sucesivos, realizando periódicamente acciones formativas sobre su utilización y potencialidades. En la medida en que se avance en la consecución de estos objetivos será posible generar una secuencia temporal de resultados para cada gobierno local. Tal secuencia constituirá un instrumento fundamental que permitirá a los responsables de las áreas de cultura tanto conocer su posición en relación al estado general de

la cultura y de la política cultural, como analizar su trayectoria.

5. Agradecimientos

Los autores agradecen la financiación recibida de la *Federación Española de Municipios y Provincias* así como la colaboración del *Ministerio de Cultura* de España.

6. Referencias

- Carrasco-Arroyo, Salvador.** "Medir la cultura: una tarea inacabada". *Periférica*, 2006, n. 7.
<http://www.uv.es/carrascscs/PDF/medir%20la%20cultura.pdf>
- Ciudades y Gobiernos Locales Unidos (UCLG). *Agenda 21 for culture*. United Cities and Local Governments, Committee on culture, 2004.
<http://agenda21culture.net>
- Federación Española de Municipios y Provincias (FEMP). *Sistema de indicadores para la evaluación de las políticas culturales locales en el marco de la Agenda 21 de la cultura*. Madrid, 24 abril 2007.
http://descarga.sarc.es/Curso14/pdf_c/4/4.pdf
- Federación Española de Municipios y Provincias (FEMP). *Guía para la evaluación de políticas culturales locales*. Madrid: FEMP, 2009.
<http://www.FEMP.es/files/11-367-fichero/Guia%20Políticas%20Culturales.pdf>
- Pascual, Jordi.** *Políticas culturales para la diversidad cultural*. División de las Políticas Culturales y del Diálogo Intercultural de la Unesco, 2006.
http://www.cities-localgovernments.org/uclg/upload/newTempDoc/ES_333_report_ex_es.pdf
- Rausell, Pau.** "Economía de la cultura". *Treballs de comunicació*, 2008, n. 2, pp. 243-256.

AGENDA <http://www.elprofesionaldelainformacion.com/agenda.html>

8 de febrero de 2012

GESTIÓN DE CONTENIDOS 2012. Gestión como valor de negocio

Madrid, Hotel Palace

IDC Iberia, tel.: +34-917 872 150

<http://www.idcspain.com/events/index.jsp>

9 de febrero de 2012

DIGITAL STRATEGY

Londres

Association of Learned and Professional Society Publishers (Alpsp)

<http://www.alpsp.org>

9-19 de febrero de 2012

EVENTO CIENTÍFICO BIBLIOTECOLÓGICO 2012 (durante XXI Feria del Libro)

La Habana

Asociación Cubana de Bibliotecarios (ASCUBI) y la Biblioteca Nacional de Cuba José Martí

esetien@infomed.sld.cu, relacint@bnjm.cu

14 de febrero de 2012

THE VALUE OF UNIQUE SCHOLARLY IDENTIFIERS TO ACADEMICS, INSTITUTIONS AND COUNTRIES

Vilnius, Lituania

<http://uniqueids.org>

22 de febrero de 2012

GESTIÓN DE CONTENIDOS 2012. Gestión como valor de negocio

Barcelona, Hotel NH Constanza

IDC Iberia, tel.: +34-917 872 150

<http://www.idcspain.com/events/index.jsp>

28-29 de febrero de 2012

REUNIÓN ANUAL DE OCLC PARA EUROPA, ORIENTE MEDIO Y ÁFRICA (EMEA). "Developing a new blend of library"

Birmingham, Reino Unido

<http://www.oclc.org/uk/en/councils/emea/meetings/2012/annual/agenda/default.htm>

1-2 de marzo de 2012

EDGE2012. Pushing the boundaries of public service delivery

Edinburgh

Edinburgh City Libraries

<http://edgeconference.co.uk>

3-5 de marzo de 2012

ICSTI 2012 ANNUAL MEMBERS' MEETING AND WORKSHOP

Paris

International Council for Scientific and Technical Information (ICSTI)

<http://www.icsti.org/spip.php?rubrique49>

10-13 de marzo de 2012

IADIS INTL CONF. e-Society 2012

Berlin

<http://www.esociety-conf.org>

18 de marzo de 2012

TOOLS OF CHANGE (TOC). e-Story telling

Bologna

O'Reilly Media

<http://www.tocbologna.com>

21-23 de marzo de 2012

COMPUTERS IN LIBRARIES

Washington, DC

Information Today, Inc.

<http://www.infotoday.com/cil2012>

22-23 de marzo de 2012

DGI-CONFERENCE 2012. SOCIAL MEDIA AND WEB SCIENCE. The Web as a living space

Düsseldorf

German Society of Information Science and Information Practice.

http://www.dgi-info.de/CfP2012DGI-Konferenz_GB.aspx

25-27 de marzo de 2012

EUROCP. Policies for the future Internet

Gante, Bélgica

European Information Society Policy, IBBT, Research Foundation of Flanders

<http://www.eurocpr.org>

26-28 de marzo de 2012

UKSG 2012 ANNUAL CONF

Glasgow

United Kingdom Serials Group (UKSG)

<http://www.uksg.org/event/conference12/cfp>

28-30 de marzo de 2012

SPECIALIZED INFORMATION PUBLISHERS ASSOCIATION

Munich

Specialized Information Publishers Association (SIPA)

<http://sipaonline.com/event?eventId=85>

29 de marzo de 2011

JORNADA DE SOFTWARE DOCUMENTAL LIBRE. Modelos de negocio y de gestión de contenidos

Madrid

Baratz Servicios de Teledocumentación SA; El profesional de la información (EPI)

usua.beitia@baratz.es

29-30 de marzo de 2012

X JORNADAS APDIS. As bibliotecas das ciências da saúde na era da literacia digital

Lisboa

Associação Nacional das Farmácias (ANF), Associação Portuguesa de Documentação e Informação de Saúde (APDIS)

<http://apdis.pt>

1-5 de abril de 2012

34TH EUROPEAN CONF ON INFORMATION RETRIEVAL (ECIR)
Barcelona
British Computer Society, Information Retrieval Specialist Group
<http://ecir2012.upf.edu>

11-13 de abril de 2012

LIBRARIANS' INFORMATION LITERACY ANNUAL CONF (LILAC 2012)
Glasgow
Cilip Information Literacy Group
<http://lilacconference.com/WP/>

11-13 de abril de 2012

IFLA INTL NEWSPAPER CONF 2012: Newspaper digitization and preservation: new prospects. stakeholders, practices, users and business models
Paris
IFLA
<http://www.ifla.org/en/events/ifla-international-newspaper-conference-2012>

16-17 de abril de 2012

INTL INFORMATION CONF ON SEARCH, DATA MINING AND VISUALIZATION (II-SDV)
Niza
Dr Haxel CEM and Infonortics, Ltd.
<http://www.ii-sdv.com/home-page.html>

16-17 de abril de 2012

3RD INTL CONF ON INFORMATION MANAGEMENT AND EVALUATION (ICIME)
Ankara
Academic Conferences International
<http://academic-conferences.org/icime/icime2012/icime12-home.htm>

16-18 de abril de 2012

Digital Media Europe 2012
London
World Association of Newspapers (WAN-IFRA)
<http://www.wan-ifra.org/events/digital-media-europe-2012>

16-20 de abril de 2012

XII CONGRESO INTL DE INFORMACIÓN (INFO'2012)
IX Taller intl sobre inteligencia empresarial y gestión del conocimiento en la empresa (Intempres'2012)
La Habana
<http://www.congreso-info.cu>

16-20 de abril de 2012

21ST INTL WWW CONF
Lyon
International World Wide Web Conferences Steering Committee (IW3C2)
<http://www.conference.org/www2012>

18-21 de abril de 2012

WEBIST 2012. 8th Intl conf on web information systems and technologies
Oporto
<http://www.webist.org>

3-4 de mayo de 2012

INTL CONF FOR E-DEMOCRACY AND OPEN GOVERNMENT (CeDEM12)
Krems, Austria
<http://www.donau-uni.ac.at/en/department/gpa/telematik/edemocracy-conference/edem/vid/16842>

9-11 de mayo de 2012

VI CONGRESO DE ARCHIVOS. El derecho a saber y el deber de la privacidad: el acceso a los documentos
Valladolid
Asociación de Archiveros de Castilla y León (ACAL)
Luis Hernández-Olivera, hernandezolivera@gmail.com
<http://congresoacal.es>

10 de mayo de 2012

2ª CONFERENCIA SOBRE CALIDAD DE REVISTAS DE CIENCIAS SOCIALES Y HUMANIDADES (CRECS 2012)
Valencia
El profesional de la información (EPI); Universidad Politécnica de Valencia (UPV)
<http://www.thinkepi.net/crecs2012>

16-18 de mayo de 2012

6TH INTL CONF ON RESEARCH CHALLENGES IN INFORMATION SCIENCE
Valencia
IEEE Spain Section
<http://rcis-conf.com/rcis2012>

16-18 de mayo de 2012

EUROPEAN LIBRARY AUTOMATION GROUP (ELAG) CONF
Palma de Mallorca
<http://www.elag.org>

21-23 de mayo de 2012

INTL CONF ON INFORMATION RESOURCES MANAGEMENT
Viena
Association for Information Systems (AIS)
<http://conf-irm2012.wu.ac.at>

22-25 de mayo de 2012

4TH QUALITATIVE AND QUANTITATIVE METHODS IN LIBRARIES INTL CONF (QQML2012)
Limerick, Irlanda
<http://www.isast.org>

23-25 de mayo de 2012

V JORNADAS OS-REPOSITORIOS. Motricidad de los repositorios de acceso abierto
Bilbao
ETSI Industriales y de Telecomunicaciones, Universidad del País Vasco
<http://www.accesoabierto.net>

Deseo recibir todos los números de la revista EPI a partir del mes de enero del año 20

Suscripción Institucional Personal

Nombre Institución

(Los suscriptores individuales no han de escribir ningún nombre de institución, sólo indicar la dirección particular)

Departamento NIF institucional

Dirección Código postal Ciudad País

Teléfono Fax Correo-e

Método de pago

Tarjeta de crédito VISA Master Card American Express

Titular de la tarjeta Código de seguridad CVC2

Número de tarjeta Caducidad (mm/aaaa)

Cheque nominativo en euros a nombre de El profesional de la información

Transferencia bancaria a la cuenta de La Caixa 2100 0818 93 0200745544 **Enviar, fotocopiado o escaneado, el resguardo de la transferencia.**

Las transferencias desde fuera de España deben hacerse a **IBAN ES95 2100 0818 9302 0074 5544 - BIC/Código Swift CAIXESBBXXX**

Giro postal al apartado de correos 32.280 de Barcelona **Enviar, fotocopiado o escaneado, el resguardo del giro.**

Domiciliación en cuenta bancaria

Entidad Oficina DC Núm

Titular de la cuenta

Enviar el boletín relleno por correo electrónico, o postal al APARTADO 32.280 - 08080 BARCELONA - ESPAÑA.

Consultas: suscripciones@elprofesionaldeinformacion.com o +34 609 352 954

Precios 2012

Suscripción normal: 186,54 € + IVA = 194 €

Suscripción personal: (sólo a domicilios particulares) 86,54 € + IVA = 90 €

Gastos de envío fuera de España: Europa: 40 € Américas: 60 €

Suscripción sólo online: 95,19 € + IVA = 112,32 €

Número suelto: 31,74 € + IVA = 33 €

Gastos de envío fuera de España: Europa: 9 € Américas: 15 €

Las suscripciones van por años naturales, de enero a diciembre

NORMAS PARA LOS AUTORES

El profesional de la información tiene dos secciones principales:

ARTÍCULOS: Trabajos de investigación y temas analizados en profundidad.

ANÁLISIS: Experiencias, estudios de casos, análisis de productos, reseñas, etc.

Las contribuciones han de ser originales e inéditas, no pueden haberse publicado previamente en soporte papel o electrónico. El tamaño ideal es de 3.500 palabras, aunque en algunos casos la Redacción puede autorizar una mayor extensión.

El texto ha de enviarse en Word, rtf u odt. Las tablas deberán ir pegadas en el mismo documento. Todos los materiales gráficos (diagramas, fotografías, capturas de pantalla, etc.) deben pegarse en el Word y además enviarse en ficheros independientes (en formatos xls, jpg, pdf, etc.). Las imágenes jpg deben tener una resolución de al menos 300 pp (unos 200 KB cada una).

El texto debe presentarse completamente plano, sin autoformatos ni automatismos de Word (subsecciones, viñetas, citas enlazadas, pies de página, sangrías, tabulaciones, colores, etc.), pero debe seguir el estilo de EPI en cuanto a **negritas** (nombres de los autores citados), *cursivas* (instituciones, títulos de revista, marcas) y mayúsculas. Los urls deben estar sin hipervínculo.

Las citas bibliográficas en el texto se realizarán de la forma: (**Apellido**, año) o (**ApellidoAutor1**; **ApellidoAutor2**, año).

La redacción debe ser concisa y precisa, evitando la retórica.

Los trabajos deben incluir: a) título, b) resumen de 100-150 palabras, c) 5-10 palabras clave, d) title, e) abstract de 100-150 palabras, f) 5-10 keywords.

Aparte se incluirá el nombre de los autores, su lugar de trabajo y dirección (postal y electrónica), su foto tipo carnet superior a 100 KB en jpg, y un currículum de unas 70 palabras.

EVALUACIÓN

Los trabajos son revisados según el sistema tradicional "peer review" en doble ciego por al menos dos expertos en el tema, del Consejo Asesor de la revista y/o externos. La revista se compromete a informar del resultado a los autores.

REFERENCIAS BIBLIOGRÁFICAS

Ordenadas alfabéticamente por autor, se limitarán a las obras citadas en el artículo. No se acepta bibliografía de relleno.

Artículos de una publicación periódica:

Apellido, Nombre; Apellido2, Nombre2. "Título del artículo".

Título de la publicación periódica, año, mes, v., n., pp. xx-yy. DOI

Dirección url iniciada en nueva línea.

Ponencia presentada en un congreso:

Apellido, Nombre; Apellido2, Nombre2. "Título de ponencia".

En: *Nombre del congreso*, año, pp. xx-yy. DOI

Dirección url iniciada en nueva línea.

Monografías e informes:

Apellido, Nombre; Apellido2, Nombre2. *Título del trabajo.*

Lugar de publicación: editor, fecha, ISBN. DOI.

Dirección url iniciada en nueva línea.

Capítulo de una monografía:

Apellido, Nombre; Apellido2, Nombre2. "Título del capítulo".

En: *Apellido, Nombre; Apellido2, Nombre2. Título de la monografía.*

Lugar de publicación: editor, fecha, pp. xx-yy. ISBN.

Dirección url iniciada en nueva línea.

Recurso en línea:

Apellido, Nombre; Apellido2, Nombre2. *Título del recurso.*

Dirección url iniciada en nueva línea.

Todas las contribuciones se tienen que enviar a la sección EPI de la plataforma OJS del Repositorio Español de Ciencia y Tecnología (Recyt) de la Fecyt: <http://recyt.fecyt.es/index.php/EPI/index>

Previamente los autores deben registrarse en:

<http://recyt.fecyt.es/index.php/EPI/user/registerJournal>