

El papel de las bases de datos institucionales en el análisis de la actividad científica de las universidades

Daniela de Filippo*, Elías Sanz-Casado*, Cristóbal Urbano Salido**,
Jordi Ardanuy**, Isabel Gómez-Caridad***

Resumen: Ante las dificultades que tienen las universidades para conocer con precisión la actividad científica que realizan, debido fundamentalmente a la ausencia de fuentes de información nacionales e internacionales que reúnan de manera exhaustiva toda su producción investigadora, se propone la utilización de las bases de datos institucionales producidas por la propia universidad. Para ello, se ha realizado una amplia encuesta a las universidades españolas, con el fin de conocer la existencia o no de estas bases de datos en cada una de ellas, así como determinar su interés y utilidad en aquellas universidades que ya las tenían. Asimismo, se realizó un análisis en profundidad de la información contenida en las bases de datos institucionales de dos universidades, la Universitat de Barcelona y la Universidad Carlos III de Madrid. Los resultados obtenidos han permitido conocer la situación de este tipo de recurso de información en el sistema universitario español, así como la capacidad que brinda para determinar con precisión las características de la investigación en las universidades y la posibilidad de realizar perfiles científicos de cada una de ellas. Finalmente se aporta una serie de recomendaciones para la mejora de este recurso institucional.

Palabras clave: bases de datos institucionales, evaluación científica, universidades, indicadores.

The role of institutional databases in the analysis of universities' scientific activity

Abstract: *Given the difficulties faced by universities for precisely gauging their scientific activity, due mainly to the lack of comprehensive national and international information sources containing all their research output, we propose the use of the universities' own institutional databases. An extensive survey of Spanish universities was performed, in order to discover which universities have these databases and to determine their potential interest and usefulness. Also, an in-depth analysis was performed of the information*

* Departamento de Biblioteconomía y Documentación. Laboratorio de Estudios Métricos de la Información (LEMI). Universidad Carlos III de Madrid. Unidad Asociada IEDCYT-LEMI. CSIC-Universidad Carlos III de Madrid. Correo-e: elias@bib.uc3m.es; dfilippo@bib.uc3m.es.

** Departament de Biblioteconomia i Documentació. Universitat de Barcelona. Correo-e: urbano@ub.edu; jordi_ardanuy@ub.edu.

*** Instituto de Estudios Documentales sobre Ciencia y Tecnología (IEDCYT). (CSIC). Unidad Asociada IEDCYT-LEMI. Correo-e: isabel.gomez@cchs.csic.es.

Recibido: 23-09-2010; aceptado: 03-12-2010.

contained in the databases of two universities, the University of Barcelona and the Carlos III University of Madrid. The results obtained reveal the state of these information resources in the Spanish university system as well as their ability to accurately determine the characteristics of research carried out in the universities and to prepare scientific profiles of each. Finally we provide several recommendations for improving this institutional resource.

Keywords: *Institutional databases, scientific evaluation, universities, indicators*

1. Introducción

En general las publicaciones constituyen una de las principales vías de comunicación para transmitir los resultados de la investigación. En el ámbito de las universidades, además, existen otros mecanismos para difundir el conocimiento, como son la enseñanza formal, la movilidad de alumnos y docentes y la producción de resultados tecnológicos. El desarrollo de los medios de comunicación electrónica también ha contribuido a la difusión de conocimiento, mejorando la accesibilidad a diferentes archivos y documentos y promoviendo la visibilidad de los resultados de investigación.

Las bases de datos bibliográficas internacionales, como el *Web of Knowledge* (WOK) y más recientemente Scopus, han sido ampliamente usadas como fuentes para analizar los resultados científicos de las Universidades, evaluar su productividad y producir indicadores bibliométricos (van Raan, 2004; Bonaccorsi y otros, 2007). Estas bases de datos se han mostrado eficaces en el análisis de publicaciones académicas destinadas a una audiencia internacional, principalmente de habla anglófona (Laredo, 2003) y en el estudio de la producción de grandes unidades como países, instituciones y áreas de investigación. Asimismo, la inclusión de indicadores de citación las ha posicionado como instrumentos poderosos para trazar un mapa de la ciencia, su evolución y dinámica (van den Besselaar y otros, 2007). A pesar de estas ventajas, su utilización ha sido criticada con frecuencia debido a las limitaciones que presentan en cuanto a sesgo temático a favor de la ciencia básica, lingüístico hacia las publicaciones en lengua inglesa y a las dificultades para los estudios a nivel micro. Estudios previos también subrayan la cobertura limitada de publicaciones que no sean artículos de revista, como libros, informes, comunicaciones, tesis, así como de publicaciones en lenguas nacionales diferentes del inglés (Hicks, 2004). Algunos de estos tipos de publicaciones, aunque no inciden en la visibilidad internacional de una universidad, tienen gran importancia en el caso de instituciones con orientación regional o nacional y constituyen una importante vía de interacción de la academia con la administración y las PYME.

Las limitaciones de las bases de datos bibliográficas internacionales resultan mucho más evidentes cuando se utilizan estas fuentes para el análisis bibliométrico en las Ciencias Sociales y Humanidades (Gómez y otros, 2009; Hicks, 2004; Nederhof, 2006; Iribarren y otros, 2009). Aun cuando estas áreas están integradas

por un conjunto heterogéneo de disciplinas, pueden ser caracterizadas por algunos rasgos comunes, como las preferencias por la difusión de los resultados científicos en ciertos tipos documentales —distintos a los artículos de revista— principalmente publicados en lenguas nacionales y con una orientación más local.

Con el objetivo de conseguir una visión más completa de la actividad científica de las universidades, se ha explorado el empleo de otras fuentes para desarrollar indicadores capaces de caracterizar sus resultados de investigación como son: bases de datos de currícula, catálogos de acceso público en línea (OPAC) y otras bases de datos públicas, como Google, sistemas de archivo abiertos o bases de datos disciplinares gratuitas como Medline (Cañibano y Bozeman, 2009; Lepori y otros, 2008; Torres Salinas y Moed, 2009). Otra iniciativa interesante son los repositorios institucionales en acceso abierto (Open Access) creados por universidades en el marco de la Declaración de Berlín. El objetivo principal de estos archivos es permitir la difusión y la disponibilidad de producción científica de interés, tanto para el colectivo científico como para un público más amplio. Estas bases de datos incluyen diversos tipos de publicaciones: artículos, capítulos de libro, conferencias, documentos de trabajo, tesis doctorales y otros tipos documentales denominados globalmente «literatura gris». Las universidades han ido implantando estos repositorios de acceso abierto, pero todavía están en una fase inicial, con frecuencia con escaso contenido, en parte debido a problemas de *copyright* y al desconocimiento que aún tiene el profesorado sobre el interés de estos repositorios para aumentar la visibilidad de su investigación.

Finalmente, también las bases de datos de gestión interna y evaluación de las actividades de investigación de las instituciones de enseñanza superior son importantes fuentes para evaluar las publicaciones. En España fue la Universitat de Barcelona la primera en crear una completa base de datos institucional con un primer objetivo de gestión interna (Bellavista y otros, 1992). Actualmente, una gran parte de las universidades españolas disponen de bases de datos internas más o menos desarrolladas, que han adoptado distintas estructuras. Destacan, entre otros, el sistema GREC, desarrollado por la Universitat de Barcelona y que siguen otras universidades catalanas; UNIVERSITAS XXI, empleado en la Universidad Carlos III de Madrid así como en otras 10 universidades españolas; Fénix utilizado por la Universidad Politécnica de Catalunya y otras universidades catalanas que no usan GREC; y el Sistema de Información Científica de Andalucía (SICA), desarrollado por la Consejería de Ciencia, Investigación y Empresa de la Junta de Andalucía para regular la actividad investigadora de las instituciones y grupos de investigación.

Las bases de datos institucionales han sido concebidas para dar respuesta a diferentes situaciones, como son: la mejora de los procesos de evaluación institucionales, la recogida de datos completos sobre publicaciones, la supervisión interna de las actividades de los Departamentos, el aumento de la difusión y visibilidad del conocimiento producido por la universidad. Aunque pueden presentar limitaciones debido al intenso debate que generan temas relacionados con

la propiedad intelectual de los productos, sin embargo contienen información muy exhaustiva sobre las actividades de publicación de la universidad y a menudo estos datos están vinculados con otros sobre personal, financiación, proyectos de investigación, etc.

Todos los actores del proceso de investigación y sus productos, es decir artículos, libros, informes, conferencias, tesis, y la diversidad de comunidades de investigación dentro de las instituciones académicas, sus comportamientos y prácticas, pueden ser «representados» a través de los archivos institucionales. Además, las bases de datos institucionales pueden hacer visible la actividad científica que va más allá de la publicación tradicional, y que refleja la transferencia de conocimiento a la sociedad. La difusión de productos como software que está en la frontera entre publicaciones y resultados tecnológicos; u otros como artículos de prensa, documentos de trabajo, cartas, material educativo, sólo son posibles a través de estas fuentes, que se presentan como poderosas herramientas para transmitir los resultados científicos producidos en las instituciones de enseñanza superior.

Teniendo en cuenta la importancia del uso de las bases de datos institucionales de las universidades como fuente para la obtención de indicadores de una actividad científica amplia y variada, recientemente se ha desarrollado el proyecto europeo UNIPUB (*Methodologies for the characterization of the publication output of higher education institutions using institutional databases*), dentro de la red de excelencia PRIME (*Infrastructures on Policies for Research and Innovation in the Move towards the European Research Area*). El objetivo de dicho proyecto es valorar la utilidad de las bases de datos institucionales para la construcción de perfiles de actividad de las diferentes universidades, tanto a nivel global de la propia institución como por área científica. No se trata de evaluar u obtener indicadores de «resultados» sino intentar, a través de estas bases de datos, recoger información que permita conocer otros aspectos de la actividad de las universidades, como la transferencia de conocimiento a la sociedad, su orientación local o internacional, el tipo de conocimiento producido (básico, aplicado, tecnológico, educativo, divulgativo) la colaboración con diferentes actores y el uso que hacen de esta información diferentes usuarios: políticos, gestores, docentes-investigadores, y público en general.

El equipo del proyecto UNIPUB ha estado compuesto por 5 grupos de investigación: CERIS del Consiglio Nazionale di Ricerca (Italia), la Università della Svizzera Italiana (Suiza), el IEDCYT, del Consejo Superior de Investigaciones Científicas (España) y las universidades Carlos III de Madrid y Universitat de Barcelona, ambas también españolas.

El estudio empírico se desarrolló seleccionando para el análisis las bases de datos utilizadas por dos universidades de cada uno de los países participantes: por España la Universidad Carlos III de Madrid y la Universitat de Barcelona, por Italia la Università degli Studi di Milano Bicocca y la Università degli Studi di Napoli Federico II, y por Suiza la Università della Svizzera Italiana y la Universidad de Zurich (Reale y otros, 2009a). En este trabajo presentamos una parte de los resultados correspondientes a la participación española.

2. Objetivos

Los objetivos que se persiguen en este estudio son, por un lado, determinar la importancia, características y uso de las bases de datos institucionales de las universidades españolas a través de una encuesta enviada a los responsables de las OTRI (Oficinas de Transferencia de los Resultados de la Investigación), y por otro, conocer los indicadores de actividad científica que se pueden obtener de estas bases de datos, tanto a nivel general de la universidad como por áreas temáticas, que permitan identificar sus perfiles de actividad de acuerdo con sus misiones, objetivos y orientación regional/nacional/internacional. Estos indicadores se calcularán sobre dos universidades, la Universitat de Barcelona (UB) y la Universidad Carlos III de Madrid (UC3M).

3. Metodología

3.1. Encuesta a universidades

Se redactó un cuestionario con el que se pretendía conocer, en primer lugar, la existencia o no de base de datos (BD) institucional de actividades de investigación realizada en las universidades, así como su origen, objetivos, usuarios, información que contiene, y también sus limitaciones. Dicha encuesta (en anexo) se envió a las 48 universidades públicas españolas, y concretamente a las OTRI de cada universidad, entre los meses de mayo y junio de 2009, repitiéndose el envío en septiembre. La tasa de respuesta obtenida fue del 38%.

3.2. Aspectos de la investigación que se pueden detectar a través de las bases de datos institucionales

Para el estudio en profundidad de las bases de datos institucionales se eligieron dos universidades de diferentes características: una con una trayectoria centenaria, grande y generalista, la Universitat de Barcelona (UB), y otra con 20 años de existencia, pequeña y especializada en Ciencias Sociales, Humanas e Ingeniería, la Universidad Carlos III de Madrid (UC3M). Ambas universidades emplean respectivamente GREC y UNIVERSITAS XXI para gestionar sus bases de datos institucionales. La buena cobertura de los datos relacionados con las actividades de investigación (más del 70% en ambos casos, según la información aportada por los responsables de la OTRI de cada universidad), así como la actualización permanente de la información, han sido características muy valiosas a la hora de seleccionar las unidades de análisis.

GREC es una herramienta informática desarrollada por la UB que consta de diversas bases de datos (CV, proyectos, publicaciones). Se inició en 1987 como plataforma para la gestión, seguimiento y evaluación de la actividad científica

(Bellavista y otros, 1992). **UNIVERSITAS XXI** es una plataforma para la gestión de recursos y actividades de la universidad, creada por OCU (Oficina de Cooperación Universitaria) y estructurada en 10 módulos, uno de ellos, «Universitas XXI Investigación», está dedicado a recoger toda la investigación realizada por el profesorado de la Universidad. En ambos casos, los resultados de la actividad investigadora los introducen los profesores, pero mientras que en el caso de la UB es la OTRI la responsable de supervisar el contenido, en la UC3M lo validan los departamentos. Existen diversos niveles de acceso: el del profesorado, que accede a sus propios datos; los gestores, normalmente de la OTRI, que acceden a toda la información para su análisis; los administradores (el equipo informático) que tiene todo el acceso, puede modificar y analizar resultados. Los usuarios externos sólo acceden a parte de la información, con frecuencia presentada como Memoria Anual, que recoge las principales actividades científicas.

Ambas bases de datos contienen *inputs* y *outputs* de investigación acerca de: artículos de revista, libros, capítulos de libro, comunicaciones a congresos, tesis, documentos de trabajo, actividades de difusión, movilidad, proyectos y financiación, patentes, cursos, etc. En este momento no se dispone del texto completo de las publicaciones, aunque se prevé en un futuro inmediato enlazarlas con un repositorio institucional. Se recogen los nombres de todos los autores, y se relacionan con bases internas de gestión de recursos humanos. En el caso de autores externos no figuran sus centros de trabajo, lo cual impide el estudio de redes de colaboración.

En este estudio se trabaja sobre el período 2005-2006. Las bases de datos proporcionan la información a nivel de departamento, pero se han reagrupado los departamentos en 10 áreas temáticas de acuerdo con la clasificación de la OST (Rapport de l'OST 2008) para la presentación de los indicadores. Dadas las diferencias de tamaño entre ambas universidades estudiadas, se analiza la población total de UC3M (26 departamentos) y una selección de 29 departamentos de los 106 de la UB, que representan a todas las áreas científicas OST. Si bien es cierto que dentro de cada departamento pueden existir grupos cuya actividad científica puede diferir de la del departamento en general, en este trabajo se ha seguido este criterio para poder conocer las «tendencias» por grandes áreas temáticas y poder hacer estudios comparativos, evitando así un estudio micro que se alejara de los objetivos del presente proyecto.

Se obtuvieron los siguientes indicadores:

1. Producción científica por tipo documental y área, basada en 5 tipos de documentos: artículos, libros, capítulos de libro, contribuciones a congresos y otros (incluye tesis, documentos de trabajo, productos multimedia).
2. Productividad. Producción científica total por profesor (EJC) y área temática.
3. Orientación nacional/internacional de la producción científica (excluyendo «otros»). Se tiene en cuenta el idioma, país de edición y ámbito nacional/internacional de la publicación. Todo ello se realizó por área científica.

4. Tamaño de los equipos de investigación. Colaboración entre autores por área científica. Se cuantifica el número de autores que firman las publicaciones.
5. Comunidad de referencia a la que se dirige la publicación o principal grupo de usuarios: orientación académica o profesional. Se asignan a la comunidad científica los artículos, libros, capítulos de libro y tesis; a la comunidad profesional o a la sociedad las comunicaciones a congresos, documentos de trabajo y artículos en la prensa. En este sentido, y respecto a la asignación de los congresos al ámbito profesional, somos conscientes de la distinta tipología de los mismos, y que un determinado porcentaje de ellos estaría más vinculado con la orientación académica. Una asignación más precisa requeriría realizar la clasificación de las publicaciones de manera individual.

Mediante el apoyo de herramientas externas a las bases de datos institucionales, se ha tratado de determinar para los artículos de revista dos características importantes: su carácter básico o aplicado, y si tienen o no evaluación por pares. Para el primer caso se ha empleado una clasificación de las revistas del *Science Citation Index (SCI)* y *Social Sciences Citation Index (SSCI)* realizada por la *National Science Foundation* (iPIQ, 2008) que distribuye las revistas en 4 niveles (del 1 al 4), siendo 1 muy aplicado y 4 muy básico.

Para determinar si se trata de revistas que utilizan el filtro de la evaluación por pares como medida de control de calidad no se dispone de información fiable en las bases de datos institucionales. Como primera aproximación se empleó el *Journal Citation Reports (JCR)* de Thomson Reuters elaborado a partir del SCI y del SSCI, considerando que todas las revistas que analizan están empleando el *peer review*.

4. Resultados

4.1. Resultados de la encuesta

Se enviaron 48 cuestionarios y se obtuvo una tasa de respuesta del 38%. Entre las universidades que respondieron las había generalistas, especializadas y politécnicas, con diferentes sistemas de gestión de sus BD institucionales. Sin embargo, la distribución geográfica de las 19 universidades que respondieron se concentró fundamentalmente en Cataluña y Madrid; de Andalucía sólo se obtuvo respuesta de un único centro y ninguna de Aragón (figura 1).

A la pregunta de si dispone de base de datos institucional, la inmensa mayoría respondió afirmativamente, sólo en un caso la respuesta fue negativa. El origen de dichas bases de datos fue en más del 80% una iniciativa interna de la institución, y su objetivo principal la gestión y evaluación de las actividades de investigación. La base de datos más antigua la implantó la Universitat de Barce-

lona en 1987, mientras que la mitad de las universidades que contestaron la encuesta las desarrollaron con posterioridad al año 2003. Los sistemas de gestión empleados son, en su mayoría, GREC y UNIVERSITAS XXI. A partir de 2004 la cobertura de las BD se considera «media» o «alta».

El uso más frecuente que se da a las bases de datos institucionales es para la gestión (100%), seguimiento (89%) y evaluación (78%) de las actividades de investigación. El objetivo de fomentar la visibilidad de la universidad lo señaló el 67% de las respuestas.

FIGURA 1

Distribución de las respuestas al cuestionario



El acceso a la información recogida en estas BD varía en función del tipo de usuario: para los gestores de los servicios de investigación de las universidades el acceso es total, para los profesores es parcial, y también parcial para las agencias de financiación y usuarios externos. Cada tipo de usuario tiene intereses diferentes: mientras los gestores de la universidad utilizan las BD para la evaluación *ex post* (78%) y la gestión (94%), los profesores la emplean para la elaboración de sus CV y para dar visibilidad a los resultados de su investigación.

La actualización de la base de datos recae, en general, en cada profesor, y no tiene fecha fija sino que es permanente. También se señala la necesidad de validación de los datos por parte del responsable de la BD. Las limitaciones no se consideran importantes, destacando las limitaciones técnicas y las que derivan de la actualización, falta de depuración de los datos y control de calidad.

Respecto de la cobertura, todas las BD recogen información bibliográfica sobre artículos de revista, libros, capítulos de libro, actas de congresos, documentos de trabajo y tesis doctorales. En general, la cobertura de las BD oscila entre el 60% y 70% de los resultados de investigación de las universidades. Se incluye el nombre de todos los autores: si se trata de autores pertenecientes a la universidad se dispone de información general adicional, cosa que no ocurre con los autores externos. Las universidades que emplean UNIVERSITAS XXI y GREC también recogen información de proyectos, contratos, financiación y otras actividades, tanto del profesorado de forma individual como de los grupos de investigación. Es posible interrelacionar esta información acerca de actividades de investigación con otras BD de los profesores, en las que figuran edad, sexo, categoría profesional y especialidad.

En cuanto a los proyectos de futuro declarados por las universidades, en algunas instituciones se contempla la creación de un repositorio de acceso abierto vinculado a esta BD, mientras que en otras existe la intención de relacionar la BD con otras herramientas para la gestión de proyectos y la elaboración de CV adaptados a las exigencias de las diversas convocatorias nacionales e internacionales (en el caso de GREC algunas de estas propuestas ya están vigentes). Destaca por su interés la propuesta de una universidad politécnica que ya conecta la BD con otros ficheros de personal y salarios, junto con algunos análisis comparativos orientados a personal de la universidad y colaboradores externos. La mayoría de las universidades que contestaron la encuesta tienen la intención de ampliar la difusión para dar a conocer a las empresas los recursos tecnológicos de la propia Universidad a través de plataformas web adecuadas.

4.2. Indicadores de actividad científica obtenidos

4.2.1. Producción científica

Se presentan a continuación los indicadores obtenidos de las BD institucionales de las dos universidades seleccionadas: Universitat de Barcelona (UB) y Universidad Carlos III de Madrid (UC3M). El número de documentos de la UB asciende a 7.743 en el bienio 2005-2006, procedentes de los 29 Departamentos seleccionados, mientras que en el caso de la UC3M este número es de 5.946 en el mismo período.

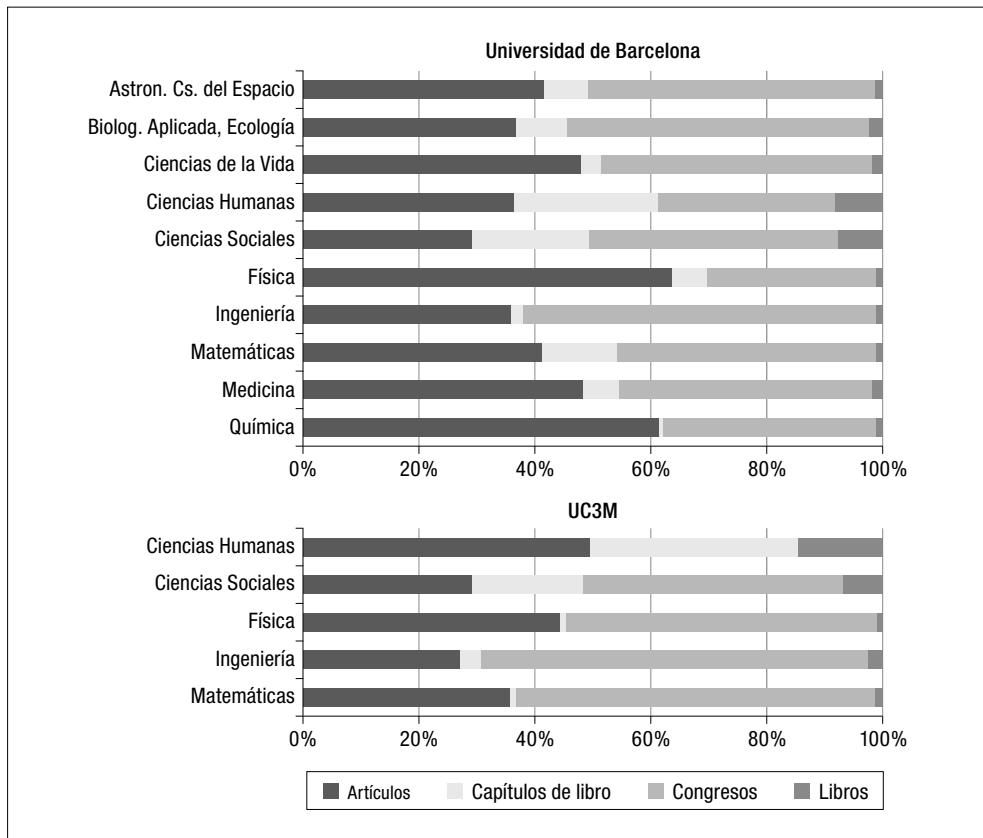
En la tabla I se observa que las publicaciones en la UB son en su mayoría artículos (41,20%), y en segundo lugar contribuciones a congresos (40,51%); mientras que la UC3M presenta un porcentaje mayor de contribuciones a congresos (51%) que de artículos (27,61%).

En la figura 2 se presenta la distribución de la producción en ambas universidades por tipo documental, agrupando la producción de los departamentos por área temática (ver Metodología). En la UB se observa el predominio de los artículos de revista en las áreas de ciencia básica. En ingeniería el mayor porcen-

TABLA I
Producción por tipo de documento

Universidad	Artículos	Libros	Capítulos de libro	Congresos	Otros	Total
Universitat de Barcelona						
Nº publicaciones 2005-2006	3.190	300	836	3.137	280	7.743
Nº publicaciones (promedio anual)	1.595	150	418	1.568,5	140	3.871,5
Porcentaje	41,20	3,87	10,80	40,51	3,62	100
Universidad Carlos III de Madrid						
Nº publicaciones 2005-2006	1.642	293	761	3.038	212	5.946
Nº publicaciones (promedio anual)	821	146,5	380,5	1.519	106	2.973
Porcentaje	27,61	4,9	12,8	51,0	3,6	100

FIGURA 2
Producción científica por tipo documental y área temática



taje corresponde a las contribuciones a congresos, mientras que en ciencias sociales y humanas destacan los libros y capítulos de libro.

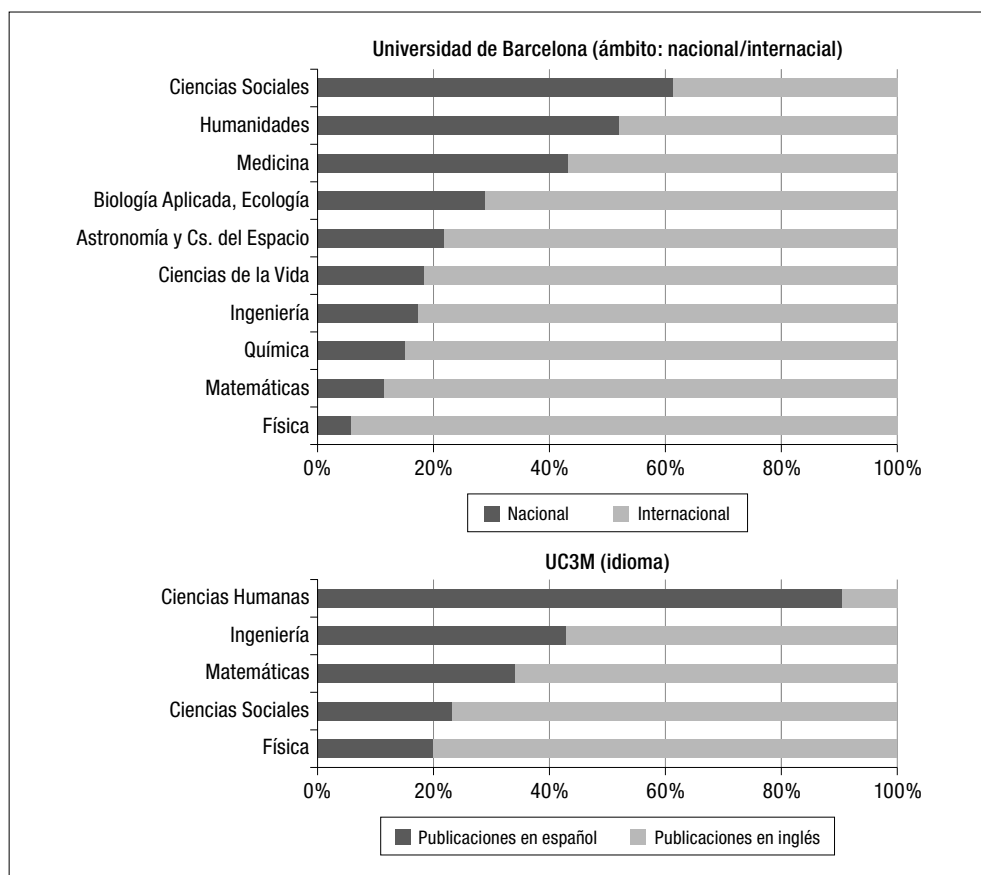
La UC3M, debido a su especialización, no tiene producción en todas las áreas. El porcentaje de congresos en ingeniería es elevado, y también en matemáticas y física, que proceden de departamentos de la Escuela Politécnica Superior. En ciencias sociales y humanidades destacan, como en el caso anterior, los libros y capítulos de libro.

4.2.2. Orientación nacional/internacional

No siempre se dispone de información acerca de los tres criterios propuestos para la clasificación de la producción según su orientación nacional/internacional: idioma, país de edición y ámbito nacional/internacional de la publicación. En el caso de UC3M sólo para artículos de revista se dispone de información acerca

FIGURA 3

Orientación nacional/internacional por área temática



de los tres parámetros. En la UB no se dispone del idioma, y sí del país de edición, así como de la orientación nacional/internacional de las publicaciones.

En la UB se muestra la orientación predominantemente internacional de las ciencias experimentales, frente al predominio de la orientación nacional en ciencias sociales y humanas (figura 3). En la UC3M se observa el predominio de las publicaciones nacionales en humanidades (basado en el idioma), frente al mayor peso de las internacionales en física y ciencias sociales.

4.2.3. Tamaño de los equipos

El estudio de la colaboración se ha realizado a través del número de autores firmantes de cada documento. En ninguna de las BD de las dos universidades se dispone de información del centro de trabajo ni del país de los colaboradores externos a la propia universidad. En la tabla II se muestra el número medio de autores por tipo de documento y área temática. En general, en ambas universidades el número más bajo de autores se observa en humanidades y ciencias sociales, mientras que el más elevado se produce en la UB en ingeniería, ciencias de la vida y medicina, mientras que en la UC3M los valores más altos se observan en ingeniería y física.

TABLA II

Número de autores por documento según área temática y tipo documental

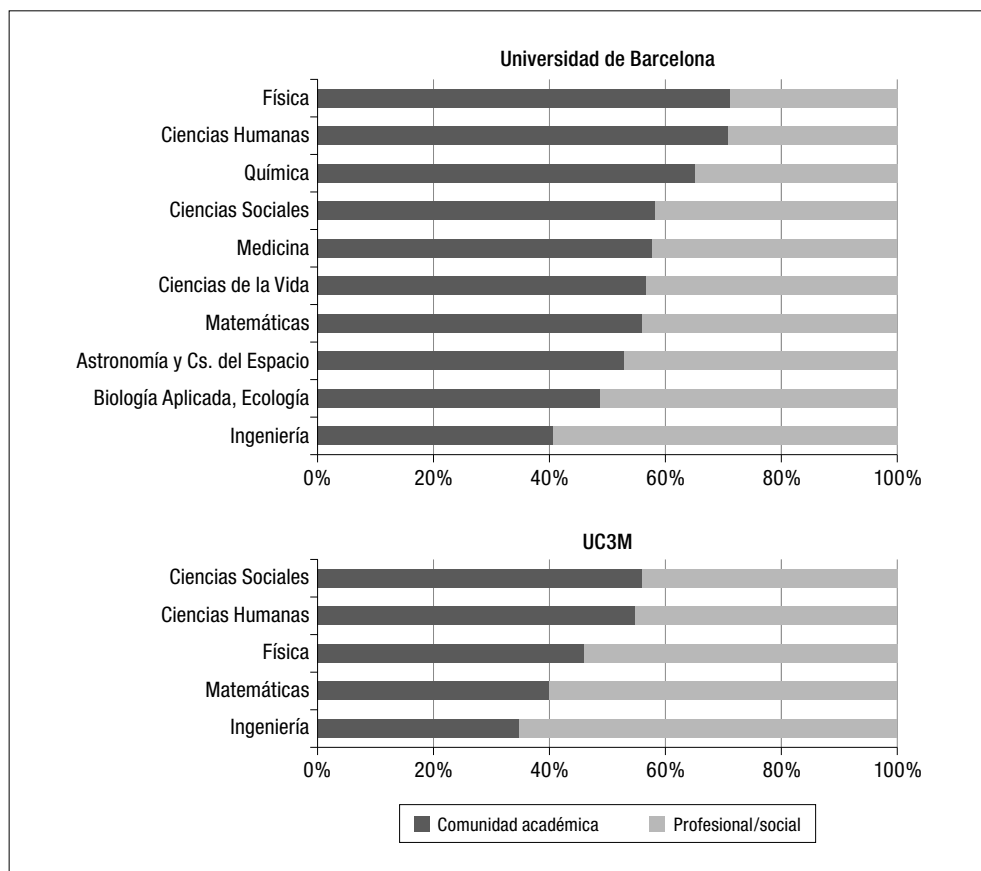
Universidad	Artículos	Libros	Capítulos de libro	Congresos	Promedio
Universitat de Barcelona					
Ciencias de la Vida	6,57	1,46	4,18	5,64	5,95
Investigación Médica	5,84	4,13	2,64	5,09	5,29
Biología Aplicada, Ecología	2,87	5,67	3,86	4,29	3,76
Química	5,26	3,00	3,25	4,51	4,94
Física	3,22	3,63	2,65	3,45	3,26
Astronomía y Ciencias del Espacio	4,39	2,00	4,06	3,33	3,81
Ingeniería	5,87	4,29	5,00	6,23	6,05
Matemáticas	2,69	2,00	3,55	2,52	2,72
Humanidades	1,81	1,59	1,82	1,69	1,76
Ciencias Sociales	1,82	3,67	1,60	1,93	1,96
Universidad Carlos III de Madrid					
Ingeniería	3,65	3,00	3,13	3,63	3,14
Humanidades	1,34	1,50	1,65	1,46	1,39
Matemáticas	2,83	1,67	2,50	2,35	2,41
Física	4,23	2,00	4,00	4,19	2,88
Ciencias Sociales	1,53	1,71	1,61	1,39	1,64

4.2.4. Comunidad de referencia

Se agrupa la producción en cuanto a la comunidad de referencia, es decir, que el destinatario de la investigación sea el sector académico o profesional/social. En la UB destacan las humanidades y la física por su orientación preferentemente académica, frente a la ingeniería en que predomina la orientación profesional (figura 4). En la UC3M se observa también que la ingeniería está más orientada a la comunidad profesional, mientras que en las humanidades y las ciencias sociales predomina la orientación académica. Sin embargo, hay que tener cierta precaución con la interpretación de estos resultados, puesto que el carácter de los congresos puede estar más orientado al mundo profesional o al académico dependiendo de la disciplina de que se trate.

FIGURA 4

Comunidad de referencia por área temática



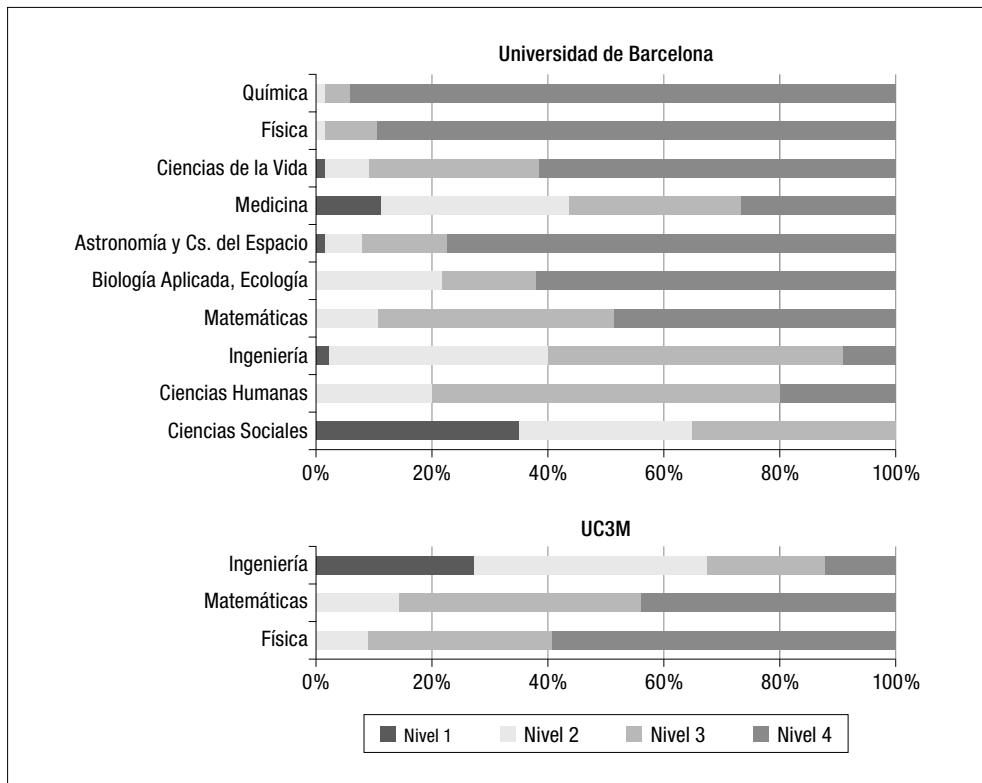
4.2.5. Carácter básico/aplicado de las publicaciones

Para la obtención de este indicador se utiliza la clasificación de revistas SCI y SSCI realizada por la *National Science Foundation*. Por ello, sólo puede aplicarse a artículos de revista cubiertos por dichas bases internacionales. En el caso de UB se aplica la clasificación aproximadamente a la mitad de los 3190 artículos totales, mientras que en UC3M sólo al 18% de los 1642 artículos de revista (tén-gase en cuenta que esta universidad no abarca las disciplinas de mayor orienta-ción internacional como las incluidas en ciencias de la vida).

En la UB se analizan 17 departamentos (de los 29 que posee esta universidad) con artículos en revistas incluidas en las bases de datos (figura 5). Las áreas de esta universidad con mayor número de artículos de investigación de nivel básico son la química y la física, mientras que el carácter aplicado se da en sociales, medicina e ingeniería. En la figura 5 se observa también el importante peso de los niveles 1 y 2 en ingeniería en UC3M frente al predominio de los niveles 3 y 4 (básicos) en física y matemáticas.

FIGURA 5

Artículos según nivel básico/aplicado, por área temática



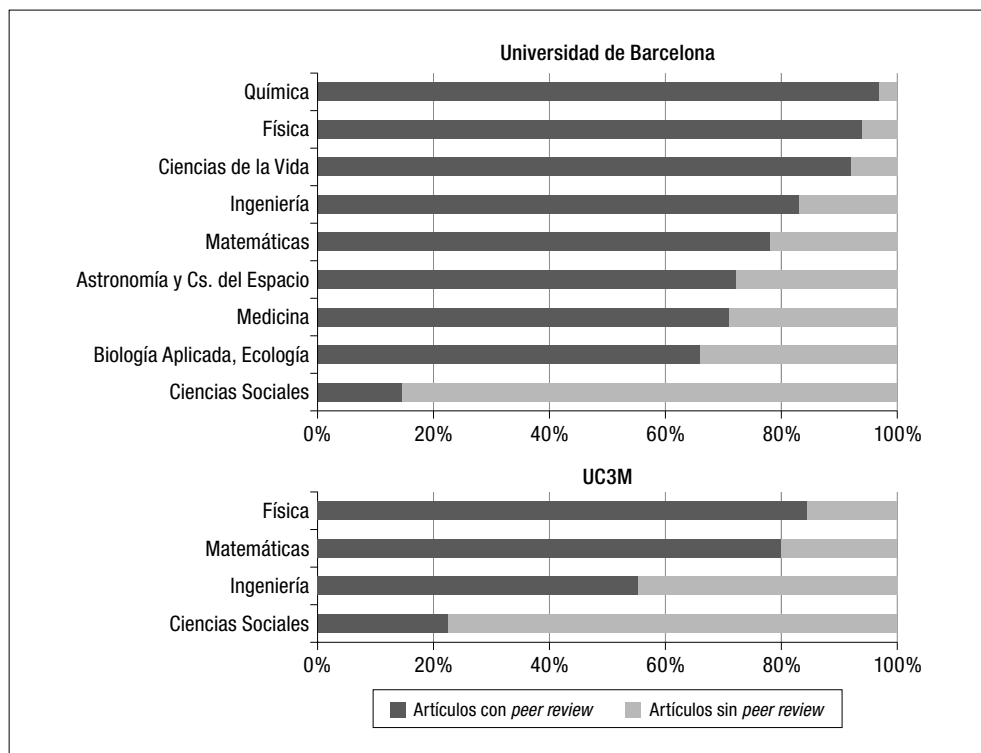
4.2.6. Existencia de un proceso de evaluación por pares

Las bases de datos institucionales no aportan información actualizada y verificable sobre este aspecto a día de hoy, por lo que en este trabajo se consideraron como revistas con sistema de evaluación por pares sólo aquellas cubiertas por el JCR de SCI y SSCI. Ha de tenerse en cuenta que el resultado real deberá ser mayor, pues hay revistas que siguen dicho proceso de selección de originales y no están cubiertas por estas bases de datos. Los artículos de humanidades publicados en revistas incluidas en A&HCI no se han considerado en este estudio ya que se trabajó con el JCR como tabla de validación de títulos.

En la UB se observa que en química, física y ciencias de la vida más del 90% de los artículos se publican en revistas SCI. En las áreas más aplicadas estos valores disminuyen, y en ciencias sociales solo alrededor del 15% de artículos están recogidos en revistas incluidas en el SSCI (figura 6). En la UC3M se observa que mientras que en física y matemáticas alrededor del 80% de sus artículos han sido publicados en revistas de JCR, en ingeniería este porcentaje disminuye hasta el 56%, y en ciencias sociales sólo alcanza algo más del 20%.

FIGURA 6

Artículos publicados en revistas cubiertas en los JCR por área temática



5. Discusión y recomendaciones

Los datos obtenidos en el estudio a nivel español, como los que resultan de la agregación de los diversos países participantes en el proyecto UNIPUB, ponen de manifiesto la utilidad de las BD institucionales para caracterizar los diferentes resultados de la investigación de las universidades, así como el amplio margen de mejora y normalización que tienen dichas BD, de cara a una mayor transparencia y comparabilidad de la actividad investigadora de las universidades.

Uno de los objetivos principales de estos instrumentos es la autoevaluación, es decir, la evaluación interna de la calidad de la investigación académica permitiendo construir «indicadores de posicionamiento» que promuevan la especialización de las universidades de acuerdo con su entorno social y geográfico (Lepori y otros, 2008). Se trata de poder situar a cada universidad en cada una de las dimensiones de su actividad, por lo que no se consideran únicamente *inputs* y *outputs*, sino también la orientación de cada universidad, su relación con el entorno a través de colaboraciones y lazos establecidos con diferentes actores. Hemos analizado diversas dimensiones, como son la orientación nacional/internacional por área temática en cada universidad, su especialización temática; el tamaño de los equipos de investigación; el tipo de conocimiento que genera (básico/aplicado); los usuarios principales de la producción científica (académicos, profesionales o la sociedad).

Según la encuesta realizada, el 95% de las universidades españolas que respondieron cuenta con bases de datos institucionales, frente al 85% de las italianas y al 83% de las suizas. Los datos de los 3 países muestran que la creación de estas BD institucionales ha tenido un importante impulso en los últimos años, especialmente a partir de 2005, y han sido desarrolladas principalmente por iniciativa propia de cada universidad. El objetivo de la creación de las BD españolas ha sido facilitar la gestión de la actividad investigadora, mientras que en países como Suiza se han desarrollado para lograr mayor visibilidad institucional con gran hincapié en la iniciativa Open Access (Reale y otros, 2009b).

Una consecuencia del origen institucional de estas bases de datos es la falta de homogeneidad de criterios entre las universidades, lo que se traduce en una cierta dificultad para obtener indicadores comunes. En la mayoría de los casos, los profesores suelen ser los encargados de introducir sus propios resultados de investigación, que no siempre están completos y validados, aunque tienden a estarlo en los años más recientes. Esta falta de consistencia de los datos se agrava con la falta de portabilidad de los mismos cuando un investigador cambia de institución, cuando participa en más de un entorno automatizado de gestión de currícula y cuando ha de presentarse a convocatorias con requisitos diversos de verificación de méritos de investigación.

Queremos poner de manifiesto el interés que presentan los indicadores de producción científica por área temática, especialmente en aquellas áreas que no son de ciencia básica y que, en consecuencia, están peor cubiertas por las BD internacionales. Las Ciencias Sociales y Humanidades difunden una gran parte

de sus investigaciones en forma de libros y capítulos de libro; en Ingeniería destaca el fuerte peso de las contribuciones a congresos, que son especialmente importantes en las áreas de ciencias aplicadas. En este sentido, a falta de la circulación de estos *outputs* de investigación en bases de datos internacionales, se plantea como fundamental la validación que, respecto a estas aportaciones, daría la conexión de las BD de gestión institucional de la investigación con repositorios a texto completo y catálogos de bibliotecas que localicen las obras, tanto en términos de transparencia como de facilitación de la evaluación cualitativa a terceros.

Sólo se hizo una primera aproximación de la comunidad de referencia a la que se orientan los resultados de investigación, y aún así se vislumbra el interés que podrían aportar estas bases de datos institucionales de cara a potenciar la tercera misión de la universidad, tanto en cuanto a su relación con el sector empresarial como con la sociedad en general. Somos conscientes de que los congresos (la tipología documental más relevante) pueden tener un carácter más orientado al mundo profesional o al académico, según la disciplina de que se trate. En ciertos casos presentan resultados preliminares que se pueden desarrollar en artículos de revista. De cualquier forma, la asignación de la orientación de los congresos sería una posible pregunta a incluir en las propuestas para mejorar las bases de datos institucionales.

En cuanto al estudio de los autores de los trabajos, cuando se trata de autores de la propia universidad se puede obtener información complementaria de los mismos a través de otras BD internas. No es así cuando se trata de autores externos, en cuyo caso hay una total falta de información. Sería interesante poder conocer su procedencia institucional, pues permitiría analizar las redes de colaboración establecidas con otros centros de investigación o con la empresa: el enriquecimiento de estos datos por medio del control de autoridades de los nombres de autores externos exigiría una mayor inversión que sólo parece viable en aquellos casos de autores presentes en bases de datos bibliográficas internacionales de indización y resumen; pese a esta limitación, esta solución de control de calidad y de verificación de datos de las BD institucionales por medio de bases de datos bibliográficas externas apunta claramente al requisito técnico de interoperabilidad de las BD.

En este sentido, si la introducción de datos por parte de profesores e investigadores se pudiera beneficiar de la importación de datos externos ya validados bibliográficamente, el grado de actualización de las BD institucionales mejoraría substancialmente y los costes de normalización y verificación se podrían reducir, aunque se presentarían problemas de derechos de autor frente a los productores de las bases de datos, en caso que se pretendiera una difusión pública vía web de la propia BD institucional.

Sin duda, pues, la necesidad de emplear información complementaria de las BD institucionales indica la conveniencia de incluir más información contrastada y validada. En particular, el hecho de que las revistas utilicen o no el sistema de evaluación externa para seleccionar sus trabajos aporta un valor añadido intere-

sante como indicador de «calidad esperada» de las publicaciones. Podrían utilizarse, junto a los JCR, otras fuentes de información que facilitarían indicios de calidad acerca de las revistas, por ejemplo estar indexadas en las bases de datos SCOPUS, en el *European Reference Index for the Humanities*, o su presencia en el catálogo LATINDEX.

A partir del análisis tanto de los datos de las universidades que contestaron la encuesta como de los obtenidos en el estudio de casos de las dos universidades consideradas, algunas de las recomendaciones generales que se podrían plantear con el fin de mejorar la utilidad de las BD institucionales para la obtención de indicadores de actividad científica serían las siguientes:

- Incluir el tipo de comunidad de referencia. Es decir, especificar claramente cuál es el destinatario de la investigación (sector académico o profesional/social). Esta información sería de gran interés a la hora de canalizar los resultados de investigación hasta su usuario final.
- Incluir la información sobre la existencia o no del proceso de evaluación por pares en la evaluación de las revistas.
- Incluir la vinculación de las publicaciones con los proyectos de investigación financiados. Esto permitiría hacer un seguimiento *ex post* de la repercusión de la investigación financiada.
- Incluir información de autores externos tanto la relativa a su género como a su lugar de trabajo. Esta información ampliaría considerablemente la utilidad de estas bases de datos, porque permitiría realizar estudios, tanto de género como de colaboración científica con otras instituciones.
- Tratar de armonizar las BD institucionales de las distintas universidades españolas, con el fin de facilitar la realización de estudios comparativos de su actividad investigadora.
- Favorecer la integración entre las bases de datos institucionales y los repositorios de acceso abierto con los textos completos de los trabajos, con el fin de verificar las publicaciones, así como de potenciar la visibilidad de la investigación que realizan las universidades.
- Incluir la participación de profesionales de la documentación en los procesos de normalización y validación de los datos introducidos por el profesorado. La formación y experiencia que tiene este tipo de profesionales en el tratamiento y gestión de la información puede ser de gran ayuda para incrementar la calidad de la información de las bases de datos institucionales.

Es evidente que tanto el interés como la utilidad de estas bases de datos institucionales van a depender, en gran medida, de la exhaustividad y fiabilidad de la información que contengan, y para ello, se debería ofrecer incentivos a los profesores, con el fin de que completen y validen la información que introducen de sus resultados de investigación. En este sentido, en las convocatorias de los programas propios de investigación de determinadas universidades, como son

las que incluyen la financiación de estancias de investigación o ayudas a proyectos, solo consideran aquella actividad científica de los solicitantes que esté incluida en sus bases de datos institucionales. Asimismo, el CSIC evalúa anualmente a sus institutos y gratifica a los investigadores en función de los resultados obtenidos. Ambas iniciativas pueden constituir un buen camino, para completar la información que contienen estas bases y, por tanto, para reforzar su papel en la institución.

En definitiva, los requisitos de transparencia y rendimiento de cuentas a los que están sujetas las universidades y centros de investigación, junto al servicio de soporte que han de prestar a sus investigadores, obligados a presentar multitud de formatos de curriculum a las diversas convocatorias y procesos de evaluación, representan una oportunidad para dar un impulso a nivel europeo y español en la mejora de la actualización, la calidad, la normalización y la interoperabilidad de las BD de gestión interna de la investigación.

6. Agradecimientos

Manifestamos nuestro agradecimiento a la Red de excelencia europea PRIME por la financiación del proyecto UNIPUB. También queremos agradecer la participación de las universidades españolas que respondieron a la encuesta, y especialmente a UB y UC3M.

7. Bibliografía

- Bellavista, J.; Escribano, L.; Grabulós, M.; Guardiola, E.; Iglesias, C., y Viladiu, C. (1992). La organización de la investigación en la universidad. *Revista española de investigaciones sociológicas*, 59, 301-309.
- Bonaccorsi, A.; Daraio, C.; Lepori, B., y Slipersaeter, S. (2007). Indicators for the analysis of Higher Education Systems: some methodological reflections. *Research Evaluation*, 16 (2), 66-78.
- Cañibano, C., y Bozeman, B. (2009). Curriculum vitae method in science policy and research evaluation: the state of the art. *Research Evaluation*, 18 (2), 86-94.
- Gomez, I.; Bordons, M.; Fernández, M. T., y Morillo, F. (2009). Structure and Research Performance of Spanish Universities, *Scientometrics*, 79 (1), 131-146
- Hicks, D. (2004). The Four Literatures of Social Science. En: Moed, H., Glänzel, W. y Schmoch, U. (eds.), *Handbook of Quantitative Science and Technology Research* (pp. 473-496). Dordrecht: Kluwer Academic Publishers.
- Iribarren-Maestro, I.; Lascurain-Sánchez, M. L., y Sanz-Casado, E. (2009). The use of bibliometric techniques in evaluating Social Sciences and Humanities. En: Åström, F.; Danell, R.; Larsen, B.; Schneider, J. W. *Celebrating Scholarly Communication Studies: A Festschrift for Olle Persson at his 60th Birthday*, pp. 25-37. Eds. International Society for Scientometrics and Informetrics.

- IPIQ (2008). SCI/SSCI Journal Classification File. Prepared by ipIQ for NSF under NSF Contract No. SRS0651343 on September 3, 2008.
- Larédo, Ph. (2003). University Research Activities: On-going Transformations and New Challenges. *Higher Education Management and Policy*, 15 (1), 105-123.
- Lepori, B.; Barré, R., y Filliatreau, Gh. (2008). New Perspectives and Challenges for the Design of S&T Indicators. *Research Evaluation*, 17, 33-44
- Nederhof, A. J. (2006). Bibliometric Monitoring of Research Performance in the Social Sciences and the Humanities: a review. *Scientometrics*, 66 (1), 81-100.
- Rapport de l'OST 2008. Classification de disciplines scientifiques. Correspondance A4-3 (DISC-9).
- Reale, E.; De Filippo, D.; Gómez, I.; Lepori, B.; Potì, B.; Primeri, E.; Probst, C., y Sanz, E. (2009a). Methodologies for the characterization of the publication output of higher education institutions using institutional databases, ISSI International Conference, Rio de Janeiro.
- Reale, E.; De Filippo, D.; Gómez, I.; Lepori, B.; Probst, C.; Potì, B.; Primeri, E., y Sanz, E. (2009b). Methodologies for the characterization of the publication output of higher education institutions using institutional databases, Final Report, PRIME NoE, <http://www.prime-noe.org>.
- Torres-Salinas, D. L., y Moed, H. F. (2009). Library Catalog. Analysis as a tool in studies of Social Sciences and Humanities: An exploratory study of published book titles in Economics. *Journal of Informetrics*, 3 (1), 9-26
- van den Besselaar, P.; Edler, J.; Heimeriks, G.; Henriques, L.; Laredo, P.; Luukkonen, T.; Nedevea, M.; Schoen, A., y Thomas, D. (2007). Toward ERA configurations. An experiment on Chemistry, Workshop «Beyond the dichotomy of national vs. European science systems. Configuration of knowledge, institutions and policy in European research», Bonn, 30 May.
- van Raan, A. (2004). Measuring Science. En: Moed, H., Glänzel, W. y Schmoch U. (eds.), *Handbook of Quantitative Science and Technology Research*, Dordrecht: Kluwer Academic Publishers, pp. 19-50.

ANEXO

ENCUESTA A LAS UNIVERSIDADES ESPAÑOLAS

DATOS GENERALES

Fecha: _____

País: _____ **Universidad:** _____

1. ¿Su universidad cuenta con alguna base de datos institucional de actividades de investigación?

Sí (ir a la pregunta 4) No (responder solamente las preguntas 2 y 3)

2. Si no hay una base de datos institucional de actividades de investigación, ¿a qué se debe?

- Problemas administrativos.
- Problemas organizativos.
- No es una herramienta útil.
- No es necesaria.
- Otros (especificar) _____

3. Si no hay una base de datos institucional de actividades de investigación, ¿están pensando en crearla?

Sí No

4. Si la Universidad tiene una base de datos institucional de actividades de investigación ¿cuál es su origen?

- Una iniciativa regional o gubernamental.
- Una iniciativa interna (de la propia universidad).
- Ambas.
- Otras.

Comentarios: _____

5. Año de implantación de la base de datos _____

6. Nombre de la plataforma utilizada para la gestión de datos y dirección web: _____

7. ¿Qué tipo de implantación es?

- Web.
- Cliente servidor.
- Otra.

8. Objetivos de la base de datos (puede seleccionar más de uno).

- Evaluación.
- Gestión.
- Seguimiento.
- Visibilidad.
- Open access.
- Otros (especificar) _____

Comentarios: _____

9. Acceso a la base de datos.

	Contenido completo	Contenido parcial	No accesible
Gestores de la propia universidad			
Profesores, investigadores, PAS			
Gobierno y agencias de financiación			
Otros usuarios externos			
Otros (especificar)			

Comentarios: _____

10. Usuarios de la base de datos (seleccionar los objetivos más importantes para cada usuario).

Objetivos	Gestores universitarios	Gobierno	Agencias de financiación	Personal docente investigador ¹
Evaluación Ex-post				
Evaluación de la financiación de proyectos				
Gestión				

¹ Personal docente-investigador incluye personal funcionarios, contratados y becarios de investigación.

Objetivos	Gestores universitarios	Gobierno	Agencias de financiación	Personal docente investigador ¹
Elaboración de informes				
Fuentes de financiación				
Visibilidad				
Elaboración de CV (creación, mantenimiento, actualización)				
Otros (especificar)				

Comentarios: _____

11. Cobertura de la base de datos:

(Alta = la base de datos incluye todos los resultados relacionados con la actividad de los docentes-investigadores; Media = la base de datos incluye una parte significativa de los resultados del personal docente-investigador. Baja = la base de datos incluye menos del 30% de los resultados del personal investigador.)

2008	Alta	Media	Baja
2007	Alta	Media	Baja
2006	Alta	Media	Baja
2005	Alta	Media	Baja
2004	Alta	Media	Baja

Comentarios: _____

12. Actualización de la base de datos:

- Semestral.
- Anual.
- Cada 2 años.
- Actualización permanente (depende de cada docente-investigador).
- Otros (especificar) _____

Comentarios: _____

13. Limitaciones de la base de datos (indicar en que grado afecta cada caso)

Problemas	Alto	Medio	Bajo
Baja cobertura			
Escasa motivación por parte de los investigadores a cargar los datos			
Limitaciones técnicas			
Falta de depuración de los datos			
Escasa actualización de los datos			
Bajo control de la calidad de los datos			
Falta de privacidad			
Problemas con el Copyright			
Otros (especificar)			

Comentarios: _____

14. ¿Qué tipo de publicaciones incluye la base de datos?

- Artículos.
- Libros.
- Capítulos de libro.
- Actas de congresos (*proceedings*).
- Presentaciones a congresos (comunicaciones orales y póster que no estén en acta).
- Documentos de trabajo.
- Tesis doctorales.
- Artículos de divulgación científica.
- Software.
- Otros (especificar) _____

15. Información bibliográfica incluida en la base de datos

Datos sobre los autores	Sí	No
1. Autor(es) de la propia universidad		
Nombre y apellido		
Afiliación institucional (facultad/escuela, centro, departamento)		
Área científica de investigación del autor		
Género		
Edad		
Categoría académica		

Datos sobre los autores	Sí	No
2. Autor(es) externo(s)		
Nombre y apellido		
Afiliación institucional (país, institución, centro, departamento)		
Área científica de investigación del autor		
Género		
Edad		
Categoría académica/profesional		

Datos sobre las publicaciones	Sí	No
Título		
Tipo de publicación		
Lugar de publicación		
Editorial		
Número de página		
Nombre de la publicación		
Disciplina científica		
Colaboración		
Con autores de la propia universidad		
Con autores externos		
Número de autores		
Número de autores de la propia universidad		
Número de autores externos		
Con evaluación/sin evaluación		
Factor de impacto		
Año de publicación		
Orientación de la publicación (nacional/internacional)		
Resumen		
Texto completo		
Open access (permite descargar el archivo completo)		
Otros (especificar)		

Comentarios: _____

16. Uso real e implementaciones previstas de la base de datos en los próximos años

Comentarios: _____