

Normalización de campos en bibliometría

Acciones de la Fecyt

Por José-Antonio Ontalba-Ruipérez

Resumen: Esta contribución detalla la participación de varios grupos de trabajo en el Seminario de la Fecyt (Fundación Española para la Ciencia y la Tecnología) sobre “Normalización de instituciones para usos bibliométricos” celebrado en Barcelona el 4 de junio de 2007. Las organizaciones intervinientes fueron Cindoc, Universidad Alcalá de Henares, Universidad Carlos III, Universidad de Extremadura, Universidad de Granada y Universidad Politécnica de Valencia – Grupo Ciepi-, así como la propia Fecyt. Se presentó una herramienta informática capaz de gestionar los nombres de entidades, una propuesta para su clasificación y unas recomendaciones para registrar los nombres de organizaciones. Finalmente se propuso la creación de la Plataforma Nacional de Cienciometría, como estructura estable de trabajo.

Palabras clave: Bibliometría, Normalización, Nombres de instituciones, Investigación, España, Fecyt

Title: Field standardization for bibliometrics studies. Fecyt actions

Abstract: This contribution details the participation of several research groups in the Fecyt (Spanish Foundation for Science and Technology) Workshop “Standardisation of institution names for scientometric uses”, held in Barcelona on the 4th of June 2007. Besides Fecyt, the participating institutions were Cindoc, Universidad Alcalá de Henares, Universidad Carlos III, Universidad de Extremadura, Universidad de Granada and Polytechnic University of Valencia –Ciepi Group-. The topics covered were: technological tools to manage institution names, a proposal for the classification of research organisations and some recommendations to standardise the registration of organisation names. Finally, the creation of a permanent National Platform on Scientometrics was suggested.

Keywords: Bibliometrics, Standardization, Institution names, Research, Spain, Fecyt

Ontalba-Ruipérez, José-Antonio. “Normalización de campos en bibliometría. Acciones de la Fecyt”. En: *El profesional de la información*, 2007, julio-agosto, v. 16, n. 4, pp. 381-383.

DOI: 10.3145/epi.2007.jul.14



José-Antonio Ontalba-Ruipérez, Universidad Politécnica de Valencia y Grupo de investigación Ciepi

EL 4 DE JUNIO DE 2007 la Fecyt (Fundación Española para la Ciencia y la Tecnología)¹ celebró en el Parque de Investigación Biomédica de Barcelona (Prbb)² un Seminario de trabajo sobre normalización de campos: instituciones y nuevas herramientas de uso bibliométricos.

La medición de la actividad científica española se basa, parcialmente, en indicadores de tipo documental. Las bases de datos que contienen la información bibliográfica que permiten este tipo de análisis conllevan un índice de errores que distorsiona los resultados de los estudios sobre la ciencia en España. La falta de normalización de la información en los campos de autoría y filiación de los autores condiciona y dificulta el trabajo de recopilación,

tratamiento y análisis de datos en los estudios bibliométricos. Para aumentar la fiabilidad de estos indicadores se reunió en Barcelona el grupo de normalización de campos bibliométricos: Cindoc, Universidad de Alcalá de Henares, Universidad Carlos III de Madrid, Universidad de Extremadura, Universidad de Granada y Universidad Politécnica de Valencia, coordinados por la

“La falta de normalización condiciona y dificulta el trabajo de recopilación, tratamiento y análisis de datos en los estudios bibliométricos”

Fecyt. Esta jornada sobre normalización estuvo dedicada principalmente a la identificación y clasificación de los nombres de las instituciones con producción científica.

En primer lugar, **Tomàs Baiget** tuvo ocasión de presentar el trabajo del grupo Ciepi³: el proyecto *IraLIS* (International Registry for Authors in Library and Information Science)⁴ que como campo de pruebas de momento se dirige sólo a los nombres de los autores del área de Información y Documentación. Surge de anteriores iniciativas en las que los miembros del grupo se encuentran involucrados (el repositorio *E-LIS*⁵, el directorio *Exit*⁶ y la revista *EPI*⁷, así como el *Grupo ThinkEPI*⁸ y la lista *IweTel*⁹). *IraLIS* pretende registrar una forma del nombre de los



José Manuel Báez, Jordi Camí y Félix de Moya

“José Manuel Báez propuso la constitución de una plataforma de bibliometría con el propósito de ayudar a la normalización de campos para el mundo de la investigación”

“IraLIS pretende registrar una forma del nombre de los autores españoles que se asemeje a las costumbres anglosajonas, imperantes en las bases de datos internacionales”

autores españoles que se asemeje a las costumbres anglosajonas, imperantes en las bases de datos internacionales. En coincidencia con las recomendaciones publicadas hace unas semanas por la propia *Fecyr*¹⁰, el sistema propone un formato de firma para evitar la multiplicidad de variantes con que se firman los trabajos. En el caso español, se ha constatado la poca disposición de los autores para adoptar una forma normalizada de su firma. Con todo esto se aspira a la recuperación exhaustiva de bibliografías personales gracias a los registros xml del sistema, compatibles con el esquema de metadatos de autoridades *Metadata authority description schema (Mads)*¹¹ propuesto por la *Library*

of Congress; a facilitar los estudios estadísticos de tipo bibliométrico, de indicadores socio-profesionales, científicos, o por sexo; y, de cara a un futuro inmediato, se conectará automáticamente con las listas de autoridades de la *Biblioteca Nacional*, y se incorporarán (de manera semiautomática) las diferentes variantes de los nombres.

Seguidamente el grupo *Scimago*¹² expuso sus avances en la normalización de nombres de instituciones científicas. **Félix de Moya** presentó una plataforma tecnológica que permite la explotación de la información que contienen las bases de datos del *Web of science (WoS)*. El desarrollo informático es capaz de procesar de forma automática los datos del campo “address”, que contienen la filiación de los autores, manteniendo la conexión con los propios documentos. Estos procesos deben ser refinados mediante la intervención humana, pero suponen una gran ayuda para analizar este tipo de información. Las pruebas realizadas reducen a unos 6.000 registros las instituciones españolas identificadas, si bien las incidencias

de *WoS* en el registro de los datos harían aumentar la cifra. Asimismo, **Elena Corera** realizó una demostración en directo de las funcionalidades de la plataforma: destaca la facilidad para editar las instituciones, asignar la comunidad autónoma a que pertenecen, vincular entidades relacionadas o subordinadas, así como para gestionar la calidad de este registro, añadiendo variantes de nombres, abreviaturas, notas explicativas... Cabe señalar que se trata de un desarrollo tecnológico adaptable a los criterios que el resto de grupos han de proponer.

Isabel Gómez Caridad (*Cindoc*) y **Jordi Camí** (*Prbb*) presentaron su propuesta de clasificación de sectores institucionales que permita estructurar homogéneamente los trabajos bibliométricos realizados en España. Para el caso de entidades subordinadas propusieron diferentes grados en la identificación de las instituciones. Si, por ejemplo, para la universidad española se escoge registrar únicamente a una universidad, ello impediría microanálisis por departamentos, algo que es habitual en el área biblio-



Aspecto característico del edificio del Parc de Recerca Biomèdica de Barcelona, de los arquitectos Manel Brullet y Albert de Pineda



Vista de la zona interior del PRBB

métrica. Por esta razón consideran fundamental que el *Cindoc* elabore un índice de autoridades clasificado por sectores de actividad (administración pública, educación superior, OPIs...) que incluya entidades hasta un tercer nivel de detalle. Finalmente, debería crearse un web donde se vayan registrando tanto los diferentes nombres y sinónimos de cada centro, como la evolución histórica de esa entidad, que permita los análisis diacrónicos. Los aspectos clasificatorios generaron un interesante debate pues las formas jurídicas que adoptan las diferentes instituciones no son útiles para asignarlas a un sector de actividad. En este momento existen tanto modelos de estructura arborescente, con una entidad madre única, como modelos en estrella, donde la colaboración se produce al mismo nivel y en varias comunidades autónomas simultáneamente. Por último, cabe reseñar que su propuesta no se limitaba a las entidades que aparecen en *WoS*, y propusieron una estructura flexible que permita una evolución futura.

Florencio Núñez (*Fecyt*) presentó la sección de entidades del proyecto CVN² que se aplicará al

Plan Nacional de I+D. Para este desarrollo técnico se han solicitado códigos EAN (*European article numbering*), que permitirán el intercambio de la información contenida en los campos.

A continuación, **Rafael Ruiz** (*EC³, Universidad de Granada*) completó lo expuesto por el grupo *Scimago* con una propuesta de criterios para la formulación normalizada de nombres de instituciones y centros españoles. Insistieron en la base que existe en el control de nombres catalográficos, y recomendaron recoger el nombre completo de cada centro, su acrónimo, su nombre en inglés, su URL, sede y otros datos.

Finalmente, **José Manuel Báez** (*Fecyt*) propuso la constitución de una plataforma de bibliometría con el propósito de ayudar a la normalización de campos para el mundo de la investigación. Se plantea, entre otros objetivos específicos, recoger y aglutinar las iniciativas y trabajos de los diferentes grupos españoles bibliométricos, además de ayudar a la normalización de las publicaciones científicas en España. Esta plataforma la constituyen los miembros del grupo de trabajo reunido en Barce-

lona, y está abierta a integrar otros grupos de bibliometría que trabajen de manera relevante en el campo de la normalización o cualquier otra tarea relacionada.

Se espera que en la próxima reunión del grupo se concrete la propuesta de una lista de autoridades dirigida al análisis bibliométrico de la producción científica española.

Notas

1. <http://www.fecyt.es>
2. <http://www.prbb.org>
3. <http://www.ciepi.org>
4. <http://www.iralis.org>
5. <http://eprints.relis.org>
6. <http://www.directorioexit.info>
7. <http://www.elprofesionaldelainformacion.com>
8. <http://www.thinkepi.net>
9. <http://www.rediris.es/list/info/iwvetel.es.html>
10. http://www.accesowok.fecyt.es/normalizacion_nombre_autor.pdf
11. <http://www.loc.gov/standards/mads/>
12. <http://www.scimago.es/>
13. <http://cv.normalizado.org>

José-Antonio Ontalba-Ruipérez.
Dpto. de Comunicación Audiovisual,
Documentación e H^a del Arte, Uni-
versidad Politécnica de Valencia, Ca-
mino de Vera s/n. 46071 Valencia.
Tel.: +34-963 879 392
joonrui@har.upv.es

ANUARIO

Think

EPI

2007

Análisis de tendencias en
información y **documentación**

El profesional de la

información

SOLICITE UN EJEMPLAR DEL ANUARIO THINKEPI 2007

Rellene este formulario (o su fotocopia) y envíelo:

- al apartado de correos 32.280 de 08080 Barcelona, o

- por correo-e a epi@sarenet.es

Apellidos:

Empresa:

Calle:

CP:

Ciudad:

Email:

Nombre:

NIF:

Número:

Piso:

Provincia:

Teléfono:

Marque con una cruz la forma de pago elegida:

Ingreso o transferencia bancaria por valor de 90 euros a la cuenta de ThinkEPI en
La Caixa, número 2100 0818 93 0200745544

Cheque de 90 euros enviado al apartado de correos 32.280 de 08080 Barcelona

Pago con tarjeta de crédito número: _____ - _____ - _____ - _____, caduca ___/___

¡Gracias! En breve recibirá un ejemplar del Anuario.

<http://www.thinkepi.net>

Think
EPI