

¿Cómo medimos la ciencia? Segundo Workshop de la Fundación General CSIC

Los talleres científicos patrocinados por la Fundación General del CSIC (www.fgcsic.es) tienen por objeto promover debates acerca de diversas actividades de interés para la Fundación. Este segundo taller dedicado a la medición de la producción científica, tuvo lugar el 20 de septiembre de 2011 en la sede de Caixa Forum de Madrid.

Durante la sesión se presentaron 6 ponencias seguidas de coloquio, para finalizar con una mesa redonda formada por los ponentes y moderada por Victoria Ley, directora de la Agencia Nacional de Evaluación y Prospectiva (ANEP), donde se analizaron y discutieron los aspectos más significativos apuntados durante la jornada en debate abierto con todos los asistentes.

Los ponentes aportaron sus experiencias como investigadores y como expertos en evaluación. El contar con ambos puntos de vista enriqueció notablemente las presentaciones y discusiones posteriores y proyectó una visión muy realista de la problemática de la medición de los resultados científicos.

Comenzaron la jornada Javier Rey y José Luis de Miguel, Director y Subdirector, respectivamente, de la Fundación General del CSIC, con una ponencia común: *«¿Cómo medir la producción científica y la transferencia del conocimiento?»*. El primero se centró en explicar la valoración que hace el CSIC cada cuatro años de la productividad científica y tecnológica de cada uno de sus centros, vinculada al cumplimiento de los objetivos (PCO) determinados según una planificación estratégica previa. La valoración positiva de la productividad supone un incremento salarial para todo el personal del centro (complemento de productividad).

La evaluación y seguimiento de la planificación estratégica se realizan por medio de paneles de expertos externos. Para la primera se tienen en cuenta los siguientes indicadores: financiación obtenida en proyectos competitivos; producción científica (número de publicaciones ISI, no ISI, congresos, libros, etc.); impacto científico (citas recibidas); producción tecnológica (patentes solicitadas y concedidas); ingresos por contratos con empresas; formación (tesis doctorales dirigidas, cursos impartidos, y otros).

Se explicaron con detalle las ventajas y desventajas de algunos de los indicadores bibliométricos empleados en la evaluación, sobre todo los relacionados con las citas recibidas por los investigadores, como son los Índices h y g, y los correspondientes al cómputo de citas obtenidas por las revistas, como el Índice de Impacto (IF), el Eigenfactor (EF), el Article Influence Score (AI), y el Scimago Journal Rank (SJR).

El segundo ponente dedicó su tiempo a analizar los indicadores de la transferencia del conocimiento al entorno empresarial, y precisó que no tienen tanta aceptación como los indicadores de ciencia al no estar tan bien establecidos;

además, los tiempos estimados para su medición son más largos que para los indicadores científicos. Téngase en cuenta que uno de los sistemas más útiles de transferencia de conocimiento es la divulgación de éste mediante las publicaciones, pero ese conocimiento lo tienen que asimilar las empresas y se requiere mucho tiempo para ello. En la transferencia de conocimiento se asocia «valor» con «precio», por lo que la transferencia se puede medir por: *a)* los contratos de I+D establecidos entre los centros de investigación y las empresas, tanto de I+D en colaboración como contratada; *b)* las consultorías y servicios tecnológicos realizados; *c)* el volumen de la financiación captada; *d)* la fidelización de los clientes; *e)* la duración y recurrencia de los contratos con las empresas (mejor menos contratos y más caros), y *f)* las solicitudes de patentes (nacionales, internacionales, o triádicas: Estado Unidos, Japón y Europa); las patentes concedidas; las patentes en explotación (difícil obtener información sobre ellas, una vez que entran en el ámbito de una empresa), las regalías, etc.

El mejor indicador de que se ha producido transferencia de conocimiento es el número de empresas «spin-off» o empresas de base tecnológica creadas. Éstas surgen de grupos de investigadores de centros públicos o universidades que actúan como promotores de la empresa, que es el medio para la explotación de los resultados de la investigación o «know-how». El problema es que se suelen crear pocas empresas de este tipo.

Se presentaron algunos datos estructurales y de producción científica y técnica para la comparación entre el CSIC y otras instituciones de investigación extranjeras que abarcan también todas las áreas del conocimiento. Éstas fueron: CNRS (Francia), CNR (Italia), MPG (Alemania), y CSIRO (Australia). En la comparación entre ellas el CSIC alcanza datos muy favorables de todos los indicadores cuando se relacionan éstos con los recursos humanos y la financiación.

A continuación, Ramón López de Mántaras, del Instituto de Investigación en Inteligencia Artificial del CSIC, se ocupó del tema: «*Singularidades del área de Computer Science*». Demostró que aunque la evaluación bibliométrica está de moda y es más barata que la evaluación por pares, cuando se emplea en ciencias de la computación no resulta adecuada. Las evaluaciones se basan en las bases de datos ISI de Thomson, que analizan unas 350 revistas de informática, pero muy pocos proceedings de conferencias y congresos. La singularidad de la informática supone que la gran mayoría de sus publicaciones están recogidas en actas y proceedings de congresos, que son básicos para transmitir los nuevos conocimientos en esta especialidad y tienen enorme importancia, ya que son tan significativos como las buenas revistas, pero la mayoría de ellos no están incluidos en ISI.

Hay que tener en cuenta que una buena parte de la investigación en informática no produce «papers» sino dispositivos (software, sistemas, etc.), que tienen el mismo interés que los artículos para los profesionales. Además, se trata de un campo muy dinámico, y las revistas pueden tardar hasta 2 años en publicar un artículo, que puede quedar obsoleto en ese tiempo; en cambio, es habitual que en 6 meses se publiquen los proceedings de un congreso.

El ponente expuso un ejemplo de los ganadores de los prestigiosos premios Turing (premio de Ciencias de la Computación, otorgado por la Asociación para la Maquinaria Computacional, ACH), que recibieron muy pocas citas en ISI, y demostró que para ISI la disciplina informática se entiende, casi exclusivamente, como bioinformática.

Todo lo anterior demuestra que la cobertura para informática en ISI es muy escasa, habiendo otras bases de datos más adecuadas para evaluar los trabajos de esta disciplina, como son Google Scholar y Microsoft Academia Search, por ejemplo.

Itziar Laka, de la Universidad del País Vasco, habló sobre: «*Singularidades de las áreas de Humanidades y Ciencias Sociales*». Argumentó que no hay una unidad de medida de la ciencia, por lo que se pregunta qué deben reflejar los indicadores. Al medir la producción científica hay que tener en cuenta la singularidad de las humanidades y ciencias sociales, que agrupan campos muy diversos y prácticas académicas muy distintas, algunas más cercanas a las ciencias experimentales y otras muy alejadas de ellas.

Para la cuantificación de la calidad e impacto de la investigación en humanidades se imponen criterios importados de las ciencias naturales, lo que produce rechazo, dado que las herramientas bibliométricas habitualmente empleadas para las ciencias experimentales no funcionan igual para las humanidades, ya que no se han creado pensando en ellas, por lo que dichas disciplinas resultan infravaloradas. Un ejemplo es el caso de las enciclopedias, frecuentes en las disciplinas de humanidades, que no se pueden evaluar con los métodos tradicionales, pero cuya elaboración supone un enorme esfuerzo de investigación de grandes equipos, y un plazo temporal muy largo.

La ponente indicó algunos de los programas y bases de datos de índices de citas en los que ha trabajado (Publish or Perish, SCOPUS, etc.), pero resaltó, una vez más, su poca fiabilidad para humanidades y ciencias sociales. El Índice Europeo de Humanidades y Sociales (ERIH), clasifica las revistas según su calidad, pero no proporciona citas.

Por tanto, el reto es desarrollar métricas específicas con criterios de calidad fiables para las disciplinas de humanidades y ciencias sociales, dado que la evaluación por pares puede resultar sesgada.

A continuación, Carlos Miravittles del Instituto de Ciencias de Materiales de Barcelona (CSIC), presentó la ponencia: «*Una visión del evaluado en producción científico técnica*», e indicó la necesidad de la evaluación en la ciencia, aunque ello conlleve gastar mucho tiempo y recursos. Puede ser una evaluación individual o colectiva, aunque son muy distintas. Siempre se pueden evaluar mejor los «inputs» (financiación, recursos materiales y humanos) porque son tangibles, mientras que los «outputs» (progreso científico, formación, etc.) son más difíciles de valorar al ser intangibles.

Ofreció datos de su Instituto: recursos humanos, distribuidos por categoría científica y edad; artículos publicados por años; patentes y tesis doctorales dirigidas; impacto medio de las publicaciones; citas recibidas por año; índice h de los investigadores, y artículos en colaboración con otras instituciones.

Presentó tablas comparativas entre todos los institutos españoles del CSIC de Ciencias de Materiales (Aragón, Madrid, Barcelona, Sevilla y País Vasco), en relación con: número de investigadores de cada Instituto; presupuesto (sin contar salarios); contratos con compañías privadas; proyectos europeos concedidos en cuatro años; número de publicaciones, tesis doctorales y patentes, también en cuatro años. Para ganar precisión en los indicadores los relacionó con el número de investigadores de los institutos y el presupuesto de cada instituto.

Así, por ejemplo, se comprueba que en el año 2004, el Instituto de Aragón alcanza el porcentaje mayor de dinero recaudado por contratos con las empresas, en relación con su presupuesto, y en comparación con los otros institutos, a pesar de que la cifra absoluta recaudada por contratos es menor que la del instituto de Barcelona. Lo mismo ocurre con el número de publicaciones por investigador en el período 2000-2004: a pesar de que el Instituto de Madrid es el que tiene más publicaciones en número absoluto, es el instituto de Barcelona el que presenta mayor número de publicaciones por investigador.

Isidro Aguillo, del Centro de Ciencias Humanas y Sociales del CSIC, presentó la comunicación: *«Futuro de la publicación científica y su medida en la era web»*, indicando que las iniciativas de acceso abierto están ampliando la importancia de la web considerada la principal herramienta de comunicación científica, ya que permite el acceso inmediato a los artículos de calidad que han sido revisados por pares. Considera el ponente que la evaluación por pares seguirá siendo la principal característica de las revistas científicas y no se puede perder, aunque otros sistemas abiertos pueden proporcionar valor añadido a la evaluación. Indica que la publicación electrónica informal y no reglada adquirirá cada vez más importancia, aunque nunca suplantarán a las revistas científicas.

En relación con lo anterior, y según una normativa europea, los resultados de los proyectos financiados por la Unión Europea (Programas Marco) tienen que ser depositados en un repositorio de acceso abierto. Por tanto, para su evaluación, aparte de indicadores bibliométricos se pueden utilizar también los webométricos como el cómputo de las visitas, distribuidas geográficamente, por ejemplo, y de las descargas considerando los documentos más volcados.

Como indicador de evaluación propone sustituir el Factor de Impacto de la revista (citas esperadas) por el número total de citas recibidas por el artículo. En los trabajos de autoría múltiple propone hacer conteo fraccionado para cada autor individual. Sugiere también que para evaluar la excelencia (indicadores de excelencia) se exija que el artículo esté entre el 10 % de artículos más citados, incluso entre el 1 % en algunos casos, y para medir la diversidad (nuevos indicadores de diversidad), recomienda medir el conjunto de actividades llevadas a cabo por el investigador.

Dado que la Web of Science es muy costosa, y tiene muchos sesgos y limitaciones, para medir impacto y la visibilidad han surgido nuevas herramientas, como los motores de búsqueda Google y Yahoo, y las bases de datos Google Scholar Citations, Microsoft Academic Search, etc.

El ponente prevé la desaparición de la edición en papel de las revistas, por su alto costo, y sugiere que se incorporen a la versión electrónica de las mismas nuevos servicios de valor añadido, propios del entorno web, así como las versiones en idioma local e inglés.

A pesar de que la web es el futuro, la mayoría de los profesores e investigadores españoles no tienen página personal propia, y las universidades y centros de investigación españoles no incluyen la página web en sus planes estratégicos.

Por último, presentó un ranking web de universidades del mundo donde las primeras de la lista son norteamericanas.

A continuación, Alonso Rodríguez Navarro, de la Universidad Politécnica de Madrid, se centró en la ponencia: «*Evaluación con los índices Z y X de los trabajos muy citados*». Mantiene que para lo que podíamos llamar «ciencia normal» España ofrece buenos indicadores, comparables con los de otros países desarrollados. No es así para la «ciencia revolucionaria» o «de excelencia», que es mucho más importante y que, según el ponente, habría que medirla por el número de premios Nobel recibidos.

Si bien la ciencia revolucionaria no puede existir sin ciencia normal, ésta se puede mantener indefinidamente sin que de lugar a importantes descubrimientos. Hay que empezar a valorar a los científicos contabilizando sus trabajos que estén entre el 50 % y el 10 % de los más citados. Algunos propugnan que estén entre el 1 % de los más citados. Thomson predice con anticipación los futuros premios Nobel por los trabajos revolucionarios de excelencia científica que publican a lo largo de su carrera, que son, precisamente, los que tienen mayor impacto y reciben el número más alto de citas durante muchos años.

El ponente presentó los índices X y Z calculados para medir la excelencia en función del número de trabajos nacionales y el número de trabajos que están entre el 1 % más citado. Estos índices se correlacionan, en general, con el número de premios Nobel concedidos. En los ejemplos que puso, EE.UU., Suiza, Reino Unido, Alemania, Canadá e Italia, tienen altos y positivos los índices X y Z, mientras que España, China, Brasil y Corea, los tienen negativos. En cuanto a las instituciones, el Massachusetts Institute of Technology y las Universidades de Stanford, y Oxford y Cambridge, son las que tienen los índices X y Z más altos y ocupan los primeros lugares. En España, la Universidad de Barcelona presenta unos índices muy bajos y la Universidad Complutense de Madrid los tiene negativos.

Para finalizar la jornada, se constituyó una mesa redonda con todos los ponentes, presidida por Victoria Ley, que hizo un resumen del día e invitó a los asistentes a un debate abierto sobre lo tratado.

En el debate se constató que las agencias de evaluación utilizan los mismos indicadores para valorar todas las disciplinas, y el personal sometido a evaluación se adapta a lo establecido por las agencias, aunque le perjudique (caso de la utilización de las bases de datos ISI de Thomson para evaluar humanidades, sociales o informática). Se afirmó que siempre se considera imprescindible la evaluación por pares, pero es muy costosa y lenta. Los evaluadores de las agencias

no tienen tiempo para leer todos los artículos de los investigadores, por lo que hay que confiar en otros indicadores. En este sentido se constató la dificultad de evaluar los temas multidisciplinares, ya que es preciso disponer de expertos en varias materias. Por otra parte, se resaltó que, en general, se deberían agilizar los procedimientos de evaluación en la ANEP, ya que actualmente resultan muy tediosos.

Se anunció que Thomson va a lanzar el «Book Citation Index», sólo para colecciones en serie de libros y monografías, de ciencias, humanidades y sociales, que se unirá a la Web of Science. Aumentarán así tanto la visibilidad de los trabajos de investigación en humanidades y sociales, que se publican habitualmente como libros, como la posibilidad de obtener citas.

Se precisó que si en España no hay premios Nobel recientes es porque no se acometen trabajos de alto nivel de excelencia, y esto es así por falta de tiempo, entre otras razones, ya que los investigadores se ven presionados por la Administración para publicar la mayor cantidad posible de trabajos aunque sean de menor nivel, para obtener financiación. Además, para conseguir el Nobel es necesario partir de una base importante de muy buenos científicos en el país y promocionar decididamente a los que se supone pueden llegar a alcanzar el premio. Además, habría que hacer lobby en la Academia de Ciencias sueca.

Se sugirió la conveniencia de disponer anualmente y de una manera sistemática de datos oficiales de resultados de las investigaciones en España, lo más consensuados y fiables posible, y que fueran comparables entre todos los organismos implicados.

Enlaces de interés mencionados durante las sesiones:

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

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

<http://scholar.google.com/>

<http://www.eigenfactor.org/>

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

<http://www.oepm.es/>

<http://ep.espacenet.com/>

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

<http://www.webometrics.info/>

<http://academic.research.microsoft.com/>

Rosa Sancho y Ramón B. Rodríguez

Departamento de Publicaciones (CSIC)