

El «eigenfactor»: un nuevo y potente instrumento bibliométrico para evaluar la investigación¹

Antonio Villar

Universidad Pablo de Olavide (Sevilla)

Este trabajo se ocupa de describir un nuevo indicador bibliométrico recientemente aparecido, conocido como “el *eigenfactor*”, que permite medir la influencia de las publicaciones de forma mucho más aquilatada que el tradicional *índice de impacto*. Los rasgos distintivos de este indicador son: (a) Toma como referencia un periodo más amplio de citas admisibles (cinco años en lugar de dos). (b) Pondera la importancia de las citas recibidas por una revista por la importancia de las revistas que la citan. (c) Toma en cuenta los patrones de citas de los distintos campos ponderando por las citas medias de cada revista. (d) Genera una clasificación temática endógena y unívoca en 87 categorías, de modo que cada revista está en una y sólo una de las categorías. (e) Usa la misma base de datos del índice de impacto (más de 7.000 revistas del Journal of Citations Reports). (f) Es accesible de forma inmediata y gratuita a través de la web.

Palabras clave: Índices bibliométricos, eigenfactor, índice de impacto.

The “eigenfactor”: a new and powerful bibliometric tool. We discuss here new bibliometric index, recently appeared and known as “the eigenfactor”, that provides a measure of the influence of research articles much more accurate than the traditional *impact factor*. The key differential features of this new index with respect to the impact factor are: (a) It considers a five-year citation window (instead of two); (b) Citations are weighted by the relevance of the citing journal; (c) It takes into account the citation patterns of the different research fields; (d) It generates and endogenous classification into 87 different and exclusive fields; (e) It uses the same database that the impact factor (some 7,000 journals listed in the JCR); and (f) It is free and immediately accessible through the web.

Keywords: Bibliometric indices, eigenfactor, impact factor.

La evaluación de los resultados de la investigación se ha ido convirtiendo poco a poco en parte de nuestra cultura universitaria. Hemos recorrido un largo camino desde que la Agencia Nacional de Evaluación y Prospectiva (ANEP) comenzara la evaluación formal y sistemática de los proyectos de investigación, a principios de los años 80, hasta la actual actividad de la ANECA en los procesos de acreditación. Pasando por la

evaluación de los “tramos de investigación” por parte de la Comisión Nacional de Evaluación de la Actividad Investigadora (CNEAI). Este camino no ha estado presidido siempre por la comprensión de los investigadores ni por el acierto en el diseño de los mecanismos de evaluación.²

A pesar de todo, y aun cuando muchos estimamos que este recorrido está jalonado de incomprensibles pasos atrás –y el actual proyecto de estatuto del PDI es un buen ejemplo–, la evaluación de la investigación parece que ha venido para quedarse. Es una de esas tareas habituales y necesarias que siempre suscitan debate sobre su naturaleza

Fecha de recepción: 7-3-2011 • Fecha de aceptación: 8-4-2011
Correspondencia: Antonio Villar
Departamento de Economía
Universidad Pablo de Olavide
Carretera de Utrera, Km 1. 41003 Sevilla (España)
e-mail: avillar@upo.es

y alcance. Cuando la evaluación se refiere a colectivos amplios, se recurre con frecuencia al uso de medidas bibliométricas que proporcionen algún tipo de soporte a la estimación de la calidad de las publicaciones científicas, que son el elemento central de la evaluación de la investigación.

El objeto de estas páginas es describir un nuevo indicador bibliométrico recientemente aparecido, conocido como “el *eigenfactor*”, que permite definir un indicador de la influencia de las publicaciones que mejora sustantivamente el tradicional *índice de impacto*, a la hora de dar apoyo a la evaluación de la investigación. Para presentarlo haremos una recapitulación previa sobre las medidas bibliométricas y, en particular, sobre el índice de impacto que quedaría obsoleto en buena medida con este indicador.

El uso de medidas bibliométricas

El recurso a medidas bibliométricas puede ser una excelente ayuda para la evaluación indirecta, siempre que seamos conscientes del alcance y la naturaleza de lo que estas medidas suponen. Hay buenas razones para ello.

Cada año miles de investigadores en todo el mundo publican decenas de miles de artículos en las revistas científicas de su especialidad. Al escribir estos artículos toman decisiones sobre cuáles son las referencias bibliográficas pertinentes para apoyar su investigación, citando trabajos de otros investigadores publicados con anterioridad. Como consecuencia, las publicaciones científicas definen una amplia y compleja red de interrelaciones que se refleja en las citas cruzadas que aparecen en las referencias bibliográficas y en las notas a pie de página. Es el resultado de esas decisiones autónomas de miles de investigadores sobre qué trabajos resultan más relevantes en su campo de especialidad. Por tanto, esa red de citas contiene una información muy rica sobre la estructura mediante la que se generan y se difunden los nuevos conocimientos. Los indicadores bibliométricos tratan de extraer información de esta red para convertirla en un instrumen-

to de evaluación sobre la relevancia de las revistas científicas y de los artículos que en ellas se publican.

Para poder realizar una evaluación indirecta de la investigación a partir de las publicaciones hay que tomar dos decisiones previas. La primera se refiere a la selección del conjunto de revistas que constituyen nuestro universo de referencia porque cualifican como *publicaciones científicas*. La segunda consiste en la determinación de la importancia relativa (relevancia) de estas revistas mediante la selección de un criterio de comparación concreto. Desde luego ambas decisiones no son independientes. En particular, si queremos recurrir al uso de criterios bibliométricos para la clasificación de revistas deberemos elegir aquellas de entre los listados que disponen de este tipo de mediciones.

La fuente más completa de mediciones bibliométricas para publicaciones científicas es el *Journal of Citations Reports* (JCR en lo sucesivo). El JCR es una publicación del *Institute for Scientific Information* (ISI) de Filadelfia, que cada año proporciona los datos bibliométricos básicos de las principales revistas científicas en todos los campos del saber. El *Institute for Scientific Information* agrupa las publicaciones consideradas en tres grandes áreas: Ciencias, Ciencias Sociales y Humanidades, y luego divide estas grandes áreas en campos del saber (*subject fields*). Se trata de una base de datos *abierto* de la que entran y salen publicaciones según el cumplimiento o incumplimiento de ciertos patrones (evaluación anónima, internacionalidad y regularidad, por citar las tres más importantes).³

Los dos procedimientos usuales para determinar la relevancia de las revistas son las *opiniones de expertos* y el uso de uno o varios *indicadores bibliométricos*. En ocasiones se combinan ambas aproximaciones en un proceso iterativo (el conocido como “método Delphi”), con objeto de afinar algunas valoraciones e incorporar en las clasificaciones revistas sobre las que no se dispone de este tipo de datos. El uso de clasificaciones de revistas basadas en indicadores bibliométricos

tricos es mayoritario en casi todos los campos del saber y es la aproximación que seguiremos aquí. Aunque también se han elaborado valoraciones basadas únicamente en la percepción de los investigadores, hay evidencia de una correlación positiva estadísticamente significativa entre este tipo de ranking y el derivado de los indicadores bibliométricos (Axarloglou y Theoharakis, 2003; Masson, Steagall y Fabritius, 1997).⁴

El índice de impacto

El *Journal of Citation Reports* ofrece una variada colección de indicadores que ayuda al propósito de clasificar la relevancia de las revistas científicas en los diversos campos de especialidad, a partir del cómputo del número de total de citas recibidas por cada revista y del número de artículos publicados. Los principales indicadores son el “índice de impacto”, el “total de citas”, el “índice de vida media” y el “índice de inmediatez”. Cada uno extrae información de la matriz de citas cruzadas, pero lo hace enfocando diferentes aspectos.

De todos estos indicadores el *índice de impacto* es sin duda el más conocido y utilizado, no siempre con la prudencia necesaria. En algunas disciplinas el índice de impacto se toma, contra la advertencia explícita del propio *Institute for Scientific Information*, como el *indicador* de calidad por antonomasia. Como señalan Amin y Mabe (2000), el índice de impacto es seguramente la medida bibliométrica más utilizada y peor entendida de todas ellas. Porque hay que tener en cuenta que se trata de un indicador muy sensible al campo de especialidad, con una variabilidad aleatoria importante de año a año.⁵

Definición y características

El *índice de impacto* es un indicador del promedio de citas que, a corto plazo, reciben los artículos de una determinada revista en relación con el total de artículos publicados. Es importante especificar que el número de citas consideradas se refiere a las recibidas por esa revista en el año tomado como refe-

rencia, relativas a trabajos publicados en los dos años anteriores. Estamos así identificando así el “impacto” de una revista con el seguimiento de sus publicaciones en los dos años inmediatos a la aparición de los trabajos.

La fórmula de este índice, relativa a un año concreto, puede describirse como sigue: se suman todas las citas válidas recibidas en ese año por la revista cuyo impacto queremos medir, por parte de todas las demás revistas del grupo, y ese número se divide por el total de artículos publicados por esa revista en ese año. Por “citas válidas” se entien- de las que se refieren a trabajos publicados en los dos años anteriores al que tomamos como referencia y excluyendo las auto-citas. Así pues el índice de impacto mide la relevancia de una revista como el promedio de citas que recibe en relación al número de artículos que publica.

Aunque se trata de una idea muy razonable, esta construcción presenta una serie de limitaciones que hay que tener en cuenta a la hora de usar este criterio como indicador de la calidad de las publicaciones. La simple inspección de los valores medios de los índices de impacto en los diferentes campos del saber muestra una notable disparidad. La mediana de estos índices tiene un rango de variación de uno a seis, aproximadamente. De modo que mientras en Ciencias de la Vida un índice de impacto de 4 no llama la atención, en el campo de las Ciencias Sociales un índice de 2 resulta una cifra excepcional.

Las tradiciones de publicación en los diferentes campos del saber explican en buena medida estas diferencias. Se ha comprobado, por ejemplo, que el número medio de autores de los artículos afecta de manera estadísticamente significativa a valor medio de los índices de impacto (a más autores, más citas y mayores valores medios de los indicadores). Es también obvio que el número medio de citas que se realiza en cada campo del saber afecta al valor medio del índice. Junto a estos aspectos encontramos la inevitable variabilidad estadística, que hace que los índices de impacto puedan cambiar de un año a otro, la existencia de “modas” (hot topics) o “redes” de citas, etc.

Hay otros elementos adicionales que afectan a los valores de este indicador que merecen un comentario más detallado. Son los siguientes.

La “ventana” por la que miramos

Uno de los aspectos más importantes que afectan a los valores del índice de impacto se refiere al periodo tomado como referencia para computar las citas, aspecto al que a veces se alude como el tamaño de la “ventana” por la que miramos. El índice de impacto computa únicamente las citas relativas a trabajos publicados en los dos años anteriores a los que se refiere el cómputo.

Esto característica del índice afecta de manera sustancial a los resultados por dos razones complementarias.

Primera: Los patrones de difusión del conocimiento tienen velocidades distintas en los diferentes campos del saber. Suelen ser muy rápidos en las ciencias de la vida y mucho más lentos en las ciencias sociales. Pero conviene no confundir esto como una diferencia entre “ciencias y letras”. Un artículo publicado en una de las revistas líderes en el campo de la biología es citada en promedio 20 veces en los dos años sucesivos a su publicación. Un artículo publicado en una de las revistas líderes en matemáticas recibe una media de 2 citas en el mismo periodo. En la biología el mayor número de citas se produce en los tres años sucesivos a la publicación. En matemáticas entre el tercero y el sexto.

Segunda. El periodo medio que requiere la publicación de un trabajo de investigación desde que es remitido a la revista hasta que aparece impreso. En algunos campos del saber este periodo puede ser de tres meses mientras que en otros supera los dos años. En este último caso, buena parte de las citas que aparezcan en el artículo resultarán no computables porque se han quedado fuera de la ventana de cálculo.

Por ello se ha ido generalizando en los últimos años el uso complementario de un índice de impacto relativo a una *ventana de cinco años*. Es decir, computando las citas a los artículos publicados hasta cinco años an-

tes del periodo en que se calcula el índice. Este nuevo indicador, que muchas revistas toman ya como referencia, cambia sustancialmente la medida de la relevancia de las publicaciones en aquellos ámbitos en que la velocidad de difusión es menor y/o la rapidez en el procesamiento de originales más reducida (por ejemplo en el campo de la Economía).

No todas las citas valen igual: Indicadores recursivos

Uno de los aspectos más discutibles que presenta el uso del índice de impacto para la valoración de las revistas científicas es que todas las citas admisibles tienen el mismo valor, con independencia de la importancia de la revista que cita. Por ello se han elaborado indicadores más sofisticados que aplican el principio de que “la importancia de una cita depende de la relevancia de la revista que la realiza”. En consecuencia, dos revistas con idéntico número de citas y de artículos publicados serán clasificadas de forma distinta si una de ellas es citada mayoritariamente por revistas “importantes” y la otra no.

Para determinar una clasificación de las revistas aplicando este principio se requiere un procedimiento iterativo que permita, partiendo de una distribución inicial de citas, ir ajustando el impacto por la clasificación de la revista que cita, lo que a su vez modifica la relevancia de las revistas y de las citas que proporcionan (básicamente se trata de calcular el índice de cada revista mediante un algoritmo de punto fijo).⁶

La fórmula de elaboración de estos índices puede describirse en los siguientes términos. Partimos del valor del índice de impacto en la primera iteración. En la segunda iteración recalculamos el índice de impacto pero ahora cada cita tiene un valor determinado por el índice de impacto. Obtenemos así un nuevo índice de impacto modificado al dar pesos distintos a las citas, según el índice de impacto inicial de la revista que cita. La tercera iteración repite el procedimiento tomando como referencia los valores del índice de impacto obtenidos en la segunda ite-

ración. Y así sucesivamente hasta que nuevas iteraciones no alteran más los valores de los índices de impacto modificados. El proceso iterativo converge siempre (dada la naturaleza de la matriz de datos) y suele hacerlo muy rápidamente.

Además del aspecto “relevancia”, estos índices recursivos pueden introducir algunas correcciones adicionales, como la ponderación por el *promedio de citas realizadas* (se pondera el total de citas que una revista realiza en promedio para ajustar su influencia: cuantas menos citas efectúe una revista, más importante resulta ser citado por ella).

Este tipo de indicadores proporcionan valoraciones muy aquilataadas y acordes con la apreciación de los investigadores de la relevancia de las diferentes revistas. Hasta ahora tenían el inconveniente de que “no se dejaban manejar” fácilmente por los investigadores, dado el costoso proceso de extracción de información, elaboración y cálculo requerido. Esta dificultad impedía, entre otras cosas, su aplicación a amplios listados de revistas y su actualización.⁷

Varias clasificaciones

Otro de los inconvenientes que surge frecuentemente al usar el índice de impacto deriva del hecho de que hay revistas que figuran en más de un campo del saber y que estos campos del saber presentan características diferenciadas, tanto en términos de valores medios como del propio tamaño del campo considerado (número de revistas que agrupa). Como consecuencia, podemos encontrarnos con que la valoración relativa del impacto de la revista puede cambiar sustancialmente según el campo en el que la consideremos. Análogamente, la revista en cuestión puede ocupar un puesto destacado en el ranking en un campo y uno muy discreto en otro.

Dado que la asignación de revistas a uno o varios campos es producto de una convención y que los campos son muy heterogéneos, estas adscripciones múltiples generan dificultades en la evaluación.

A modo de resumen: el uso de los índices de impacto como estimaciones de la ca-

lidad de la investigación en la valoración de individuos, grupos o instituciones, o en la realización de un ranking de revistas científicas debe tomarse con cautela.

El “eigenfactor”: la influencia de las revistas y de los artículos

Recientemente han comenzado a publicarse con regularidad unos indicadores recursivos, elaborados a partir de los mismos datos que usa el índice de impacto (revistas del catálogo del Journal of Citations Reports), que tienen importantes ventajas con respecto a los indicadores convencionales. Son también indicadores bibliométricos basados en el cómputo del número de citas, pero incorporan mejoras sustantivas que resultan especialmente relevantes en aquellos campos en los que el proceso de publicación es largo y la velocidad de difusión lenta. Se trata de los indicadores vinculados al concepto de eigenfactor, disponibles de forma rápida y sencilla en la página web: <http://www.eigenfactor.org/>.

Hay dos índices diferenciados, obtenidos a partir de los mismos datos. Uno se refiere a la *influencia de la revista* (el *eigenfactor score*, *EF*). El otro mide la *influencia de un artículo de la revista* (el *article influence score*, *AI*), que simplemente consiste en ponderar el valor anterior por el número de artículos publicados. En ambos casos la citada página web proporciona información tanto de los valores numéricos de estos índices para cada revista como del lugar que ocupan (percentil) en la clasificación en su campo del saber (véanse los gráficos 1 y 2 más adelante). Haremos primero una breve descripción de estos índices y expondremos luego algunos detalles de su construcción.

El índice que mide la influencia de la revista, el *eigenfactor* propiamente dicho, *EF*, toma en cuenta las citas totales recibidas por cada revista en el año de referencia, relativas a las publicaciones que caen en la ventana predeterminada, ponderadas por el número medio de citas de cada una de esas revistas y ajustadas por su importancia (el propio índice de influencia de las revistas que citan).

Nótese que no hay aquí una ponderación por el número de artículos publicados, a diferencia de lo que ocurre con el índice de impacto, de modo que revistas que publiquen más tenderán a tener índices más altos (esa ponderación se introduce, precisamente, en el otro indicador). Para fijar la escala de este indicador se sigue el criterio de dar valor 100 a la suma de valores *EF* de todas las revistas del catálogo.

El *índice de influencia de un artículo* de una revista, *AI*, corresponde a la estimación de la relevancia de una revista ponderada por el número de artículos que publica (el *eigenfactor* de cada revista dividido por el número de artículos publicados en el año de referencia). Para este índice se fija una escala distinta: se toma la media de este indicador (entre todas las publicaciones del catálogo) igual a la unidad. De este modo, si el índice de influencia de un artículo toma el valor 4 significa que cada artículo publicado en esas revistas tiene en promedio una influencia cuatro veces mayor que la media de los artículos publicados en todas las revistas.

Este segundo indicador es referente a tomar en cuenta en la evaluación porque evita tratar por igual revistas que publican poco más de una docena de artículos por año con otras que publican centenares. Se trata de la alternativa metodológica al índice de impacto (ya que éste sí toma en cuenta la diferencia en el tamaño de las revistas, como hemos visto).

Dado que el índice de influencia de un artículo es una medida derivada del *eigenfactor*, nos centraremos en la elaboración de este último para ilustrar las mejoras que supone este tipo de aproximación.

Una ventana más grande

El eigenfactor toma en cuenta las citas que cada artículo realiza de los trabajos aparecidos en los cinco años anteriores a su publicación. Este aspecto es clave dado que el índice de impacto sólo computa las citas a los artículos aparecidos en los dos años anteriores, lo que desvirtúa sustancialmente su sentido en aquellos campos en los que el proceso de publicación es largo o la difusión de la investigación se produce con lentitud.

Cada oveja con su pareja

El eigenfactor genera una clasificación temática endógena, excluyente y exhaustiva. Las revistas aparecen clasificadas de forma que cada una pertenece a un único campo del saber, a diferencia de lo que ocurre con el índice de impacto donde encontramos revistas que figuran en distintos listados, con diferentes órdenes de relevancia. Esta clasificación, que consta de un total de 87 categorías, se genera de forma endógena a partir de la red de relaciones establecida entre las distintas revistas. Como se señala en su página web, “esta clasificación por campos científicos no parte de nociones preconcebidas sobre cuál debe ser la estructura de estos campos del saber, sino que dicha clasificación resulta determinada por los propios datos –los patrones de citas–. Se clasifican los campos científicos de acuerdo a lo que los investigadores hacen, no de acuerdo a lo que dicen o con qué campos se identifican”.

La página web proporciona las valoraciones de las revistas por campos del saber, de modo que la posición relativa de cada revista en su campo está determinada de forma unívoca. Los datos se refieren tanto a los valores numéricos de los indicadores como al *percentil que ocupan dentro de su campo*.⁸ (Véase el Gráfico 1 más adelante).

No todas las citas valen lo mismo ni todos los campos son iguales

Se pondera la relevancia de las citas recibidas en función de la importancia de la revista que cita y se toman en cuenta los diferentes patrones de citas de los distintos campos del saber. Como en el caso del índice de impacto recursivo que hemos descrito con anterioridad, la relevancia de la revista se determina mediante un algoritmo en el que el valor de la revista depende del valor de las revistas que la citan. Es interesante indicar que este procedimiento puede describirse también como el resultado de aplicar un sencillo conjunto de axiomas que determina este tipo de solución (Palacios y Volij, 2004).

En este caso el punto de partida viene dado por los valores de citas que cada revista recibe en relación al total de citas que realizan las revistas que la citan. De este mo-

do se incorpora directamente en el procedimiento de cálculo la idiosincrasia de los distintos campos del saber, que se refleja en los valores medios de citas que producen.

Revistas y más

El *eigenfactor* usa una base de cómputo mucho más amplia que el índice de impacto. Aunque todos los datos sobre las citas se toman del *Journal of Citations Reports*, que incluye información completa de unas 7.000 revistas, se efectúan evaluaciones de más de

100.000 ítems de los que aparecen citados en estas revistas.

Gratis, inmediato y contrastable

Un aspecto importante de estos indicadores es que están disponibles *on line* de forma gratuita, de modo que resulta inmediato hacer comparaciones entre revistas y seguir su evolución en el tiempo (incluyendo la comparación con el índice de impacto).⁹

Los gráficos 1 y 2 ilustran el tipo de información que puede obtenerse y la forma

eigenfactor.org - ranking and mapping scientific journals

02/03/11 14:14

Journal Name	Percentile	Eigenfactor™ Score	Article Influence™ Score
1. ANNUAL REVIEW OF ANTHROPOLOGY ISSN: 0084-6570	EF [62.48] AI [89.25]	0.00458364	1.65455
2. EVOLUTIONARY ANTHROPOLOGY ISSN: 1060-1538	EF [51.14] AI [89.31]	0.00297534	1.47528
3. CURRENT ANTHROPOLOGY ISSN: 0011-3204	EF [71.56] AI [86.83]	0.00680841	1.31283
4. JOURNAL OF HUMAN EVOLUTION ISSN: 0047-2484	EF [78.32] AI [86.58]	0.00983263	1.29725
5. PUBLIC CULTURE ISSN: 0899-2363	EF [47.72] AI [75.60]	0.00263031	0.899109
6. AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY ISSN: 0002-9483	EF [82.68] AI [72.66]	0.0125815	0.829956
7. JOURNAL OF MODERN AFRICAN STUDIES ISSN: 0022-278X	EF [41.31] AI [69.57]	0.00206045	0.768343
8. CULTURAL ANTHROPOLOGY ISSN: 0886-7356	EF [33.69] AI [64.48]	0.00153423	0.685404
9. AMERICAN ETHNOLOGIST ISSN: 0094-0496	EF [48.87] AI [64.47]	0.00274939	0.685387
10. AMERICAN JOURNAL OF HUMAN BIOLOGY ISSN: 1042-0533	EF [66.92] AI [64.26]	0.00553219	0.682017

Gráfico 1. Las 10 revistas más importantes de antropología

NEW ENGLAND JOURNAL OF MEDICINE		
ISSN: 0028-4793		
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008		
ISI Field Designations	MEDICINE, GENERAL & INTERNAL	Percentiles
Group	Science	<i>Eigenfactor</i> Score Overall <input type="text" value="99.90"/>
<i>Eigenfactor</i> ™ Category	MEDICINE	<i>Article Influence</i> Score Overall <input type="text" value="99.93"/>
<i>Eigenfactor</i> Score	0.680294	<i>Impact Factor</i> Overall <input type="text" value="99.99"/>
<i>Article Influence</i> ™ Score	18.7626	
Total Articles	1636	
Publisher	MASSACHUSETTS MEDICAL SOC/NEJM	
Year First Published	0	
Profit Status	Non-Profit	
Price	3400	
Price per <i>Eigenfactor</i>	4733.4	

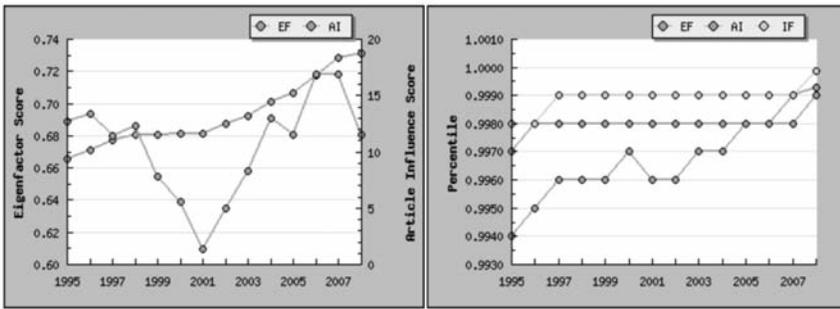


Gráfico 2. Información detallada sobre el New England Journal of Medicine

en que es presentada en la página web del *eigenfactor.org*.

El gráfico 1 nos da una parte de la pantalla que aparece cuando le pedimos información sobre el área de antropología. En la parte superior derecha figura el número de revistas que incluye este campo (65). Debajo aparece la información de las revistas articulada en cuatro columnas. En la primera figura el nombre de la revista junto con su número de identificación en el registro ISSN. En la segunda columna aparece una doble entrada para cada revista, incluyendo los valores de los percentiles que ocupan en la distribución dentro de este campo del saber en términos del *eigenfactor* y del *índice de influencia de los artículos*. Las columnas tercera y cuarta proporcionan los valores del *eigenfactor* y del *índice de influencia de los artículos* publicados en esas revistas. Puede observarse que la ordenación corresponde al valor de este último indicador.

El gráfico 2 se obtiene cuando hacemos click sobre una revista concreta (columna de la izquierda de la pantalla anterior). Se abre una nueva pantalla que nos da información mucho más detallada sobre esa revista, desde la evolución temporal de los indicadores *EF*, *AI* y su comparación con el índice de impacto, *IF* (en términos de percentiles), hasta el número de artículos que publica al año o su precio. En este caso hemos “pinchado” sobre una revista del área de medicina (New England Journal of Medicine).

Comentarios finales

El *índice de influencia de los artículos* es un indicador bibliométrico obtenido a partir del cálculo del *eigenfactor* que supone un criterio de evaluación de revistas científicas mucho más rico y consistente que el tradicional índice de impacto. Entre las ventajas que presenta esta nueva medida cabe destacar las siguientes:

- Toma como referencia un periodo más amplio de citas admisibles (cinco años en lugar de dos).
- Pondera la importancia de las citas recibidas por una revista por la importancia de las revistas que la citan.
- Toma en cuenta los patrones de citas de los distintos campos ponderando por las citas medias de cada revista.
- Genera una clasificación temática endógena y unívoca en 87 categorías, de modo que cada revista esté en una y sólo una de las categorías.
- Usa la misma base de datos del índice de impacto (más de 7.000 revistas del Journal of Citations Reports).
- Es accesible de forma inmediata y gratuita a través de la web.

Una llamada de atención

A pesar de todas estas ventajas que hacen de estos indicadores poderosos instrumentos para la evaluación, conviene no perder de vista que la evaluación de la investigación es un proceso que requiere decisiones académicas y no sólo administrativas, porque tiene mucho de cualitativo y aproximado.

Como indicaba en un trabajo anterior (Villar, 2003), la investigación científica es una actividad multidimensional cuyo output está compuesto fundamentalmente por las publicaciones científicas, pero también por otras tareas como la dirección de Tesis Doctorales y proyectos de investigación, las labores de evaluación de investigación, la edición de publicaciones científicas, etc. La forma precisa de evaluar la investigación dependerá de cuál sea el propósito del análisis y el ámbito de referencia. En particular si se trata de individuos, grupos o instituciones, de cuántos investigadores y campos de investigación haya implicados en la evaluación, qué bases de datos estén disponibles, y qué objeto tenga el ejercicio de valoración.

La evaluación de la investigación de un individuo aislado por un grupo de especialistas de su mismo campo se realiza generalmente mediante una *valoración directa*. El uso de clasificaciones de revistas o indica-

dores bibliométricos puede servir como apoyatura, pero lo esencial es el juicio de los especialistas que pueden analizar su currículo, leer su obra, escuchar una exposición de su trabajo o de sus líneas de investigación y evaluar así su relevancia. Es lo que sucede, por ejemplo, en la valoración de Tesis Doctorales y en la evaluación de los currícula de los jóvenes Doctores que compiten por entrar en los Departamentos que reclutan sus miembros competitivamente en el mercado (típicamente los Departamentos de Universidades Norteamericanas, pero también muchas europeas y algunas españolas).¹⁰

Conforme la evaluación abarca más individuos y más temas de investigación, el recurso a la *valoración indirecta* se va haciendo más necesario. Sucede así en la valoración de candidatos a acreditaciones o concursos de plazas de plantilla en España, donde una comisión tiene que evaluar uno o varios investigadores cuyos campos de especialidad pueden ser diversos. Desde luego la evaluación de la investigación a través de la relevancia asignada a las revistas científicas resulta prácticamente inevitable cuando se tiene que evaluar instituciones o grupos de investigación. Aunque no sea el único elemento a considerar.

En este contexto estos nuevos indicadores bibliométricos suponen una importante ayuda para la evaluación porque permiten aplicar criterios muy aquilatados para juzgar la relevancia de las contribuciones científicas. Junto a otros, como el “índice H” (el mayor número x de trabajos de un autor que son citados al menos x veces) o las citas recibidas por cada contribución individual, que pueden obtenerse directamente de diversas fuentes (v.g. Scholar Google).

Disponer de instrumentos bibliométricos consistentes para formar un juicio en el proceso de evaluación es sin duda una gran ventaja que puede dar solidez y transparencia a los procedimientos. Pero sin olvidar que la evaluación es un juicio informado elaborado por especialistas y no la aplicación automática de un baremo que pueda hacer un programa informático. Como señala Alan Fersht (2009):

El legado terrible del índice de impacto es que está siendo usado para evaluar científicos, en lugar de revistas, lo que supone una preocupación creciente para muchos de nosotros. El juicio sobre los individuos debe derivar, obvio es decirlo, de un análisis en profundidad realizado por investigadores expertos en su área. Sin embargo, muchos burócratas prefieren el uso de un simple baremo. Mi experiencia como evaluador en comités internacionales me dice que el recurso al índice de impacto es más frecuente cuanto menos conocimiento se tiene para realizar una evaluación científica independiente. (p. 6883)

Finale: ¿Eigen qué?

La expresión *eigenfactor* alude al término matemático de autovector de una matriz (del alemán *eigenvector*). Dada una matriz cuadrada A de orden n , se busca el vector n -dimensional x que verifique la ecuación $Ax = \lambda x$, para algún número λ . Este vector tiene propiedades particulares relacionadas con la estabilidad de un sistema dinámico asociado (convergencia del proceso iterativo) y con la invarianza de la transformación lineal deri-

vada de la ecuación que la define (el vector x es transformado por la matriz A en un vector idéntico, salvo un cambio de escala).

En el contexto de los indicadores bibliométricos, la matriz A describe la distribución de las citas de cada revista entre todas las demás (con ponderaciones relativas a las citas totales de cada revista). El autovector propio de la matriz de citas determina de forma endógena el valor que se da a las citas de cada revista, en función de su posición en el ranking (que va cambiando al cambiar el peso asociado en cada iteración). Véase el Apéndice para mayor detalle.

El tipo de algoritmo que usa el eigenfactor tiene muchos elementos en común con el que usa Google para ordenar las páginas web en su buscador. En ambos casos se han hecho algunas críticas de base muy técnica sobre la robustez del algoritmo para los casos de citas muy bajas (Serrano, 2004).

Hay además, obvio es decirlo, otras propuestas valorativas de interés de las que no podemos ocuparnos aquí. Mencionemos en todo caso las desarrolladas por Albarrán, Crespo, Ortuño y Ruíz-Castillo (2010) y Albarrán, Ortuño y Ruíz-Castillo (2011a, 2011b).

Referencias

- Albarrán, P., Crespo, J., Ortuño, I. y Ruíz-Castillo, J. (2010). A Comparison of the Scientific Performance of the US and the European Union at the Turn of the XX1st Century. *Scientometrics*, 85, 329-344.
- Albarrán, P., Ortuño, I. y Ruíz-Castillo, J. (2011a). High and Low-impact Citation Measures: Empirical Applications. *Journal of Infometrics*, 5, 122-145.
- Albarrán, P., Ortuño, I. y Ruíz-Castillo, J. (2011b). The Measurement of Low- and High-impact in Citation Distributions: Technical Results. *Journal of Infometrics*, 5, 48-63.
- Amin, M. y Mabe, M. (2000). Impact Factors: Use and Abuse. *Perspectives in Publishing*, 1, 1-6.
- Axaroglou, K. y Theoharakis, V. (2003). Diversity in Economics: An Analysis of Journal Quality Perceptions. *Journal of the European Economic Association*, 1, 1402-1423.
- Fersht, A. (2009). The Most Influential Journals: Impact Factor and Eigenfactor. *Proceedings of the National Academy of Sciences*, 106, 6883-6884.
- Kalaitzidakis, P., Mamuneas, P. T., y Stengos, T. (1999). European Economics: an Analysis Based on Publications in the Core Journals. *European Economic Review*, 43, 1150-1168.
- Laband, D. N. y Piette, M. J. (1994). The Relative Impacts of Economics Journals: 1970-1990. *Journal of Economic Literature*, 32, 640-666.
- Liebowitz, S.J. y Palmer, J. C. (1984). Assessing the Relative Impacts of Economics Journals. *Journal of Economic Literature*, 22, 77-88.
- Masson, P., Steagall, J. y Fabritius, M. (1997). Economics Journal Rankings by Type of School: Perception versus Citations. *Quarterly Journal of Business and Economics*, 36, 69-79.
- Palacios-Huerta, I. y Volij, O. (2004). The Measurement of Intellectual Influence. *Econometrica*, 72, 963-977.

- Rosvall, M. y Bergstrom, C. T. (2007). An Information-Theoretic Framework for Resolving community Structure in Complex Networks. *Proceedings of the National Academy of Sciences*, 104, 7327-7331.
- Serrano, R. (2004). The Measurement of Intellectual Influence: The Views of a Sceptic. *Economics Bulletin*, 1, 1-6.
- Villar, A. (2003). La Evaluación de la Investigación en Economía. *Revista Valenciana de Economía y Hacienda*, 8, 97-133.

Notas

- ¹ Para quienes estén interesados en el detalle formal de este indicador, existe una versión que describe las ecuaciones correspondientes en la página web del autor, http://web.me.com/antonio.villar/Sitio_web/Welcome.html
- ² La propia creación de la CNEAI a finales de los años 80, como un organismo diferenciado de la ANEP, surge en parte de la necesidad de preservar esta última frente a la resistencia que existía en muchos colectivos de profesores que se negaban a ser evaluados. El riesgo de que la creación de estos complementos de productividad pudiera generar una presión insostenible que se llevara por delante lo ya conseguido por la ANEP, llevó a la creación de esta nueva institución. En cuanto a los problemas de diseño, baste señalar que ninguno de los dos Premios Nobel de Física de este año habría conseguido acreditarse en nuestro país ni siquiera como profesor titular de Universidad, porque les faltaría gestión, cursos de innovación docente y horas de clase (y, por supuesto, carecerían de todos los papeletos que certifican su actividad en congresos y reuniones científicas).
- ³ Para una descripción del proceso de selección de revistas que entran en esta base de datos, véase <http://www.isinet.com/isi/hot/essays/1999701>.
- ⁴ En el año 2003 el *Journal of the European Economic Association* dedica un número monográfico a la evaluación de la investigación económica en Europa, que incluye numerosos ejercicios clasificatorios de revistas, cuya consulta recomendamos. Los economistas han hecho muchas contribuciones sobre estos temas, que resultan de aplicación general.
- ⁵ Se advierte aquí un aspecto de *profecía que se auto-realiza*, en el siguiente sentido. Si en una disciplina se toma el índice de impacto como el criterio de valoración esencial, entonces los mejores investigadores tenderán a publicar sus trabajos punteros en las revistas de mayor impacto, lo que reforzará el impacto de estas revistas. Y ocurrirá lo contrario cuando el índice de impacto no sea un referente importante.
- ⁶ Véase el trabajo de Kalaitzidakis, Mamuneas y Stengos (1999). Este índice supone una extensión y actualización del elaborado en Laband y Piette (1994), siguiendo la metodología propuesta anteriormente por Liebowitz y Palmer (1984).
- ⁷ Para obviar esta dificultad se han construido algunos indicadores combinados que se aplicaron a la evaluación de Departamentos universitarios. Se basa en la combinación de algunos de los índices básicos que proporciona el JCR. Por ejemplo, un indicador consistente en multiplicar el índice de impacto por el número de citas totales recibidas por cada revista en el año tomado como base, y normalizando al valor 100 la revista número uno del ranking, usado para evaluar los Departamentos de Economía en Bélgica. Véase la discusión en Villar (2003).
- ⁸ Para quienes prefieran referirse a la clasificación del JCR existe la posibilidad de buscar ese tipo de agrupación en la búsqueda avanzada.
- ⁹ Todos los algoritmos usados para hacer los cálculos están descritos con detalle en los documentos técnicos y son abiertos y replicables -véase Rosvall y Bergstrom (2007)-.
- ¹⁰ En ciertos casos los indicadores bibliométricos pueden servir para dar pistas a los jóvenes investigadores sobre el nivel de requerimiento exigido para una determinada actividad, en términos de publicaciones. En otras palabras, puede servir como indicación de las revistas objetivo para publicar los resultados de la investigación. Pero en general este conocimiento se debe adquirir en el proceso de formación.