

Breve introducción a la bibliometría

Dr. Jordi Ardanuy

Departament de Biblioteconomia i Documentació

Universitat de Barcelona

Barcelona, abril 2012



Contenido

Contexto y definición	3
Producción científica y bases de datos bibliográficas	5
Leyes bibliométricas	9
Ley de la productividad de los autores	9
Ley de dispersión de la bibliografía científica.....	12
Ley de crecimiento exponencial	14
Ley de obsolescencia de la bibliografía científica	15
Indicadores bibliométricos	16
Indicadores personales	16
Indicadores de producción	16
Indicadores de dispersión	17
Indicadores de visibilidad o impacto	17
Indicadores de colaboración	21
Indicadores de obsolescencia.....	23
Indicadores de forma y contenido	23
Referencias	24

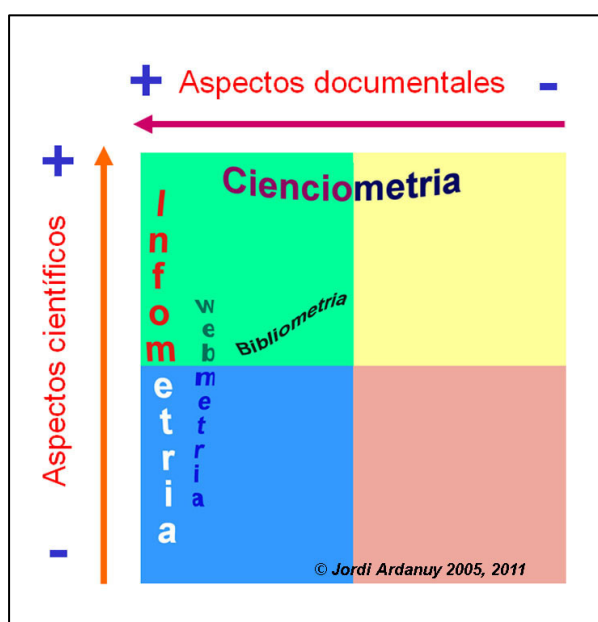
Contexto y definición

La ciencia es una actividad intelectual que tiene por finalidad dar respuesta a preguntas y desde una perspectiva más práctica, la resolución de problemas. Como consecuencia, el grado de desarrollo científico se manifiesta en la profundidad de las preguntas o problemas con los que enfrenta. La manera sistemática de dar respuesta satisfactoria a los problemas planteados es la investigación que, por tanto, resulta inextricablemente ligada a toda actividad científica, y cuyos resultados posibilitan el crecimiento del caudal de conocimiento científico.

La ciencia no puede considerarse una actividad independiente y desligada de los procesos sociales. Quizás se podría haber contemplado de esta manera en las incipientes etapas, en las que era un divertimento intelectual y no tenía ninguna incidencia práctica en la sociedad. Pero a partir del siglo XVII el desarrollo científico perdió esa inocencia, de modo que sus aplicaciones fueron fundamentales en la revolución industrial. A partir de ese momento, la realidad socioeconómica y la ciencia interactúan influenciándose mutuamente. A partir del siglo XIX, el fuerte desarrollo industrial que se produce en los países llamados occidentales convierten la ciencia y su derivada, la tecnología, en una actividad social clave, vinculándola al proceso productivo de una manera rotunda, al mismo tiempo de su incorporación al sistema educativo (Merton 1977; Medina 1982; Merton 1992).

Aunque uno de los puntales de la ciencia es el uso de técnicas cuantitativas, hasta tiempos relativamente cercanos no comenzó a aplicarse para estudiar su naturaleza y realidad social. La medida de magnitudes sociales como: los presupuestos científicos, la cantidad de investigadores, las publicaciones científicas, etc., precisan de una técnica de

análisis sociológico cuantitativa que corresponden a la disciplina de la **cienciometría**, aunque no existe unanimidad en el uso de tal término. La **bibliometría**, por su parte, se centra esencialmente en el cálculo y en el análisis de los valores de lo que es cuantificable en la producción y en el consumo de la información científica (López Piñero 1972; Spinak 1996).



Los estudios bibliométricos no han estado exentos de ver cuestionada su validez. Buena parte de estas críticas han sido motivadas por carencias en la preparación científica y estadística de los investigadores que los han realizado: utilización de técnicas cuantitativas demasiado elementales; falta de un planteamiento previo que justifique la utilización de un método cuantitativo concreto; ausencia de valoración de los datos de las que se parte y de las bases de datos utilizadas... Otras críticas tienen su origen en el abuso que se ha hecho de los resultados de algunos estudios bibliométricos utilizados como único criterio válido a la hora de valorar o dar sentido a los resultados científicos. Igualmente, están los que cuestionan la validez de los resultados bibliométricos, dado que conllevan el

cumplimiento de unas hipótesis que son, en algunos casos, discutibles. Así, por ejemplo, se acostumbran a considerar supuestos, entre otros, que toda contribución científica deja huella en la literatura científica; que la información de las bases de datos es suficientemente fiable; o que el número de citas que recibe una publicación es una medida aceptable del su valor. Esta falta de marco teórico es una de las críticas más persistentes (Moravcsik 1989; Torres Alberó 1994). Finalmente, también están los que consideran que se descuida la individualidad de los genios científicos, considerados así motores verdaderos del avance científico.

Aunque en algunos casos las hipótesis bibliométricas han estado poco justificadas, existe consenso en que no es razón suficiente para menospreciar estos estudios. Interpretados con cautela considerando sus limitaciones, con un uso de indicadores realmente estandarizados y con una correcta normalización de resultados, son bastante útiles a la hora de valorar la actividad de investigación científica y la planificación de unidades de información, lo que se ha puesto reiteradamente en evidencia (Smith 1981; Sancho 1990; Urbano 2000; Wallin 2005).

Aunque no existe consenso en cuanto al inicio de los métodos bibliométricos, se suele considerar el trabajo publicado 1917 de F. J. Cole y Nelly Eales en *The history of comparative anatomy* en que se analizan las publicaciones de historia de la anatomía aparecidas entre 1543 y 1860, el primer análisis con las características de la bibliometría moderna.



Breve resumen crítico de la historia de la bibliometría y su alcance: Urbano (2000), p. 68-76.

Producción científica y bases de datos bibliográficas

Price (1963) definió la ciencia como aquello que se edita en las publicaciones científicas y al científico como la persona que ha colaborado escribiendo alguna de esas publicaciones.

Aunque es una definición muy abierta lo cierto es que la investigación científica queda incompleta sin su publicación, puesto que ésta es la que proporciona el proceso de conexión con la comunidad científica que puede evaluarla y dar o no su asentimiento. Consecuentemente, si la publicación constituye el producto final de la actividad científica, la productividad científica de los investigadores puede ser evaluada mediante una descripción cuantitativa de su producción.

La revista científica o académica es, en la mayor parte de disciplinas, el medio especializado donde la investigación científica puede publicarse. Como consecuencia de ello encuentra publicidad y existencia social, permitiendo además la conservación y el archivo de conocimiento.

La evaluación de los resultados de la investigación científica puede realizarse mediante dos modos complementarios: la evaluación por pares (*peer review*) y la evaluación a posteriori mediante la valoración de las publicaciones en las que se recojan los resultados.

La evaluación por pares suele aplicarse fundamentalmente en tres ámbitos: procesos de incorporación de investigadores en universidades y organismos de investigación; procesos competitivos de asignación de recursos para la investigación; y en la evaluación previa de los trabajos candidatos a ser publicados en la mayor parte de revistas científicas.

Los artículos en revistas, así como las monografías y otras contribuciones académicas, son elementos de un sistema de relaciones que se establecen entre el conjunto de documentos de una disciplina e incluso con toda la producción científica. Para hacer aflorar estas relaciones puede recurrirse a la manifestación explícita de los autores, expresadamente tradicionalmente mediante las citas y más modernamente a través de los enlaces web. Así, si un

documento es mencionado o enlazado por un autor es necesario suponer que a su parecer existe una relación entre el documento que redacta y aquel que cita o enlaza. El análisis de citas y en análisis de enlaces es la parte de la bibliometría que estudia tales relaciones.

Para el análisis bibliométrico se necesita disponer de cuantiosa información bibliográfica. Generalmente se suele recurrir para ello a una base de datos bibliográfica. Estas bases de datos están constituidas por un conjunto de registros con información bibliográfica (autor, título de la contribución, de la publicación, fecha de la publicación, editorial,...), almacenado y gestionado mediante sistemas informáticos. Muchas de estas bases de datos contienen descriptores, palabras claves y resúmenes. Y unas pocas contienen citas (índices de citas).

Durante muchos años los únicos índices de citas disponibles –**Science Citation Index** (SCI), **Social Sciences Citation Index** (SSCI) y **Arts and Humanities Citation Index** (AHCI)– fueron los elaborados por el Institute for Scientific Information (ISI) fundado por Eugene Garfield en 1960. En la actualidad estos índices están integrados en el **Web of Science** (WOS) –suministrado por Thomson Reuters a través del la plataforma Web of Knowledge (WOK). El WOS indexa unas 9.300 de las revistas de investigación más prestigiosas. De hecho, que una publicación periódica esté indexada en esta base de datos se considera internacionalmente un criterio significativo de calidad.

Otro de los productos ofrecidos por el WOK es el **Journal Citation Reports** que evalúa el impacto de las revistas a partir de las citas. Existen dos ediciones, la *Science Edition*, que ofrece datos sobre el impacto de 5.900 publicaciones periódicas en ciencias; y el *Social Sciences Edition* que ofrece datos sobre 1.700 publicaciones en ciencias sociales y algunas de humanidades.



Información sobre los productos del WOK:
Science español
(2012).

El año 2002 Elsevier comenzó a publicar **Scopus** – oficialmente SciVerse Scopus–, una base de datos bibliográfica con citas con voluntad de ser mucho más exhaustiva que el WOS. En la actualidad indexa unos 19.500¹ títulos. Sin embargo mientras que el SCI recogen documentación que se remonta a 1900 y el SSCI desde 1956, Scopus solamente recoge resúmenes des de 1960 y citas des de 1996.

A finales de 2004 Google Inc lanzó **Google Scholar** –o *Google Académico*. Fiel a la filosofía de Google, a diferencia de las bases de datos tradicionales, no vacía los contenidos de revistas sino que explora sistemáticamente la Web. Pero en este producto convergen dos servicios, un buscador de publicaciones científicas y un índice de citas que permita conocer el impacto de los trabajos publicados. Sin embargo, el uso del rastreo automático indiscriminado conlleva una importante dosis de errores. Por otra parte se indiza una amplia variedad de tipos de documentos, algunos de los cuales distan de lo que se considera académico (guías académicas, documentos administrativos, bibliografía de asignaturas, libros de divulgación...). Este punto es importante ya que no es lo mismo ser citado en un documento científico que en otro tipo de documento.

Más recientemente Google Inc ha incorporado dos nuevos productos **Google Scholar Citations** (GSC) y **Google Scholar Metrics** (GSM). El primero recopila la producción científica de un investigador y la ofrece agregada en una página web, añadiendo información sobre el número de citas de cada referencia. En cambio GSM ofrece el impacto de las revistas científicas a partir de los recuentos de citas.



Información sobre
SciVerse Scopus:
About Scopus (2012).



Análisis de Google
Scholar como
herramienta de
evaluación científica:
Torres-Salinas et al.
(2009)



Análisis de GSC y
GSM: Cabezas-
Clavijo; Torres-
Salinas (2011);
Cabezas-Clavijo,;
Delgado-López-Cózar
(2012)

¹ En lo que sigue, el punto se utiliza para la separación de los miles, mientras que la coma se reserva para indicar decimales.

Leyes bibliométricas

El desarrollo de la bibliometría como disciplina científica se fundamenta en la búsqueda de comportamientos estadísticamente regulares a lo largo del tiempo en los diferentes elementos relacionados con la producción y el consumo de información científica. Las explicaciones globales a los fenómenos observados se consiguen mediante la formulación de las leyes bibliométricas.

Ley de la productividad de los autores

Esta ley enuncia de una relación cuantitativa entre los autores y las contribuciones producidas en un campo dado a lo largo de un periodo de tiempo. Esta distribución de probabilidades discreta resulta desigual puesto que la mayor parte de los artículos proceden de una pequeña porción de autores altamente productivos. Originalmente propuesta por Lotka en 1926 como una ley de cuadrado inverso la **ley de Lotka** se expresa en la actualidad de manera más general en forma de potencia inversa generalizada (Cuadro 1)

Cuadro 1. Ley de Lotka

La ley de Lotka original se expresa como:

$$A(x) = \frac{A_0}{x^2}$$

donde $A(x)$ representa el número de autores con x trabajos publicados. A_0 es el número de autores con 1 trabajo publicado.

La expresión actual

$$A(x) = \frac{C}{x^n}$$

donde C y n deben ser estimados a partir de los datos observados.

Veamos la representación gráfica de este tipo de funciones matemáticas en un ejemplo concreto. Supongamos que se hubieran obtenido los datos de la tabla 1. Así, en nuestro ejemplo ficticio, se habrían encontrado 1214 autores con una única contribución, 258 autores con 2 trabajos publicados y así sucesivamente.

Tabla 1.

Número de autores	Número de contribuciones
1214	1
258	2
92	3
52	4
32	5
26	6
19	7
14	8
10	9
6	10
5	11
5	12
3	13
1	14
1	15
1	16
1	17
1	20
1	22
1	23
1	30
Total de autores 1744	Total de contribuciones publicadas 3236

La figura 1 reproduce los resultados gráficamente. Sin embargo se observa que debido a su rápido decrecimiento,

no es demasiado útil por lo que suele recurrirse a una representación logarítmica (figura 2) en que el gráfico se vuelve lineal (una recta) aproximadamente. La calidad del ajuste puede medir por el estimador R^2 .

Figura1.

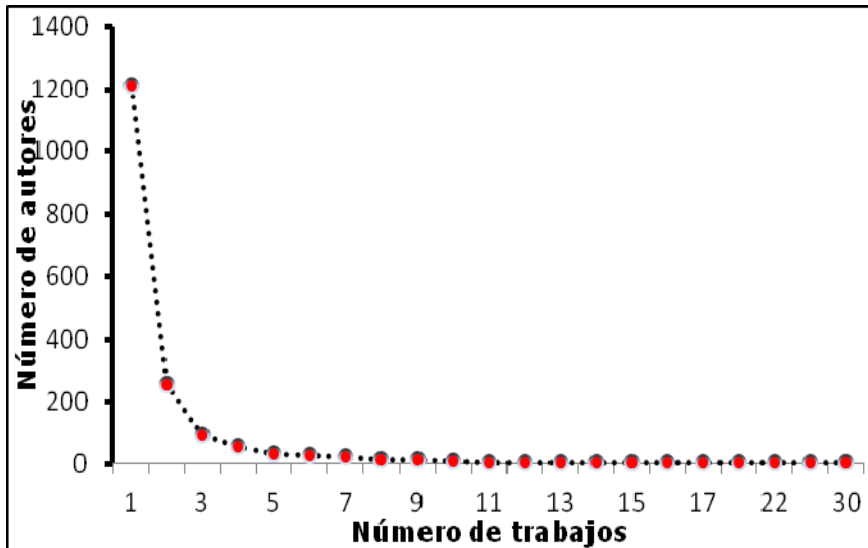
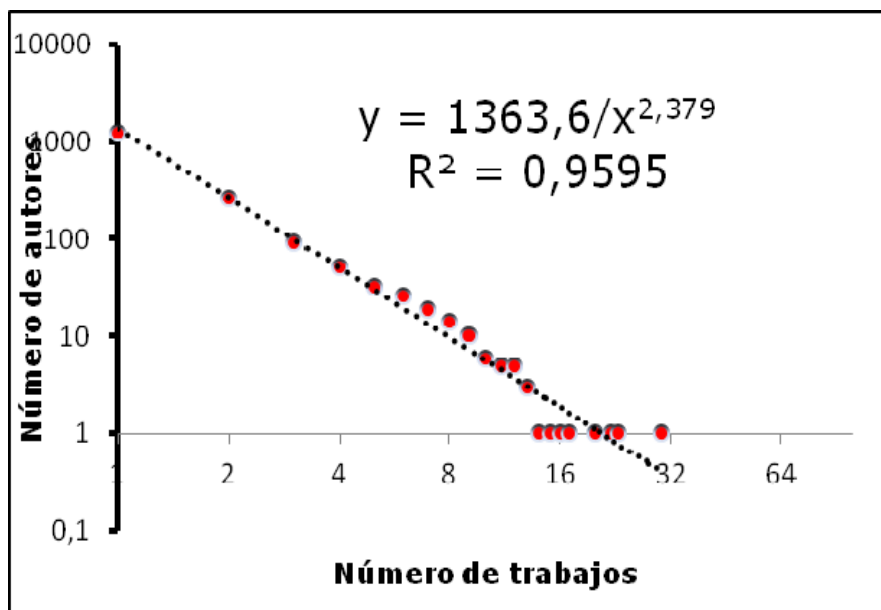


Figura 2.



En este caso la expresión empírica obtenida por mínimos cuadrados sería aproximadamente de

$$A(x) = \frac{1364}{x^{2,4}}$$

La ley de Lotka es equivalente al llamado principio de Pareto en economía, conocido también como la regla del 80-20 y que viene a decir que el 20 % de la población dispone del 80% de los recursos. Y que, por otra parte, el grupo mayoritario, formado por un 80% de población, solamente puede acceder al otro 20%.



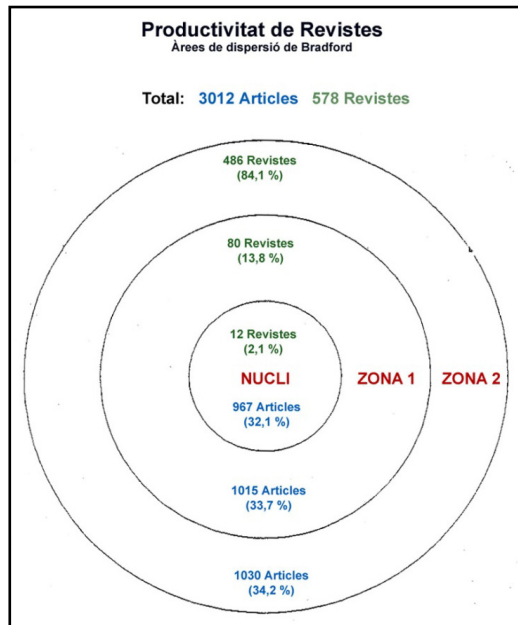
Un método más exacto de cálculo y programa de cálculo Rousseau; Rousseau (2000).

Ley de dispersión de la bibliografía científica

En 1934 Bradford publicó un artículo en que evidenciaba la concentración en un reducido número de títulos de revista del porcentaje mayoritario de la bibliografía pertinente en una materia lo que implica una caída rápida del rendimiento de ampliar la búsqueda de referencias fuera de un núcleo reducido. Una consecuencia inmediata es que no todas las publicaciones son igual de consultadas sino que de hecho unas pocas acumulen la mayor parte del consumo. Este consumo puede medirse por el volumen de acceso a los documentos o a partir de las citas que cada uno reciben.

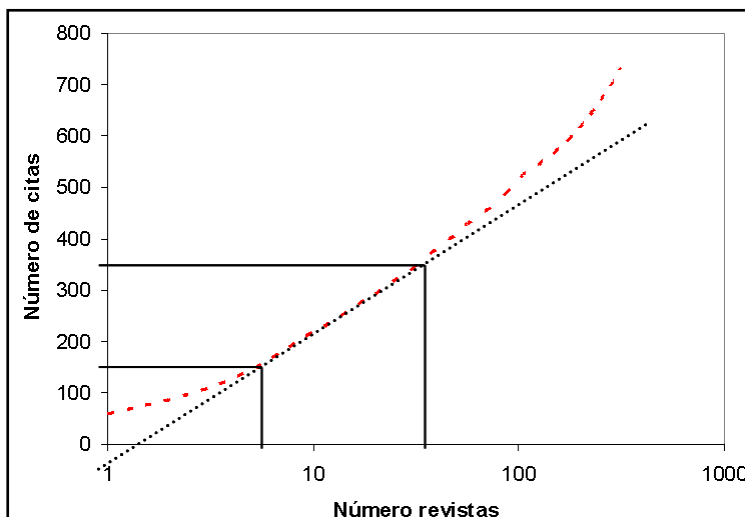
El enunciado textual de la **ley de Bradford** viene a decir que si las revistas científicas se ordenan en una secuencia decreciente de productividad de artículos sobre un campo específico, éstas pueden dividirse en un núcleo de revistas que abordan en particular el tema (núcleo de Bradford) y varios grupos o zonas que contienen aproximadamente el mismo número de artículos que el núcleo, donde el número de revistas en el núcleo y en las zonas sucesivas está en una relación de relación de 1: n: n² (figura 3).

Figura 3. Ejemplo de las zonas de Bradford



Sin embargo el cálculo de las zonas de Bradford se realiza habitualmente de forma gráfica: «se representa en una escala semilogarítmica el número acumulado de títulos por orden decreciente de citas frente al número de citas acumulado, el punto en el que la curva se transforma en una recta proyectado sobre el eje de abscisas determina el número de títulos que forman el núcleo» (Urbano 2000: 145). La figura 4 muestra un ejemplo.

Figura 4



En muchas ocasiones el esfuerzo de determinación exacto de las zonas de Bradford no tiene interés práctico y se calcula la dispersión clasificando las revistas por orden decreciente de frecuencias de consumo, acceso o citas y se calcula el núcleo a partir del 50 % acumulado tal y como muestra el ejemplo de la tabla 2. En este caso el núcleo de la disciplina estaría formado por las 3 revistas con más citas.

Tabla 2

Revista	Número de citas	Número acumulado de citas	% acumulado de citas
1	22	22	21,8%
2	17	39	38,6%
3	13	52	51,5%
4	9	61	60,4%
5	8	69	68,3%
6	7	76	75,2%
7	6	82	81,2%
8	5	87	86,1%
9	4	91	90,1%
10	3	94	93,1%
11	2	96	95,0%
12	2	98	97,0%
13	1	99	98,0%
14	1	100	99,0%
15	1	101	100,0%



Ley de crecimiento exponencial

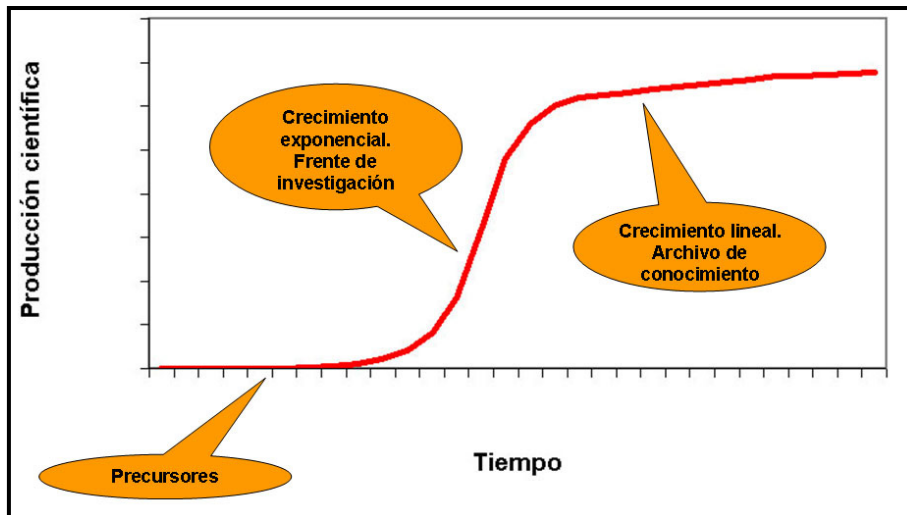
Derek J. de Solla Price (1956) constató que el crecimiento de la información científica era exponencial y se producía a un ritmo tan rápido que cada 10-15 años la información global existente se duplicaba (ley de Price). Sin embargo cada disciplina sufre su evolución propia, pasando por diversas etapas (figura 5):

Precursores: primeras publicaciones en un campo de investigación

Crecimiento exponencial: El campo se convierte en un frente de investigación.

Crecimiento lineal: El crecimiento se ralentiza. Las publicaciones tienen como finalidad primordial la revisión y el archivo de conocimiento.

Figura 5



Ley de obsolescencia de la bibliografía científica

Price también constató que la literatura científica pierde actualidad cada vez más rápidamente, pero el envejecimiento de la bibliografía no es uniforme para todas las disciplinas científicas.

Así, en las ciencias experimentales y en la tecnología el incesante crecimiento de la bibliografía conlleva también un rápido envejecimiento, mientras que, en las humanidades es mucho más lento. Incluso dentro de las ciencias experimentales, las diferencias pueden ser significativas. Por ejemplo la obsolescencia en bioquímica es mucho más rápida que en botánica.

Indicadores bibliométricos

La bibliometría utiliza un conjunto de indicadores que permiten expresar cuantitativamente las características bibliográficas del conjunto de documentos estudiado así como las relaciones existentes entre estas características. Estos indicadores bibliométricos son datos numéricos calculados a partir de las características bibliográficas observadas en los documentos publicados en el mundo científico y académico, o los que utilizan los usuarios de tales documentos, y que permiten el análisis de rasgos diversos de la actividad científica, vinculados tanto a la producción como al consumo de información

El indicador sintetiza una característica bibliográfica o una combinación de ellas utilizando un valor numérico que toma más interés cuando se puede comparar con observaciones de otros conjuntos de documentos (de otras zonas geográficas, universidades, disciplinas, bases de datos, etc.) y estudiar su evolución con el tiempo.



Indicadores
bibliométricos y
estudios de usuarios,
(Sanz, Martin 1997).

Indicadores personales

Estos indicadores se ocupan de características de los autores del estudio tales como la edad, el sexo, la posición profesional, el país, la afiliación institucional..., obteniendo la distribución porcentual en los trabajos de una publicación en serie, una base de datos u otro conjunto de documentos. Derivados de estos indicadores se definen otros como el **índice de aislamiento** o porcentaje de referencias de una revista que corresponden al mismo país donde se publica.

Indicadores de producción

Los indicadores de producción se obtienen a partir del recuento de las publicaciones científicas. El número de

publicaciones suele considerarse una medida de la actividad científica.

Por una parte se puede medir la productividad científica que es la cantidad de publicaciones producidas por un autor, un departamento o grupo de investigación, una institución, una disciplina o un país durante un período de tiempo determinado. Esto permite establecer rankings según el nivel de producción de los autores u otros elementos obtenidos por agregación que se comparen. El **indicador de productividad** se define como el logaritmo decimal del número de publicaciones.

A partir de los datos de productividad de los autores estudiados, pueden obtenerse indicadores como el número medio de trabajos por autor o el número medio de autores por trabajo, así como los respectivos indicadores estadísticos de dispersión (desviación estándar).

También se calcula en ocasiones el índice de transitoriedad, definido como el porcentaje de autores con un solo trabajo publicado.

Indicadores de dispersión

Se trata de determinar que publicaciones constituyen el núcleo de la disciplina. Para ello frecuentemente se recurre a considerarlo aquellas que acumulan el 50% de las citas.

Indicadores de visibilidad o impacto

Los indicadores de visibilidad e impacto miden la influencia de los autores y de los trabajos publicados y son las magnitudes más conocidas de la bibliometría dado el efecto directo que tienen sobre las carreras profesionales de los investigadores.

Buena parte del sistema de promoción y asignación de recursos de la comunidad científica y académica en los

países más desarrollados se basa en la estimación de la productividad del investigador y de la influencia de su labor. Dicha estimación puede hacerse a partir del análisis de citas o de enlaces web.

Se supone razonablemente que la obtención por parte de un autor de un mayor número de citas – o de enlaces – responde a una mayor influencia y, por tanto, que su trabajo es más valioso. Igualmente, una frecuencia relativa más alta de citas de una publicación implica un mayor prestigio, tanto para los editores como para los autores que escriben en ella. El indicador más simple es el del **número total de citas** recibidas. En ocasiones se calcula su logaritmo, conocido como **índice de Platz**. Otro indicador utilizado es el **promedio de citas** que obtiene cada contribución de un autor.

La importancia del número de citas recibidas depende del contexto en que se aplica. Así, en el caso de una publicación en serie como una revista, el impacto real de esta publicación depende también del número de artículos que produce durante el tiempo analizado, dando lugar al llamado **factor de impacto (FI)** que en esencia es la razón entre las citas recibidas y los artículos publicados en una revista.

La primera vez que se mencionó el factor de impacto fue Eugene Garfield en un artículo publicado en la revista *Science*. Los factores de impacto se calculan cada año para las revistas que están indexadas en el **Journal Citation Reports (JCR)** de Thomson Reuters.

El JCR define el *FI* de una publicación como el cociente de dividir las citas obtenidas en un año por una publicación en los artículos publicados los dos años anteriores entre el total de artículos publicados durante el mismo periodo.

Este cálculo prima la bibliografía que tiene un gran nivel de obsolescencia, lo que motiva que se utilicen también ventanas temporal más amplias –por ejemplo 5 años en



Texto original de
Garfield (1955).
Historia y problemas
del FI, Garfield (2005)

lugar de 2. En otros casos interesa utilizar el **índice de inmediatez**, que se calcula a partir de las citas recibidas el mismo año de publicación de los artículos y el número de artículos publicados.

El *FI* beneficia a las publicaciones con pocos artículos y de gran extensión, y aquellas que poseen un valor informativo inmediato (los frentes de investigación) más que de archivo del conocimiento.

Una característica importante del *FI* es que no se pueden comparar los valores de diferentes disciplinas. Así, por ejemplo, en la edición del JCR de 2010 el *FI* más elevado corresponde a la revista médica *CA: A cancer journal for clinicians* con un valor de 94,333 mientras que en el campo del derecho el *FI* más alto (4,741) es el de *Harvard law review*.

Sin embargo, el *FI* puede verse afectado por prácticas inadecuadas de autocita. Aunque es absolutamente legítimo que aun autor cite sus trabajos anteriores si es necesario, o que los trabajos de una revista incluyan referencias a la misma publicación en serie, la práctica fraudulenta de la autocita puede modificar artificialmente los *índices de Platz* o el *FI*. **Él índice de autocitas de una revista** o publicación seriada se define como el porcentaje de referencias a la propia publicación independientemente del autor. Paralelamente se define el **índice de autocitas de un autor** como el porcentaje de referencias citadas por un autor a trabajos anteriores suyos, independientemente de la publicación en la que se hubieran realizado. Naturalmente, a partir de la información de las autocitas de los autores, nada impide extrapolar la definición a una institución u otro nivel de agregación que resulte conveniente.

El *FI* es un valor que se calcula para publicaciones en serie. No obstante no informa realmente sobre el impacto concreto de un trabajo. Un autor puede publicar trabajos en revistas



Ejemplo de estudio del impacto de las autocitas en CC.SS. Campanario; Candelario (2010)

con un *FI* importante sin que se citen apenas sus contribuciones. Por tanto no puede trasladarse, al menos directamente, el impacto de una publicación a la de sus autores.

Para solucionar este problema se han propuesto diversos indicadores. Uno de los más utilizados es el llamado **índice *h*** o de Hirsch. Un científico tiene un *índice h*, si de todos sus trabajos *h* reciben al menos *h* citas cada uno, y el resto tiene como máximo *h* citas.

La tabla 2 muestra dos ejemplos de cálculo del *índice h* a partir de los 5 trabajos publicados por dos científicos ficticios y ordenados de mayor a menor número de citas.


 Texto original sobre el índice de Jorge Hirsch (2005).

Tabla 3

Autor 1		Autor 2	
Trabajos publicados ordenados por número de citas recibidas	Número de citas	Trabajos publicados ordenados por número de citas recibidas	Número de citas
1	8	1	610
2	7	2	290
3	5	3	2
4	4	4	1
5	3	5	1
Total obra = 5	Total de citas= 27	Tol obras = 5	Total de citas= 904

El autor 1 tiene un *índice h* igual a 4 ya que tiene 4 trabajos con 4 citas o más. No tiene 5 trabajos con 5 citas o más. El autor 2, por su parte, tiene 2 trabajos con 2 o más citas. Por eso su *índice h* es 2. Para llegar a 3 necesitaría tener una 3 obra con al menos 3 citas.

El ejemplo también permite observar uno de los defectos atribuidos al *índice h*: el *índice h* tiende a penalizar a los autores que priman la calidad frente a la cantidad y no publican un número de documentos elevado, pero sus

contribuciones terminan teniendo mucho impacto. En nuestro caso, ambos autores han publicado la misma cantidad de trabajos. Pero aunque el autor 2 recibe globalmente muchas más citas, al acumularlas solamente en dos trabajos, alcanza un *índice h* igual a 2, mientras que al autor 1 con un impacto global mucho menor atendiendo al número total de citas, tiene un índice mayor ($h = 4$).

El índice *h* adolece de otras limitaciones. Por una parte muestra una alta correlación positiva con el número total de citas recibidas y documentos publicados por los investigadores, de manera que tiende a favorecer a los científicos con carreras más dilatadas y perjudica a los más noveles por el bajo número de publicaciones. Por otro lado presenta problemas para discriminar entre investigadores situados en niveles intermedios. Además, no permite comparar investigadores de áreas diferentes debido a los diferentes hábitos de publicación y de citación en cada campo. Por ejemplo, los investigadores más relevantes internacionalmente en biología doblan en sus valores de *índice h* a los de áreas como la física.

Pese a todo esto, puede ser un indicador útil si se huye de la pretensión de evaluar a un investigador a partir de un solo indicador, puesto que la investigación es un proceso multidimensional.

El *índice h* se puede calcular también para valorar las revistas, las instituciones o los países siguiendo igual criterio. Por otra parte se puede recurrir a índices *h* sucesivos. Por ejemplo, puede calcularse un índice *h* de segundo orden igual a h_2 si la institución tiene h_2 investigadores cuyo *índice h* es al menos h_2 .



Índices *h* sucesivos:
Arencibia-Jorge
(2007).

Indicadores de colaboración

Miden las relaciones existentes entre los productores científicos que han terminado con la publicación conjunta de

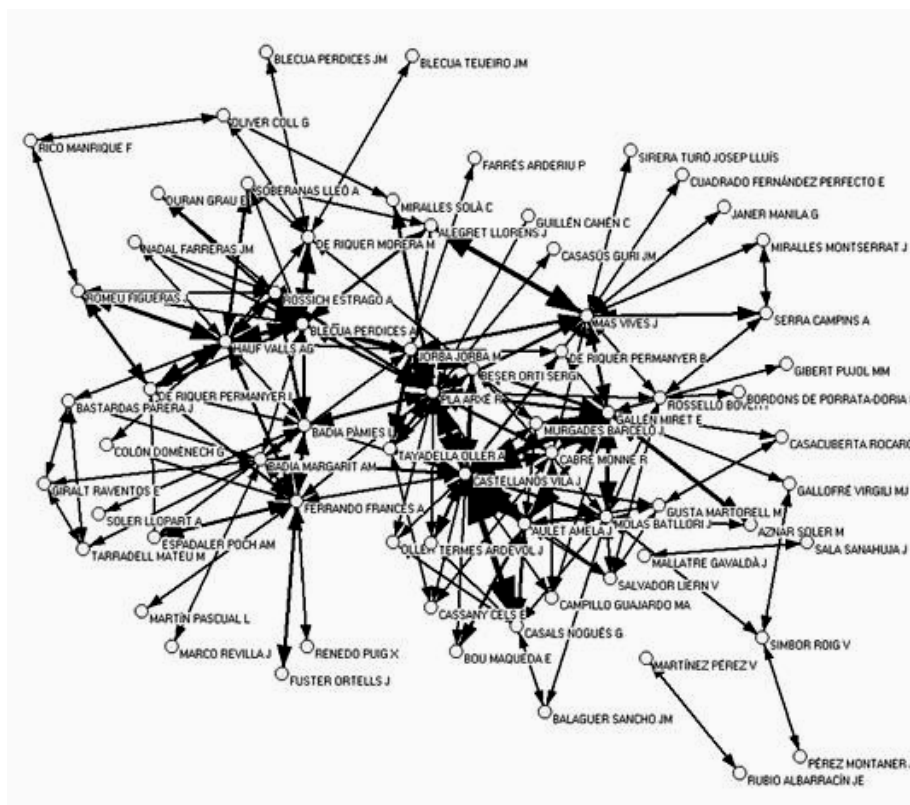
resultados. Puede contabilizarse la **proporción de trabajos con dos, tres o más autores** e identificarse aquellos individuos o instituciones que más han publicado conjuntamente.

Una opción más compleja es el análisis de la red social que se establece entre los autores que publican conjuntamente. Este análisis puede hacerse naturalmente con otros niveles de agregación tales como instituciones o países.

El análisis de redes sociales permite también estudiar la **cocitación**. Se trata del fenómeno por el cual dos o más autores son citados conjuntamente. Cuando esto ocurre con una frecuencia significativa cabe suponer que trabajan en una misma área de conocimiento aunque no estén necesariamente colaborando.

El análisis de redes sociales también se aplica a membresías de tribunales académicos y otras situaciones similares.

Figura 6. Red de coparticipación en tribunales de tesis doctorales de literatura catalana



Indicadores de obsolescencia

El actual desarrollo científico lleva parejo un envejecimiento veloz de la literatura científica. En otras palabras, existe una fuerte tendencia en muchas disciplinas a que las publicaciones científicas caigan en desuso con relativa rapidez.

Por una parte, en las áreas con un elevado nivel de producción, los documentos son substituidos por otros con información más novedosa. En otros casos, la información es válida, pero existe un decreciente interés en tales campos del conocimiento.

Los indicadores de obsolescencia miden el envejecimiento de las publicaciones. Entre los más utilizados se encuentra el **semiperíodo de Burton y Kebler** que se define como la mediana de las referencias ordenadas por su antigüedad. Otra medida es el **índice de Price** o porcentaje de referencias que tienen menos de 5 años de antigüedad.

Indicadores de forma y contenido

Puesto que la producción científica puede realizarse mediante canales diferentes, se suele considerar el porcentaje de documentos de cada **tipología documental**: libros y capítulos; comunicaciones de congresos, tesis doctorales y especialmente los artículos de revista –que a su vez también presentan variantes tales como artículos de investigación, reseñas, editoriales, notas y comunicados, ...

Otro aspecto a considerar es el **soporte documental**, centrado tradicionalmente en detectar la importancia del uso de soportes diferentes al papel.

También se consideran la distribución porcentual de **temas** y de *lenguas* utilizadas.

Referencias

Todos los enlaces estaban accesibles el 15 de abril de 2012.

- *About Scopus* (2012). <<http://www.info.sciverse.com/scopus/about>>.
- Aleixandre-Benavent, R.; Valderrama-Zurián, J.C.; González-Alcaide, G. (2007). "El factor de impacto de las revistas científicas: limitaciones e indicadores alternativos". *El profesional de la información*, 16(1): 4-11. <<http://eprints.rclis.org/bitstream/10760/9489/1/4m1m9wt6vv600fr.pdf>>.
- Arencibia-Jorge, Ricardo (2009). "Nuevos indicadores de rendimiento científico institucional basados en análisis de citas: los índices H sucesivos". *Revista española de Documentación Científica*, 32(3): 101-106
- Cabezas-Clavijo, Á.; Torres-Salinas, D. (2012). "Google Scholar Citations y la emergencia de nuevos actores en la evaluación de la investigación". *Anuario ThinkEPI*, 6. <<http://ec3noticias.blogspot.com.es/2011/12/thinkepi-google-scholar-citations-y-la.html>>.
- Cabezas-Clavijo, A; Delgado-López-Cózar, E. "Scholar Metrics: el impacto de las revistas según Google, ¿un divertimento o un producto científico aceptable? ". *EC3noticias*, 2012. <<http://ec3noticias.blogspot.com.es/2012/04/scholar-metrics-el-impacto-de-las.html>>.
- Campanario, J M.; Candelario, A. (2010). "La influencia de las autocitas en el aumento del factor de impacto en revistas de Ciencias Sociales". *Revista Española de Documentación Científica*, 33(2): 185-200. <<http://redc.revistas.csic.es/index.php/redc/article/view/551/625>>.
- Garfield, Eugene (1955). "Citation indexes to science: a new dimension in documentation through the association of ideas". *Science* 122(3159): 108-111. <<http://garfield.library.upenn.edu/essays/v6p468y1983.pdf>>.
- Garfield, Eugene (2005). "The Agony and the Ecstasy—The History and Meaning of the Journal Impact Factor" [en línea]. *International Congress on Peer Review And Biomedical Publication Chicago, September 16, 2005*. Accesible en: <garfield.library.upenn.edu/papers/jifchicago2005.pdf>.
- Hirsch, J. E. "An index to quantify an individual's scientific research output". *Proceedings of the National Academy of Sciences of the United States of America*. 2005, 102(46):16569-15572. <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1283832/?tool=pubmed>>.
- López Piñero, José María (1972). *El análisis estadístico y sociométrico de la literatura científica*. València: Facultad de Medicina.
- Medina, Esteban (1982). "Teorías y orientaciones de la sociología de la ciencia". *Revista española de investigaciones sociológicas*, 20: 7-58.
- Merton, R.K. (1977). *La sociología de la ciencia*. Madrid: Alianza.
- Merton, R.K. (1992). *Teoría y estructuras sociales*. México: Fondo de Cultura Económica.

- Moravcsik, M. J (1989). "¿Cómo evaluar la ciencia y a los científicos?". *Revista Española de Documentación Científica*, 12 (3): 313-325.
- Price, D. J. S. (1963). *Little science, big science*. New York: Columbia University Press. [Existe una traducción en español, Barcelona: Ariel, 1973].
- Rousseau, B; Rousseau, R. (2000). "LOTKA: A program to fit a power law distribution to observed frequency data". *Cybermetrics*, 4(1), paper 4. Accesible en: <<http://cybermetrics.cindoc.csic.es/articles/v4i1p4.html>>.
- Sancho, R (1990). "Indicadores bibliométricos utilizados en la evaluación de la ciencia y la tecnología". *Revista Española de Documentación Científica*, 13 (3-4): 842-865.
- Sanz, Elías; Martín Moreno, Carmen (1997). "Técnicas bibliométricas aplicadas a los estudios de usuarios". *Revista general de información y documentación*, 7 (2): 41-68. <<http://revistas.ucm.es/index.php/RGID/article/view/RGID9797220041A/10878>>.
- *Science español* (2012). Disponible en <<http://ip-science.thomsonreuters.com/es>>.
- Smith, Linda Cheryl (1981). "Citation analysis". *Library Trends*, 30 (1): 83-106.
- Spinak, Ernesto (1996). *Diccionario enciclopédico de Bibliometría, Cienciometría e Infometría*. Caracas: UNESCO.
- Torres-Albero, Cristóbal (1994). *Sociología política de la ciencia*. Madrid: CIS-Siglo XXI.
- Torres-Salinas, D.; Ruiz-Pérez, R.; Delgado-López-Cózar, E. (2009). "Google Scholar como herramienta para la evaluación científica". *El profesional de la Información*, 18(5), 501-510. <ec3.ugr.es/publicaciones/d700h04j123154rr.pdf>.
- Urbano, Cristóbal (2000). *El análisis de citas en publicaciones de usuarios de bibliotecas universitarias: estudio de tesis doctorales en informática de la Universidad Politécnica de Cataluña, 1996-1998. Vol 1. [Tesis doctoral]* Barcelona: Universitat de Barcelona. <<http://www.tdx.cat/bitstream/handle/10803/2891/urbano-tesis1.PDF?sequence=1>>.
- Wallin, Johan A. (2005). "Bibliometric Methods: Pitfalls and Possibilities". *Basic & Clinical Pharmacology & Toxicology*, 97 (5): 261-275.