

APLICACIÓN DE ALGUNAS TÉCNICAS BIBLIOMÉTRICAS A LA INVESTIGACIÓN
EN HISTORIA DE LA CIENCIA

Carlos López Fernández
Manuel Valera
Pedro Marsset
Universidad de Murcia

1. Aspectos metodológicos generales

El análisis bibliométrico constituye un método documental que ha experimentado un importante crecimiento durante las dos últimas décadas. Surgido como un producto del contacto interdisciplinar, su desarrollo ha dependido de dos fuentes distintas: la Documentación Científica y el conjunto de disciplinas que integran la denominada "ciencia de la ciencia".

Los objetivos del análisis bibliométrico pueden incluirse en dos grandes apartados: a) estudio del tamaño, crecimiento y distribución de la literatura científica y b) la investigación de la estructura y dinámica de los grupos que producen y consumen dicha literatura, así como de la información contenida en la misma. Una característica peculiar de este tipo de análisis es la utilización de modelos matemáticos, distinguiéndose así de la estadística bibliográfica tradicional, que constituye solamente su punto de partida (Ferrada, 1983).

Como ya se indicaba en la primera obra publicada en nuestro país sobre el análisis bibliométrico aplicado a la literatura científica, entre las primeras disciplinas en servirse del mismo, además de la propia Documentación Científica, estarían la historia y la sociología de la ciencia, pues podrían utilizarlo "como instrumento para analizar objetivamente muchos aspectos confiados hasta ahora a la especulación arbitraria" (López Piñero, 1972).

Algunas técnicas bibliométricas

Sin embargo, la aportación de los historiadores de la ciencia no ha sido todo lo abundante que sería de esperar y desear en este terreno, puesto que, como también señalaba el propio López Piñero en la obra anteriormente mencionada, el análisis bibliométrico carece de interés para la mayor parte de los historiadores inscritos en lo que podemos denominar historiografía tradicional de la ciencia.

Efectivamente, dentro de la historiografía externalista, de carácter esencialmente sociológico y humanista, es irrelevante este tipo de análisis, pues a lo sumo en esta historiografía se llega a realizar de forma auxiliar, y a comentar después intuitivamente, algunas tabulaciones muy concretas. Tampoco tiene interés para la historia interna, la cual, centrada exclusivamente en el estudio abstracto de la evolución de las ideas y el pensamiento científicos, se basa casi siempre en el estudio filosófico y/o técnico de los textos de grandes figuras, estudio que aparece impregnado en la mayoría de los casos de la filosofía de la ciencia neopositivista y sus desarrollos posteriores.

Con todo, hay que señalar que la contribución de algunos historiadores de la ciencia ha tenido una gran importancia en el desarrollo del nuevo campo en su conjunto. De hecho, quizá la obra que más ha pesado y también la más discutida, ha sido la del historiador recientemente fallecido D.J. de Solla Price, Little Science, Big Science.

En general, y fuera ya del estricto ámbito de los investigadores histórico-científicos, ha habido un claro rechazo del análisis bibliométrico por parte de aquellos que estiman como únicos acercamientos aptos a la realidad científica, el ensayo, la erudición histórica y la reflexión filosófica y que consideran, por tanto, como no válidos o superfluos los intentos de aproximación al análisis de dicha realidad utilizando los mismos métodos de la propia ciencia. Paradójicamente, este punto de vista es ampliamente compartido por muchos científicos profesionales, que critican o cuestionan la validez de la historia de la ciencia como disciplina autónoma, con sus mé-

todos de investigación específicos y sus presupuestos epistemológicos propios, que no han de ser necesariamente, como señala canquilhem, los de la propia ciencia de la cual se hace historia (Canquilhem, 1968).

Evidentemente, tampoco parece correcta la posición de los que suponen haber encontrado en los métodos y técnicas bibliométricas la panacea que permita analizar científica y objetivamente, al margen de todo tipo de prejuicios ideológicos, la propia realidad científica, punto de vista que lleva en sí mismo una fuerte carga ideológica. Además, la utilización de unos determinados métodos, en este caso los derivados de la bibliometría, no tiene por qué estar en contradicción con la utilización de otros, y en especial de prescindir de los métodos y resultados fruto de los acercamientos tradicionales. Creemos que ya es hora de ir desterrando de la práctica mental y científica la idea de la existencia de un método particular. El Método, susceptible de ser aplicado en todo tiempo y lugar a cualquier campo de investigación como único capaz de revelarnos La Verdad de ese campo. Cualquier problema, cualquier objeto de investigación, es susceptible de ser abordado desde diversos puntos de vista, utilizando distintos métodos de análisis y de acercamientos, que no tienen por qué ser exclusivos ni excluyentes. Quizá de una postura más abierta, menos dogmática, se obtendrían resultados más fructíferos en un campo como el de la historia de la ciencia en el que antes de haber conseguido una plena institucionalización, se pretende a veces pontificar sobre la metodología, incluso por personas ajenas a las propias tareas de investigación en la disciplina.

Con este planteamiento, el grupo de investigación nucleado en la Cátedra de Historia de la Medicina de la Universidad de Murcia, hemos venido trabajando desde hace unos años en historia de la ciencia contemporánea en nuestro país, mediante la utilización del análisis bibliométrico. En especial, la utilización del mismo nos ha resultado del mayor interés a la hora de iniciar la investigación en un campo científico concreto del que se desconocen los parámetros generales de comportamiento. Así, mediante el análisis

bibliométrico hemos podido caracterizar los aspectos más destacados de la producción recogida en la literatura científica de ese campo investigativo, los autores de mayor producción, las revistas de mayor incidencia en la materia, los principales grupos de trabajo existentes en la misma, las áreas de influencia geográfico-idiomáticas y un largo etcétera. Es decir, todas aquellas características de la producción científica que permiten situar rápidamente los aspectos globales más destacables de la investigación en un campo científico dado, y que resultan ser un requisito previo imprescindible antes de abordar en profundidad el análisis pormenorizado y sistemático de los contenidos de la misma.

El presente trabajo tendrá como objeto sistematizar algunos de los métodos y técnicas más elementales utilizados en este tipo de análisis por nuestro equipo de trabajo, así como mostrar algunos aspectos que consideramos de mayor generalización y que pueden ser utilizados en cualquier campo del análisis histórico-científico. Esta sistematización será ilustrada a continuación con algunos ejemplos concretos que recogerán determinadas aplicaciones de la misma.

2. Técnicas bibliométricas

Una vez reflejado someramente el papel que, a nuestro juicio, han de desempeñar las técnicas bibliométricas en las investigaciones histórico-científicas, pasamos ahora a dar una visión de conjunto de aquellas que hemos tenido ocasión de manejar con más profución en las investigaciones de nuestro equipo durante los últimos años.

La utilización que hemos hecho de las mismas ha ido siempre dirigida al estudio de la producción de una comunidad científica a lo largo de una época determinada. Normalmente, hemos accedido a dicha producción a través de los órganos de expresión de la comunidad en cuestión.

El primer paso para acercarnos al estudio de lo que es y lo que representa auténticamente esa producción científica, no puede ser otro que una correcta cuantificación de la misma y el análisis de la evolución cronológica de este fenómeno. Para ello, hemos comenzado siempre por elaborar una ta-

Algunas técnicas bibliométricas

bulación anual del número de artículos publicados, la cual puede ya darnos una visión de conjunto sobre las principales características generales, así como las distintas etapas en que puede dividirse la producción global de la comunidad, que servirá para estructurar el análisis posterior. En ocasiones, al analizar la producción durante una época dilatada se hace imprescindible agrupar la misma en periodos triales o quinquenales para una correcta delimitación de esas etapas productivas.

Tras esta primera caracterización, hay que recurrir a otros indicadores bibliométricos que permitan profundizar en el análisis, lo que generalmente se consigue mediante un estudio comparativo del ritmo de variación de la producción a lo largo de los distintos periodos cronológicos establecidos. Los indicadores más frecuentemente utilizados para este objetivo son: la evolución cronológica de la tasa de crecimiento del número de artículos (mediante números índices), la media anual de artículos para cada periodo y el porcentaje que representa la producción de esos periodos sobre el total de la misma.

Conseguida ya una adecuada cuantificación de la producción científica, se procede al análisis de otros aspectos fundamentales de ella, como sería una aproximación a su calidad y a su grado de institucionalización. Para ello, y siempre a fin de mantenernos en un plano rigurosamente objetivo, pondremos en juego una doble gama de nuevos indicadores bibliométricos.

Por lo que respecta al estudio de la calidad de la producción, en nuestras investigaciones nos hemos acercado a la misma estudiando la evolución cronológica, en valores absolutos y porcentuales, de los siguientes indicadores:

- Trabajos realizados en equipo (distinguiéndose los firmados por uno, dos y tres o más autores).
- Trabajos portadores de referencias bibliográficas.
- Cuantía global de referencias y número de referencias por artículo.
- Índice de obsolescencia de las referencias.
- Índice de impacto asignable a la producción de los autores más relevantes de la comunidad estudiada.

Por su parte, la aproximación al grado de institucionalización de la producción científica, con lo que hacemos referencia a la posible existencia de equipos estables de trabajo, la hemos llevado normalmente a cabo con estos indicadores:

- Trabajos en equipo.
- Cuantía global de autorreferencias y número de autorreferencias por artículo.
- Número de autorreferencias sobre el total de referencias.
- Índices de productividad de los autores, efectuando una clasificación de los mismos de acuerdo al logaritmo de su número de artículos.

Como es natural, el estudio de la evolución cronológica, absoluta y porcentual, de estos indicadores es completada con el estudio del ajuste de los datos recogidos respecto a las distribuciones estadísticas artículos-auteur más utilizadas, en particular la de A. Lotka y sus rectificaciones o modificaciones posteriores. (Lotka, 1926, Rao, 1982). Finalmente, otro de los recursos usuales en este terreno es la elaboración de redes de citas (Garfield, 1964).

Además de los dos aspectos fundamentales de la producción científica indicados (calidad e institucionalización), habría otros también muy interesantes, que no podemos detallar por motivo de espacio. Citemos, a título de ejemplo, la determinación de las áreas de influencia geográfico-idiomática que actúan sobre la comunidad científica estudiada. Para ello, basta poner en juego un nuevo indicador bibliométrico: la distribución idiomática de las referencias bibliográficas.

La utilización óptima de todas las técnicas bibliométricas que hemos ido reflejando se consigue, a nuestro juicio, con un doble uso de las mismas. Primero, deben ser aplicadas al total de la producción científica estudiada, y posteriormente han de volver a aplicarse a todas y cada una de las grandes áreas temáticas de investigación que se determinen dentro de dicha producción.

Algunas técnicas bibliométricas.

De inmediato se comprende que dada la complejidad y laboriosidad que requiere la evaluación de la mayor parte de los indicadores bibliométricos que hemos propuesto, se hace imprescindible el recurso a la informática y a las técnicas estadísticas de muestreo, inferencia y análisis de grupos, dentro de las cuales hemos tenido ocasión de aplicar el análisis factorial de c componentes principales (López Fernández, 1983) con objeto de caracterizar globalmente y clasificar estadísticamente las distintas áreas temáticas cultivadas por una comunidad.

3. Aplicaciones concretas.

Al objeto de ilustrar, cuando menos parcialmente, la utilización de las técnicas bibliométricas anteriormente citadas, en el presente apartado veremos de caracterizar algunos de los aspectos de la investigación y de la producción científica en física durante los periodos inmediatamente anterior y posterior a la guerra civil española y recogidos en los Anales de la Real Sociedad Española de Física y Química. (Un estudio más general sobre esta cuestión puede verse en un trabajo anterior que abarca el periodo 1903-1975 de Marset, López Fernández, Valera, 1981). Mostraremos aquellas variables que no resultan aparentemente afectadas por la guerra y las que sí lo son, incidiendo así sobre los aspectos que sugieren interpretaciones o que ponen sobre la pista de "anomalías" que exigen un estudio más detallado.

Según las líneas metodológicas antes señaladas, nuestro estudio debe comenzarse por el análisis de la evolución del número de artículos publicado (Tabla 1), del cual se deriva claramente la idea de que la guerra civil, salvados los años 1937 y 1940 como anómalos por la proximidad del conflicto, no afectó de manera significativa a los niveles cuantitativos de producción. Situación que coincide con la indicada por Price (Price, 1973) en estudios más generales, en el sentido de que las grandes contiendas mundiales no afectaron al desarrollo cuantitativo de la ciencia, a lo más lo congelaron durante una cierta etapa. Asimismo, a través de nuestros datos, cabe idéntica afirmación con respecto a la distribución por grupos de autores del mis-

mo nivel de productividad (Tabla 2), pues la coincidencia de porcentajes entre los correspondientes a cada uno de los tres niveles determinados por el índice de productividad es prácticamente completa. Por su parte, la coincidencia de los porcentajes de artículos realizados por estos grupos de autores es también bastante grande.

Sin embargo, esta sensación de continuidad comienza a romperse cuando revisamos la distribución por periodos quinquenales de trabajos realizados en equipo (Tabla 3), pudiéndose apreciar cómo el porcentaje de los mismos disminuye algo más de diez puntos nada más terminar la guerra, tardándose unos diez años en recuperar las cotas de preguerra, si bien ya a partir de ahí éstas son claramente rebasadas. De todo esto parece desprenderse que hay una cierta ruptura a nivel de equipos de investigación a raíz de la contienda civil, la cual no quedaría enmendada de forma definitiva hasta que, ya entrados los años cincuenta, toman cuerpo nuevos equipos de trabajo.

Dicha idea queda sólidamente reforzada al analizar otras magnitudes bibliométricas de notable significado. Así, es de apreciar cómo uno de los indicadores más óptimos para medir la pervivencia de equipos de trabajo, el porcentaje de autorreferencias sobre el total de referencias (Tabla 4), experimenta una baja sensible a lo largo de toda la década de los cuarenta. Otra sensación de ruptura se experimenta al observar la distribución idiomática de esas mismas referencias (Tabla 5), donde llama la atención la notable baja del inglés, en beneficio casi exclusivo de un idioma como el italiano que antes de la guerra carecía de la menor significación.

Continuando con esta cuestión de la ruptura provocada por la guerra, al estudiar la procedencia institucional de los artículos (Tabla 6), se ve que aunque las instituciones dominantes en los periodos de preguerra y postguerra son administrativamente homologables, (JAE y CSIC respectivamente), a partir del conflicto se detecta en una primera fase, un desplazamiento hacia la investigación en instituciones diversas y, en una segunda, hacia los centros universitarios. Volvemos a abundar en la sensación de ruptura cuan-

Algunas técnicas bibliométricas

do revisamos la evolución del número de autores que publican por vez primera en los Anales y su porcentaje respecto al total de los mismos en cada período (Tabla 7). Vemos cómo dicho porcentaje, que en el quinquenio 1931-35 es ligeramente inferior al 70%, pasa a más del 81% en el período inmediatamente posterior a la guerra civil, lo que es un síntoma claro de la discontinuidad producida en los autores, máxime si consideramos que ya el porcentaje correspondiente a los años treinta es superior a lo normal debido al incremento producido por la puesta en funcionamiento del Instituto Nacional de Física y Química que tuvo lugar en esos años.

Fuera ya de la idea de que la guerra no parece conllevar una ruptura a nivel cuantitativo de producción pero sí en los aspectos más cualitativos de ésta, incidiremos, aunque sea de una manera muy superficial, en otro aspecto importante, como sería el aproximarnos a la calidad de producción. Sobre esto tenemos ya un indicador que nos ha dado un mensaje negativo, el de los artículos realizados en equipo, pues bien, ese mensaje podríamos reforzarlo con el análisis de otro indicador de acalidad, el porcentaje de artículos con referencias bibliográficas, para el que también se aprecia (Tabla 8) una baja al término de la guerra de casi diez puntos. Posteriormente se recupera, pero durante los tres primeros quinquenios de postguerra sólo consigue alcanzar los mismos niveles que al final de los años veinte.

Desde luego, muchos otros aspectos pueden ser discutidos a partir de los datos suministrados que permitiese ofrecer un análisis más detallado y riguroso sobre el impacto de nuestra guerra civil en la investigación española en física (Valera, 1982; López Fernández, 1984; Maset, Valera, López Fernández, 1981), pero pensamos que son suficientes para ilustrar sobre la aplicación de algunas técnicas bibliométricas a la investigación en historia de la ciencia que es el objetivo del presente trabajo.

No queremos terminar sin insistir en dos aspectos contenidos implícitamente en el trabajo, pero que quizá convenga aclarar algo más. En primer lu-

gar, hemos utilizado estas técnicas sólo para estudios dedicados a la época contemporánea en la que la producción científica y, más en general, la actividad científica, sigue unas pautas de funcionamiento muy diferentes a las existentes en épocas anteriores. Ello no quiere decir que el análisis bibliométrico no pueda ser utilizado como auxiliar de la investigación histórico-científica para tales épocas, pero las técnicas concretas que hemos expuesto no serían aplicables en el sentido indicado. Por otra parte, los diferentes modelos matemáticos aplicados a las magnitudes bibliométricas consideradas y las leyes o regularidades establecidas (Lotka, Bradford, Prive, etc.) son utilizados generalmente en el sentido de contrastar el grado de acercamiento de la producción analizada a tales leyes. Es decir, el uso de este tipo de técnicas no se efectúa con el ánimo de acomodar los datos manejados a los modelos existentes, pretendiendo así explicar la realidad subyacente a dichos datos, sino más bien con el objeto de detectar los componentes específicos de la realidad científica concreta, que explique tanto las concordancias como las divergencias con los modelos establecidos.

BIBLIOGRAFIA

- Canguilhem, G. (1968): "L'objet de l'histoire des sciences" Etudes d'histoire et de philosophie des sciences, Paris, Vrin, 9-23.
- Gardfield, E. (1964): "Science citation index: a new dimension in indexing", Science, 144, 649-654.
- Loosjes, Th.P. (1973): On documentation of scientific literature, London, Butterworths.
- Lotka, A.P. (1926): "The frequency distribution of scientific productivity", J. Washington Acad. Sciences, 16, 317-332.
- Lopez Fernández, C. & Valera, M. (1983): "Estudio bibliométrico-multivariante de los artículos de física publicados en los Anales de la RSEFQ durante el periodo franquista", Llull, 6 (num.10-11), 39-56.
- López Fernández, C. (1984): La producción española en física durante el franquismo (1940-1975) a través de los Anales de la RSEFQ, Murcia, Tesis de Doctorado (pendiente de lectura).
- López Piñero, J.M. (1972): El análisis estadístico y sociométrico de la literatura científica, Valencia, Centro de Documentación e Informática Médica de la Fac. de Medicina.
- Marset, P. & Valera, M. & López Fernández, C. (1981): "Repercusiones de la guerra civil española (1936-1939) en la producción científica en física a través de los Anales de la RSEFQ (1903-1975)", Dynamis, 1, 179-202.

Algunas técnicas bibliométricas

- Mijailov, A.I. & Guiliarevskii, R.S. (1974): Curso introductorio de Informática-Documentación, Caracas, Instituto Venezolano de Productividad.
- Mijailov, A.I. & Chernii, A.I. & Guiliarevskii, R.S. (1973):
Fundamentos de informática (2 vols), Moscu,
La Habana, Nauka, Ac. de Ciencias de Cuba.
- Price, J.D.S. (1973): Hacia una ciencia de la ciencia, Barcelona, Ariel.
- Ravichandra Rao, J.K. (1980): "The distribution of scientific productivity and social change", J. Am. Soc. for Information Science, 31, 111-123.
- Terrada, M.L. (1983): La Documentación Médica como disciplina, Valencia, Centro de Documentación e Informática Biomédica.
- Valera, M. (1982): La producción española en física a través de los Anales de la SEFQ (1903-1937). Murcia, Tesis de Doctorado.
-

TABLA 1

ARTICULOS Y AUTORES DE FISICA APARECIDOS EN LOS ANALES DE LA
REAL SOCIEDAD ESPAÑOLA DE FISICA Y QUIMICA DESDE 1932 HASTA
1945

<u>AÑOS</u>	<u>ARTICULOS</u>	<u>AUTORES</u>
1932	21	15
1933	17	20
1934	23	21
1935	17	19
1936	19	17
1937	8	8
.....
1940	9	8
1941	20	14
1942	16	18
1943	23	22
1944	26	25
1945	29	22

TABLA 2

DISTRIBUCION DE AUTORES Y ARTICULOS POR INDICES DE PRODUCTIVIDAD. COMPARACION ENTRE LOS PERIODOS DE 1903-1937 y 1940-1975

a) Período 1903-1937

<u>Indice Prod.</u>	<u>AUT</u>	<u>ART</u>	<u>% AUT</u>	<u>%ART</u>
$I_p = 0$	60	60	47,2	11,5
$0 < I_p \leq 1$	59	244	46,4	47,0
$I_p > 1$	8	216	6,3	41,5

b) Período 1940-1975

<u>Indice Prod.</u>	<u>AUT</u>	<u>ART</u>	<u>% AUT</u>	<u>% ART</u>
$I_p = 0$	375	375	47,7	13,6
$0 < I_p \leq 1$	364	1.426	46,3	51,6
$I_p > 1$	47	964	6,0	34,9

TABLA 3

ARTICULOS REALIZADOS INDIVIDUALMENTE Y EN COLABORACION
DISTRIBUCION POR QUINQUENIOS (1921 A 1955)

QUINQUENIOS	ART IND	ART COLAB	% IND	% COLAB
1921-1925	39	8	83,0	17,0
1926-1930	45	10	81,8	18,2
1931-1935	46	41	53,5	47,1
.....
1940-1944	59	35	62,7	37,3
1945-1949	80	64	55,6	44,4
1950-1954	66	107	38,2	61,8

TABLA 4

VALORES PORCENTUALES DE AUTORREFERENCIAS SOBRE REFERENCIAS
DISTRIBUCION QUINQUENAL

<u>QUINQUENIOS</u>	<u>% AUTORREF / REF</u>
1921-1925	17,6
1926-1930	14,7
1931-1935	11,6
.....
1940-1944	9,9
1945-1949	6,9
1950-1954	18,6

TABLA 5

DISTRIBUCION IDIOMATICA DE LAS REFERENCIAS

<u>QUINQUENIOS</u>	<u>ESP</u>	<u>PRA</u>	<u>ING</u>	<u>AL</u>	<u>OTROS</u>
1921-1925	13,2	13,6	30,3	42,8	0,0
1926-1930	14,7	17,7	33,4	33,4	0,7
1931-1935	14,8	9,5	40,1	34,3	1,2 "
.....					
1940-1944(1)	13,9	15,2	28,4	29,1	13,1 (2)
1945-1949(1)	17,1	11,2	43,3	27,2	1,1
1950-1954(1)	18,3	6,7	60,2	8,6	6,0

(1): Datos obtenidos mediante muestreo

(2): Del porcentaje 13,1, algo más del 11% corresponde
al italiano.

TABLA 6

DISTRIBUCION ARTICULOS POR CENTROS DE TRABAJO

<u>QUINQUENIOS</u>	<u>JAE-CSIC</u>	<u>UNIV</u>	<u>OTROS</u>
1921-1925	70,2	2,1	27,7
1926-1930	63,6	12,7	23,6
1931-1935	81,6	4,6	13,8
.....			
1940-1944	52,7	11,8	35,5
1945-1949	62,2	17,1	17,7
1950-1954	64,1	24,1	11,8

TABLA 7

AUTORES NUEVOS POR PERIODOS QUINQUENALES Y PORCENTAJES QUE SUPONEN SOBRE EL TOTAL EN CADA PERIODO

<u>QUINQUENIOS</u>	<u>AUT NUEVOS</u>	<u>% AUT NUEVOS/TOTAL AUT</u>
1926-1930	14	50,0
1931-1935	32	69,6
.....
1940-1944	48	81,3
1945-1949	72	66,6

TABLA 8

EVOLUCION QUINQUENAL DE ARTICULOS CON REFERENCIAS BIBLIOGRAFICAS

<u>QUINQUENIOS</u>	<u>% ART REF DIA</u>
1921-1925	70,2
1926-1930	89,1
1931-1935	81,6
.....
1940-1944 (1)	72,7
1945-1949 (1)	86,9
1950-1954 (1)	88,4

=====

(1) : Datos obtenidos mediante muestreo