

# La Bibliometría-Informetría como herramienta al servicio de la archivística

**Dra. MERCEDES DE LA MONEDA**

**Dra. ROSARIO RUIZ-BAÑOS**

F. de Biblioteconomía y Documentación. U. de Granada

El objeto de la presente comunicación es analizar las posibles aplicaciones de la bibliometría en el campo de la gestión archivística y esbozar un pequeño ejemplo práctico.

## Introducción

La archivística ha evolucionado en las últimas décadas de un modo espectacular, pasando de una concepción monopolizada por los aspectos relacionados con la conservación de documentos a una concepción en la que el valor dominante es el de servicio e integración en las instituciones a las que sirve, con una concepción integral de la gestión de los documentos y de la unidad de información. Este proceso evolutivo está muy ligado a las nuevas realidades posibilitadas por los cambios tecnológicos, que nos han permitido ampliar la capacidad de gestión, pasando del control y gestión de los datos a la gestión de información (años ochenta) y de esta a la gestión del conocimiento (años noventa). En cierto modo se ha producido un cambio conceptual que amplía los servicios del clásico archivo, con la consiguiente mudanza en la terminología utilizada. Hoy los conceptos en alza se vinculan a gestión de la información y del conocimiento, documento electrónico, sistemas de información, auditoría de la información, etc., como se mostrará en un trabajo de próxima aparición.

En cualquier caso no podemos olvidar que un sistema de gestión del conocimiento se basa en un sistema de información estructurado, que a su vez gestiona datos contenidos en documentos, de distintas tipologías y soportes. Por tanto para que una entidad, sea cual sea su forma jurídica, pueda acceder a las ventajas competitivas derivadas de la gestión del conocimiento, debe contar de base con un sistema de gestión de la información y por tanto de gestión de documentos, que garantice el control, el acceso y la conservación de esto. Pero además necesitamos contar con herramientas metodológicas que nos permitan obtener información sobre las característi-

cas y evolución de la información, y por tanto de los documentos que la soportan, dentro de la entidad, siempre con el objeto de mejorar el rendimiento de la información concebida como recurso intangible y fundamental de la organización. Es en este punto donde la bibliometría-informetría puede prestar un servicio importante a la gestión de la información y de los documentos en una entidad.

La bibliometría-informetría se ha utilizado básicamente al servicio de la política científica, aplicándose a la evaluación de instituciones y de la producción/difusión científica, fundamentalmente de unidades de investigación o de áreas temáticas concretas, siendo muy amplia tanto la literatura teórica al respecto como las aplicaciones prácticas. Podemos decir que se trata de una metodología de evaluación de los procesos de producción y consumo de la actividad científica, con el consiguiente conjunto de herramientas y con un elevado nivel de aceptación por parte de la comunidad internacional, lo que ha quedado reflejado en la publicación de distintos manuales para la aplicación de indicadores, como la familia de Manuales de Frascati (OCDE, 1976, 1993, 2002), El Manual de Oslo (OCDE, 1992, 1997), el Manual de Canberra (OCDE, 1995) o las directrices de Y. Okubo para la elaboración de indicadores bibliométricos (Okubo, 1997)

Dado que nos encontramos ante un corpus de leyes y herramientas consolidado y ampliamente experimentado (Egghe, 1990), creemos que es plausible la exportación del mismo a ámbitos de la gestión de información distintos a los vinculados a la actividad científica. La afirmación anterior se basa en la existencia de un amplio factor común a la producción y consumo de información científica y a la producción y consumo de información interna en una entidad con otro tipo de fines. En esencia las variables fundamentales son básicamente las mismas en los dos campos, estribando la diferencia

en la unidad de observación elegida.

Si analizamos la literatura vemos que son escasísimas las aplicaciones de la bibliometría-informetría al mundo de la archivística. Una búsqueda en Lisa incluyendo los conceptos (record\* or archiv\*) and (bibliometri\* or informetri\* or scientometri\*) nos lleva a recuperar escasos documentos pertinentes. Se han publicado algunos trabajos sobre la producción científica en archivística (Ruiz Rodríguez, 1996), pero prácticamente no ha habido publicaciones centradas en la aplicación de las técnicas bibliométricas a la gestión y evaluación de archivos, Lisa solo recoge un estudio sobre el archivo privado de Getulio Vargas en Brasil (Bomeny, 1978). Resultados parecidos, en cuanto a su escasez obtenemos al centrar la búsqueda en gestión de la información y bibliometría.

Sin embargo y a pesar de la escasa literatura creemos que no es solo viable, sino también conveniente, la aplicación de las técnicas bibliométricas-informétricas a la gestión de los archivos, con independencia de la tipología de estos. En las siguientes páginas adelantaremos algunas reflexiones sobre las posibilidades teóricas de aplicación de las leyes bibliométricas a la gestión de las unidades archivísticas, así como de su rentabilidad para mejorar el sistema de información de una entidad.

### ALGUNAS REFLEXIONES SOBRE LA BIBLIOMETRÍA-INFORMETRÍA Y SUS POSIBLES APLICACIONES

**La ley de Bradford** (Egghe, 1990) básicamente nos informa sobre las tasas de concentración de información sobre un asunto dado en distintos canales de distribución.

Nos planteamos una posible aplicación de la ley sustituyendo el concepto de revista como canal de distribución por el de serie documental. Los resultados nos permitirán determinar el núcleo de series con mayor información sobre un tema dado, sobre una entidad, la mayor presencia de una tipología documental, etc, etc.

Si relacionamos esta información con otras variables podemos obtener información útil para distintos momentos del proceso de gestión de documentos. Por ejemplo:

- la concentración de las distintas tipologías documentales en las series generadas por la empresa, nos aportará información valiosa para auditar los procesos de gestión de documentos y tomar decisiones a la hora de priorizar tipologías en el proceso de normalización de documentos, indicándonos por que series y por que tipologías debemos comenzar en función de su presencia.
- Para diseñar la política de expurgo será conveniente

estudiar cuestiones como la relación entre la variable obtenida como consecuencia de la aplicación de la ley de Bradford a las tipologías documentales y su carácter de original o copia o su valor legal.

- En el archivo histórico será de gran utilidad para los investigadores conocer el grado de concentración de la información sobre los distintos temas, localizaciones geográficas, etc en las distintas series documentales.

**La ley de Lotka** (Egghe, 1990) y el conjunto de indicadores derivados del estudio de la productividad, también parecen tener una aplicación fácil y rentable en el mundo de la producción archivística.

- Si tomamos como unidad de producción el documento y como autor a la unidad administrativa emisora del mismo, obtendremos el ranking de volumen de producción de las distintos departamentos de la entidad, información de gran utilidad a la hora de auditar el sistema de gestión de la entidad.
- Si tomamos como unidad de producción el expediente y como autores los distintos departamentos que intervienen en su resolución obtendremos valiosos datos sobre la productividad fraccionaria de los distintos departamentos.

**La ley de Brookes** (Egghe, 1990), que en su origen aporta información sobre el envejecimiento de las publicaciones, también puede ser de gran utilidad en la gestión de los documentos de una entidad. Veamos algunas posibilidades:

- Si nuevamente sustituimos el concepto de revista por el de serie documental o simplemente por el de recurso de información y a su vez sustituimos el concepto de cita recibida por el de consulta de documentos o recursos, obtendremos información muy fiable sobre la vida media de las distintas tipologías documentales o recursos. Esta información puede ser utilizada como un variable determinante a la hora de definir la política de transferencia y expurgo de la documentación.

**La ley de Zipf** (Egghe, 1990), nos aporta información sobre el comportamiento de las palabras en un corpus lingüístico. Su comprobación y los indicadores que de ella se derivan pueden tener también distintas aplicaciones en la gestión de la información interna generada por una entidad, por ejemplo:

- El conocimiento del Corpus lingüístico manejado por la entidad y su comportamiento nos ayudará a definir las características de la cultura de la empresa y la evolución de sus valores dominantes.
- El conocimiento de los términos más utilizados y su

frecuencia de uso es un buen punto de partida para la elaboración de herramientas de control terminológico como los thesauros.

- Con vistas a la investigación en fondos histórico la ley de Zipf nos permitirá determinar los temas más usuales y los términos más frecuentes bajo los que estos aparecen.

### DISTRIBUCIÓN DE DESCRIPTORES DE LOS DOCUMENTOS DEL ARCHIVO MUNICIPAL DE ALMERÍA (SIGLOS XV Y XVI)

A título de ejemplo ilustrativo del potencial de la bibliometría aplicada a archivos, se ha considerado el caso de los descriptores dados a los documentos del Archivo Municipal de Almería de los siglos XV y XVI (Alcoer Martínez, 1986). En el Anexo **Error! Unknown switch argument.** se recoge la frecuencia de los descriptores por orden descendente. Esta forma de representar los datos de frecuencias frente a rangos se denomina en bibliometría-informetría "Distribución de Zipf".

La primera información que nos aporta la distribución de Zipf es poner de manifiesto, ordenadamente, qué descriptores son los más significativos. El más abundante es el término "Cargos", que aparece en 88 documentos, seguido de una serie de descriptores con frecuencia progresivamente más baja. Se observa que hay muy pocos descriptores de alta frecuencia y muchos de frecuencia bajísima: la mitad de ellos, desde la posición 56 hasta la 115, aparecen tan solo una o dos veces. Esto ya es una indicación de que hay una concentración de documentos sobre una temática muy similar, representada por los descriptores más abundantes. Mas adelante, con la distribución Bradford, se podrán delimitar las tres zonas típicas de concentración.

La distribución del "Anexo **Error! Unknown switch argument.**" puede correlacionarse con la denominada "Ley de Zipf", que dice que la frecuencia,  $F(r)$ , es inversamente proporcional a una potencia del rango,  $r$ .

$$F(r) = \frac{k}{r^m}$$

**Ecuación Error! Unknown switch argument.**

Donde  $k$  es una constante y  $m$  el llamado exponente de Zipf. En el caso particular en que  $m$  sea igual a la unidad, la Ecuación **Error! Unknown switch argument.** se transforma en la siguiente:

$$F(r) \times r = k$$

**Ecuación Error! Unknown switch argument.**

que indica que, en algunos casos, la frecuencia por el rango es una constante. En el caso general, será constante la frecuencia por una potencia del rango.

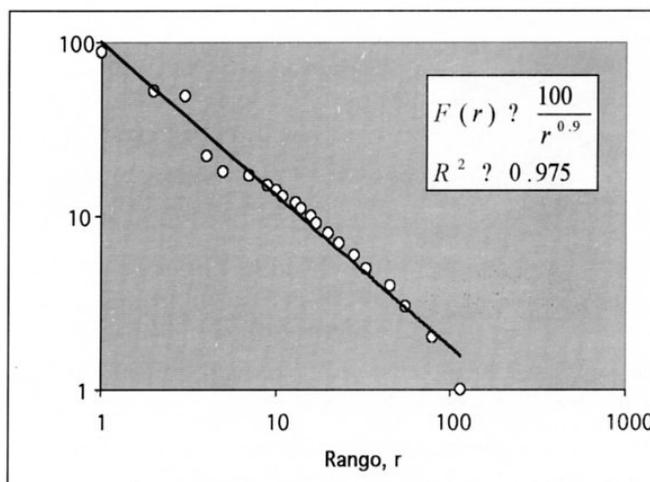


Figura **Error! Unknown switch argument.**.-Ajuste a la Ley de Zipf

Mediante regresión lineal, se han ajustado los valores empíricos de la distribución de Zipf del Catálogo del Archivo Municipal de Almería en los siglos anteriormente citados, según la indización realizada por Adela Alcoer Martínez y se ha encontrado que se ajusta perfectamente a una Ley de Zipf (Figura **Error! Unknown switch argument.**). Esto es muy positivo porque confirma que la indización realizada presenta un comportamiento regular, garantía de un trabajo acertado.

Como ya se ha adelantado, representar los datos según una "distribución de Bradford", nos va a permitir definir un núcleo de descriptores, una zona intermedia (también llamada zona recta) y finalmente una zona periférica denominada inflexión de Groos. Esto se lleva a cabo haciendo una representación de las frecuencias acumuladas frente al logaritmo del rango. En el caso que nos ocupa, el resultado es el de la "Figura **Error! Unknown switch argument.**". La frecuencia acumulada total es de 646. Por regresión lineal, se han ajustado los datos centrales a una "ley de Brookes-Ferreiro", que nos delimita entonces las tres zonas. El núcleo son los descriptores más abundantes que se separan de la recta de regresión y la inflexión de Groos son aquellos que también se separan pero en el otro extremo de muy baja frecuencia absoluta (elevada frecuencia acumulada, o "cola final" de la distribución)

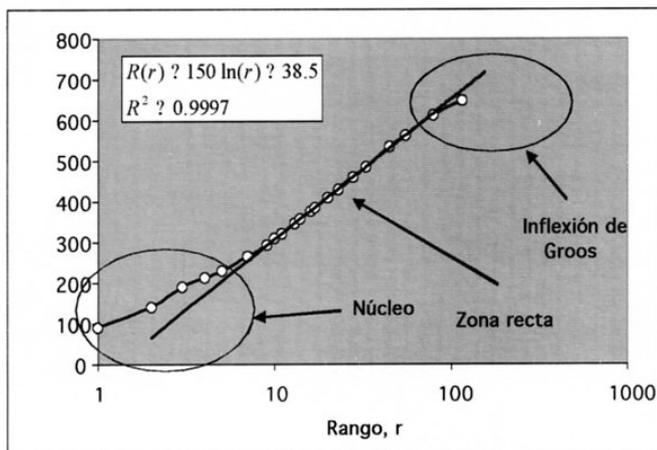


Figura Error! Unknown switch argument..- Distribución de Bradford

Con este análisis se encuentra que el núcleo está formado por los 7 primeros descriptores, que son los mostrados en la Tabla Error! Unknown switch argument.. Teniendo en cuenta que en total se han dado 115 descriptores, el núcleo está formado por el 6% de todos ellos, pero acumula el 41% de las frecuencias. Por ello, se deduce que este pequeño grupo de términos es capaz de representar las principales preocupaciones de los ciudadanos de Almería en los Siglos XV y XVI.

Tabla Error! Unknown switch argument..- Descriptores del núcleo

Rango	Descriptores	Frecuencia
1	Cargos	88
2	Regidores	52
3	Nombramientos	49
4	Propios	22
5	Ordenanzas	18
6	Pleitos	17
7	Abastecimiento	17
<b>Frecuencia acumulada núcleo 263 (41%)</b>		

La zona recta va desde el descriptor de rango 8, "Pan", hasta el descriptor del rango 55, "Aljibes". Por tanto, está formada por 47 descriptores (41% del total de ellos) que acumulan 299 ocurrencias en documentos (un 46% de la frecuencia total acumulada). Este grupo es casi 7 veces más grande que el núcleo, pero acumula más o menos la misma frecuencia (46% frente a 41%).

Por último, la inflexión de Groos está conformada por 60 descriptores (52% de todos ellos), pero que por su baja frecuencia, acumulan tan solo el 13% de las frecuencias. Evidentemente, son descriptores menos representativos de los documentos de los siglos XV y XVI de Almería.

## CONCLUSIONES

La bibliometría puede ser una buena herramienta de gestión al servicio de los archivos, como ha quedado de manifiesto en el ejemplo del Catálogo Documental del Archivo Municipal de Almería.

En nuestras reflexiones iniciales se proponen numerosas líneas de actuación en que la bibliometría-informetría podría ser de gran utilidad. Nuestra intención es abrir las primeras brechas en este campo de investigación que esperamos sea muy fructífero.

## REFERENCIAS

- Alcozer Martínez, A. (1986). Catálogo documental del archivo municipal de Almería. Siglos XV y XVI. Almería: Ayto., 1986.
- Bomeny, R. H. D.; Vargas, G. (1978). Estudio bibliométrico aplicado al Archivo Privado de Getulio Vargas. A bibliometric study of the private archives of Getulio Vargas. Ciencia-da-Infomacao. 7 (1) 37-42.
- Egghe, L.; Rousseau, R. (1990). Introduction to informetrics: quantitative methods in library, documentation and information science. Amsterdam: Elsevier.
- OCDE (2002). La mesure des activités scientifiques et techniques. Méthode type proposée pour les enquêtes sur la recherche et de développement experimental. Manuel de Frascati 2002- París: OCDE
- OCDE, EUROSTAT (1997). Principes directeurs proposés pour le recueil et l'interprétation des données sur l'innovation technologique. Manuel d'Oslo. París: OCDE
- OCDE, EUROSTAT (1995). Manuel sur la mesure des ressources humaines consacrées à la science et à la technologie. Manuel de Canberra. París: OCDE
- Okubo, Y. (1997). Indicateurs bibliométriques et analyse des systèmes de recherche: méthodes et exemples. OCDE/GD(97). París: OCDE
- Ruiz Rodríguez, A.A.; Peis, E. Moneda, M. de La (1996). Análisis de las publicaciones periódicas en materia de archivos: la hemeroteca técnica. En: Actas de las IX Jornadas Bibliotecarias de Andalucía: El bibliotecario ante la revolución tecnológica. Granada, pp. 476-484

Anexo Error! Unknown switch argument..-Distribución de frecuencias de los descriptores

Rango, r	Descriptor	Frecuencia, F(r)
1	Cargos	88
2	Regidores	52
3	Nombramientos	49
4	Propios	22
5	Ordenanzas	18
6	Pleitos	17
7	Abastecimiento	17
8	Pan	15
9	Impuestos	15
10	Obras	14
11	Salarios	13
12	Moriscos	12
13	Ganado	12
14	Trigo	11
15	Jurisdicción	10
16	Franquicia	10
17	Rentas	9
18	Murallas	8
19	Milicia	8
20	Elecciones	8
21	Seda	7
22	Censos	7
23	Armamentos	7
24	Pesca	6
25	Lugartenientes	6
26	Bastimentos	6
27	Alcaldes	6
28	Agua	6
29	Viviendas	5
30	Repartimiento de los Reyes Católicos	5
31	Mayordomos	5
32	Jurados	5
33	Gastos	5
34	Procuradores	4
35	Ofensas	4
36	Monasterios	4

37	Limpieza de sangre	4
38	Libramientos	4
39	Hierbas	4
40	Escribanos	4
41	Concordia	4
42	Bulas	4
43	Alguaciles	4
44	Alcaides	4
45	Acequias	4
46	Prohibición	3
47	Navíos	3
48	Guerra	3
49	Gobierno	3
50	Fieles ejecutores	3
51	Esparto	3
52	Corregidores	3
53	Aposentamiento	3
54	Almojarifazgo	3
55	Aljibes	3
56	Venta	2
57	Terremotos	2
58	Síndico	2
59	Recaudación	2
60	Rebelión	2
61	Presos	2
62	Precios	2
63	Pósito	2
64	Murcianos	2
65	Mudéjares	2
66	Moros	2
67	Libre circulación	2
68	Hermandad	2
69	Heredades	2
70	Galeras	2
71	Fuentes	2
72	Feria y fiestas	2
73	Fallecimientos	2
74	Escrituras públicas	2
75	Diputados	2

76	Bienes comunales	2
77	Almotalifes	2
78	Alifes	2
79	Alcazaba	2
80	Visita	1
81	Viajes	1
82	Velas	1
83	Testamento	1
84	Tesoreros	1
85	Tabernas	1
86	Salteadores	1
87	Rogativas	1
88	Puertas	1
89	Prendas de vestir	1
90	Pobres	1
91	Piratería	1
92	Penas de cámara	1
93	Penados	1
94	Pena de muerte	1
95	Noria	1

96	Juez de residencia	1
97	Juboneros	1
98	Inundación	1
99	Guardas de costas	1
100	Fuero eclesiástico	1
101	Fuero	1
102	Escuderos	1
103	Emperador	1
104	Dehesa	1
105	Cristianos viejos	1
106	Clérigos	1
107	Casa-concejo	1
108	Casa-audiencia	1
109	Cárcel	1
110	Capitulaciones	1
111	Capitán general	1
112	Caminos	1
113	Calceteros	1
114	Armada Real	1
115	Alférez mayor	1