

AREAS DE INVESTIGACION A DESARROLLAR EN EL CAMPO DE LA INFORMACION Y DOCUMENTACION CIENTIFICA Y TECNICA

ROSA SANCHO

Colaborador Científico. Instituto de Información y Documentación en Ciencia y Tecnología (C.S.I.C.)

En este artículo, se presenta una panorámica de las posibilidades que ofrece la Ciencia de la Información y Documentación, como posible campo de investigación en sus dos vertientes:

a) Investigación aplicada de tipo finalista, indispensable para desarrollar mejores sistemas de tratamiento, almacenamiento y transmisión de la información científica y técnica.

b) Investigación básica acerca del fenómeno de la información científica y técnica como disciplina, para deducir las leyes que rigen el tamaño y crecimiento de la Ciencia. Así como de las necesidades y empleo de la información por parte de las comunidades científicas, para deducir indicadores de actividad y producción científica, modelos de conducta, etc. de dichas comunidades.

Se indican las tendencias en esta disciplina en los países más desarrollados así como en España.

Básicamente el fenómeno de la información científica y técnica comprende tres estados: generación, transmisión y utilización, los cuales hay que contemplarlos dentro de un proceso dinámico e interactivo de transferencia de conocimiento, como se muestra en el siguiente modelo básico de comunicación de información.

generación de
información

mensajes
informativos

utilización de
la información

(fuentes)

(canales de información)

(receptores)

La generación de la información científica, se produce durante el proceso de investigación y desarrollo científico y técnico. Su comunicación es en esencia un fenómeno de transmisión de conocimiento, ya que los resultados de cualquier investigación deben hacerse fácilmente disponibles para la comunidad científica, que es la receptora, y a su vez creadora de información.

Los vehículos utilizados para la transmisión de información científi-

ca y técnica, son principalmente las publicaciones periódicas primarias, y las actas de congresos, simposia, informes, tesis, etc. Los artículos aparecidos en dichas publicaciones, no son subproductos o epifenómenos de la actividad científica, sino el producto final de dicha actividad.

El continuo incremento de la ciencia (crecimiento exponencial), da como resultado enormes cantidades de publicaciones primarias cuyo número se duplica aproximadamente cada 10 años.

El tratamiento de la ingente cantidad de información que se genera, contempla las vertientes teórica y práctica. La primera, consagrada al estudio de las propiedades de la información en si misma, y la segunda al desarrollo de métodos para hacerla asequible y de rápida difusión.

Esta última, comprende el área que se denomina *Tecnología de la Información*, la cual ha experimentado un enorme desarrollo en los últimos tiempos gracias a los grandes avances producidos en la Informática y la Telecomunicación, sin cuyo apoyo, sería impensable abordar hoy día cualquier investigación en este campo.

El análisis del contenido, clasificación, indización y resumido de la literatura científica, para la creación de publicaciones secundarias y bases de datos automatizadas, a partir de la misma; así como nuevas posibilidades de procesado y acceso a dicha literatura; creación de soportes cada vez más adecuados para el almacenamiento de la información, etc., ofrece un amplio campo de investigación cuyo objetivo sería el de proporcionar los medios técnicos necesarios para aumentar la eficacia en el manejo de la información, y por tanto, la capacidad de comunicar los resultados de la investigación científica en cualquier rama.

A continuación citaremos dos importantes tipos de estudios, que se pueden llevar a cabo dentro de la disciplina que nos ocupa y que son creadores, a su vez, de ciencia informativo documental.

1. ESTUDIOS DE LOS SISTEMAS DE ALMACENAMIENTO Y RECUPERACIÓN AUTOMATIZADA DE INFORMACIÓN

La investigación en este terreno ocupa un puesto principal en la mayoría de los países industrializados, donde se ha creado una potente «Industria de la Información», con fuerte apoyo económico, y que origina importantes beneficios, por lo que, dichos países no escatiman esfuerzos para conseguir sistemas informativos cada vez más eficaces.

La consolidación de la citada industria de la información, es una de las características fundamentales de la segunda mitad del siglo XX, que ha transformado profundamente a la sociedad científica.

De forma esquemática citaremos en este apartado las investigaciones y estudios sobre:

- a) Creación de bases y bancos de datos (bibliográficos, factuales, de texto completo, etc.)
- b) Desarrollo de logiciales adaptados a la recuperación automatizada a partir de bases y bancos de datos.

- c) Desarrollo de lenguajes de recuperación versátiles y potentes.
- d) Estudios sobre las diferentes posibilidades de recuperación y almacenamiento de texto completo.
- e) Estudios para facilitar la interconexión entre diferentes sistemas de recuperación.
- f) Diseño de bases de datos para «hardware» especiales.
- g) Creación de nuevos soportes para el almacenamiento de la información (video discos ópticos, ópticos digitales, etc.).
- h) Desarrollo de las llamadas «revistas electrónicas» (publicaciones editadas en-línea, y de acceso también en-línea).

2. ESTUDIOS SOBRE INTELIGENCIA ARTIFICIAL Y LINGÜÍSTICA

Los grandes esfuerzos dedicados a la investigación y desarrollo en estos campos, han producido un gran número de proyectos bien conocidos en la literatura.

Esquemáticamente podemos incluir en este apartado, investigaciones y estudios acerca de:

- a) Indización automática de documentos.
- b) Procesado y análisis del lenguaje natural.
- c) Resumido automático de documentos.
- d) Conversión automática de textos.
- e) Traducción automática de textos.
- f) Desarrollo de sistemas de reconocimiento de voces, lenguajes, etc., para abordar estas tareas, es necesaria una profunda investigación lingüística sobre estructura de las lenguas, determinación estadística sobre frecuencia de palabra de valor informativo, etc.

Pasemos ahora a considerar la vertiente teórica de la investigación en este campo, o lo que podíamos denominar *Ciencia de la Información y Documentación*.

Trataremos este apartado algo más ampliamente, por considerar que el fenómeno de la Información constituye una disciplina por derecho propio, que ofrece problemas de gran interés para el investigador, y que cuenta con sus propios métodos de investigación.

Según Resnikoff (1), la ciencia de la información se ocupa de la estructura, comportamiento y propiedades de la información, mientras

(1) Resnikoff, H. Information Science and Technology at the National Science Foundation. Inform. Proc. Managem., v. 16, pp. 243-250, 1980.

que la tecnología de la información se ocupa de los medios técnicos disponibles para usarla en un intercambio efectivo.

Es un hecho comprobado, que la literatura científica tiene la propiedad de mostrar un comportamiento regular, habiéndose establecido una serie de modelos matemáticos (normalmente denominados leyes) verificados repetidamente que expresan aproximadamente los aspectos del comportamiento de dicha literatura.

Es por esto que sometiendo a un análisis riguroso (diversas aproximaciones matemáticas) a la literatura científica que se publica, estaremos en condiciones de resolver problemas generales de la ciencia misma, si la consideramos como entidad medible.

Podemos pues afirmar que la Ciencia de la Información y Documentación (*), es una ciencia auxiliar de todas las demás, y que ejerce una influencia directa en el progreso de las mismas.

Taylor R. S. (2), define la Ciencia de la Información como: «La ciencia que estudia las propiedades y comportamiento de la información, las leyes que gobiernan el flujo y el proceso de la información, para conseguir una accesibilidad y uso óptimo de esta información».

Según López Yepes, J. (3), esta ciencia estudia los procesos de comunicación científica, tendentes a establecer las bases de nuevos conocimientos, es decir, investiga los fundamentos informativos de la propia investigación.

Con ayuda de esta disciplina, se podrán deducir leyes que rigen el tamaño y crecimiento de la ciencia a gran escala, así como el modelo teórico de las necesidades y empleo de la información por parte de la comunidad científica, pudiéndose de esta forma diseñar diversos indicadores de la actividad científica para cada comunidad, (indicadores de la producción científica, del consumo de información y de los modelos de conducta de la comunidad científica).

Los resultados procedentes de estas investigaciones representan una gran ayuda para la toma de decisiones en política científica con un criterio fiable.

Podemos dividir la Ciencia de la Información y Documentación en tres grandes capítulos:

-
- (*) «Information Science», según la denominación anglosajona. «Informatika», según la Unión Soviética. «Documentation und Information-wissenschaft», según la terminología alemana.
 - (2) Taylor, R. S.
Professional aspects of Information Science and Technology. Annual Review of Information Science and Technology. C. A. Cuadra (Ed.). v. 1, p. 19, 1966.
 - (3) López Yepes, J.
Teoría de la Documentación.
Ediciones Universidad de Navarra, S. A. Pamplona, 1978.

1. *Cienciometría Documentaria*

Solla Price, D. J. (4), profesor de Historia de la Ciencia de la Universidad de Harvard, y uno de los precursores de este campo, denominó a esta disciplina la «Ciencia de la Ciencia». Introdujo la aplicación de métodos científicos para la investigación del rápido crecimiento de la ciencia, y de la función de las publicaciones científicas, así como de la evolución de las organizaciones científicas.

Nalimov, V. V. y Mulchenko E. M. (5), definen esta actividad como la investigación sobre el desarrollo de la ciencia como proceso informativo, dado que el desarrollo de las ciencias incide directamente en la calidad y cantidad de literatura producida. Lo que significa que estudiando esta literatura, podremos valorar esa ciencia y a los científicos que la producen.

Este capítulo cubre las siguientes áreas:

- a) Estudio sobre la producción y productividad de los autores científicos, organismos de investigación, escuelas científicas, etc. (Índice de Lotka).
- b) Determinación del «impacto» y «visibilidad» de los trabajos científicos en el mundo científico internacional. (Índice de Platzl).
- c) Determinación de los «colegios invisibles» y otras relaciones entre científicos (redes de citas, colaboraciones en los trabajos, transmisión de ideas científicas, etc.)
- d) Indicadores sobre relaciones entre científicos, instituciones y trabajos a través de las conferencias impartidas. (Índice de Kessler).
- e) Indicadores sobre relaciones entre científicos, instituciones, y trabajos a través de las co-citas. (Índice de Small).

Estos estudios sobre la evaluación de la actividad investigadora (seguimiento de grupos de científicos), se basan principalmente en el recuento y análisis de citas que tienen por objeto proporcionar una medida para determinar la importancia relativa de las publicaciones, y al mismo tiempo, la valoración de los trabajos publicados.

Se han producido numerosas listas de clasificación de revistas en orden de su importancia, de acuerdo con la frecuencia de citas. Se ini-

(4) De Solla Price, D. J.
Little Science, Big Science.
Columbia University Press, 1963.
Traducido al castellano: Hacia una Ciencia de la Ciencia. Editorial Ariel, p. 183, 1973.

(5) Nalimov, V. V. y Mulchenko, E. M.
Study of the development of Science as an information process. Moskva, Nauka, p. 192, 1969 (trad. inglés AD 735634).

ciaron con Gross P. L. K. y Gross E. M. (6), más tarde seguidos por Brown C. H. (7), Garfield E. (8) (9), Martín J. (10), etc.

Sin duda, la práctica del análisis de citas, recibió un estímulo considerable con la aparición en 1963 de las publicaciones periódicas de índices de citas: «Science Citation Index», y el repertorio de indicadores bibliométricos «Journal Citation Report», del Institute for Scientific Information de Filadelfia.

2. Sociometría Documentaria

Considerando que el proceso informativo se desarrolla dentro de un contexto social, este capítulo se ocupa de las características y estructura social de los grupos productores (autores), y consumidores (usuarios), que pueden ser científicos, académicos, industriales, gestores, etc.

Según un reciente informe de Wilson T. D. (11), el estudio de los usuarios ocupa actualmente el segundo lugar (después de la recuperación de información), en cuanto a esfuerzos dedicados a investigación en información, en los países industrializados.

Dentro de este capítulo se estudian:

- a) Necesidades de información de distintos grupos de usuarios.
- b) Relación entre la información que reciben o generan estos grupos, y la investigación que están llevando a cabo; es decir, valor de la información para la solución de problemas y en el desarrollo de la ciencia.
- c) Análisis de las relaciones entre las variables de los distintos servicios de información, y la eficacia o efectividad de los individuos y organizaciones que utilizan esos servicios de información.

-
- (6) Gross, P. L. K. y Gross, E. M.
College Libraries and Chemical Education.
Science, V. 66, pp. 385-389, 1927.
 - (7) Brown, C. M.
Scientific Serials: Characteristics and list of the most cited publications in Mathematics, Physics, Chemistry, Geology, Physiology, Botany, Zoology and Entomology.
ACRL Monograph n.º 16. Chicago Association of College and Research Libraries, 1956.
 - (8) Garfield, E.
Citation Analysis as a tool in Journal Evaluation.
Science, 178 (4060), P. 471-479, 1972.
 - (9) Garfield, E.
Significant Journals in Science.
Nature, 264, pp. 609-615, 1976.
 - (10) Martyn, J.
Citation Analysis.
J. Doc. v. 31, pp. 290-297, 1975.
 - (11) Wilson, T. D.
On User Studies and Information Needs.
J. Doc. V. 37, n. 1, pp. 3-15, 1981.

- d) Interacciones existentes entre los grupos productores y consumidores de literatura científica.
- e) Sistemas de información más utilizados por los científicos.

3. *Bibliometría*

Esta disciplina se ocupa del rendimiento de los fondos de publicaciones científicas, así como de la selección y consumo por parte de los usuarios.

Se estudian aquí:

- a) Análisis del tamaño, crecimiento y envejecimiento de la literatura científica. Indicadores de «Obsolescencia» y «vida media» de la literatura. (Índice de Burton y Kebler).
- b) Dispersión de la literatura científica. Núcleos de publicaciones más utilizadas. (Ley de Bradford).

Estos tres capítulos presentados, se encuentran grandemente interrelacionados entre sí. López Piñero, J. (12), en su libro «Análisis estadístico y sociométrico de la literatura científica», presenta una revisión sistemática de estos temas.

La investigación de los tres campos señalados, puede ser de carácter descriptivo o analítico, y se dirige hacia la obtención de nuevos procedimientos y conceptos, en busca siempre de una mayor eficacia en las bases de partida para el establecimiento de las adecuadas políticas de adquisición de fondos bibliotecarios, políticas de seguimiento científico, de decisión sobre nuevos campos científicos a desarrollar, etc.

Como hemos visto a lo largo de este trabajo, la disciplina que hemos venido llamando Ciencia de la Información y Documentación, tiene un carácter multidisciplinar. Intervienen en ella: Matemáticas, Lógica, Estadística, Lingüística, Psicología, Sociología, Tecnología de las comunicaciones, Cibernética, etc.

Es pues imprescindible conocer y aplicar adecuadamente las materias arriba indicadas, para conseguir un mayor desarrollo de la Ciencia de la Información.

Precisamente, el retraso que padece esta disciplina respecto a otras ciencias sociales (Sociología, Pedagogía, Economía, etc.), puede ser debido, en parte, a que en ella las matemáticas se empezaron a aplicar con rigor científico mucho más tarde que en estas otras. En líneas generales, se admite que la matematización de un campo del saber, es paralelo a su científicación, al hacer más racional y objetiva la investigación en ese campo.

(12) López Piñero, J. M.

Análisis estadístico y sociométrico de la literatura científica.

Cuadernos de Documentación e Informática Médica. Fac. Medicina, Valencia, p. 82, 1972.

Recientemente, en un trabajo publicado por Nacke O. (13), se propone la matematización de la Ciencia de la Información, integrando la ciencia matemática a los problemas de la Ciencia de la Información en una subordinación denominada *Informetría*.

La National Science Foundation (14), da prioridad a este campo desde un punto de vista más amplio y trata de desarrollar modelos análogos a los econométricos, que harán posible predecir el impacto de los campos económicos, tecnológicos, etc., en la componente «información» de la sociedad.

INVESTIGACIÓN EN INFORMACIÓN Y DOCUMENTACIÓN EN EL MUNDO

En el encuentro que tuvo lugar en el año 1979, en la Technical School of Information and Computer Science, de Georgia (EE.UU.), sobre «programas de investigación nacional en Ciencias de la Información», y en la que estuvieron representados los países más desarrollados, (Alemania, Francia, Japón, Unión Soviética, Inglaterra y EE.UU.), se analizaron los programas de investigación en esta ciencia en los últimos 15 años, y se presentó la situación de los programas nacionales en ese momento.

Se diferenciaron claramente tres períodos de tiempo:

—Década de los 60 y primeros años 70.

La investigación en materia de Información y Documentación, se inclinaba de una parte, a la organización y estructura de sistemas y servicios orientados a disciplinas concretas y en base a las necesidades de los usuarios, y de otra parte, a la experimentación de nuevas técnicas de tratamiento de la información.

—Década de los años 70.

En esta época la investigación en este campo se centraba en la instalación de grandes sistemas informatizados de tratamiento de la información y documentos.

—Años 80.

Por parte principalmente de EE.UU., se presentan planes de investigación que enfatizan por primera vez la naturaleza científica de los problemas de la información, y sin abandonar la investigación de tipo finalista, orientado al diseño y desarrollo de sistemas de información, se preparan para realizar una investigación básica en «Información Science», que se centra en la naturaleza puramente científica de los problemas de información.

(13) Nacke, O.
Informetrie, Ein neuer Name für eine neue Disziplin.
Nachr. f. Dokum. v. 30, n. 6, p. 219, 1979.

(14) National Science Foundation.
Information Science and Technology.
Nat. Sci. Found. Program. Report v. 3, n. 5, 1979.

En un reciente artículo de Méndez A. (15), se presenta una panorámica de la investigación en Ciencias de la Información y Documentación en los países más desarrollados. Todos ellos:

- a) Dedicar a esta disciplina un sustancioso porcentaje de la cantidad destinada a investigación en general, gracias a lo cual, la información científica puede desarrollar su propia investigación.
- b) Mantienen una infraestructura adecuada y poseen un potencial humano muy importante dedicado a estas tareas.
- c) Este tipo de investigación, a diferencia de la que se lleva a cabo en otros campos científicos, es impulsada a nivel gubernamental.

En relación con esto, han aparecido a lo largo del tiempo, gran cantidad de revistas periódicas de gran prestigio, especializadas en investigación en esta disciplina, tanto desde el punto de vista de la tecnología de la información, como de la teoría de la Ciencia de la Información. Citaremos sólo a modo de ejemplo: *Journal of Documentation*; *American Documentation*; *Le Documentalist*; *Information Processing Management*; *Information Storage and Retrieval*; *Journal Chemical Information and Computer Science*; *On-line Review*, etc. Recientemente, ha aparecido *Scientometrics*, que se subtitula: «Revista internacional para todos los aspectos cuantitativos de la "Ciencia de la Ciencia" y de la política científica».

INVESTIGACIÓN EN ESPAÑA SOBRE INFORMACIÓN Y DOCUMENTACIÓN

La participación española en este campo es muy escasa, tanto en lo que se refiere a investigación de tipo finalista como a investigación básica. Respecto a esta última podemos citar los siguientes ejemplos:

El grupo del Profesor López Piñeiro, del Instituto de Información y Documentación en Biomedicina de Valencia, que ha sido pionero en los estudios de investigación estadística y sociometría de la literatura científica española, en las áreas de Medicina y Farmacia, habiendo realizado varias tesis doctorales sobre estos temas. Citaremos a modo de ejemplo, y por haber sido la primera, la de Terrada M. L. (16).

Del mismo modo, desde la Cátedra de Documentación de la Facultad de Ciencias de la Información, de la Universidad Complutense de Madrid, el Profesor López Yepes, está potenciando también la investigación en este sentido.

Aparte de los trabajos de estos grupos, los Institutos de Información y Documentación en Ciencia y Tecnología (ICYT), y en Ciencias Sociales

(15) Méndez, A.

La Información y Documentación Científica, en busca de identidad como Ciencia. *Mundo Científico*, n. 24, pp. 417-422, 1983.

(16) Terrada, M. L.

La literatura médica española contemporánea. Estudio estadístico y sociométrico. Cuadernos de Documentación e Informática Médica. Fac. Medicina. Valencia, 1973.

y Humanidades (ISOC), del C.S.I.C., han venido publicando, con carácter esporádico, algunos estudios de investigación en esta disciplina, entre los que cabe citar: productividad de los autores del C.S.I.C., producción bibliográfica española en diversos campos de la ciencia, etc. Citamos sólo a modo de ejemplo, (17) (18) (19).

A nivel gubernamental, la CAICYT, responsable de la política científica, ha financiado dos trabajos dentro del área de la Sociometría de la Ciencia, sobre productividad científica española en el campo de la Física.

Por otra parte, esta misma institución, para realizar la evaluación de los resultados de los proyectos de investigación por ella financiados, ha decidido aplicar algunos indicadores bibliométricos, como son: «factores de impacto», en base a la frecuencia relativa de citas recibidas por los artículos publicados, «índice de vigencia» de las publicaciones, etc.

También, el Ministerio de Cultura, ha otorgado dos ayudas de investigación en el campo de la Información Científica. Una de ellas se refiere a la aplicación de las técnicas bibliométricas en la evaluación de la literatura científica, y otra a la elaboración de un tesoro sobre nomenclatura de productos farmacéuticos.

Es obvio que estas pequeñas aportaciones demuestran el escaso interés que las Ciencias de la Información y Documentación suscitan en nuestro país. Lo realizado hasta ahora en este campo, no es suficiente para mantenerle a un nivel adecuado.

Es significativo que la Comunidad Económica Europea, encargara a Méjico la realización de un programa sobre terminología científica en español para ser usado por la Comunidad. Asimismo, la UNESCO, encargó también a Méjico el desarrollo de la parte española del glosario multilingüe sobre terminología de información y documentación.

-
- (17) Ferreiro, L., Mendez, A.
La productividad de los autores del C.S.I.C. durante el año 1977. *Rev. Esp. Doc. Cient.* v. 3, n. 1, 1980.
- (18) Rey Guerrero, A., Gutiérrez Muñoz, F.
Producción bibliográfica española en el campo de la Ciencia y Tecnología de Alimentos. *Rev. Agroquim. Tecnol. Alim.*, v. 21, n. 3, 1981.
- (19) Vázquez Valero, M., Sancho Lozano, R.
Estudios de la producción científica española sobre polímeros en el período 1974-79. *Rev. plásticos Mod.* v. 294, pp. 713-720, 1980.