

Revista Ciencias de la Información Vol. 33, No. 2, Mayo - Agosto, 2002

Tabla de contenidos

Artículos

Juan Antonio Moreiro-González, Juan Lloréns-Morillo, Miguel Ángel Marzal-García-Quismondo, Jorge Morato-Lara, Sonia Sánchez-Cuadrado, Pilar Beltrán-Orenes. [Utilización de estructuras verbales en la identificación de relaciones y descriptores en tesauros](#). Vol. 33, No. 2, mayo - agosto, 2002, p. 3-14

Rubén Urbizagástegui-Alvarado, María Teresa Cortés-Contreras. [La productividad de autores en la Revista Geológica de Chile](#). Vol. 33, No. 2, mayo - agosto, 2002, p. 15-25

Marlery Sánchez-Díaz, Juan Carlos Vega-Valdés. [Algunas consideraciones teórico-conceptuales sobre la Bibliotecología y la Ciencia de la Información](#). Vol. 33, No. 2, mayo - agosto, 2002, p. 27-37

Conferencia

Leonid P. Pavlov. [La literatura gris de Rusia en el mundo de la información digital](#). Vol. 33, No. 2, mayo - agosto, 2002, p. 39-43

Utilización de estructuras verbales en la identificación de relaciones y descriptores en tesauros¹

José Antonio Moreiro González

Juan Lloréns Morillo

Miguel Ángel Marzal García-Quismondo

Jorge Morato Lara

Sonia Sánchez Cuadrado

Pilar Beltrán Orenes

RESUMEN

Se parte del desarrollo que los tesauros automatizados han proporcionado a las llamadas relaciones no-clásicas de Van Slype, fomentando una eclosión de las relaciones de asociación, agrupadas en categorías por Tudhope, Alani, Jones y engrosadas por el traspaso de relaciones antes de jerarquía por la American Library Association (ALA). Frente a la estaticidad y clasificación taxonómica de los tesauros clásicos, inadaptados al dinamismo, transversalidad, exigencia de mayor abstracción de la información en hipermedia, se proponen las formas verbales como conceptos dinámicos, relacionales, conceptualizadores de la acción, muy útiles para representar el contexto mediante relaciones circunstanciales, a partir de los precedentes establecidos por A. L. Tharp y SYNTOL. Se indican como sus aplicaciones inmediatas la determinación de géneros de Swales, la tipología de las secciones del documento y unas funciones tesaurales más eficaces en red. Se concluye con futuros desarrollos de investigación en las categorizaciones asociativas, desde las experiencias aplicativas de WordNet.

ABSTRACT

The background of the work presented in this paper is based on automatic thesauri, and the improvement they provided to the "so called" not-classical relationships proposed by Van-Slype. We intend to enhance the usage of association relationships, grouped by Tudhope, Alani y Jones and emphasized by the American Library Association (ALA) relationships translation. We fight the static approach, as well as the taxonomical classification presented in classical thesauri, not allowing them to represent dynamics, transversal relationships, and hypermedia navigation. This paper propose "verbal forms" as dynamic concepts, origin to relationships, essentials of action, and very useful to represent context by circumstantial relationships (based on the antecedents of A. L. Tharp and SYNTOL). The main application areas of our work are gender determination of Swales, document sections typology and a more effective thesauri functions in the net. Future research developments include associative categorization from WordNet applicative experiences.

¹ Este trabajo ha sido realizado dentro del marco del Proyecto financiado por la CICYT (Comisión Interministerial de Ciencia y Tecnología), titulado "Desarrollo de un tesoro de verbos para entornos de información dinámica. Aplicación del estándar ISO/ICE: 13250:1999, del Plan General del Conocimiento. TIC 2000-2003.

Introducción

Revisión de las relaciones tesaurales

La norma ISO 2788 (1986)² marca que las relaciones básicas de un tesoro son de tres tipos: equivalencia, jerárquica y asociativa. En el presente documento se hace una revisión de los principales cambios que han tenido lugar desde la publicación de la norma en las relaciones empleadas en los tesauros, para a continuación proponer un enfoque que emplea relaciones diferentes, atendiendo a categorías gramaticales distintas de las recogidas por dicha norma. La categoría gramatical en la que nos centraremos es la de los verbos. Se revisan, así, las posibles ventajas que pueda tener la inclusión de formas verbales como términos documentalmente relevantes a la hora de contextualizar la información y automatizar la construcción de los tesauros.

Van Slype [1] amplía el número de relaciones presentes de facto en los tesauros hasta cinco, incluyendo las relaciones que él denomina de pertenencia y de equivalencia interlingüística. Ahora bien, en la propuesta presentada por Van Slype aparece un sexto grupo de relaciones que él denomina no-clásicas por no estar presentes en la mayor parte de los tesauros en uso [2]. Este grupo de relaciones englobadas por él bajo el epígrafe único no-clásicas son: parte-todo, colocación, relación paradigmática, taxonomía y sinonimia.

Si analizamos con detenimiento estas últimas relaciones, podemos concluir que no son más que reformulaciones de relaciones que ya estaban contempladas en las relaciones que Van Slype denomina clásicas. Así, por ejemplo, la relación no-clásica parte-todo, que él ejemplifica con “cuerno PART vaca”, no deja de ser un caso concreto de la relación jerárquica partitiva, en la que un descriptor

es una parte particular de otro descriptor superior. En los ejemplos de este autor, “Francia y Europa” o “motor y vehículo”, responden ambos a la pregunta ¿es uno parte de otro? [3].

Las relaciones que están experimentando una mayor expansión son las de asociación. La expansión se está produciendo a dos niveles: por una parte se están definiendo nuevas relaciones de asociación y, por otra, se están considerando como relaciones de asociación algunas relaciones que antes eran consideradas sólo de jerarquía.

En la actualidad, esta duplicidad esbozada por Van Slype está siendo concretada para tesauros automatizados. El trabajo de D. Tudhope, H. Alani y C. Jones [4] sobre la expansión de las relaciones en los tesauros automatizados actuales es una buena muestra de ello. Para estos autores, las relaciones que están experimentando una mayor expansión son las de asociación. La expansión se está produciendo a dos niveles: por una parte se están definiendo nuevas relaciones de asociación y, por otra, se están considerando como relaciones de asociación algunas relaciones que antes eran consideradas sólo de jerarquía. Basándose en un estudio presentado por la American Library Association (ALA) [5], Tudhope, Alani y Jones hablan de nueve grandes

2. Documentación – directrices para el establecimiento y desarrollo de Tesauros monolingües (parte I)

3. La norma ISO 2788 (1986), en su apartado 8.4.1., resume a dos los posibles tipos de términos que puede conectar la relación asociativa:

a) Aquellos que pertenecen a la misma categoría;

b) Aquellos que pertenecen a categorías diferentes.

Dentro de los que pertenecen a la misma categoría, señala como único grupo el de aquellos términos emparentados a través de su significado (apartado 8.4.2.1.). En cuanto a los que pertenecen a categorías diferentes, la enumeración de los tipos se extiende a diez: una disciplina o campo de estudio y los objetos o fenómenos estudiados, una operación o proceso y su agente o instrumento, una acción y el resultado o producto de tal acción, una acción y su sujeto pasivo, conceptos y sus propiedades, conceptos relacionados con sus orígenes, conceptos ligados por una dependencia causal, objetos y sus contraagentes, conceptos y unidades de medida, frases sincategoremáticas y los sustantivos implicados.

grupos de relaciones de asociación que engloban las once que presenta la norma ISO 2788 (1986),³ y que, con sus correspondientes subtipos, aumenta el número de forma considerable el número de posibles relaciones de asociación de un tesoro.

Estos diez subtipos de relaciones de primer nivel son:

- 1) Ideas combinadas.
- 2) Términos relacionados conceptualmente.
- 3) Contigüidad.
- 4) Relaciones asociativas por definición.
- 5) Relaciones asociativas con diferente jerarquía o facetadas.
- 6) Relaciones asociativas traslapadas por significado.
- 7) Relaciones asociativas con idéntica jerarquía.
- 8) Cuestiones de finalidad.
- 9) Relaciones asociativas sin especificar.

Los niveles llegan hasta seis en profundidad, lo que da un número total de relaciones de asociación de cientoveintitrés. Esto supone un aumento de ciento doce relaciones de asociación. Por otra parte, un estudio de las relaciones de jerarquía que ofrecen estos dos autores, basado también en los estudios de la ALA, nos lleva a la conclusión de que muchas relaciones que eran recogidas en el conjunto de las denominadas de jerarquía han pasado a ser clasificadas como relaciones de asociación. Algunas de ellas coinciden con las que Van Slype llamaba no-clásicas, como por ejemplo las relaciones del tipo parte-todo, que están empezando a ser consideradas como de asociación, aunque siguen apareciendo como de jerarquía, y son las que se duplican en la nueva caracterización de la ALA (aparecen como relaciones de asociación y de jerarquía, como por ejemplo las de parte-todo de sistemas u órganos anatómicos).

El principal problema que enfrentan los tesauros clásicos en la actualidad no es otro que el de la obsolescencia.

Este cambio en las relaciones de un tesoro viene dado, a nuestro entender, por el principal problema con el que se enfrentan los tesauros clásicos en la actualidad que no es otro que el de la obsolescencia. Y la razón fundamental de la obsolescencia de los tesauros hay que buscarla, a nuestro juicio, en la nota característica de los tesauros tradicionales: la estaticidad de la indización y de las relaciones establecidas que les es propia⁴ Las estructuras que se encuentran definidas en un tesoro clásico hacen que sólo pueda recuperar los documentos que posean aquellos términos y relaciones estáticas con las que trabaja y que han sido definidas a priori. Esta limitación, que conlleva la obsolescencia, provoca el silencio o el ruido en los casos en que la búsqueda se haga sobre los mismos términos y relaciones predefinidas, pero en los que hayan cambiado su significado debido a nuevos usos. Esta clase de problemas surgían ya cuando el universo de discurso documental es el que vamos a calificar de clásico,⁵ esto es, cuando se resume a unos tipos de documentos concretos muy bien encuadrados en un ámbito cultural y científico, o en ambos, perfectamente delimitado. En este caso, el de los documentos clásicos, hay que señalar, además, que, fuera cual fuera la temática concreta de los documentos tratados y su forma de presentación, siempre contaban con una presentación física (tridimensional), que permitía su localización espacio-temporal en unas coordenadas concretas. Es decir, los tesauros clásicos se diseñaban pensando en un universo documental muy bien parcelado temáticamente (salvando excepciones) y que contaba con una plasmación física concreta de los documentos a los que se orientaba.

4. Hemos decidido dejar fuera de este contexto otro tipo de problemas que pueden influir en la eficacia de un tesoro por referirse más al ámbito de los recursos humanos como pueden ser: la especialización, o falta de ella, de aquellos que llevan a cabo un tesoro concreto, los errores tipográficos y conceptuales debidos al factor humano, por mencionar algunos ejemplos.

5. Llamamos documentos clásicos a los documentos existentes antes de lo que se empieza a conocer como la era de la Sociedad de la Información, que va acompañada por el fenómeno de la globalización.

6. Documentos como pueden ser páginas web, que sólo tienen vida mientras que están en ella, pero que desaparecen definitivamente si su autor o autores deciden en un momento dado retirarlas de la web.

El universo documental actual ha cambiado sensiblemente con la aparición de la denominada Sociedad de la Información

Sin embargo, el universo documental actual ha cambiado sensiblemente con la aparición de la denominada Sociedad de la Información, y lo ha hecho en los dos aspectos señalados. Por una parte, la creciente especialización en los diferentes ámbitos culturales, en concreto los que están relacionados con los campos científicos surgidos desde la misma rama común, hace mucho más difícil la tarea de distinguir en clases diferentes los textos que son de un área científica y los de otras que se encuentran muy próximas a ella [6]. Además, los contenidos de determinadas áreas como las relacionadas con la Informática, están ofreciendo documentos expresados en lenguajes mucho más abstractos que los usados hasta el segundo tercio del siglo XX, con gran cantidad de gráficos y formatos de diseño, en resumen, menos alfabetizados. Por otra parte, las formas de presentación de los documentos (y no nos referimos sólo a los diferentes soportes posibles) ya no se limitan a las coordenadas espacio-temporales, sino que existen documentos (muchos de los contenidos en la red Internet) que no son localizables mediante estas coordenadas.⁶ La consecuencia directa de estos cambios es la imposibilidad de llevar a cabo un tesoro específico, fiable y actualizado, que cumpla bien las funciones de indización precisa, pero sobre todo, que sea útil en el momento de la recuperación. Así pues, los problemas concretos que presenta un tesoro clásico se traducen en la práctica en dos graves inconvenientes:

- 1) *Dificultad para describir campos con un nivel de abstracción alto* (el ámbito de la Informática, por ejemplo), o en aquellos en los que, por el contrario, los documentos no estuvieran estructurados en ninguna medida (como puede ser el caso de los textos de libre formato).
- 2) *Un elevado costo, ya sea en la creación o en el mantenimiento*, que sólo se podría eliminar

con la automatización o semiautomatización de ambas tareas. Lamentablemente, esta automatización difícilmente puede ser llevada a cabo de forma eficaz y eficiente desde la caracterización actual de los tesauros estáticos, debido a la riqueza de las estructuras semánticas de los textos (habría que definir una infinidad de nuevas relaciones cada vez que nos enfrentásemos a nuevos textos).

El dinamismo en la producción de información, propio de esta nueva Sociedad, cuenta con una nítida proyección epistemológica y apunta directamente al conocimiento

Hay que tener en cuenta, además, que el dinamismo en la producción de información, propio de esta nueva Sociedad, cuenta con una nítida proyección epistemológica y apunta directamente al conocimiento. Si bien durante la modernidad el acceso a la observación y la comprensión del mundo eran objetivas, es decir, se encontraban fuera del individuo, en el postmodernismo, sin embargo, tanto una como otra no existen fuera del individuo, ya que la tecnología permite al individuo diseñar su propio acceso a la información y, en última instancia, al conocimiento. La repercusión es inmediata para la organización del conocimiento, pues desaparecen las referencias absolutas y, por consiguiente, la clasificación y ordenación de temas deben referirse más a conveniencias, o posibles conveniencias, constituyendo de este modo una realidad que satisface en un momento y para una finalidad determinados.⁷

Se produce, lógicamente, una disarmonía entre el dinamismo en la creación de la información -y, por ende, del conocimiento- y los actuales lenguajes de clasificación del conocimiento, fundamentados en campos académicos normalizados. La clasificación por disciplinas es un sistema de clasificación del

7. Arguye el dinamismo del conocimiento en el postmodernismo Miska, F. L. *The DDC, The universe of knowledge and the postmodern library*. Forest Press Albany (NY), EE. UU., 1998.

conocimiento íntimamente referido a un método de investigación racionalista y pragmático. Este sistema está perdiendo utilidad debido, en gran parte, a los nuevos métodos y formas de producir, disseminar y usar la información que constituye, en última instancia, el receptáculo natural del conocimiento. Las nuevas comunidades discursivas cooperan en la producción de documentos que no se adaptan a las estructuras académicas [7], en tanto que las búsquedas se orientan cada vez más a un material no disciplinar. La automatización y la técnica hipertextual han dado respuestas en la recuperación documental, pero su eficacia está determinada por la necesidad de una nueva clasificación del conocimiento para optimizar la información. A este respecto, se ha desarrollado el concepto de disciplinaridad compartida en la creación del conocimiento, lo cual trae consigo que, para autores como D. W. Langridge [8], la clasificación deba distinguir entre disciplinas (formas de conocimiento) y fenómenos (objetos de conocimiento).

Las formas verbales como conceptos dinámicos

Como se comentó al inicio del presente documento una de las corrientes principales a la hora de denominar nuevas relaciones en tesauros se basa en el empleo de formas verbales en los tesauros. Aunque la carga conceptual de los textos recaiga en los sustantivos, éstas constituyen los conectores naturales de esos conceptos de los que los sustantivos son depositarios, y son las que dan un sentido concreto a esa carga conceptual,⁸ que estaría incompleta si las formas verbales no completasen su contenido. Es decir, aunque el contexto lo aporten los sustantivos de forma habitual, las formas verbales son las que declaran las acciones y, consecuentemente, las que dan una orientación concreta a un determinado texto, cuando no son las que dotan al texto de la carga conceptual necesaria para clasificarlo.

Una de las corrientes principales a la hora de denominar nuevas relaciones en tesauros se basa en el empleo de formas verbales

La inclusión de los tesauros de formas verbales que complementen, sin eliminarlos, a los tesauros estáticos tradicionales parece la solución que mejor recoge la problemática actual. Como descriptores, las formas verbales presentan evidentes ventajas en determinados casos como, por ejemplo, a la hora de la realización de búsquedas de documentos en los que se recojan acciones y para la indización de materiales especiales. Ejemplos concretos de estos usos son: la indización de imágenes de video mediante la identificación del verbo en gerundio que expresa la acción, o la utilización de formas verbales en los diagramas de clases tanto en la denominación de los métodos como en las asociaciones de clases. Otro ejemplo clarificador puede verse con los documentos de análisis de requisitos, este tipo de documentos cuentan con estructuras lingüísticas sencillas unidas por formas verbales (por ejemplo, afirmaciones como “El parámetro B debe ser mayor que 18”), y en las que lo importante es cuál es la conexión que se establece entre los elementos relacionados, y no ellos mismos.

La unión de un tesoro estático clásico de sustantivos con uno dinámico de formas verbales permitiría una indización automática más flexible

8. Desde el punto de vista lingüístico toda proposición debe tener un verbo, es más, todo verbo, incluso aunque esté aislado, constituye una proposición, y no lo es la expresión que carezca de él.

Por otra parte, la identificación del papel de una asociación mediante un verbo permite un abanico de relaciones mucho más adaptable a dominios concretos. Además, existe la posibilidad, en caso de realizar una indización automatizada, de grabar para su posterior uso ciertos modos y tiempos verbales como matices de la relación con la entidad, que pueden ser muy útiles en las tareas de identificación y localización de un documento a través de la lematización. Este aserto halla su más concreta constatación en la documentación científica, donde se emplean determinadas formas verbales para plantear la hipótesis (condicional), el estado de la cuestión (presente), método (futuro, potencial), conclusiones (pasado), como también un jugoso empleo del gerundio en la indización para los documentalistas de unidades de información audiovisual, tal como ya se señaló.

En resumen, esta nueva fórmula, que supone la unión de un tesoro estático clásico de sustantivos con uno dinámico de formas verbales, permitiría, a nuestro entender, una indización automática más flexible, es decir, la dinamización del tesoro eliminaría en gran medida los problemas en la representación de algunos dominios y aumentaría, al mismo tiempo, la precisión y eficacia de los mismos.

Justificación documental de la utilización de formas verbales en la recuperación automática de información

Dado que las principales funciones de un tesoro son la indización y la recuperación desde el punto de vista de la carga conceptual de los documentos, parece lógico que la atención a la hora de su diseño debe centrarse en cómo conseguir que el tesoro detecte la carga conceptual relevante en cada documento, atendiendo a su estructura lingüística en sus cuatro vertientes: léxica, morfológica, sintáctica y semántica.

Desde una perspectiva puramente documental, se podrían dividir en dos tipos de redes los resultados de la clasificación automática de la información. Por un lado, estarían las redes que crean las relaciones con carácter permanente, esto es, lo que se representa en los tesauros clásicos. En este tipo de red tiene mucha importancia contar con un número limitado de relaciones de asociación, equivalencia y jerarquía. Estas relaciones son percibidas por todos como obvias en determinado dominio (o, si se prefiere, son inherentes al conocimiento compartido por una comunidad

determinada) y son analizados, independientemente del vocabulario empleado en los corpora del dominio (ya que son creados a priori con respecto al proceso de indización). Estas relaciones que denominamos permanentes son las que aparecen en los tesauros clásicos de manera habitual.

Por otra parte se encuentra la red de relaciones circunstanciales, en la que sí es importante tener un número elevado de formas verbales para representar todas y cada una de las relaciones que se puedan dar. Las relaciones que tenemos en una red de relaciones circunstanciales son, principalmente, del tipo “asociado con”, y, en este caso, necesitamos, además, un conjunto de subtipos dependientes del dominio para subdividirlas, esto es, necesitamos una serie de especificaciones dadas por el dominio concreto y que sólo rigen en él. Es decir, el por qué de la necesidad de un gran conjunto de tipos de relaciones circunstanciales (representados por formas verbales) se debe a que el contenido (la semántica) de cierto documento distinto al de los recogidos hasta ese momento es, por su misma novedad, impredecible. Una red de relaciones circunstanciales es la herramienta idónea para representar el contenido de una colección documental (y para, posteriormente, facilitar la recuperación).

Actualmente los trabajos se orientan a utilizar las gramáticas semánticas y sintácticas para establecer relaciones

De cara a la automatización, es la red circunstancial (la correspondiente a los subtipos del dominio) la determinante de la mayoría de sistemas automáticos de clasificación. En estos sistemas se selecciona un corpus de documento y sobre la base de la frecuencia de coaparición de las palabras en los documentos, y entre los documentos, se hace una clasificación de tipo estadístico (redes neuronales, kmeans axiales, coocurrencia de términos, algoritmos genéticos, etc). Es decir, se parte de la indización de un corpus documental del dominio para construir el tesoro. De hecho, tradicionalmente se han creado redes de relaciones no sólo circunstanciales mediante herramientas estadísticas; actualmente los trabajos en ese sentido se orientan a utilizar las gramáticas semánticas y sintácticas para establecer relaciones, sin que esto suponga que no se pueda mejorar los resultados o acotar los listados de términos con herramientas estadísticas clásicas.

Una cuestión que queda pendiente es si es posible pasar de un tesauro de relaciones permanentes a uno de relaciones circunstanciales, basándonos exclusivamente en parámetros estadísticos. Parece ser que sí, pero siempre que se parta de un conjunto de documentos lo suficientemente grande, pero esto es, precisamente, lo que se pretende eliminar con la automatización.

Antecedentes del uso de formas verbales en los tesauros

Desde un punto de vista histórico, nuestra propuesta no constituye el primer intento en el que se usan las formas verbales como elemento en la recuperación automática o manual de la información; así como tampoco es la primera vez que se usan para la fundamentación teórica del tipo de herramientas que deben ser usadas en estas tareas. Ya en la década del 60 y del 70 existen algunos antecedentes en la aplicación automática y manual de los verbos para la mejora de la recuperación de información y como elementos claves a tener en cuenta en la fundamentación teórica.

Entre los antecedentes estrictamente teóricos que proponen al verbo como elemento principal en la indización y la recuperación de los documentos cabe destacar la propuesta de Alan L. Tharp. En un escrito de 1973 titulado *Using verbs to automatically determine text descriptors*, Tharp [9] hacía una defensa, desde el punto de vista teórico, de este uso de los verbos en el contexto correspondiente a las tareas de recuperación automatizada de la información. Según Tharp, los métodos de indización automática que se basan exclusivamente en estadísticas, permutaciones, citaciones o asociaciones no permiten la exactitud en la búsqueda y la recuperación, puesto que lo hacen atendiendo a criterios extrínsecos. A su juicio, este tipo de métodos indizan sin tener en cuenta el contexto de cada uno de los términos.

Una cuestión que queda pendiente es si es posible pasar de un tesauro de relaciones permanentes a uno de relaciones circunstanciales, basándonos exclusivamente en parámetros estadísticos

La idea de fondo que mantiene Tharp a lo largo de su argumentación es que un buen sistema de recuperación de la información debe seguir los esquemas de memoria del ser humano. Y, según sus propias palabras, “la simulación de la memoria humana conduce a usar el verbo como el elemento significativo en la determinación de los descriptores del texto” [10, p. 247]. Esta afirmación se apoya, a su vez, en estudios psicológicos que demuestran que es una característica de la memoria humana el ser imprecisa para recuperar de manera estática, y, sin embargo, muy fiable en la recuperación del movimiento, esto es, de acciones.

Tal como sostiene Tharp, conviene recordar que todos los sistemas de indización, clasificación y recuperación automática se basan en la medida de diversos factores (estadísticas, permutaciones, citaciones y asociaciones, por ejemplo). A juicio de Tharp, el hecho de que los verbos significativos, es decir, los que transmiten una mayor cantidad de información contextual, aparezcan en un número muy pequeño de veces, los haría desaparecer de la lista de posibles términos del documento si se utilizan los anteriores parámetros, lo cual resulta contraproducente para recoger la carga conceptual

9. Un completo estudio se encuentra en Halliday and Hasan, M. A. K. *Language, context, and text: aspects of language in a social-semiotic perspective*. Universidad de Oxford, 1989.

completa del mismo. Siguiendo los mismos estudios propuestos para los sustantivos en el caso de los verbos (y sorteando los problemas de las formas flexionadas de los verbos), tendríamos que esos verbos con poca frecuencia de aparición, pero muy representativos desde el punto de vista contextual, dejan de ser candidatos para representar al texto como los identificadores del mismo a la hora de la búsqueda y recuperación del documento. Es lógico pensar que cualquier tipo de filtro o normalización implica una pérdida de semántica, por ejemplo, podemos normalizar los términos niñas, niña, niño y niños a una misma forma canónica, por ejemplo, niño. Pero ¿qué ocurre si en el futuro queremos sólo recuperar documento que traten de niñas pero no de niños? Es por el mismo hecho por lo que hacer desaparecer en la indización términos con carga conceptual relevante, en este caso los verbos, puede limitar en el futuro las posibilidades de búsqueda y recuperación.

Un ejemplo de plasmación concreta de este tipo de desarrollos automatizados o semiautomatizados se puede encontrar en el SYNTOL [11]. El SYNTOL fue un estudio auspiciado por la Comunidad Europea de la Energía Atómica (EURATOM) para la elaboración de un tesoro especializado, que intentaba sortear la caducidad de los lenguajes de indización del tipo de los tesoros al uso en ese momento [12]. La implantación del SYNTOL fue un fracaso en su momento por su carestía; ahora bien, el elevado coste estaba relacionado directamente con la limitación de los ordenadores con los que se contaba en ese momento. Los equipos actuales abaratarían notablemente el gasto en este tipo de implementaciones, y esto tanto desde el punto de vista crematístico como por el tiempo que las máquinas necesitarían emplear actualmente para llevar a cabo las rutinas de indización y recuperación.

Sobre la utilización de formas verbales aplicada a la contextualización de la información

La utilización de formas verbales aplicada a la contextualización de la información ha sido estudiada en tres vertientes concretas: determinación de géneros, determinación de la tipología de las secciones del documento y la determinación temática.

En cuanto a la determinación de géneros, Swales [13] caracterizó los géneros según los formas

verbales presentes en cada documento. Se centró en dos factores: cuáles son los verbos prototípicos y en qué flexión (tiempo/modo/número) se encuentran más frecuentemente. Por ejemplo, en un estilo no académico destaca la disminución de formas verbales en pasiva y presente simple aumentando la del pasado simple.⁹ Si es posible esta caracterización es de suponer que el proceso contrario también es posible. Es decir, buscando ciertos verbos y flexiones verbales podemos contextualizar el género en el que se produce, lo que solucionaría en gran medida la problemática de los documentos que se pueden localizar en Internet, que no es otro que la mezcla de distintos géneros con calidad variable.

En el caso del vocabulario científico, por ejemplo, prevalecen los verbos de tipo informativo como: demostrar, establecer, argumentar,... Así como también son frecuentes en dicho ámbito los verbos negativos que nos indican las carencias que intentamos salvar con la documentación: faltar, subestimar, adolecer, eludir, etc.

En lo relacionado con la determinación de la tipología de las secciones del documento, con un enfoque similar se ha determinado el posicionamiento en las distintas secciones de los documentos. Un buen ejemplo es el Heslot [14], quien diseñó el esquema de distribución de formas verbales que aparece en la tabla 1.

Tanto la determinación de géneros como de secciones de un documento difieren en su modo y tenor dependiendo del campo [15] al que pertenezca el documento, lo que nos lleva, nuevamente, a la necesidad de aplicar también una red semántica en la indización y establecimiento de relaciones de cara a la recuperación. En este punto conviene recordar que ya se han empleado este tipo de herramientas en la contextualización de la información, como puede ser el intento de Haas [16] para determinar de modo automático artículos de carácter experimental, sin embargo, tampoco el trabajo de Haas cuenta con un tesoro de formas verbales que complementa a los clásicos de sustantivos.

Por último, en la determinación temática una de las vías de investigación más exitosas en los últimos años es WordNet [17]. La forma de afrontar el problema es utilizar los términos raíces de los verbos para acotar tanto el género como el tema [18]. Esta red semántica de WordNet presenta, sin embargo, los mismos problemas que señalábamos en el caso anterior, es decir, no cuentan con un tesoro de formas verbales que sirva de complemento al de

Tabla 1. Esquema de distribución de formas verbales de J. Heslot

CARACTERÍSTICA	INTRODUCCIÓN	MÉTODOS	RESULTADOS	DISCUSIÓN	AUTORES
<i>Comentarios autor</i>	Alto	Muy bajo	Muy bajo	Alto	Adams Smith
<i>Pasiva</i>	Bajo	Alto	Variable	Variable	Heslot
<i>Pasado</i>	Bajo	Alto	Alto	Bajo	Heslot
<i>Presente</i>	Alto	Bajo	Bajo	Alto	Heslot
<i>Información</i>	Alto	bajo	Bajo	Alto	West

sustantivos, ya que se limita a ser un tesauro de verbos sin ninguna relación semántica con el de sustantivos.

Papel de la categoría de las formas verbales en los tesauros

Aunque en la ISO R1087 se comenta que “... los conceptos pueden ser no sólo cosas (expresadas mediante sustantivos), sino que en un sentido más amplio, también se puede tratar de cualidades (adjetivos y sustantivos), de acciones (como por ejemplo, verbos y sustantivos) e incluso lugares, relaciones o situaciones (expresadas mediante adverbios, preposiciones, conjunciones y sustantivos) [19, p. 129]”, no es menos cierto que, según la ISO 2788 (1986), en la construcción de un tesauro se deben evitar los usos como términos de indización de determinadas categorías gramaticales. Entre las categorías excluidas se encuentra la de los verbos [20], así como de adjetivos aislados – salvo determinadas excepciones-[21], y adverbios [22].

En cuanto a lo que atañe a la categoría gramatical de los verbos es preceptivo, según la segunda de las normas citadas, que se “presenten como sustantivos o formas verbales sustantivadas [23, p. 10]”. La interpretación estricta de la anterior afirmación ha sido la que más se ha extendido y, como consecuencia, el resultado es que las formas verbales sólo han sido consideradas como términos de cara a la indización y la recuperación cuando funcionan como sustantivos en el sentido gramatical. Sin embargo, siguiendo la norma ISO R1087, el tipo de conceptos que representan los verbos (formas verbales) es el de las acciones, que pueden ser concretadas en sustantivos o verbos, según dicha norma. Los sustantivos que responden a este tipo de conceptos, el de las acciones, suelen ser los que comparten su semántica con un verbo, de quien la reciben. Por ejemplo, el sustantivo “navegación” recibe su carga semántica del verbo “navegar”,

“aceleración” de “acelerar”, y así sucesivamente. Así pues, parece que nos encontramos frente a tipos conceptuales distintos, a los cuales hay que tratar, a nuestro juicio, desde su diferencia, y no como se ha hecho tradicionalmente, esto es, asimilándolos todos a la categoría conceptual de los sustantivos.

Mantener la idiosincrasia de los conceptos que representan acciones conlleva la inclusión en el tesauro de la categoría que los refleja en los documentos, sean estos sustantivos o formas verbales

Mantener la idiosincrasia de los conceptos que representan acciones conlleva la inclusión en el tesauro de la categoría que los refleja en los documentos, sean estos sustantivos o formas verbales. Así, por ejemplo, si nos encontramos con la frase “La aceleración del cohete espacial ha sido correcta”, tendremos que buscar la carga conceptual de la misma, además de en el sustantivo que representa la cosa de la que trata la frase, “cohete espacial”, en el sustantivo que representa la acción que concreta el contexto en el que se encuentra inmersa, “aceleración”. Y, en última instancia, tendremos que remitirnos al verbo del que recibe su carga conceptual, “acelerar”, que en este caso no aparece en la frase, es decir, el término “aceleración” no se refiere a una cosa, sino que está representando una acción. A nuestro entender, sólo así podremos recoger la carga conceptual de la frase, que quedaría incompleta si nos limitásemos al término “cohete espacial”. Esta posibilidad de precisar aún más el contenido conceptual de un texto delimita mucho

más la caracterización tipológica del documento, y hace que la indización sea más exhaustiva y, consecuentemente, la recuperación más exacta.

Futuros desarrollos

La creación de este tesoro de formas verbales, como complemento al tesoro clásico de sustantivos pasa por un proceso que no es meramente automático, sino que incluye otro proceso de naturaleza más lingüística en el que se puedan encajar las categorías verbales a modo de relaciones facetables. Este planteamiento nos separa de la mayoría de proyectos de este tipo, ya que casi toda la literatura trata de localizar un conjunto de términos (a los que denominan clusters), que resultan relevantes por su frecuencia de coaparición en determinado contexto.

Existen tres grandes corrientes en la que la inclusión de formas verbales en la indización y construcción de los tesauros:

- 1) *Utilización de clasificaciones verbales*, para mejorar la recuperación. La mayoría de los trabajos se basan en las aportaciones hechas por Levin [24]. Por ejemplo, destacan los trabajos de Green [25] en el que se crea una ampliación de la red semántica WordNet con los verbos de Levin.
- 2) *Utilización de WordNet* [26]. Se trata de una red semántica multidisciplinar en inglés. Su calidad y disponibilidad lo han convertido en la herramienta idónea en lingüística durante los últimos años. Su uso ha estado centrado desde un principio muy vinculado a la desambiguación conceptual. Dentro de esta tendencia una de las líneas que está dando mejores frutos es el empleo para desambiguar mediante verbos. Un ejemplo típico de desambiguación se puede encontrar en Moldovan¹⁰ en donde se esquematiza de la siguiente forma:¹¹
 - a) Se selecciona de la frase todas las parejas de sustantivo-verbo.

b) Se escoge el significado más probable del término (subproceso que Moldovan denomina Desambiguación Terminológica):

- i. *Las palabras de la frase se agrupan en parejas.*
- ii. *Se buscan en WordNet los distintos significados de cada término.*
- iii. *Se forman todos los diferentes pares de conceptos posibles.*
- iv. *Se busca cada par en Internet. Luego se ordenan los resultados según los conceptos más frecuentes, esto es, según el número de veces que en Internet nos aparezcan juntos dos conceptos determinados.*

c) *Teniendo en cuenta los conceptos más frecuentes (paso b), se seleccionan todos los sustantivos de los “glosarios” de cada verbo y sus descendientes jerárquicos.*

d) *Teniendo en cuenta los conceptos más frecuentes, se seleccionan todos los sustantivos de los “glosarios” de cada sustantivo y sus descendientes jerárquicos.*

e) *Se calcula mediante una fórmula los conceptos comunes entre los sustantivos del punto c y d.*

f) *Se ordenan todas las parejas de conceptos de sustantivo-verbo según el resultado de la fórmula.*

3) *Utilización de las formas verbales de los documentos de determinado dominio* [27]. En esta vía se pretende automatizar la construcción de tesauros basados en las concurrencias de descriptores del tesoro con formas verbales en determinado corpus documental. Esta última vía es la que estamos desarrollando, aunque empleando algunas de las tareas señaladas en las dos anteriores, y es por eso por lo que las incluimos como futuros desarrollos.¹²

10. Moldovan también ha propuesto otros algoritmos de desambiguación basados en WordNet, véase Sanda M. Harabagiu & Dan I. Moldovan. *Enriching the WordNet Taxonomy with Contextual Knowledge Acquired from Text. En Natural Language Processing & Knowledge Representation. Ed. Lucja M. Iwńska and Stuart C. Shapiro. Menlo Park (CA): AAAI Press, MIT Press, 2000. pp. 301-333.*

11. El autor denomina al método desambiguación semántica. *Using WordNet and lexical operators to improve Internet searches. Moldovan-DI; Mihalcea-R. IEEE-Internet-Computing. 4 (1):34-43, Jan.-Feb. 2000. Con este sistema Moldovan afirma que obtiene una considerable mejora frente a Yarowsky (1995) que desambiguaba el 94% de los sustantivos (no otros términos) y a Stetina (1998) y este trabajo, desambiguar el 80% de todas las partículas.*

12. En el próximo Congreso INFO 2002, que se celebrará en Cuba, presentaremos los primeros resultados de nuestro trabajo.

Referencias

- 1) Van Slype, G. *Los lenguajes de indización. Construcción y utilización de los sistemas documentales*. Fundación Sánchez Ruipérez. Madrid, España, 1991. 198 p.
- 2) Van Slype, G. Op. Cit., p. 58.
- 3) Van Slype, G. Op. Cit., p. 53.
- 4) Tudhope, Douglas, Harith Alani y Christopher Jones. "Aumenting Thesurus Relationships: Possibilities for Retrieval" [en línea]. Journal of Digital Information (1), febrero 2001. <<http://jodi.ecs.soton.ac.uk/Article/v01/i08/Tudhope>>. [Consulta: 27 de abril del 2001].
- 5) Appendix B (part 2) *Taxonomy of Subject Relationships compiled by Dee Michel with the assistance of Pat Kuhr* [en línea]. June 1996 draft (hierarchical display), en <<http://ala.org/alcts/organization/ccs/sac/appendxb.html>>. [Consulta: 8 de febrero del 2002].
- 6) Niiniluoto, Ilkka. *The emergence of Scientific Specialities: six models, Poznan Studies in the Philosophy of the Sciences and the Humanities* (44):211-223, 1995.
- 7) Clegg, S. R. y G. Palmer. *The politics of management knowledge*. Sage London, Inglaterra, 1996.
- 8) D. W. Langridge. *Classification, its kinds, systems, elements and applications*. Bowker Sauer London, Inglaterra, 1992.
- 9) Tharp, Alan L. *Using verbs to automatically determine text descriptors*. Information Storie and Retrieval (9):243-248, 1973.
- 10) Tharp, A., loc. cit., p. 247.
- 11) Cros, R. C., J. C. Gardin, y F. Lévy,. *L'automatisation des recherches documentaires. Un modèle general «LE SYNTOL»*, Gauthiers-Villars Paris, Francia, 1968, 260 p.
- 12) Cros, R. C., J. C. Gardin, y F. Lévy, *L'automatisation des recherches documentaires. Un modèle general «LE SYNTOL»*, pp. 5-6.
- 13) J. M. Swales,. *Genre analysis: English in academic and research settings*. Cambridge University Press, Cambridge, Inglaterra, 1990.
- 14) Heslot, J. *Tense and other indexical markers in the typology of scientific texts in English*. Hedt: 83-103, 1982.
- 15) Lavid, Julia. *Towards a text type taxonomy: a functional framework for text analysis and generation. Procesamiento Lenguaje Natural* (16):29-43, 1995.
- 16) Haas, S. W., J. Sugarman y H. Tibbo. *A text filter for the automatic identification of empirical articles*. Journal of the American Society for Information Science, 47(2):167-169, 1996.
- 17) Fellbaum, C. *Wordnet. An Electronic Lexical Database (Language, Speech and Communication)*. MIT Press Cambridge (Massachusetts), 1998.
- 18) Klavans, Judith y Min-Y Kan,. *Role of Verbs in Document Analysis*. En Proceedings of the Conference, COLING-ACL. Montreal (Canada): Université de Montreal, 1998. pp. 680-686.
- 19) ISO R1087, citado por Yukio Nakamura: "A Language for Knowledge Representation", en la revista *Advances in Knowledge Organization*, (4):127-133, 1994.
- 20) Norma ISO 2788 - UNE 50-106-90. *Directrices para el establecimiento y desarrollo de tesauros monolingües*, 1986, p. 10.
- 21) Norma ISO 2788 - UNE 50-106-90. *Directrices para el establecimiento y desarrollo de tesauros monolingües*, 1986, pp. 9-10.
- 22) Norma ISO 2788 - UNE 50-106-90. *Directrices para el establecimiento y desarrollo de tesauros monolingües*, 1986, p. 10.
- 23) Norma ISO 2788 - UNE 50-106-90. *Directrices para el establecimiento y desarrollo de tesauros monolingües*, 1986, p. 10.
- 24) Levin, B. *English verb classes and Alternations: a preliminary investigation*, University of Chicago Press Chicago, Ill, 1993.

25) Green, Rebecca, Lisa Pearl, Bonnie J. Dorr and Philip Resnik. *Mapping Lexical entries in Verbs Database to WordNet Senses*. En Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL'2001), Toulouse, France, July 9-11, 2001.

26) Fellbaum C., op. cit.

27) Díaz Rodríguez, S. I. *Esquemas de representación de información basados en relaciones : aplicación a la generación automática de representaciones de dominios*, Tesis doctoral, Director, Juan Lloréns Morillo. Leganés: Universidad Carlos III de Madrid, Departamento de Informática, 2001.

Recibido: 8 de febrero del 2002.

Aprobado: 10 de marzo del 2002.

José Antonio Moreiro González

*Departamento de Biblioteconomía y Documentación
y Departamento de Informática
Universidad Carlos III de Madrid
C/ Madrid, 126 28903 Getafe (Madrid) España
Correo electrónico: <jamore@bib.uc3m.es>.*

La productividad de autores en La Revista Geológica de Chile

Rubén Urbizagástegui Alvarado

María Teresa Cortés

RESUMEN

La presente investigación fue realizada analizando los datos de autoría de los artículos publicados en la Revista Geológica de Chile entre 1974 y 1997, cubriendo un período de 24 años. Fue adoptado el método del Conteo Directo, lo que significa que sólo los autores principales fueron acreditados con la contribución a un artículo. Fueron encontrados 178 autores que conjuntamente produjeron 320 artículos, y 72% de ellos, contribuyeron con un único artículo a la revista. El método de los Mínimos Cuadrados y el test Kolmogorov-Smirnov fueron usados para verificar los datos observados y esperados. Al 0.01 nivel de significancia, se encontró que el valor crítico fue 0.1222 y la D_{max} fue 0.1063. Se concluyó, entonces, que esta literatura se ajusta al modelo de Lotka.

ABSTRACT

The present research was based on data collected from articles published by the Revista Geologica de Chile between 1974 and 1997, covering a period of 24 years. Strait count was adopted and therefore only senior authors were credited with full contribution to the production of an article. In this period, 178 authors were identified as publishing a total of 320 articles in this journal; 72% of them produced only one paper. The Least Square Method and the Kolmogorov-Smirnov test were used to verify the observed and expected values. At the 0.01 level of significance the critical value was found to be 0.1222 and the D_{max} was 0.1063. Therefore, this literature did conform to the Lotka's law.

Introducción

La Revista Geológica de Chile, principal órgano de difusión científica de esta temática en el país, cuenta con escasos estudios bibliométricos, entre los cuales faltan investigaciones respecto a la productividad de sus autores. Los únicos artículos publicados hasta la fecha son aquellos de Urbizagástegui y Cortés [1, 2] y están referidos solamente a la obsolescencia de la literatura y al análisis de citas en este campo. En consecuencia, el objetivo de este trabajo es examinar en la literatura geológica chilena publicada en la Revista Geológica de Chile, posibilidades de identificar regularidades bibliométricas, tales como las propuestas por el modelo del poder inverso [3, 4].

Para establecer la proporción con que diferentes estudiosos contribuyen al progreso de la ciencia, Lotka [5] estableció los fundamentos de la Ley del Cuadrado Inverso y afirmó que el número de autores que hacen "n" contribuciones en un determinado campo científico, es aproximadamente $1/n^2$ de los que hacen una sola contribución, y que la proporción de los que hacen una única contribución es de alrededor de 60%. En general, este modelo trata de demostrar que en la producción de artículos existe una distribución estratificada, según la cual la mayoría de los artículos publicados están concentrados en una pequeña proporción de autores altamente productivos que, en relación con los autores únicos, tienen una regresión negativa igual a 2.

En la producción de artículos existe una distribución estratificada, según la cual la mayoría de los artículos publicados están concentrados en una pequeña proporción de autores altamente productivos

A pesar de las numerosas investigaciones realizadas sobre este asunto, hasta ahora los resultados no serían concluyentes y los datos no proporcionan un ajuste claro a la formulación teórica de esta Ley. Por ejemplo, Oppenheimer afirma que “*debe enfatizarse que la Ley de Lotka ha sido verificada en muchas colecciones de datos, pero el ajuste no siempre ha sido bueno* [6, p. 355]”. También Nicholls señala que “*los resultados de esos estudios son incomparables debido a las diferencias sustanciales en la forma de medición, parámetros de estimación, formas de verificación, y aún en la interpretación del modelo* [4, p. 382]”. Estos desacuerdos llevaron a Pao [3] a proponer un procedimiento uniforme para corroborar la Ley de Lotka, que fue apoyado por Nicholls [4].

Últimamente, han aparecido algunos desacuerdos relacionados con las formas de realizar el conteo para medir autoría múltiple, es decir, cuando un trabajo es firmado por dos o más autores. Un comentario sobre las causas y razones de estos desacuerdos fue hecho por Lolas [7]. Las cuestionadas formas de realizar el conteo son, *el conteo directo*, cuando solamente el autor principal (el autor nombrado en primer lugar) es acreditado con la contribución y los autores secundarios (colaboradores) son omitidos; *el conteo completo*, cuando cada autor (principal y/o secundario) es acreditado con una contribución completa; y *el conteo ajustado*, cuando cada autor (principal y/o secundario) es acreditado en partes iguales: en igual proporción con una fracción o una porción de la contribución total, i.e. si hay cinco autores de un único artículo cada uno es acreditado con 1/5 de la contribución.

Sin embargo, la experiencia parece indicar que el conteo directo y el conteo ajustado no producen diferencias sustanciales. Como bien afirman Nath y Jackson, “*esencialmente, los dos medios de conteo producen el mismo resultado y por eso no es*

necesario considerar el conteo ajustado y se debería dar mayor atención al conteo directo [8, p. 207]”.

Los autores comparten estas afirmaciones, y más aún, el propio formulador de esta Ley, en una nota de pie de página, indicó que “*en todos los casos, las contribuciones conjuntas fueron atribuidas solamente al autor principal* [5, p. 323]”. Por esa razón, en este trabajo se ha considerado suficiente utilizar el conteo directo.

También Potter en una extensa revisión de la literatura sobre la Ley de Lotka, afirma que “*cuando el período cubierto es diez años o más y la comunidad de autores es definida ampliamente, la productividad de los autores se aproxima a la distribución de frecuencias que observó Lotka, y que es conocida como la Ley de Lotka* [9, p. 36]”

De esa forma, para asegurar esta extensa cobertura, se decidió analizar los autores que publicaron en la *Revista Geológica de Chile*, desde 1974 hasta 1997.

Revisión de la literatura

Una de las primeras aplicaciones de la Ley de Lotka a los autores que publicaron en una única revista representativa de una área determinada, fue realizada por Murphy [10] quien aplicó esta Ley al área de historia de la tecnología. Al seleccionar “*artículos académicos*” de la revista *Technology and Culture* entre 1960 y 1969, encontró que 170 autores habían producido 231 artículos. Aún cuando los datos presentaban una desviación en la cola de la recta de regresión, concluyó que la Ley de Lotka se ajustaba a este campo.

Coile [11] criticó la aplicación de Murphy [10] aduciendo una mala interpretación de la fórmula de Lotka que tiene una constante y un exponente del Cuadrado Inverso. Para verificar si los datos de este autor se ajustaban a la Ley de Lotka, Coile [11] hizo algunas modificaciones para recalcular la bondad del ajuste usando el test de Kolmogorov-Smirnov y encontró que, aparentemente, la Ley de Lotka no se ajustaba a la muestra usada por Murphy [10].

Radhakrishnan y Kernizan [12] realizaron dos experimentos en la literatura de ciencia de la computación. Para el primer experimento usaron los artículos publicados, de 1968 a 1972, en las revistas *Communications of the Association for Computing Machinery (CACM)* y *Journal of the Association for Computing Machinery (JACM)*,

encontrando que 715 autores publicaron en CACM y 382 en JACM. Asumieron que un autor publicó exclusivamente en una revista científica y, en este caso, la Ley de Lotka no se ajustó muy bien a los datos observados. Sin embargo, la predicción por x/n^3 estuvo muy cerca de ajustarse a la Ley.

En el segundo experimento, se efectuó una selección aleatoria de 300 autores de ambas revistas. El número de artículos publicados por cada una de ellas fue determinado mediante el índice acumulado de autores de *Computer and Control Abstracts*. Concluyeron que el desvío de los datos, en relación a la Ley de Lotka, fue muy elevada.

Posteriormente fueron realizados estudios similares aplicando la Ley de Lotka a una única revista especializada en un área determinada, por Gisbert Tio y Valderrama Zurian [13] en el campo de la drogodependencia; por Fernández [14] en el área de la psiquiatría; por Blackwell [15] en el campo de las ciencias de la información; y por Gargantini [16] en el área de trastornos de fluencia.

Material y método

Los datos del presente estudio proceden de la *Revista Geológica de Chile* publicados entre 1974 y 1997. Esta revista comenzó a ser publicada en 1974 y es la publicación nacional de mayor prestigio en el área de la geología y la octava revista chilena incorporada a los registros del *Institute for Scientific Information* (ISI). Como una evidencia de su prestigio, está indexada por fuentes de información secundarias internacionales, como GEOREF, ULRICH'S International y Bibliography and Index of Geology. Representa, también, la publicación nacional especializada de mayor excelencia, tanto que el 25 de agosto de 1993 fue incorporada al *Science Citation Index*.

Como es sabido, "en cada disciplina se destacan algunas revistas por la calidad de los trabajos que publican, las cuales mantienen celosamente su prestigio a través de una muy estricta selección [17, p. 22]". Y este sería el caso de la *Revista Geológica de Chile*.

La decisión de comprobar el modelo de Lotka en esta revista se basó en que, hasta la fecha, no han aparecido estudios globalizadores de su bibliografía. Además, como gran parte de los profesionales preocupados por el desarrollo de la geología chilena publican en esta revista, la producción intelectual comunicada mediante ella podría constituirse en una buena muestra de la producción geológica en el país.

Los datos fueron recolectados manualmente usando fichas catalográficas normalizadas. Después contabilizadas y organizadas para finalmente proporcionar la distribución de frecuencias de la productividad de los autores que publicaron sus trabajos en la *Revista Geológica de Chile*. Los parámetros fueron calculados usando el software Mathematica 4.0. Las tablas y los trazados de los gráficos fueron realizados usando el paquete estadístico SPSS 10.0 estándar.

Gran parte de los profesionales preocupados por el desarrollo de la geología chilena publican en esta revista, la producción intelectual comunicada mediante ella podría constituirse en una buena muestra de la producción geológica en el país.

De acuerdo con las observaciones de Nath y Jackson [8], se adoptó el conteo directo y, en consecuencia, solamente los autores principales fueron acreditados con una contribución a la producción de un artículo. Para verificar la Ley fue adoptado el modelo propuesto por Pao [3, 18], de modo que para encontrar el valor de la constante "n" en el sesgo de la línea de regresión, se usó el método de los mínimos cuadrados con la ecuación siguiente:

$$n = \frac{N \sum XY - \sum X \sum Y}{N \sum X^2 - (\sum X)^2}$$

donde

N = cantidad de datos pares observados

X = log x

Y = log y

Para calcular la constante C se usó la siguiente ecuación:

$$C = \frac{1}{\frac{1}{\sum x^n} + \frac{1}{(n-1)(P^{n-1})} + \frac{1}{2P^n} + \frac{n}{24(P-1)^{n+1}}}$$

donde,

n = es la constante del sesgo de la línea de regresión,

P = es el número de pares de datos presentes en la

distribución y usados para estimar el valor de la constante C.

El test Kolmogorov-Smirnov fue usado para evaluar el ajuste de la distribución teórica a la distribución de los datos observados. Como se sabe, este test es un simple método no-paramétrico para probar si existen diferencias significativas entre las frecuencias de la distribución observada y las frecuencias de la distribución teórica. Es más poderoso que el chi-cuadrado, fácil de usar, no requiere que los datos sean agrupados, y solamente depende del máximo valor absoluto de la diferencia entre la distribución observada y esperada.

Resultados

El estudio bibliométrico de la *Revista Geológica de Chile* permite señalar que desde 1974 hasta 1977, se publicó un número anual, y desde 1978 hasta 1979 dos números anuales. En 1980 comenzó a publicar tres números anuales, pero de forma irregular, hasta que finalmente a partir de 1988, se realizó un cambio a volumen y número con periodicidad semestral, regularidad que mantiene hasta hoy.

La revista opera con un editor y con un comité de editores asociados que incluye especialistas nacionales e internacionales. Los editores que esta revista ha tenido desde 1974 son los siguientes:

1974	Álvaro Tobar
1975	Eduardo Abad
1976 - 1978	Vladimir Covacevich
1978 - 1979	Carlos Huete
1980 - 1983	Álvaro Puig y Manuel Suárez
1983 - 1985	Constantino Mpodozis
1986 - 1993	Ernesto Pérez
1994 -	Francisco Hervé

La distribución de autores y artículos son mostrados en la tabla 1. Se puede observar que del total de 178 autores identificados, 72% de ellos publicaron solamente un artículo, equivalente a 40% de la producción total. Por otro lado, 17.3% de los autores publicaron mas de tres artículos cada uno en un periodo de 24 años.

Esta cifra de baja productividad es mas alta que 60% encontrado por Russell y Galina [19] en el campo de la veterinaria, a pesar que el período estudiado

No. De Contribuciones por autor	% de Autores	No. de Autores	% de Artículos	No. de Artículos
x	y	y	x.y	x.y
1	128	71.9	128	40.0
2	19	10.7	38	11.9
3	12	6.7	36	11.2
4	5	2.8	20	6.3
5	6	3.4	30	9.4
6	1	0.6	6	1.9
7	2	1.1	14	4.4
8	2	1.1	16	5.0
9	1	0.6	9	2.8
10	1	0.6	10	3.1
13	1	0.6	13	4.0
Total	178	100.0	320	100.0

Tabla 1. Distribución de la frecuencia observada de los artículos producidos por autor

por esos autores fue de 15 años. Es también mayor que los 64.9% encontrados por López-Muñoz y otros [20] en glándula pineal, pero es menor que 80.33% encontrado por Rivas y Peiró [21] en el campo de la psicometría.

Del total de 178 autores identificados, 72% de ellos publicaron solamente un artículo, equivalente a 40% de la producción total. Por otro lado, 17.3% de los autores publicaron más de tres artículos cada uno en un periodo de 24 años.

Este alto porcentaje de autores con una única contribución indicaría migración de los investigadores a otras actividades más rentables que la investigación, y falta de continuidad en los esfuerzos de investigación y publicación de resultados, o de ambos; también podría ser indicativo de un alto porcentaje de autores transeúntes o “golondrinos” explorando este campo.

La figura 1 muestra más claramente la estratificación de la producción de artículos en la *Revista Geológica de Chile*.

Los autores más productivos (1.2%) contribuyeron

con 10 ó más artículos cada uno. La productividad media por autor es de 1.8 artículos, pero esta productividad media es igual a 1.5 artículos entre los que publicaron hasta cinco artículos cada uno y crece a 8.9 artículos entre los que publicaron 6 ó más artículos. En relación con los trabajos realizados en colaboración, i.e. con autoría múltiple, se encontró que 110 artículos (34.4%) fueron producidos por autores únicos. Esto significa que 65.6% de los trabajos fueron producidos en colaboración por dos o más autores. Se encontró, también, un caso de un artículo producido conjuntamente por ocho autores y otro por nueve autores.

Estas colaboraciones permitieron identificar un grupo de profesionales que comparten intereses comunes, se comunican entre ellos e intercambian información; es decir, tienen las características que tipifican a los participantes de un “colegio invisible”.

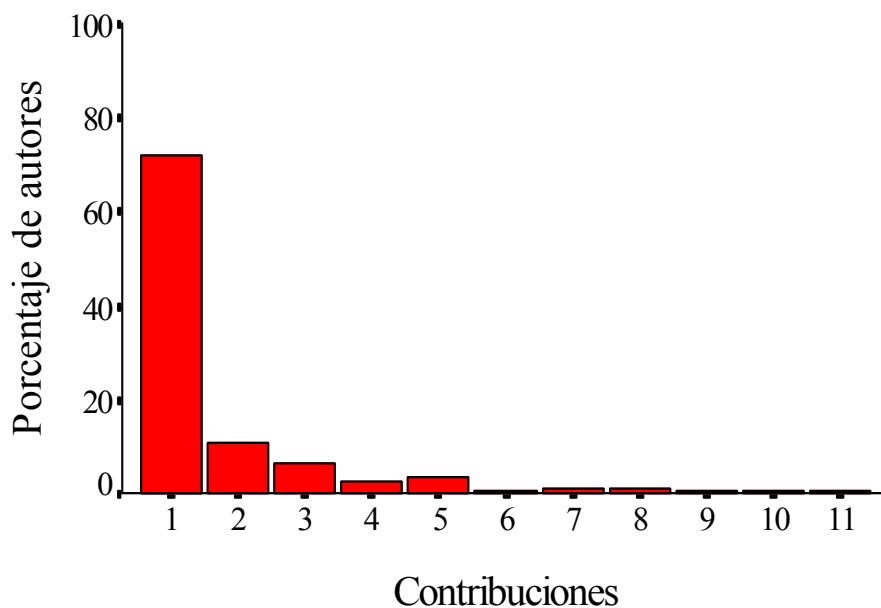


Fig. 1. Porcentaje de autores versus el número de contribuciones en la *Revista Geológica de Chile*.

Estas colaboraciones permitieron identificar un grupo de profesionales que comparten intereses comunes, se comunican entre ellos e intercambian información; es decir, tienen las características que tipifican a los participantes de un “colegio invisible”. Los nombres mas destacados y que aparecen más frecuentemente como pares, son los siguientes:

Ernesto Pérez -----Renato Reyes
Daniel Frasinetti ----- Vladimir Covacevich
José Antonio Naranjo ----- Roland Paskoff
Manuel Suárez -----M. C. Bell

La tabla 2 muestra la relación de los autores que produjeron tres y más artículos, ordenados en forma decreciente. Tres de estos autores laboran en Estados Unidos (Stern, Forsythe y Skewes) y uno en Inglaterra (Bell). Estos 31 autores publicaron, conjuntamente, 48% de los artículos, siendo 27 de ellos chilenos, de los cuales 14 laboran en el Servicio Nacional de Geología y Minería (SERNAGEOMIN). Esto significa que 53% de los investigadores más productivos pertenecen a esta institución.

Nótese también que, con la excepción de solo uno de ellos, los autores que participan del “colegio invisible” están incluidos entre aquellos que publicaron tres y más artículos.

Cabe señalar que el autor perteneciente al “colegio invisible” que no aparece entre los altos productores, es un geógrafo que ha participado como coautor de autores chilenos, en hasta ocho artículos (R. Paskoff).

Conforme al modelo propuesto por Solla Price [22, 23], se identificó a quienes forman la élite de los

Nombre del Autor	No. De Artículos
Pérez, E	13
Naranjo, J. A.	10
Hauser, A.	9
López-Escobar, L	8
Stern, C. R.	8
Mpodozis, C.	7
Suárez, M.	7
Skarmeta, J.	6
Bell, M. C.	5
Camus, F.	5
Covacevich, V.	5
Niemeyer, H.	5
Rivano, S.	5
Vergara, M.	5
Charrier, R.	4
Gana, P.	4
Hervé, F.	4
Parada, M. A.	4
Reyes, R.	4
Alfaro, G.	3
Arias, J.	3
Arratia, G.	3
Davidson, J.	3
Forsythe, R. D.	3
Frassinetti, D.	3
Godoy, E.	3
Hervé, M.	3
Martínez Pardo, R	3
Skewes, M. A.	3
Thiele, R.	3
Vivallo, W.	3

Tabla 3. Distribución de la contribución realizada por los grupos de autores en la producción de artículos en la Revista Geológica de Chile

CATEGORÍA	NO. DE AUTORES	% DE AUTORES	NO. DE ARTÍCULOS	% DE ARTÍCULOS	PRODUCTIVIDAD MEDIA
Grandes productores (10 o más Trabajos)	2	1.2	23	7.1	11.5
Productores moderados (5 a 9 trabajos)	12	6.8	75	23.5	6.25
Aspirantes (2 a 4 trabajos)	36	20.2	94	29.4	2.6
Transeúntes (1 solo trabajo)	128	71.9	128	40.0	1.0
Total	178	100.0	320	100.0	

Tabla 2. Autores que publicaron tres y más artículos

No. de Contribuciones por autor x	No. De Autores y	Log x	Log y	Log x Log y	Log x ²
1	128	0	2.10720997	0	0
2	19	0.30103	1.2787536	0.38494319	0.09061906
3	12	0.47712125	1.07918125	0.51490031	0.22764469
4	5	0.60205999	0.69897	0.42082187	0.36247623
5	6	0.69897	0.77815125	0.54390438	0.48855906
6	1	0.77815125	0	0	0.60551937
7	2	0.84509804	0.30103	0.25439986	0.71419070
8	2	0.90308999	0.30103	0.27185717	0.81557152
9	1	0.95424251	0	0	0.91057877
10	1	1	0	0	1
13	1	1.11394335	0	0	1.24086979
Total	178	7.67370639	6.54432607	2.39082678	6.45602919

Tabla 4. Distribución de los mínimos cuadrados de los datos observados

autores que publican en la *Revista Geológica de Chile*. Esta élite está formada por los 14 primeros autores de la lista mostrada en la tabla 2. Ellos representan 8% de la población total (178 autores) y son responsables de 1/3 de la productividad total de artículos en el período investigado. La media de artículos publicados por esta élite es de siete artículos por autor.

Para resaltar, una vez más, la estratificación de los autores de la *Revista Geológica Chilena* según su productividad, la tabla 3 muestra su categorización en cuatro niveles bien diferenciados.

Obsérvese la alta tasa de autores transeúntes y aspirantes, la que alcanza en conjunto a casi 92%

de la población estudiada, con apenas 8% de moderados y grandes productores. Obsérvese, también, cómo la productividad media casi se duplica conforme se progresa de nivel en nivel. La relación de producción media entre transeúntes y aspirantes es del orden de 1 a 3; esa tasa se reduce de 1 a 2 entre aspirantes y moderados, así como entre moderados y grandes productores.

La tabla 4 muestra la distribución de los mínimos cuadrados usados para calcular las constantes *n* y *c*. Se encontró que «*n*» fue igual a 1.97 y «*c*» igual a 0.6128.

La figura 2 muestra la línea de regresión que mejor se ajusta a esta distribución. El *r*² encontrado fue

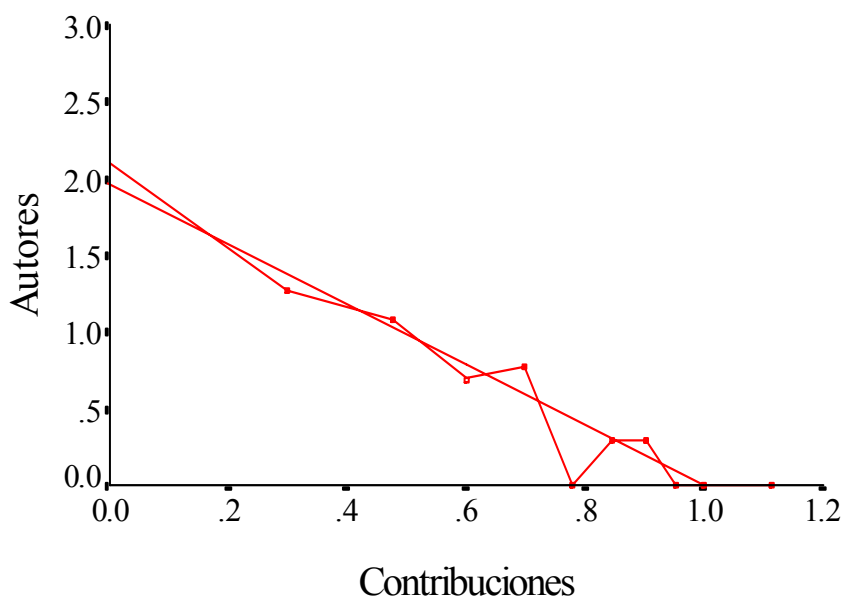


Fig. 2. Recta de regresión de autores *versus* las contribuciones en escala logarítmica.

igual a 0.9277 indicando una fuerte correlación entre las variables artículos producidos y autores. El r^2 ajustado fue igual a 0.9197, $F = 115.5$ al 0.0001 nivel de significancia.

La tabla 5 muestra los valores de la distribución teórica usados para verificar el ajuste al test Kolmogorov-Smirnov. El valor crítico encontrado

índice de crecimiento se tomó como base 1988, cuando se realizó un cambio a volumen y periodicidad semestral.

No. de Contribuciones por autor x	No. de Autores y	$y_x / \Sigma y_x$	$\Sigma (y_x / \Sigma y_x)$	$C (1/x^n)$	$\Sigma (C (1/x^n))$	D_{max}
1	128	0.71910	0.71910	0.6128	0.6128	0.1063
2	19	0.10674	0.82584	0.15642	0.76922	0.05662
3	12	0.0674	0.89324	0.07037	0.83959	0.05365
4	5	0.02809	0.92133	0.03993	0.87952	0.04181
5	6	0.03371	0.95504	0.02572	0.90524	0.0498
6	1	0.00562	0.96066	0.01796	0.9232	0.0434
7	2	0.01124	0.97190	0.01326	0.93646	0.03544
8	2	0.01124	0.98314	0.01019	0.94665	0.03649
9	1	0.00562	0.98876	0.00808	0.95473	0.03403
10	1	0.00562	0.99438	0.00657	0.9613	0.03308
13	1	0.00562	1.00000	0.00392	0.96522	0.03478

Tabla 5. Test de ajuste Kolmogorov-Smirnov de la distribución de los datos observados

fue de 0.1222 y la desviación máxima (D_{max}) fue de 0.1063. Como el valor crítico es mayor que la desviación máxima, se concluye que esta distribución se ajusta al modelo del poder inverso generalizado.

La figura 3 muestra el crecimiento lineal de la literatura publicada en la *Revista Geológica de Chile* en los 24 años analizados. Para analizar el

Hay una tendencia a un crecimiento lineal en la publicación y diseminación de artículos por la *Revista Geológica de Chile*.

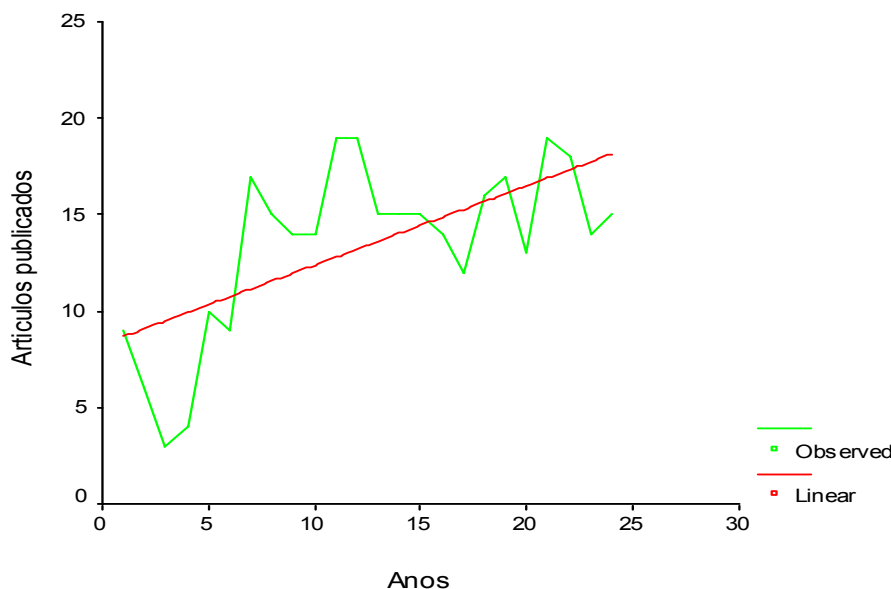


Fig. 3. Índice de crecimiento de la publicación de artículos en la Revista Geológica de Chile, entre 1974 y 1997.

Además, desde esa época, esta revista mantiene inalterada su regularidad de publicación. Se observó fluctuaciones en los años 1980, 1984-1985, 1992, y 1994-1995. La tasa de publicación en 1980 fue 13% mayor que en 1988; esa misma tasa de crecimiento se conoció en 1992 subió a 26% en 1994 y fue de 20% en 1995. A pesar de estas fluctuaciones hay una tendencia a un crecimiento lineal en la publicación y diseminación de artículos por la *Revista Geológica de Chile*. Se observó, también, que desde 1980 hay una tendencia a publicar 15 artículos por año. La correlación “r” de Pearson encontrada fue igual a .646 al .001, nivel de significancia que indica una fuerte asociación entre ambas variables. Se calculó, también, el modelo de regresión simple encontrándose que r^2 igual a 0.41763, $a = 8.275362$, $\beta = .41134$, $F = 15.78$ al 0.006 nivel de significancia. La ecuación de regresión lineal, calculada con el método de los mínimos cuadrados, que mejor expresa ese crecimiento, es:

$$Y_t = 8.275362 + .41134 t$$

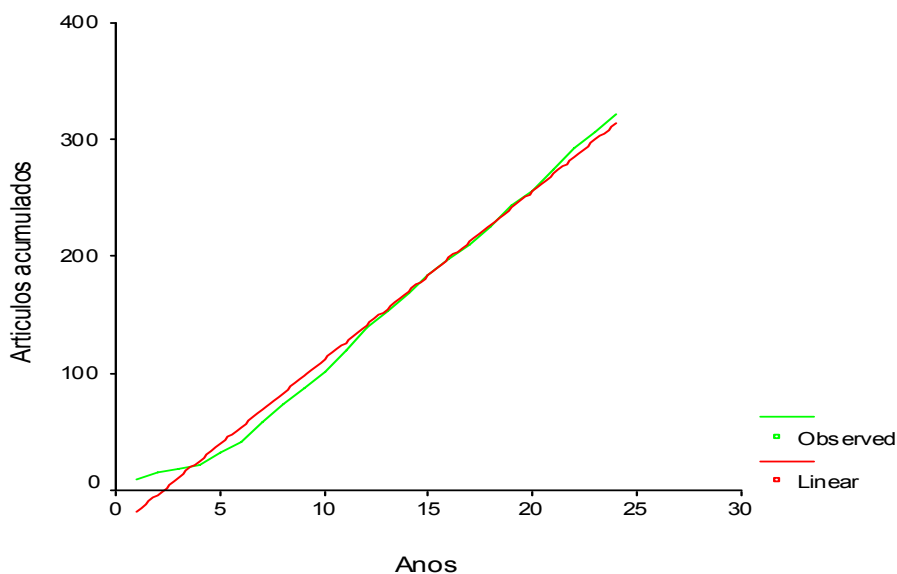


Fig. 4. Acumulado de artículos publicados entre 1974-1997.

La figura 4 muestra la versión acumulada de este crecimiento. Obsérvese cómo entre 1974 y 1980 hay un inicio tímido para a partir de 1985 afianzar su crecimiento acumulativo, con una tendencia claramente lineal ($r^2 = .99573$, $a = -32.86$, $b = 14.47$, $F = 2561.89$ al .00001 nivel de significancia).

Conclusiones

Esta investigación demuestra que el modelo del poder inverso se ajusta muy bien a la literatura producida en 24 años de existencia de la *Revista Geológica de Chile*. Adicionalmente, establece que el modelo propuesto inicialmente por Lotka constituye un instrumento efectivo para identificar los autores más productivos y su desigual distribución, en los artículos contenidos en esta revista.

Entre los logros del presente estudio se pueden señalar que el modelo permitió identificar un alto porcentaje, casi 83%, de pequeños productores responsables de poco más de la mitad (52%) de la literatura publicada. Por su parte los medianos y grandes productores que estarían haciendo progresar la ciencia geológica chilena corresponderían a 8%.

No obstante que la media de 1.8 (± 2) artículos producidos por autor, parece baja, ésta es superior a los 0.57 trabajos por autor, reportados por López-

Calafi y otros [24] en aceites lubricantes por espectroscopia de absorción atómica; a los 1.42 trabajos por autor encontrados en el campo de la psicología por Carpintero y otros [25]; muy similar a la media de 1.716 artículos por autor observado por Allison [26] en el campo de la química americana, pero menor que los 2.6 trabajos por autor en el campo de la medicina, López-Muñoz y otros [20]. Estas son pues, evidencias que esa

productividad es una constante relativa en otros campos del conocimiento y en otros países.

Los datos encontrados al utilizar el método de conteo directo, evidencian que, en este campo, existe una alta tasa de colaboración que alcanza a casi 66% de los artículos que fueron producidos por dos o más autores. Por lo tanto, sería interesante proponer la realización de un estudio en este mismo campo, que incluyera a los autores colaboradores, con lo que las cifras encontradas en el presente trabajo, posiblemente variarían.

Se evidencia la necesidad de un análisis global de la literatura geológica chilena, que incluya la compilación de la producción nacional publicada en todas las revistas de geología y ciencias afines existentes en el país, más allá de la *Revista Geológica de Chile*; la *Carta Geológica de Chile*; el *Boletín del Servicio Nacional de Geología y Minería*; *Comunicaciones del Departamento de Geología de la Universidad de Chile*; la *Revista Chilena de Historia Natural*; el *Boletín del Museo Nacional de Historia Natural*, y las revistas *Revista Geografía del Norte Grande*; *Innovación y Terra Australis*; *los Anales de la Universidad de Chile*, así como otras similares tanto nacionales como extranjeras, donde los especialistas en la geología chilena están publicando.

Este trabajo muestra evidencias de que la geología chilena es una área abierta a la exploración, usando modelos bibliométricos y cientométricos. Sin embargo, se requiere una base de datos nacional que recoja esa productividad y que facilite estudios de este tipo, orientados a medir adecuadamente la producción científica nacional en el campo de la Geología.

Agradecimientos

Los autores desean expresar sus agradecimientos a: *S. Amar (SERNAGEOMIN)*; *Gajardo (SERNAGEOMIN)* quienes realizaron importantes observaciones al manuscrito, así como también a *G. Verdugo (SERNAGEOMIN)* quien colaboró activamente en la recolección de la información primaria.

Referencias

- 1) Urbizagástegui Alvarado, Rubén y María Teresa Cortés. *Método gráfico para medir la obsolescencia de la literatura geológica: el caso de la Revista Geológica de Chile. Investigación Bibliotecológica, México, 24(12):81-98, enero-junio de 1998.*
- 2) Urbizagástegui Alvarado, Rubén y María Teresa Cortés. Análisis de citas bibliográficas de la Revista Geológica de Chile. *Revista Geológica de Chile, 25(2):265-272, dic., 1998.*
- 3) Pao, Miranda Lee. Lotka's law: a testing procedure. *Information Processing & Management, 21(4):305-320, 1985.*
- 4) Nicholls, Paul Travis. *Bibliometric modeling process and the empirical validity of Lotka's Law.* Journal of the American Society for Information Science, 40(6):379-385, 1989.
- 5) Lotka, Alfred J. *The frequency distribution of scientific productivity.* Journal of the Washington Academy of Sciences, 16(12):317-323, June 19, 1926.
- 6) Oppenheimer, C. *The use of online database in bibliometric studies.* In International On-line Information meeting. (9th. : 1985 : London, England). 9th International Online Information Meeting, London 3-5, December, 1985. Oxford, England: Learned Information, 1986. pp. 355-364.
- 7) Lolas, Fernando. *Los autores múltiples en la literatura científica.* *Vida Médica, Santiago, 38(1):22-24, 1987.*
- 8) Nath, Ravinder and Wade M. Jackson. Productivity of management information system researchers: Does Lotka's law apply?. *Information Processing & Management, 27(2/3):203-209, 1991.*
- 9) Potter, W. G. Lotka's Law revisited. *Library Trends, 31:21-39, 1981.*
- 10) Murphy, Larry J. Lotka's Law in the Humanities? *Journal of the American Society for Information Science, 24(6):461-462, Nov.-Dec. 1973.*
- 11) Coile, Russell C. Lotka's frequency distribution of scientific productivity. *Journal of the American Society for Information Science, 28(6):366-370, 1977.*
- 12) Radhakrishnan, T. and R. Kernizan. Lotka's Law and computer science literature. *Journal of the American Society for Information Science, 30(1):51-54, Jan.*

- 1979.
- 13) Gisbert Tio, Amparo y Juan Carlos Valderrama Zurian. *Estudio bibliométrico de la Revista Española de Drogodependencias, 1976-1993. En Actas de las 5as. Jornadas de Información y Documentación de Ciencias de la Salud, Palma de Mallorca, 4, 5 y 6 de mayo, 1994.* pp. 1-7.
- 14) Fernández Doctor, Asunción y Antonio Seva. *Bibliometric evaluation of "The European Journal of Psychiatry" at the end of the first ten years of its history.* Madrid?: Caja de Ahorros de la Inmaculada, 1997?
- 15) Blackwell, Lisa S. *A bibliometric analysis of RQ, vols. 31-35.* Ph. D. Dissertation, 1997.
- 16) Gargantini, Marisa B. Mendes. Autoria de artigos do "Journal of Fluency disorders". *Transinformação*, 9(3):1-7, Sept.-Dez., 1997.
- 17) Izquierdo, Francisco. La medición de la ciencia. *Cause*, Santiago, 72:22-23, 1986.
- 18) Pao, Miranda Lee. Lotka's test. *Collection management*, 4(1-2):111-124, Spring-Summer, 1982.
- 19) Russell, J. M. and C. S. Galina. Productivity of authors publishing on tropical bovine reproduction. *Interciencia*, 13(6):311-313, Nov.-Dic. 1988.
- 20) López-Munoz, Francisco, Jesús Boya, Fernando Marin and José Luis Calvo. Scientific research on the pineal gland and melatonin: a bibliometric study for the period 1966-1994. *Journal of Pineal Research*, 20(3):115-124, 1996.
- 21) Rivas, Francisco y José María Peiró. *Estado actual de la investigación psicométrica: una aproximación bibliométrica.* En *Psicología contemporánea: teoría y métodos cuantitativos para el estudio de su literatura científica*/Helio Carpintero y José María Peiró, directores; con la participación de María L. García-Merita... [et al.]. Valencia: Alfapplus, 1981. pp. 199-218.
- 22) Price, Derek John de Solla. Networks of scientific papers. *Science*, 149(3686):510-515, 1965.
- 23) Price, Derek John de Solla. Some remarks on elitism in information and the invisible college phenomenon in science. *Journal of the American Society for Information Science*, 22:74-75, Mach-April, 1971.
- 24) López-Calafi, J., A. Salvador y M. de la Guardia. Estudio bibliométrico de la literatura científica sobre la determinación de elementos metálicos en aceites lubricantes por espectroscopia de absorción atómica. *Revista Española de Documentación Científica*, 8(3):201-213, 1985.
- 25) Carpintero, Helio, José María Peiró e Ismael Quintanilla. El "Anuario de Psicología" (1969-1974): un estudio estadístico y bibliométrico. *Anuario de Psicología*, 16(1):22-34, 1977.
- 26) Allison, Paul D. Inequality and scientific productivity. *Social Studies of Science*, 10:163-179, 1980.

Recibido: 27 de febrero del 2002.

Aprobado: 6 de noviembre del 2002.

Rubén Urbizagástegui Alvarado

Associate Librarian University of California.

Riverside Science Library

P.O. Box 5900

Riverside, CA 92521 U.S.A.

Correo electrónico: <Ruben@ucrac1.ucr.edu>.

Algunas consideraciones teórico-conceptuales sobre la Bibliotecología y la Ciencia de la Información

Marlery Sánchez Díaz

Juan Carlos Vega Valdés

RESUMEN

A partir de la necesidad de concepciones teóricas y conceptuales en la Bibliotecología y Ciencia de la Información, el trabajo expone mediante los criterios de diferentes autores distintos aspectos relacionados con la epistemología de la Bibliotecología y la Ciencia de la Información. Se reseñan con relación a la Bibliotecología y Ciencia de la Información los objetos de estudio, disciplinas relacionadas, campo fenoménico de cada una y algunos otros aspectos filosóficos.

ABSTRACT

Starting from the need of theoretical and conceptual conceptions in Library and Information Science, this paper reviews different approaches of authors in some aspects related with the epistemology of Library and Information Science. The authors point out study objects, related disciplines, field phenomonic of each and some other philosophical aspects.

Introducción

El análisis epistemológico de una disciplina entraña los estudios teóricos y filosóficos, los conceptos, las terminologías, la interdisciplinariedad, las metodologías y el estudio de su objeto. Toda profesión es una mezcla de teoría y práctica, donde debe haber una relación armónica entre estos dos elementos, sin excederse ni el uno ni el otro. Cada disciplina práctica descansa en un cuerpo teórico, toda valoración práctica tiene lugar en un cuerpo teórico.

Por tanto toda disciplina científica necesita una fundamentación teórica, la Bibliotecología y Ciencia de la Información, no están exentas de esto y sin embargo en ellas se detecta escasez de criterios teóricos y metodológicos. Una de las tareas a cumplir

en la fundamentación de la Bibliotecología y Ciencia de la Información consiste en analizar y esclarecer el contenido de los conceptos centrales utilizados en dicha disciplina.

Al hacerse imprescindible la definición clara de los objetos de estudio de la Bibliotecología y Ciencia de la Información, su relación con otras disciplinas y el campo fenoménico emanado del objeto de estudio que les compete; el siguiente trabajo trata de reseñar esos aspectos. A partir de la revisión de trabajos de diferentes autores se exponen los disímiles criterios y cuestiones que tienen que ver con la Bibliotecología y Ciencia de la Información, sin pretender descubrir y proponer, o ambos, cuestiones absolutas y terminales en estos temas.

Hay que reconocer que se hacen intentos entre los profesionales de contribuir al desarrollo teórico de estas disciplinas, como por ejemplo, en Internet circula el Seminario Permanente sobre problemas teóricos y epistemológicos de la Ciencia Bibliotecológica y de la Información Documental, y se han creado grupos de trabajo referente a epistemología de la Bibliotecología en el que participan México, Colombia, Cuba y Brasil liderado por el Centro Universitario de Investigación Bibliográfica y por la Universidad Nacional Autónoma de México (CUIB/UNAM) a través del profesor Miguel Ángel Rendón.

Consideraciones acerca de la Bibliotecología

Buonocuore señala en uno de sus trabajos que la Bibliotecología “...constituye un saber especializado, un saber técnico que se propone, en última instancia el arbitrio de medios, recursos y procedimientos tendientes a lograr un fin práctico. Esa finalidad consiste en reunir, seleccionar, organizar y administrar una colección determinada de material bibliográfico para ponerla al alcance del estudioso o investigador... [1]”.

Fernández propone a la Bibliotecología como “una ciencia que abarca el conjunto de conocimientos dinámicos referentes al libro y a la biblioteca y que en función de ese dinamismo ejerce y sufre la acción recíproca del medio sobre el cual actúa... [2]”.

Llama la atención el análisis efectuado por los autores de una de las ponencias presentadas en el III Encuentro de Docentes en Argentina Conforti y Artaza de la Universidad Nacional de Mar del Plata exponen que “la Bibliotecología no constituye una disciplina científica. No cuenta con un cuerpo orgánico de leyes ni con un sólido aparato teórico... Como tampoco tiene entre sus metas la obtención de conocimientos puros desprovistos de posteriores utilidades como acontece con las ciencias básicas o puras. No constituye... una ciencia social o cultural, en la medida en que no busca conocimientos desprovistos de utilidad, ni formula enunciados generales de carácter teleológicos. Tampoco es una ciencia aplicada. Recordemos que las ciencias aplicadas tienen por objetivo central la búsqueda de conocimientos que puedan tener alguna utilidad práctica... [3]”. Plantean que “la Bibliotecología utiliza los conocimientos y la metodología de la ciencia con la finalidad de generar productos y diseñar y optimizar procedimientos. Esta circunstancia sumada a las características de su saber y de su accionar

profesional, permite encuadrarla con claridad en el ámbito de las disciplinas tecnológicas... [3]”. Así consideran la Bibliotecología dentro del grupo de las tecnologías sociales y argumentan que “...tiene un sistema de conocimientos cuyo objetivo primordial se vincula con el control, la modificación y construcción de instrumentos y de procesos: reglas de catalogación, de clasificación, instrumentos de control terminológico, etc... Es un saber que tiene por objetivo la resolución de problemas prácticos y control de un aspecto de la realidad a través de la utilización de conocimiento científico y de la metodología de la ciencia. La Bibliotecología construye y diseña procedimientos tendientes a solucionar problemas prácticos. Por lo tanto crea planes de acción, instrumentos y dispositivos que le permiten accionar la realidad... [3]”.

Estos autores delimitan a la Bibliotecología como puro instrumentalismo, como sólo la suma de procesos técnicos, como un conjunto de prácticas.

Toda profesión es una mezcla de teoría y práctica, donde debe haber una relación armónica entre estos dos elementos, sin excederse ni el uno ni el otro

Restrepo plantea que “... el objeto de estudio son los libros y las bibliotecas, concebidas hoy como Centros de Acceso a la Información con herramientas electrónicas y procedimientos diferentes a los tradicionales.” ... “la ciencia de la bibliotecología no es exclusiva para las bibliotecas. También se manejan con ella, archivos, sistemas de información gerencial,..., manejo de bases de datos [4].” Las disciplinas que la afectan son la ciencia de la administración, ciencia de la computación, ciencia de la comunicación, las matemáticas y la estadística, la sociología, la psicología, la lingüística... [4]”.

Este autor concentra el factor bibliotecológico solo a un tipo de documento: el libro, no incluye en las disciplinas que la afectan a la filosofía y la historia y además incluye en la bibliotecología a los archivos, entre otros.

García expone que “la selección de documentos, la sistematización y recuperación de los mismos como actividades básicas de las bibliotecas,... son

quizás los que convierten a las bibliotecas en común denominador de todas las disciplinas... En una sociedad o cultura dada, las bibliotecas, de cualquier tipo constituyen instrumentos documentales que posibilitan la formación de redes cuyo propósito es integrarse al sistema total de la comunicación social... la Biblioteca más que un eslabón en la cadena de la comunicación, como sistema, es una parte fundamental del proceso total del conocimiento o de la situación del saber en un tiempo determinado... [5]”.

Almanza Morales plantea que “con el transcurso del tiempo, cuando la información crece y se reproduce en forma vertiginosa, la necesidad de la información por parte de la sociedad se hace más apremiante y ante la falta de organización que prevalecía en las grandes bibliotecas, aparece la Bibliotecología como una nueva disciplina con la que se inicia una nueva etapa, en la que ni los sabios, ni el erudito serían los candidatos idóneos para continuar con las actividades bibliotecarias y estudiar el comportamiento de los elementos que en ellas intervienen como son la información, los usuarios y los métodos de almacenamiento, la recuperación y difusión, entre otras, y no sólo en función de su erudición. El proceso evolutivo que presentó esta nueva disciplina se dio de tal forma que primero surgió la necesidad de una institución llamada biblioteca y posteriormente se generó una disciplina denominada bibliotecología [6].” José Luis aborda en su trabajo la interdisciplina, expone que “...el nuevo punto de vista de la interdisciplina dentro del campo bibliotecológico se da a raíz de las necesidades de información más especializada de cada uno de los sectores sociales..., de tal forma que el bibliotecario, por sí solo se encuentra en desventaja para resolver dichas demandas; para ello es indispensable la colaboración de colegas profesionistas de otras disciplinas... [6]”.

¿Qué sucede en muchas ocasiones? El bibliotecario tiene muy débil su autoestima y en muchos casos prefiere ser un ente aislado antes de resolver los problemas a través de metodologías interdisciplinarias.

Rendón Rojas considera que “Actualmente en Bibliotecología ... el objeto tradicional de la disciplina: el libro y la biblioteca, se escabulle, se transforma y se desplaza a otros fenómenos (documentos, bases de datos, redes, centros de documentación, bibliotecas electrónicas, etc.) que distan mucho de ser semejantes a lo que se venía manejando... Por lo anterior se hace necesario replantear el paradigma con el cual se enfrenten

los cambios en la disciplina y se logre abarcar y explicar esta nueva problemática [7].” Explica en su trabajo:

La teoría sintáctica de la información desarrollada en la Teoría Matemática de la Información de Shannon y Weaver. No tiene en cuenta el contenido de los símbolos, sino únicamente la frecuencia con que éstos puedan ocurrir [7], la teoría semántica de la información creada por Bar-Hillel y Carnap. Aunque sí toma en cuenta el significado de los enunciados, interpreta ese significado de una manera extensional, debido a que se basa en la lógica clásica, es decir, el significado de las proposiciones son dos objetos abstractos: la verdad y lo falso, dejando también a un lado el sentido de esos enunciados, la intencionalidad y el contexto del sujeto [7], la teoría pragmática de la información. El mundo de la información unido al sujeto. [7].

Rendón Rojas considera que las dos primeras teorías “no son aptas para ser el eje sobre el que se debe mover la Bibliotecología... [7]”. Señala que “la fundamentación teórica de la Bibliotecología debe de tener como uno de sus pilares un concepto de información pragmática donde la dialéctica del sujeto con el mundo que le rodea sea tomada en cuenta, de lo contrario al subordinar al sujeto a una serie de tecnologías muy sofisticadas y eficientes que ... pueden combinar millones de signos,...,nunca se logrará satisfacer la necesidad que originó la actividad bibliotecológica... [7]”.

Plantea que “La Bibliotecología se ocupa de estudiar como ciertos documentos (libros), pueden acceder al mundo de la información a los lectores. Precisamente conocer las leyes que rigen esa comunicación de los lectores con la “noosfera” a través del fondo bibliotecario y las actividades que deben realizarse para que se dé esa relación es tarea de la Bibliotecología [7].”

Rendón también expone “entendemos a la ciencia bibliotecológica como la disciplina teórica que tiene por objeto de estudio el sistema de información documental dado por los siguientes elementos: la información, el documento, el usuario y la institución informativa-documental y las interrelaciones entre ellos cuando un sujeto con necesidades de información desea ingresar al mundo de la información a través de documentos proporcionados por una institución informativa... El origen, el centro y el fin de la acción bibliotecológica tiene un elemento humano-social,

de ahí que califiquemos a la ciencia bibliotecológica como una ciencia del espíritu... la función epistemológica de la bibliotecología es además de explicar, comprender;.... Dicha comprensión a su vez es el resultado de un diálogo entre sujetos... [8]". Este autor considera tres elementos centrales: sujeto, diálogo y comprensión. Rendón Rojas expone que *"aún aceptando el uso de la tecnología en la ciencia de la bibliotecología, cuestión que por demás no se puede negar, es un hecho, eso no nos conduce necesariamente al objetivismo puro, a la pérdida del sujeto, de los valores y del diálogo en la disciplina, sino sólo a una confusión que puede ser aclarada con ayuda del mismo sujeto que se pretende desplazar... con el empleo de las tecnologías de la información no se da la eliminación del diálogo, sino por el contrario, para una comprensión de ese tipo se requiere de un diálogo a otro nivel. De este modo la complejidad de la tecnología exige la complejidad del diálogo y consecuentemente del hombre... Si encontramos algo en la máquina es porque nosotros lo pusimos... [8]*". Considera que *"Para salir de esa situación efectiva pero impersonal se necesita, además de la racionalidad estratégica, reconocer la racionalidad dialógica que le regresará el ser auténtico (humano y social) al campo bibliotecológico. De esta manera... podemos constatar no sólo la posibilidad del diálogo en la esfera bibliotecológica en el contexto de las nuevas tecnologías de la información sino la necesidad del mismo para recuperar o no perder el sujeto que es quien da sentido a toda esta esfera. Así pues se habla de un diálogo que sigue existiendo potencialmente, no con las máquinas, sino entre sujetos... [8]*." El autor concluye que:

"el diálogo que aparentemente era desplazado sigue jugando un papel importante dentro del sistema de información documental, no son las máquinas quienes actúan como sujetos creativos sino que están en función de los objetivos y tareas que sus creadores les asignaron y consecuentemente la información que crean, procesan y transmiten es una información pragmática del y para el sujeto... no se debe ... olvidar el ser que los creó. no se debemos rechazar las tecnologías, sino comprenderlas... [8]".

Olga Oropeza de Ojeda expone que la Bibliotecología y la archivología *"se consideraron inicialmente como un conjunto de conocimientos teórico- técnicos y de saberes que tenían como objetivo almacenar, conservar documentos, pero con poco interés en su contenido y administrar las*

bibliotecas, los archivos y otras unidades de información contando con métodos basados más en la práctica que en la ciencia, y eso daba el carácter de disciplinas técnicas y no científicas a la Bibliotecología y a la archivología. Sin embargo, los grandes avances en el desarrollo de la informática y de nuevas tecnologías han hecho que en su evolución, la Bibliotecología haya cambiado sus objetivos, así como algunos de sus paradigmas de investigación y las técnicas de la llamada investigación aplicada: todo lo cual en la mayoría de los casos ha sido tomado de otras ciencias... [8]". Ella plantea que *"al concebir en el ámbito científico a la información como objeto de estudio de la Bibliotecología y de la archivología y a éstas como disciplinas de la Ciencia de la Información cobra singular importancia, tanto legitimar científicamente la producción intelectual en el área, como enriquecerla con la generación de nuevos conocimientos... [9]*".

La autora parece considerar la Bibliotecología como una disciplina de la Ciencia de la Información, una un caso especial de la otra.

la Bibliotecología no es una disciplina meramente técnica, pero al estar unida a una actividad práctica muy activa motiva que a veces no se tome en cuenta lo teórico

Expone que *"se observa que hasta ahora en nuestro medio, durante la carrera se han producido muchos trabajos o estudios en el área, algunos muy valiosos e interesantes, pero obviamente no pueden ser considerados como investigaciones porque carecen de un método que los conduzca a darles el carácter de investigación científica... porque en muchos de esos trabajos los resultados no siempre relacionan el fenómeno con sus causas, no se contrastan supuestos con evidencias, ni se establecen relaciones entre los factores, por lo tanto no pueden servir de base para formular generalizaciones posteriores que permitan un sustento teórico, ni generar un nuevo conocimiento..."* Considera que *"se ha podido apreciar que en países avanzados, a partir de mediados de siglo, el interés por la investigación en Bibliotecología y Ciencias de la Información (utiliza en este caso en plural ciencias) se ha incrementado. En cuanto a las teorías y métodos de investigación se considera que a pesar*

de las limitaciones,..., la bibliotecología ha estado desarrollando sus propias y peculiares técnicas para ser aplicadas a sistemas de conocimientos en casos específicos, aunque esta disciplina no tiene un solo o único método de investigación... crece la preocupación por la falta de autocrítica y de análisis sincero sobre la producción intelectual en el área, para distinguir en cuáles se han aplicado modelos, desarrollado teorías y utilizado métodos que conducen a una genuina producción de conocimientos, de aquellas actividades que se caracterizan por aplicar a la realidad los conocimientos que ya se tienen para obtener resultados útiles y eso es lo que estamos confundiendo con investigación científica... [9]".

Es destacable señalar aquí que con relación a nuestra profesión, es muy importante convertir esto que la autora considera hallazgos y actividades de búsqueda (promedios, porcentajes, cuantificaciones, etc) en partes principales de un proceso serio de investigación bibliotecológica donde se vaya más allá de estas cuestiones y se puedan proponer principios, tendencias, resultados a preguntas fundamentales, etc.

Queda claro que la Bibliotecología no es una disciplina meramente técnica, pero al estar unida a una actividad práctica muy activa motiva que a veces no se tome en cuenta lo teórico, sin dejar de reconocer que en muchos casos cuando se pretende dilucidar los principales conceptos y dar univocidad a los términos, se presentan problemas, las diferentes concepciones expuestas anteriormente dan muestra de ello, como también las diferentes vertientes que se dan con relación al nombre de la ciencia:

Librarianship	Inglaterra
Library Science and Information Science	Estados Unidos
Biblioteconomía	Brasil, Portugal, Italia
Ciencia de la Información y Ciencia de la Comunicación.	Francia
Bibliología, Bibliotecología, Documentación	España
Informática	Rusia

A partir de las diferentes concepciones descritas, es preciso comentar que el campo fenoménico de la bibliotecología es la información y su tratamiento, siendo importante las finalidades para lo cual se destina y la actividad del sujeto, esto no responde a las concepciones tradicionales (libro y biblioteca).

Es importante además plantear que a pesar de que las tecnologías han desarrollado la actividad en Bibliotecología y ha hecho que cambien concepciones y se importen términos, la tecnología necesita del propósito humano y social de la disciplina.

Consideraciones en relación con la Ciencia de la Información

Garza Mercado plantea la ciencia de la información como *"la ciencia que investiga las propiedades y el comportamiento de la información, las fuerzas que gobiernan su flujo y los medios para procesarla para su acceso y uso óptimo. El proceso incluye la generación, diseminación, recolección, organización, almacenamiento, recuperación, interpretación y uso de la información. El campo se deriva de, o se relaciona con matemáticas, lingüística, psicología, tecnología de la computación, investigación de operaciones, artes gráficas, comunicación, bibliotecología, administración y algunos otros campos... El campo de la ciencia de la información se encuentra particularmente interesado en mensajes registrados o almacenados, su creación o marcas distintivas o documentos, su propagación y uso... [10]*".

Borko afirma que *"la Ciencia de la Información constituye una ciencia interdisciplinaria que investiga las propiedades y el comportamiento de la información, las fuerzas que gobiernan el flujo y el uso de la información, así como las técnicas tanto manuales como mecánicas, del procesamiento de la información para un óptimo almacenamiento, recuperación y diseminación... [11]*".

Rendón Rojas expresa que *"la ciencia de la información conforma una ciencia multidisciplinaria que estudia la transmisión del conocimiento, su naturaleza y propiedades, los soportes en los cuales se contiene este conocimiento y las técnicas aptas para su procesamiento, almacenamiento, recuperación y difusión. Comprende tres disciplinas: la documentación..., la ciencia de la comunicación... y la informática... [12]*".

Geldres Flores considera que *"la Ciencia de la Información constituye el marco teórico de todas las disciplinas profesionales cuyo objeto es el registro, organización y comunicación de los conocimientos. La documentación y la ciencia de*

la información no son independientes, ni distintas de la Bibliotecología... [13]”.

Atherton plantea que *“las actividades de información y la ciencia de la información puede ser definidas brevemente como las disciplinas profesionales referentes a la acumulación, el almacenamiento y la transferencia de conocimientos registrados... [14]”.*

Currás expone que *“la Bibliotecología, la documentación, la ciencia de la información y la información son ciencias: ciencias que se sitúan a un mismo nivel conceptual y teórico con unas aplicaciones prácticas semejantes, todas ellas reunidas podrán formar parte de una ciencia de rango superior que llamaremos ciencia de la documentación... [15]”.*

Conforti y Artaza consideran que *“en la actualidad la denominada ciencia de la información se encuentra en un período escolástico o prepragmático. En efecto, no existe por el momento un común acuerdo (paradigma unificador), un corpus integrado de conocimientos, un conjunto de problemáticas y una clara metodología de abordaje, que sea consensuada mayoritariamente por la comunidad de especialistas. Predomina la diversidad de teorías sobre la naturaleza de la disciplina y la discusión sobre sus fundamentos, sin haberse llegado hasta el presente a un acuerdo general... [16]”.* Plantean lineamientos generales de trabajo para lograr un paradigma unificador en ciencia de la información:

“...indagar entre los distintos enfoques teóricos, algunos puntos de acuerdo tácito, tratando en lo posible de no enfatizar las diferencias conceptuales y sí de acentuar las analogías teóricas entre los distintos enfoques...”

...analizar las problemáticas principales que presenta la información y a partir de ellas alcanzar convergencias y complementariedades teóricas que permitan esbozar metodologías para resolverlos eficazmente...

...replantearse las metodologías de investigación empleadas hasta el presente e indagar cuáles son las más útiles para las problemáticas que plantea la Ciencia de la Información;

...implementarse para la resolución de una determinada problemática, la aplicación de una amplia batería de métodos diferentes de abordaje, seleccionando posteriormente los más eficaces para cada caso.

...unificación de criterios en el vocabulario y en los significados...; examinar epistemológicamente nuestra disciplina, aceptando las diferencias teóricas y evitando las posturas rígidas y dogmáticas... [16]”.

Amaral escribe que *“la Ciencia de la Información surge en el horizonte de transformaciones de las sociedades contemporáneas, que pasaron a considerar el conocimiento, la comunicación, los sistemas de significado y los usos de la lengua como objetos de investigación científica y de dominios de la intervención tecnológica... [17]”.*

Pinheiro tiene la opinión de que *“la Ciencia de la Información tiene su propio regulamento científico, como ciencia social y es por tanto interdisciplinar por su naturaleza, presentando interfaces con la Bibliotecología, Ciencia de la Computación, Ciencia cognitiva, Sociología de la Ciencia y Comunicación, entre otras áreas. Sus raíces, en principio, vienen de la bifurcación de la Documentación Bibliografía y de la Recuperación de la Información, considerándose que la propia información por si misma sea de categoría abstracta y de difícil aprehensión... [18]”.*

¿No pueden ser tema de estudio de la Bibliotecología, tomando como ejemplo lo planteado por Saracevic como áreas de la Ciencia de la Información, las búsquedas en Internet, la bibliotecas digitales, la recuperación de información, etc?

Hawkins expone que *“la Ciencia de la Información es una disciplina estructurada en importantes conceptos de un número de disciplinas próximamente relacionadas que se transformaron en un todo armónico focalizando la información [19]”.*

Miranda y Barreto plantean que *“la Ciencia de la Información es interdisciplinar por su naturaleza, está inexorablemente asociada con la tecnología*

de la información y como otras áreas del conocimiento, es una actividad participante de la evolución de la sociedad de la información [20]”.

Capurro considera que *“Información es la forma del conocimiento al fin de la modernidad [21].”* Plantea algunas características del fin de la modernidad:

- a) *“con respecto al abandono de la primacía del pensamiento racional o científico como cualitativamente superior a todos los otros tipos de discursos. La información es considerada como teniendo un carácter fragmentario o admitiendo parcelación... Al reducir el conocimiento en trozos,... el contexto original desaparece o se vuelve tácito. El conocimiento se transforma literalmente en parcial dependiente de prejuicios o del sistema de referencia del conocedor... Esta relatividad del conocimiento con respecto a un horizonte cambiante de interpretación pone en el centro de la epistemología una nueva categoría: la de verdad pero ahora, al fin de la modernidad inseparable de la categoría de relevancia.*
- b) *... con referencia al abandono de la idea de subjetividad humana como opuesta a objetividad, donde intersubjetividad y contextualidad juegan sólo rol menor. La información es descrita como teniendo una cierta comunalidad. La información es algo básicamente humano que debería ser accesible en principio a todos. El conocimiento moderno es algo común, compartido por una comunidad...*
- c) *... con referencia al abandono de la idea (platónica) del conocimiento humano como algo separado del conocedor... La información se vuelve parte de los medios, ella misma es un medio [22]”.*

Saracevic expone que *“la Ciencia de la Información como ciencia y como profesión se define por los problemas que ha enfrentado y por los métodos que ha usado en la búsqueda de soluciones a través del tiempo. Cualquier avance en Ciencia de la Información depende de si la disciplina está realmente progresando en relación a estos problemas y métodos [20]”.* Plantea tres características generales de la evolución y existencia de la Ciencia de la Información, tres áreas de

problemas con los cuales tiene que lidiar a un nivel general:

- a) *“La Ciencia de la Información es interdisciplinar por naturaleza, sin embargo las relaciones con varias disciplinas están cambiando...”*
- b) *La Ciencia de la Información está inexorablemente conectada con las tecnologías de información... impulsa y a un tiempo limita la evolución de la ciencia de la información en tanto ésta depende de la evolución de un número de disciplinas y por otra parte de la sociedad de la información como un todo.*
- c) *La Ciencia de la Información es, junto a muchas otras disciplinas un participante activo de la evolución de la sociedad de la información. Tiene una fuerte dimensión social y humana, por encima y más allá de la tecnología [22].”*

“La sociedad de la información es más un mito ideológico que una realidad histórica y una disciplina no se puede agarrar a algo que está por demostrar... [22].”

Saracevic en su trabajo considera en Ciencia de la Información tres ideas poderosas:

“La primera idea original, surgida en los años 50, es la recuperación de la información, apareciendo el procesamiento de la información basado en la lógica formal...”

La segunda emergida poco después, es la relevancia orientando y asociando directamente el proceso con las necesidades de información humanas.

La tercera, derivada de alguna otra parte unas dos décadas después, es la interacción, posibilitando intercambios directos y retroalimentación entre sistemas y personas imbuidas en el proceso de recuperación de la información... [22].”

Saracevic expone tres sentidos de la información:

“Sentido estrecho: La información es considerada en términos de señales o mensajes... [22].”

Rendón Rojas lo trata como lo sintáctico.

“Sentido amplio: La información es tratada como directamente involucrada con la comprensión y el procesamiento cognitivo...

El sentido más amplio: La información es tratada en un contexto. La información involucra no solo mensajes, que son procesados a un nivel cognitivo sino además en un contexto... La información abarca además motivaciones e intencionalidad y por ende está conectada al contexto social... [22].”

Rendón Rojas lo trata como lo pragmático.

Para Saracevic la Ciencia de la Información es “*un campo de práctica profesional e investigación científica que enfoca los problemas de la comunicación efectiva de los registros del conocimiento –“literatura”- entre los humanos en el contexto de las organizaciones sociales y las necesidades y usos de la información por los individuos... [22]*”. Para él “*el foco específico de la Ciencia de la Información es sobre los registros de conocimiento humano, como objetos portadores de información en todas sus formas, tamaños y medios. El énfasis primario es sobre el contenido de estos objetos en términos de su potencial para transmitir información... [22]*”. Este autor considera importante que ambos enfoques, el orientado a sistemas y orientado a usuarios, funcionen juntos.

El autor considera que “*El carácter interdisciplinar en la Ciencia de la Información fue introducido por dos factores. En primer lugar, los problemas abordados no pueden ser resueltos mediante aproximaciones y construcciones provenientes de una sola disciplina... En segundo lugar, fue introducida y ha sido perpetuada hasta el presente por la multiplicidad de orígenes académicos de las personas que se han ocupado de la solución de los problemas descritos... [22]*”.

Es determinante hacer algunas precisiones en cuanto a las diferencias significativas que Saracevic expone entre la Bibliotecología y la Ciencia de la Información:

“La selección de los problemas abordados y la manera de definirlos... nunca fueron ni son tema de estudio de la ciencia de la biblioteca... [22].”

¿No pueden ser tema de estudio de la Bibliotecología, tomando como ejemplo lo planteado por Saracevic como áreas de la Ciencia de la

Información, las búsquedas en Internet, la bibliotecas digitales, la recuperación de información, etc?

“Las cuestiones teóricas planteadas y los marcos de referencia tomados por la Bibliotecología, fundamentalmente basados en teorías filosóficas y de la comunicación, no tienen contrapartida en la ciencia de la información y viceversa... [22].”

¿No son incidentes tanto en la Bibliotecología como en la Ciencia de la Información por ejemplo, de la filosofía en el mundo de las grandes teorías, el positivismo; no tienen influencia ambas, de la psicología, de dos orientaciones teóricas expresiones de la revolución cognitiva, el cognitivismo y el constructivismo, etc?

“La naturaleza y grado de experimentación y de desarrollo empírico de ambas, así como los requerimientos profesionales también difieren a escala significativa... [22]”

¿Por qué?

“Las herramientas y aproximaciones empleados... [22].”

¿Cuáles son esas herramientas y aproximaciones, si la Bibliotecología puede utilizar las mismas herramientas de la Ciencia de la Información, por ejemplo para automatizar la biblioteca?

Considera que “*la Bibliotecología y la Ciencia de la Información son dos campos de conocimientos diferentes con fuertes nexos interdisciplinarios... la ciencia de la computación lidia con la manipulación de los símbolos y la ciencia de la información con la manipulación de los contenidos... [22]*”.

Conclusiones

No cabe dudas que una de las profesiones más comprometidas con el estudio de la información es la Bibliotecología y la Ciencia de la Información, disciplinas que están teniendo una mirada nueva y un enfoque epistemológico renovado, con el fin de ubicarse adecuadamente en el mundo que nos rodea.

Existen diferentes tendencias a concebir la Bibliotecología y la Ciencia de la Información, unas que las afirman como un mismo y único campo;

otras que las consideran una como caso especial de la otra; y algunas como campos diferentes. Es importante estudiar sus bases históricas de aparición y desarrollo (primero, la Bibliotecología en el siglo XIX, como respuesta a la revolución política e industrial, el desarrollo de las ciencias naturales y el surgimiento de las ciencias sociales; y después en el siglo XX la Ciencia de la Información como respuesta a la revolución científica y tecnológica y el fenómeno de la explosión de la información), sus objetivos concretos, las áreas que abordan, las herramientas que utilizan, etc.

El análisis teórico en Bibliotecología y Ciencia de la Información permite ordenar, explicar e interpretar todos sus fundamentos y además su estudio epistemológico contribuye a la visualización de problemas reales en estas áreas y conocerse a sí mismas.

Saber de nuestra realidad conceptual es una obligatoriedad para un profesional en Bibliotecología y Ciencia de la información, hay que tomar en cuenta que no es solo saber hacer sino también saber pensar, como dice el profesor Radamés Linares en sus clases.

Dilucidar los problemas teórico-metodológicos de estas área*-s es estar preparados para entender cabalmente los cambios ocurridos, los que ocurren y los que ocurrirán y algo muy importante dejar de ser profesionales pasivos y crecernos como actores de la investigación, con un pensamiento crítico y claro.

Referencias

- 1) Buonocore, Domingo. "¿Es la bibliotecología una ciencia?" *Documentación Bibliotecológica* (Bahía Blanca) No.5: 3, 1975. Citado por Conforti, Noemí y Carlos Hugo Artaza. "La Bibliotecología: ¿ciencia, técnica o tecnología?" [en línea]. III Encuentro de Docentes. Ponencias Generales. Argentina, 2001. <http://www.eubca.edu.uy/Encuentro/encdocentes/encdo_ponencias_generales/encdo_ponencias_grales_Argentina/encdo_ponencias_grales_argentina1.html>. [Consulta: diciembre del 2001].
- 2) Fernández, Stella Maris. "¿Es la bibliotecología una ciencia?". *Documentación Bibliotecológica* (Bahía Blanca) No.5: 10, 1975. Citado por Conforti, Noemí y Carlos Hugo Artaza. "La Bibliotecología: ¿ciencia, técnica o tecnología?" [en línea]. III Encuentro de Docentes. Ponencias Generales. Argentina, 2001. <http://www.eubca.edu.uy/Encuentro/encdocentes/encdo_ponencias_generales/encdo_ponencias_grales_Argentina/encdo_ponencias_grales_argentina1.html>. [Consulta: diciembre del 2001].
- 3) Conforti, Noemí y Carlos Hugo Artaza. "La Bibliotecología: ¿ciencia, técnica o tecnología?" [en línea]. III Encuentro de Docentes. Ponencias Generales. Argentina, 2001. <http://www.eubca.edu.uy/Encuentro/encdocentes/encdo_ponencias_generales/encdo_ponencias_grales_Argentina/encdo_ponencias_grales_argentina1.html>. [Consulta: diciembre del 2001].
- 4) Restrepo L., Fabio. "Fundamentos de Bibliotecología y Ciencia de la Información" [en línea]. Universidad Interamericana Recinto San Germán. Programa Graduado de Educación. Maestría en Ciencias Bibliotecarias y de la Información. <<http://www.sg.inter.edu/lisc/online/5000/semana1.pdf>>. [Consulta: diciembre del 2001].
- 5) García, Jesús Francisco. "Las bibliotecas especializadas y su incidencia en el contexto económico y social de América Latina" [en línea]. 62nd IFLA General Conference- Conference Proceeding- August 25-31, 1996. <<http://www.ifla.org/IV/ifla62/62-arc.htm>>. [Consulta: diciembre del 2001].
- 6) Almanza Morales, José Luis. "La interdisciplina en la Bibliotecología" [en línea]. Dirección General de bibliotecas de la Universidad Nacional Autónoma de México. <<http://www.dgbiblio.unam.mx/servicios/dgb/publicdgb/bole/fulltext/volVI13/interdis.html>>. [Consulta: diciembre del 2001].
- 7) Rendón Rojas, Miguel Ángel. "Hacia un nuevo paradigma en Bibliotecología" [en línea]. *Transinformacao* (Brasil) 8(3), septiembre-diciembre 1996 <<http://www.puccamp.br/~biblio/rojas83.html>>. [Consulta: diciembre del 2001].

- 8) Rendón Rojas, Miguel Ángel. "La naturaleza dialógica de la Ciencia Bibliotecológica en el contexto de las nuevas tecnologías de la información" [en línea]. *Revista General de Información y Documentación* (Madrid) 9 (1): 33-45.
<<http://dois.mimas.ac.uk/DolS/data/Articles/julqgkotgy:1999:v:9:i:1:p:33-45.html>>. [Consulta: diciembre del 2001].
- 9) Oropeza de Ojeda, Olga. "Reflexiones sobre aspectos teóricos metodológicos de la investigación en el área de Bibliotecología y Archivología" [en línea].
<<http://150.185.88.116/Humanitas/Postgrado/Akadememos/Vol1-n1/pag12.asp>>. [Consulta: diciembre del 2001].
- 10) Garza Mercado, Ario. "Enseñanza bibliotecológica: dos ensayos y un proyecto". México, Asociación Nacional de Universidades e Institutos de Enseñanza Superior, 1974, p.32. Citado por Conforti, Noemí y Carlos Hugo Artaza. "Thomas S. Khun y la ciencia de la información" [en línea]. III Encuentro de Docentes. Ponencias Generales. Argentina, 2001.
<http://www.eubca.edu.uy/Encuentro/encdocentes/encdo_ponencias_generales/encdo_ponencias_grales_Argentina/encdo_ponencias_grales_argentina2.html>. [Consulta: diciembre del 2001].
- 11) Borko, H. "Information Science: what is it?" *American Documentation*, 1968. Citado por Conforti, Noemí y Carlos Hugo Artaza. "Thomas S. Khun y la ciencia de la información" [en línea]. III Encuentro de Docentes. Ponencias Generales. Argentina, 2001.
<http://www.eubca.edu.uy/Encuentro/encdocentes/encdo_ponencias_generales/encdo_ponencias_grales_Argentina/encdo_ponencias_grales_argentina2.html>. [Consulta: diciembre del 2001].
- 12) Rendón Rojas, Miguel Ángel. "Concepto de ciencia de la información". *Trasinformación* No.3: 1-4, sept-dic 1996. Citado por Conforti, Noemí y Carlos Hugo Artaza. "Thomas S. Khun y la ciencia de la información" [en línea]. III Encuentro de Docentes. Ponencias Generales. Argentina, 2001.
<http://www.eubca.edu.uy/Encuentro/encdocentes/encdo_ponencias_generales/encdo_ponencias_grales_Argentina/encdo_ponencias_grales_argentina2.html>. [Consulta: diciembre del 2001].
- 13) Geldres Flores, D. "Bibliotecología, documentación e información". *RIDECAB* (Lima) 3(5): 19-33, 1982. Citado por Conforti, Noemí y Carlos Hugo Artaza. "Thomas S. Khun y la ciencia de la información" [en línea]. III Encuentro de Docentes. Ponencias Generales. Argentina, 2001.
<http://www.eubca.edu.uy/Encuentro/encdocentes/encdo_ponencias_generales/encdo_ponencias_grales_Argentina/encdo_ponencias_grales_argentina2.html>. [Consulta: diciembre del 2001].
- 14) Atherton, P. "Manual para sistemas y servicios de información". París, UNESCO, 1983, p. 268. Citado por Conforti, Noemí y Carlos Hugo Artaza. "Thomas S. Khun y la ciencia de la información" [en línea]. III Encuentro de Docentes. Ponencias Generales. Argentina, 2001. <http://www.eubca.edu.uy/Encuentro/encdocentes/encdo_ponencias_generales/encdo_ponencias_grales_Argentina/encdo_ponencias_grales_argentina2.html>. [Consulta: diciembre del 2001]
- 15) Currás, E. "La información en sus nuevos aspectos: ciencias de la documentación". Madrid, Parainfo, 1988, p.31. Citado por Conforti, Noemí y Carlos Hugo Artaza. "Thomas S. Khun y la ciencia de la información" [en línea]. III Encuentro de Docentes. Ponencias Generales. Argentina, 2001.
<http://www.eubca.edu.uy/Encuentro/encdocentes/encdo_ponencias_generales/encdo_ponencias_grales_Argentina/encdo_ponencias_grales_argentina2.html>. [Consulta: diciembre del 2001].
- 16) Conforti, Noemí y Carlos Hugo Artaza. "Thomas S. Khun y la ciencia de la información" [en línea]. III Encuentro de Docentes. Ponencias Generales. Argentina, 2001.
<http://www.eubca.edu.uy/Encuentro/encdocentes/encdo_ponencias_generales/encdo_ponencias_grales_Argentina/encdo_ponencias_grales_argentina2.html>. [Consulta: diciembre del 2001].

- 17) Do Amaral, Sueli Angelica. "Investigación en Ciencia de la Información: un desafío latinoamericano para construir la sociedad de la información" [en línea]. 67th IFLA Council and General Conference. August 16-25, 2001.
< <http://www.ifla.org/IV/ifla67/papers/113-173s.pdf>>. [Consulta: diciembre del 2001].
- 18) Pinheiro, Lena Vania Riberio. Campo interdisciplinar en Ciencia de Información: fronteras remotas y recientes. *En* Ciencia da Informacao, Ciencias Sociais e Interdisciplinaridade. Brasília, Instituto Brasileño de Información en Ciencia y Tecnología, 1999, p.155-182. Citado por Do Amaral, Sueli Angelica. "Investigación en Ciencia de la Información: un desafío latinoamericano para construir la sociedad de la información" [en línea]. 67th IFLA Council and General Conference. August 16-25, 2001.
< <http://www.ifla.org/IV/ifla67/papers/113-173s.pdf>>. [Consulta: diciembre del 2001].
- 19) Hawkins, Donald T. *Journal of the American Society for Information Science and Technology*. 52(1), 2001. Citado por Do Amaral, Sueli Angelica. "Investigación en Ciencia de la Información: un desafío latinoamericano para construir la sociedad de la información" [en línea]. 67th IFLA Council and General Conference. August 16-25, 2001.
< <http://www.ifla.org/IV/ifla67/papers/113-173s.pdf>>.
- 20) Miranda, Antonio y Aldo de Albuquerque Barreto. Pesquisa em Ciencia de Informação em Brasil: síntesis y perspectiva. *Revista de Ciencia da Informacao* (Brasil) 1(6), diciembre 2000. Citado por Do Amaral, Sueli Angelica. "Investigación en Ciencia de la Información: un desafío latinoamericano para construir la sociedad de la información" [en línea]. 67th IFLA Council and General Conference. August 16-25, 2001.
< <http://www.ifla.org/IV/ifla67/papers/113-173s.pdf>>. [Consulta: diciembre del 2001].
- 21) Capurro, Rafael. La hermenéutica y el fenómeno de la información [en línea]. JAI/Elsevier Inc., (19): 79-85, 2000.
<<http://www.capurro.de/herminf.html>>. [Consulta: diciembre del 2001].
- 22) Saracevic. Ciencia de la Información. Material entregado en clases por el profesor Radamés Linares en la Asignatura Bases Teóricas de la Bibliotecología y Ciencia de la Información, Maestría en Bibliotecología y Ciencia de la Información.

Recibido: 18 de junio del 2002.

Aprobado: 19 de julio del 2002.

Marlery Sánchez Díaz

Centro Nacional de Biopreparados
Carretera Beltrán Km 1 ½ Bejuca, La Habana,
Cuba.

Correo electrónico:
<infctf@biocen.colombus.cu>.

La literatura gris de Rusia en el mundo de la información digital

Leonid P. Pavlov

RESUMEN

Se da la definición de literatura gris y se destaca su nuevo papel en la sociedad de la información. El Centro de Información Científica y Técnica de Rusia (acrónimo en ruso - VNTIC), fue creado hace 35 años por decisión del gobierno, con el objetivo de centralizar las colecciones y el control bibliográfico de toda la literatura gris soviética (hoy rusa). El VNTIC tiene la responsabilidad de mantener todos los informes de I+D y las tesis (de candidato y doctor). El VNTIC realiza las dos funciones, la de archivar y diseminar la literatura gris. Se da una panorámica general de las actividades del VNTIC, así como una descripción breve de sus productos y servicios de información. Se considera a la literatura gris, como una dirección muy promisoría y fructífera en el campo de la cooperación internacional y el intercambio de información. Se dan algunas vías para eliminar obstáculos en la cooperación.

ABSTRACT

The definition of grey literature is given and its new role in information society is outlined. The Scientific and Technical Information Centre of Russia (the Russian acronym - VNTIC) was created 35 years ago by the government decision to centralise the collection and bibliographic control of all Soviet (now - Russian) grey literature. VNTIC is responsible for the maintenance of the complete repository for scientific R&D reports and dissertations (candidate and doctoral). VNTIC carries out the dual function of both archiving and disseminating grey literature. A general outline of VNTIC activities is given and a brief description of VNTIC information products and services is presented. The sphere of grey literature is considered to be a very promising and fruitful direction of international cooperation and information exchange. Some ways to overcome cooperation barriers are discussed.

La sociedad de la información y la literatura gris

No importa cuan difícil y complicada sea la situación económica en Rusia, el país está inevitablemente envuelto en el proceso internacional de entrar en la era de información digital. La sociedad de información y el desarrollo de la infraestructura es la esfera más excitante y prometedora de cooperación internacional moderna.

Lo más importante sobre la sociedad de información es que está caracterizada por la coexistencia de dos mundos: un mundo físico natural y un mundo virtual artificial creado por el hombre, y que existe en los medios de computación. El mundo virtual, o la realidad virtual es una reflexión parcial del mundo físico en la memoria de computadora. La parte

reflejada está creciendo muy rápido. Es de importancia vital que los dos mundos no solo coexistan pasivamente, sino interactúen de forma más activa e influyan uno sobre el otro: en la era de la información el mundo físico no es capaz de existir y desarrollarse por largo tiempo sin el mundo virtual, lo anterior se convierte en computadora (software) dependiente y lo último se convierte en una parte del *metabolismo* del mundo físico.

En este contexto es bastante evidente el nuevo papel de la información y la documentación digital. Somos testigos de un cambio importante en la comunicación del conocimiento que va desde los materiales impresos a los medios electrónicos de comunicación. La digitalización y la conexión de redes ofrecen muchas oportunidades para el desarrollo de nuevas formas de comunicación. En un espacio de la red global digital la literatura gris se ha convertido en una importante forma de representación de la información.

La literatura gris es un término que describe productos de información que se crean y distribuyeron para diseminar el conocimiento (las ideas, los hechos, las opiniones) y no para venderlos para obtener una ganancia. En la práctica, por consiguiente, la literatura gris puede definirse también como la información que no es comercializada y distribuida por las organizaciones de la publicación comercial. La Tercera Conferencia Internacional sobre la Literatura Gris (GL'97) fue un éxito al redefinir el término *literatura gris* antes de iniciarse un nuevo milenio. A la definición revisada se le ha llamado La Convención de Luxemburgo sobre la Literatura Gris, debido a la ciudad en que fue organizada. Ahora, la definición de literatura gris expresa "la que se produce en todos los niveles de gobierno, académicos, comerciales e industriales en formatos impresos y electrónicos, pero que no es controlado por los publicadores comerciales." Además, al constituir la base de información para la gestión de redes modernas, se conoce que la literatura gris tiene ventajas, tales como la rápida disponibilidad para los usuarios (sin los típicos retrasos que se producen al publicar), la riqueza de informaciones y el hecho de que está libre de la censura.

Centro de Información Federal de Literatura Gris Rusa

A finales de la década del 70 e inicios de la década del 80, en la Unión Soviética se estableció una

infraestructura bien desarrollada para la información científica y técnica (ICT). Esta infraestructura se basaba en el llamado Sistema del Estado para ICT (abreviado en ruso como GSNTI).

El Sistema del Estado para ICT integró una avanzada jerarquía de tres niveles de centros de información incluyendo las instituciones a nivel de toda la Unión, los centros específicos y regionales en un segundo nivel y en el tercer nivel los departamentos de ICT de diferentes organizaciones y empresas. El alcance de los centros específicos y regionales es evidente por sus propios nombres, por ejemplo Infornelectro, Informenergo y otros centros pertenecientes a los correspondientes Ministerios y comité Estatales. El Centro ICT Volgogrado, ICT Centro Khabarovsk, etc. En total existían 69 centros que cubrían las diferentes regiones geográficas.

Cada uno de los centros de toda la Unión eran responsables de un tipo particular de documento de ICT. Por ejemplo, el Instituto de Información Científica y Técnica (VINITI), que entonces era de toda la Unión y ahora abarca a toda Rusia, se ocupaba de todos los documentos de ICT publicados (artículos de revistas científicas, monográficos, etc.); la Asociación Científica e Industrial "Poisk" — patentes y solicitudes de patente; la Biblioteca Pública Estatal de Ciencia y Técnica (GPNTB) - libros, revistas y publicaciones seriadas, etc.,

Entre las organizaciones del primer nivel se encuentra una creada hace 35 años con el objetivo de coleccionar y controlar la literatura gris soviética. Este centro, ahora nombrado Centro de Información Científica y Técnica de Rusia (la sigla en ruso es VNTIC) es uno de los cuerpos de información más grandes del mundo y mantiene una inmensa colección de literatura gris rusa que refleja el estado del arte en la ciencia y tecnología rusa y abarca importantes datos no disponibles fuera de la antigua Unión Soviética (FSU).

El VNTIC es una agencia nacional (es decir, federal) de información, responsable del mantenimiento y desarrollo de una amplia colección de informes y disertaciones (de candidatos y doctores) científicas de I+D de toda Rusia. La presentación de los documentos al VNTIC es obligatoria para todas las organizaciones involucradas en la investigación y desarrollo científicos con presupuesto estatal y se lleva a cabo a tenor de la Ley Federal de la Federación Rusa "Acerca de la copia obligatoria de documentos" adoptada por el Parlamento ruso y publicada 17 de enero de 1995. La colección está

compuesta por documentos primarios (informes y disertaciones) a texto completo, guardadas en microficha, y documentos secundarios legibles por máquina, y que contienen descripciones bibliográficas y resúmenes de los documentos primarios, almacenados en una estructura de base de datos para facilitar la búsqueda y recuperación de la información.

En los 35 años de existencia del VNTIC se han adquirido unos siete millones de documentos que cubren todas las áreas del conocimiento humano. La singularidad de la colección del VNTIC radica en:

- 1) Los documentos son inéditos y no circulados; es decir existen solo en dos copias mecanografiadas o en computadora, una de las cuales pertenece al autor o la institución fuente y la otra se encuentra en el VNTIC;
- 2) la información contenida en los documentos es puesta a disposición de los usuarios con el menor retraso posible; es decir, alrededor de dos o tres meses, en comparación con los dos años necesarios para que esta información (e incluso a veces incompleta) aparezca publicada (en revistas, libros, monografías);
- 3) los expertos estiman que sólo un cuarto de la información contenida en los informes y disertaciones científicas será publicado en algún momento.

Por consiguiente, la colección del VNTIC es una parte integrante de la riqueza científica y cultural nacional de Rusia e, indudablemente posee una importancia y un valor mundial. A la larga, la información de la colección puede ahorrar miles de millones de dólares a la comunidad científica mundial al proporcionar información sobre caras investigaciones científicas (digamos, en química o física nuclear) que posiblemente ya hayan sido realizadas en la antigua Unión Soviética y Rusia. La colección y las bases de datos del VNTIC no sólo son de gran interés general, sino también para los especialistas - economistas, analistas políticos o ecólogos — que estén involucrados en la investigación sobre la transformación de Rusia y la antigua Unión Soviética.

El VNTIC proporciona los siguientes tipos de servicios a los usuarios:

- Una revista de resúmenes en 28 series;
- Diseminación selectiva de la base de datos en

formato electrónico en diferentes áreas temáticas;

- Búsqueda en línea y sin conexión en la base de datos; entrega de documentos (primarios) a texto completo;
- Traducción al inglés de resúmenes y disertaciones y producción de bases de datos en idioma inglés;
- Publicación de resúmenes en idioma inglés en diez series;
- Publicación de versiones en idioma ruso en inglés de un directorio de organizaciones de I+D;
- Producción de un CD-ROM que cubre tanto los documentos en idioma ruso como los de lengua inglesa;
- Boletines de registro que contienen los datos sobre los nuevo proyectos de I+D.

A continuación aparece una lista de bases de datos originales generada y puesta a disposición por el VNTIC:

- *Base de datos sobre informes de I+D realizados en Rusia (antes de 1991 - en la URSS)*. La base de datos es abarcadora y cubre todo los temas científicos. Cubre 1 200 000 documentos en ruso, a partir de 1982 y se actualiza una vez al mes.
- *Base de datos de las disertaciones de candidatos y doctores presentadas en Rusia (antes de 1991, en la URSS)*. Todo el conjunto de disertaciones cubre 400 000 documentos en ruso; producidos desde 1982 en todos los campos temáticos. Se actualiza una vez al mes.
- *Base de datos "Poseedores del Título de Doctor en Rusia"*, un quién es quién que abarca más de 14 000 científicos rusos. Se actualiza anualmente.
- *"Investigación y desarrollo en Rusia" (RDIR)*; una base de datos en idioma inglés que contiene 60 000 resúmenes de documentos (1989-2000) sobre los informes de I+D y las disertaciones para los diez áreas más solicitadas. Traducido del ruso.

A pesar de la muy conocida situación de crisis en la

ciencia rusa, la entrada de documentos al VNTIC sigue manteniendo un nivel relativamente alto. Se reciben varias decenas de miles de documentos anualmente (en el 2001 llegaron al VNTIC casi 9 500 informes y 21 000 disertaciones). De esa manera, el VNTIC se ocupa de procesar varios miles de páginas de informes y disertaciones diariamente. Ello se debe a que las tecnologías usadas en el VNTIC son más a escala industrial que de laboratorio u oficina.

A inicios de la década del 90, habían grandes objeciones a la existencia de un gran centro de información nacional en Rusia. Todo lo que fuera centralizado o controlado por el estado era automáticamente sospechoso y tildado de soviético. Sin embargo, la experiencia ha demostrado que los grandes centros de información a nivel federal tienen un importante papel que jugar. El VNTIC es un caso clásico en este sentido.

Superando las barreras de la cooperación

Se considera que la esfera de la ICT en general y de la literatura gris en particular es una dirección muy prometedora y fructífera de cooperación internacional e intercambio de información. Ya es algo usual que los excelentes logros científicos no sean un privilegio de una persona o una nación aislada, sino el resultado de los esfuerzos colectivos de muchas personas e instituciones del mundo.

Felizmente, la política exterior existente en Rusia elimina todas las barreras políticas y económicas, o ambas, para la cooperación internacional libre, global y regional, incluyendo el campo de la ciencia y la información.

Las próximas barreras importantes son de naturaleza legal. Está bastante claro que al entrar en el mercado mundial con los productos intelectuales, o como decimos en Rusia, con los resultados de las actividades científicas y técnicas, debe tenerse mucho cuidado con el respeto a las complicadas leyes y regulaciones acerca de los derechos de propiedad intelectual. Mucho se ha hecho últimamente en Rusia para mantener un marco legal amplio para definir y hacer cumplir los derechos de

propiedad intelectual, las leyes y regulaciones que rigen la producción, la transferencia internacional y el empleo de los conocimientos y de la tecnología.

Sólo existen barreras técnicas e idiomáticas para una cooperación internacional íntima que deben ser superadas por los oficiales ejecutivos, los científicos, los estudiosos y los especialistas. Eliminar las barreras técnicas significa asumir proyectos colectivos más eficaces en la esfera de la ICT.

Por último aparece la barrera del idioma, evidentemente una de las más importantes en la comunicación científica con Rusia. Desgraciadamente en Rusia no consideramos que adoptar el idioma inglés como el internacional sería una manera de salir de la situación. El idioma ruso científico y técnico está tan bien desarrollado y profundamente arraigado en la mentalidad de científicos e ingenieros rusos que sería poco realista esperar que todos los autores rusos pasen del idioma ruso al inglés a la hora de documentar sus trabajos de I+D. Al mismo tiempo, una traducción humana de todos los documentos de ICT rusos es prohibitivamente cara. Se deben considerar las siguientes formas de resolver el problema del idioma:

La traducción selectiva de los temas más *calientes* en la ciencia y tecnología por parte de los propios autores; para facilitar esta vía se han realizado muchos esfuerzos encaminados a mejorar las normas de aprendizaje de idiomas extranjeros en Rusia; se supone que fuera de Rusia se realicen esfuerzos recíprocos para aumentar el número de estudiantes de ruso y mejorar la calidad del aprendizaje del idioma ruso; el desarrollo y aplicación de sistemas de traducción computadorizados; en este sentido, en Rusia se están obteniendo resultados impresionantes, por ejemplo, los paquetes de programas RETRANS y ERTRANS (Ruso-Inglés y viceversa) con el diccionario de 1,5 millones de términos científicos y técnicos, buscadores multilingües para la búsqueda en lengua inglesa en las bases de dato rusas y viceversa, etc.; el uso del llamado modo de demanda retrasada: las solicitudes de los usuarios (las órdenes de la búsqueda), son enviadas, traducidas al en inglés y procesadas en ruso.

Para concluir y desde un punto de vista optimista se debe plantear que Rusia está abierta a cualquier nivel de cooperación internacional fructífera, independientemente de cuán complicadas puedan estar la situación nacional o la internacional.

Recibido:

Aprobado:

Leonid P. Pavlov

*Centro de Información Científica y Técnica de
Rusia,*

14 St. de Smolnaya, Moscú,

125993, la Rusia,

Correo electrónico: <postmaster@vntic.org.ru>

<www.vntic.org.ru>.
