

Resúmenes de documentación: análisis de la literatura de la documentación.

Parte 2: una nueva taxonomía para la documentación [i]

DONALD T. HAWKINS, SIGNE E. LARSON Y BARI Q. CATON [ii]

En la actualidad y debido a Internet, las clasificaciones por materia y los tesauros han alcanzado su momento de mayor relevancia. La clasificación de la información por materias, o taxonomías, ofrece a los usuarios la oportunidad de mejorar sustancialmente la efectividad de la búsqueda y recuperación de información. Este artículo es una continuación de una investigación cuyo objetivo era buscar una nueva definición del campo de la Documentación y crear un “mapa” que muestre cuáles son las materias principales y sus relaciones con otras materias de secundaria importancia dentro de la disciplina. Mediante un caso práctico se describe la creación de una nueva estructura de clasificación (taxonomía) para la base de datos *Information Science Abstracts* (ISA), cuyo objetivo es reflejar y adaptarse a los rápidos y continuos cambios de la tecnología y del mercado que afectan a la industria de la información hoy y con vistas al futuro. Utilizando una muestra de unos 3.000 resúmenes de ISA, un equipo de tres miembros compuesto por un editor de bases de datos, un bibliotecario de referencia y un resumidor-indizador, que representan a tres de las principales comunidades dentro del campo de la Documentación, llevaron a cabo dos experimentos de validación. En el primero, se clasificó la muestra de resúmenes de acuerdo con la nueva taxonomía propuesta. Posteriormente, tras el análisis de los resultados y la revisión de la taxonomía, ésta fue ratificada en un segundo experimento. El grado de coincidencia entre los tres indizadores fue significativamente mayor que el alcanzado en estudios previos. La taxonomía obtenida tras esta investigación parte de conceptos, definiciones y un mapa del campo de la Documentación que ya habían sido previamente desarrollados y los presenta de forma jerárquica, lo que supone una significativa contribución a esta disciplina.

Palabras claves: Vocabularios controlados. Tesauros. Documentación. Taxonomía. Análisis de áreas del conocimiento

- i. Traducción del artículo “Information science abstracts: Tracking the literature of information science. Part 2: A new taxonomy for information science”, de Donald T. Hawkins, Signe E. Larson y Bari Q. Caton, publicado en la revista *Journal of the American Society for Information Science and Technology*, Vol. 54, Número 8 (2003), pp. 771-781, con la debida autorización del editor.
- ii. Traducido por Lozano Palacios, A. y Cifuentes Pérez, M.A..

INFORMATION SCIENCE ABSTRACTS: TRACKING THE LITERATURE OF INFORMATION SCIENCE. PART 2: A NEW TAXONOMY FOR INFORMATION SCIENCE

Subject classifications and thesauri have become more important than ever in the Web environment. Efforts made to organize information into subject classifications, or *taxonomies*, offer users the opportunity to substantially improve the effectiveness of their search and retrieval activities. This article continues earlier research on the development of a new definition of the field of information science and the creation of a “map” of the field showing subjects central to it and their relationships to those on the periphery. A case study describes the creation of a new classification structure (taxonomy) for the *Information Science Abstracts* (ISA) database, aiming to reflect and accommodate the rapid and continued technological and market changes affecting the information industry today and into the future. Based on a sample of some 3,000 ISA abstracts, two validation experiments were conducted by a three-member team comprising a database editor, a reference librarian, and an abstractor-indexer, who represent three of the major communities within the information science field. In the first experiment, the sample of abstracts was classified according to the proposed new taxonomy; after analysis of the data and revision of the taxonomy, it was revalidated and fine tuned in a second experiment. Indexer consistency measures obtained in this study were significantly higher than those found in previous studies. The taxonomy resulting from this research employs the concepts, definition, and map of information science previously developed. It presents them in an organized hierarchical view of the field and thus makes a significant contribution to information science.

Keywords: Controlled vocabularies. Thesauri. Information science. Taxonomy. Domain analysis.

1. INTRODUCCIÓN

Desde hace tiempo los servicios de indización y resumen ofrecen a los usuarios ayudas para la recuperación de la información de sus bases de datos. Los primeros productos de muchos de estos servicios bibliográficos eran publicaciones impresas organizadas en secciones divididas de forma similar a los capítulos de un libro y que contenían índices de autor y de materia de gran utilidad. Cada sección contenía resúmenes organizados bajo epígrafes muy generales que los usuarios consultaban para encontrar artículos de su interés. Los índices de materias de estas publicaciones se asignaban generalmente a partir de vocabularios controlados de términos creados con un coste y un esfuerzo considerable por parte de los profesionales de la información con experiencia en la disciplina objeto de la publicación. Con frecuencia, estos vocabularios estaban organizados jerárquicamente a modo de tesoro y los títulos de las secciones de estas publicaciones eran los términos más genéricos o “encabezamientos principales” de los tesauros.

Cuando el contenido de estas publicaciones pasó a formar parte de las bases de datos en línea, los distribuidores de estas bases de datos incluyeron la posibilidad de hacer búsquedas a través de estos términos y encabezamientos principales. Estas funciones de valor añadido suponían una inestimable ayuda para los investigadores, especialmente antes de que estuvieran accesibles en línea los textos completos de los documentos. Muchos productores de bases de datos emplearon gran cantidad de recursos en la preparación de los usuarios (que en aquel entonces eran principalmente los profesionales de la información que actuaban como intermediarios) en el

uso de sus tesauros. (Uno de los productores, la Biblioteca Nacional de Medicina de los Estados Unidos, exigía en sus inicios la participación en un curso preparatorio antes de facilitar la contraseña de entrada a su sistema MEDLINE)

TAXONOMÍAS EN EL ENTORNO DE INTERNET

Una de las cuestiones a las que debe enfrentarse un productor de bases de datos es si le merece la pena dedicar el dinero y el esfuerzo que suponen producir una taxonomía para indizar la información haciendo uso de ella. Cuando aparecieron las primeras bases de datos comerciales en línea, los profesionales solicitaron la inclusión de tesauros y clasificaciones por materia que les ayudaran a formular sus búsquedas. Los productores de bases de datos respondieron favorablemente, y la mayoría de las primeras bases de datos en línea incluían campos de vocabulario controlado (descriptores) extraídos de los tesauros, además de otros campos para los sistemas de clasificación. Las experiencias de los últimos años en el entorno de Internet ha puesto de manifiesto que todavía merece la pena el esfuerzo de incluir este tipo de datos.

Cuando se pudo acceder a través de Internet a las primeras bases de datos de información (especialmente a texto completo), muchos usuarios pensaban que los tesauros y las clasificaciones por materias ya no eran necesarios y que llegarían a desaparecer. Después de todo, según la teoría, si todo estaba disponible en línea a texto completo, uno simplemente necesitaría escribir los términos apropiados en un buscador de Internet y de este modo se recuperaría la información deseada. Los usuarios con poca experiencia pronto se dieron cuenta de la falsedad de este presupuesto sobre la recuperación de la información cuando tuvieron que enfrentarse a millones de páginas web, de las cuales muy pocas contenían la información deseada.

Pronto se reconoció que, lejos de desvanecerse en el entorno de la red, las clasificaciones temáticas y los tesauros habían cobrado más importancia que nunca y que organizar la información por temas, o *taxonomías*, ofrecía a los usuarios una importante herramienta para la recuperación. Uno de los primeros ejemplos de uso de taxonomías fue Yahoo!, que contrataba a profesionales de la información con la formación necesaria para organizar y clasificar las direcciones de las páginas web. Actualmente, además de ofrecer la posibilidad de hacer búsquedas simples, muchos buscadores disponen de una taxonomía opcional que puede usarse para la recuperación. El uso de las taxonomías está también muy extendido en las intranets de muchas compañías.

Consideraciones iniciales: En un artículo previo, Hawkins (2001) describió parte de la historia de Information Science Abstracts (ISA) [iii] y presentó una nueva

iii. ISA, una de las principales bases de datos de indización y resumen dedicadas a la documentación, publicada por Information Today, Inc. Se puede conseguir en papel directamente del editor o en

definición de la Documentación así como un “mapa” de las materias principales y sus relaciones con otras materias de importancia secundaria. El trabajo que se describe aquí es una aplicación práctica de la definición de Documentación y las materias relacionadas con esta disciplina que se esbozaron en el artículo de Hawkins. Hemos usado su descripción de la disciplina para desarrollar una nueva taxonomía para la Documentación (y para ISA en particular) que refleje fielmente el campo tal como es hoy. Estos conceptos se han usado igualmente para seleccionar artículos relevantes para su inclusión en ISA.

Tanto el sistema de clasificación (es decir, la taxonomía) como el vocabulario controlado usado por ISA experimentaron la última modificación en 1993, introduciéndose sólo pequeñas revisiones. Desde entonces, la taxonomía se ha quedado obsoleta, debido a que el campo de la información ha experimentado gran cantidad de cambios drásticos. Contenía muchos términos desfasados y no estaba preparada para adaptarse a los rápidos cambios tecnológicos y de mercado que afectan a la industria de la información. Los editores y los indizadores de ISA la encontraron difícil de usar; por lo que podemos deducir que los usuarios tendrían aún más dificultades. Un defecto aún mayor en la taxonomía previa fue que cada sección de nivel superior contenía una subsección que se llamaba “General”. Con el paso del tiempo, muchos resúmenes se habían incluido en estas subsecciones debido a que el campo había avanzado y la taxonomía no había recogido estos cambios. El resultado fue que la taxonomía prácticamente perdió su utilidad como herramienta para la recuperación de información.

El Gráfico 1 contiene datos obtenidos de la base de datos original de ISA, e ilustra estas observaciones anteriores. Nos muestra el número de registros incluidos en cada sección entre 1992 y 2001. Obsérvese la enorme variación en el número de registros contenidos en las distintas secciones. La sección con el mayor número de registros (5.11, Búsqueda y Recuperación) tenía más de 2.700 referencias y la que tenía menos (5.03, Macroordenadores) contaba con sólo cuatro referencias. Doce secciones tenían más de 1.000 referencias cada una y 15 tenían menos de 100 referencias. Estos datos reflejaban la clara necesidad de una nueva taxonomía para la Documentación y se consideró que la definición y el mapa desarrollados previamente eran un buen punto de partida.

línea a través de varios distribuidores. Para más información, véase <http://www.infoday.com/isa/default.htm>.

GRÁFICO 1 DISTRIBUCIÓN DE LOS REGISTROS SEGÚN EL SISTEMA DE CLASIFICACIÓN - TAXONOMÍA ANTERIOR



FILOSOFÍA Y METODOLOGÍA SEGUIDA EN LA CREACIÓN DE LA TAXONOMÍA

Dado que son los propios usuarios los que utilizan los buscadores de Internet, la estructura de una taxonomía debe ser clara y lógica además de ser fácil de usar si se desea que tenga una buena aceptación. Debe también ser dinámica, fácil de actualizar y capaz de reflejar los rápidos cambios y los avances tecnológicos. Estos principios son especialmente importantes para aquellos usuarios que utilizan la versión impresa de una publicación de resúmenes en busca de aquellas referencias que les puedan interesar.

Si la taxonomía está bien construida, los usuarios de Internet pueden beneficiarse de los encabezamientos de materia utilizándolos como términos generales para limitar las búsquedas. De hecho, como ya hemos señalado, muchos de los distribuidores de bases de datos incluyen índices invertidos de los encabezamientos principales con este propósito. (Por ejemplo, el sistema Dialog incluye el equivalente numérico de los encabezamientos principales en el campo *SC* (*Código de la Sección*) o en el campo *MC* (*Código del Encabezamiento Principal*) y el encabezamiento propiamente dicho en el campo *SH* (*Encabezamiento de la Sección*) o en el campo

MH (Encabezamiento Principal). Estos términos de búsqueda son muy prácticos cuando se desea recuperar una gran cantidad de términos relacionados en el ámbito general y usarlos como criterio limitador de otros aspectos de la búsqueda siguiendo un método de búsqueda del tipo “fracciones sucesivas” (véase Hawkins y Guajetas (1983) para el análisis de ésta y otras técnicas de búsqueda).

Tomando como punto de partida estos principios orientativos, recogimos los posibles términos iniciales para la taxonomía. Se utilizaron las siguientes fuentes:

1. Mapas previos del campo de la Documentación desarrollados por Hawkins (2001) y otros (véase el artículo de Hawkins como referencia).
2. La lista de descriptores de ISA vigente en aquel momento.
3. Las listas de descriptores usados por otras dos bases de datos del campo de la Documentación: la *Library and Information Science Abstracts* (LISA) de R.R. Bowker, ahora propiedad de *Cambridge Scientific Abstracts*, y la *Library Literature* de H.W. Wilson.
4. El *ASIS Thesaurus of Information Science and Librarianship* compilado por Milstead (1998).
5. La cobertura de dos de las principales revistas de Documentación, el *Journal of the American Society for Information Science* (JASIS) y el *Annual Review of Information Science and Technology* (ARIST).
6. Las secciones pertinentes de los tesauros usados por otras bases de datos que no son específicas de la Documentación pero que contienen términos sobre Documentación (por ejemplo la base de datos INSPEC).

Los términos posibles se agruparon y organizaron en una taxonomía preliminar que contenía 13 encabezamientos principales. Antes de ser aceptados, muchos términos fueron comprobados mediante búsquedas en la base de datos ISA, observando el número de referencias obtenidas así como el contexto en el que se usaban dichos términos.

VALIDACIÓN DE LA TAXONOMÍA

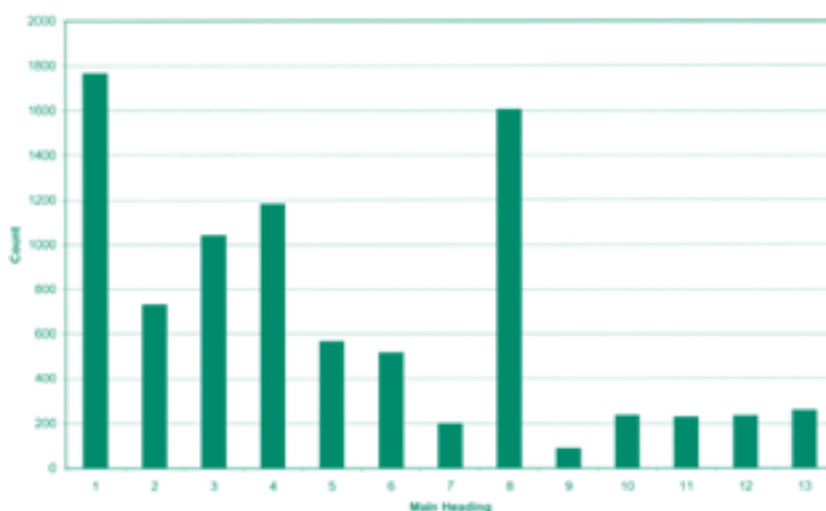
Cada uno de los autores de este artículo indizó independientemente las referencias bibliográficas aparecidas en ISA entre diciembre de 1998 y mayo de 1999 usando esta taxonomía preliminar. El total de referencias fue de 3.004 a las que se les asignaron un total de 9.000 términos. Cada referencia fue asignada a la clasificación temática que el indizador creyó más relevante. Para simplificar y agilizar este proceso, a cada referencia se le asignó sólo un número de clasificación (aunque, como todos sabemos, en la vida real a muchas referencias se les asignan varias clasificaciones). El formato del número de clasificación era x, y , siendo x el

“encabezamiento principal” e y el “subencabezamiento”. Además, se indicaron específicamente aquellas referencias para las que, a juicio de los indizadores, no existía una clasificación adecuada.

La clasificación así creada fue introducida en una base de datos de Microsoft Access. Si como mínimo dos de los indizadores coincidían en la clasificación de un resumen, entonces el número de la clasificación asignado se aceptaba como su descriptor correcto. Aquellos resúmenes que habían sido clasificados de modo diferente por los tres indizadores así como aquellos para los que no existía una clasificación adecuada se revisaron y bien se intentó reconciliar los diferentes puntos de vista o bien se introdujeron las modificaciones pertinentes en la taxonomía.

Usando la base de datos de Access, pudimos obtener rápidamente datos sobre la distribución de los resúmenes en cada una de las secciones (véase el gráfico 2). Como resultado, pudimos identificar ciertas categorías “raras” y mejorar la taxonomía combinando categorías afines. La distribución por secciones que se creó, como muestra el gráfico 2, se dividió en tres grandes grupos: el grupo formado por las secciones 1 y 8, relativas a la investigación básica en Documentación y a los sistemas de información electrónica, recibió el 35% de las asignaciones de descriptores. Un segundo grupo de secciones (de la 2 a la 6), recibió en total alrededor del 45% y el restante 20% se asignó a las demás secciones. Esta tarea fue extremadamente útil y reveladora ya que al haber resúmenes que no se pudieron indizar según la taxonomía usada y debido a la distribución desproporcionada de algunas secciones, como muestra el gráfico 2, se vio la necesidad de matizar la clasificación.

GRÁFICO 2
RESULTADO DEL PRIMER TEST DE VALIDACIÓN



COINCIDENCIA EN LA INDIZACIÓN

La labor de validación nos permitió también medir la coincidencia entre los tres indizadores, como se muestra en la tabla 1.

TABLA 1
COINCIDENCIA ENTRE LOS INDIZADORES

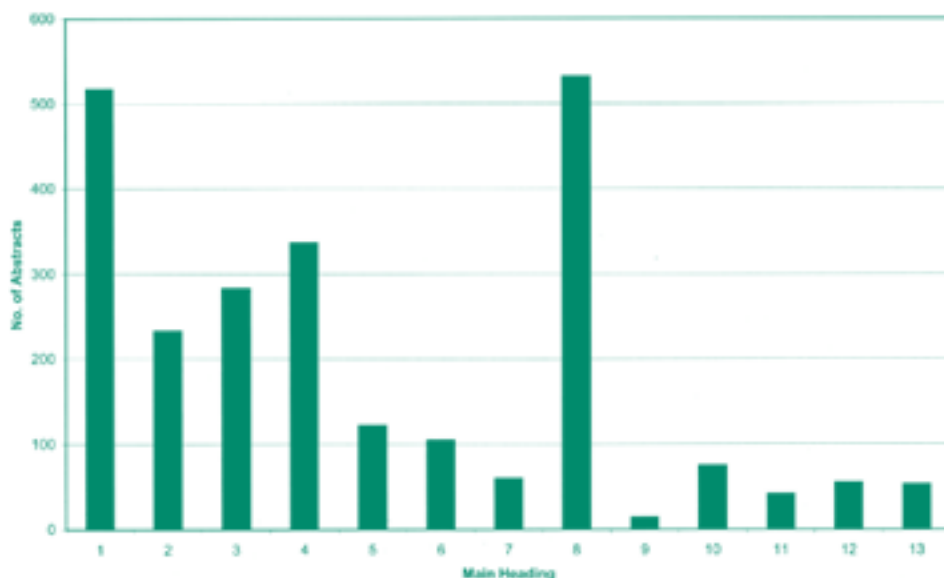
Coincidencia entre indizadores	Encabezamiento y subencabezamientos		Sólo encabezamientos	
	Total	% del Total	Total	% del Total
Los 3 coinciden	793	26,4	1069	35,5
A y B coinciden	402	13,3	443	14,7
A y C coinciden	424	14,1	434	14,4
B y C coinciden	484	16,1	482	16,0
Sólo 2 coinciden	1310	43,6	1359	45,2
Ninguno coincide	901	30,0	576	19,2

Teniendo en cuenta la clasificación completa de cada resumen, compuesta tanto por el encabezamiento principal como por los subencabezamientos, al menos dos de los indizadores coincidieron en la clasificación del 70% de los 3.004 resúmenes tomados de ISA. Ya que la distinción entre muchos de los subencabezamientos puede ser bastante pequeña, también estudiamos la coincidencia teniendo en cuenta sólo los encabezamientos principales. Como consecuencia, la coincidencia aumentó significativamente, obteniéndose una coincidencia del 81% entre los tres indizadores. El gráfico 3 muestra la distribución resultante de la clasificación de 2.428 resúmenes donde al menos dos indizadores coincidieron en los encabezamientos principales. Esta distribución confirmaba los resultados del gráfico 2, mostrando que algunas de las secciones se podían unir y que otras debían subdividirse. (Por ejemplo, el gráfico 3 muestra que con el transcurso del tiempo la sección 9 probablemente no contendría suficientes resúmenes como para justificar su existencia y que debería fundirse con otra sección relacionada).

Varios estudios han indicado que la coincidencia entre indizadores no excede generalmente el 50%, situación que no ha cambiado de forma significativa en los últimos 30 años (Leininger, 2000). Sievert y Andrews (1991) descubrieron que 71 pares de registros duplicados en ISA tenían los mismos descriptores aproximadamente en el 48% de los casos. Reich y Biever (1991) compararon los descriptores asignados por diferentes indizadores que usaban el mismo tesoro en dos bases de datos de agricultura y hallaron que en varias muestras la media de coincidencia oscilaba entre el 24% y el 45%. Nuestros resultados muestran unas coincidencias significati-

vamente más altas que las halladas previamente ya que en el 35% de los casos los tres indizadores coincidimos en la asignación y en el 45% de los casos coincidimos dos de nosotros. Sólo en el 19% de los casos ninguno de los tres coincidimos en la asignación de descriptores a los resúmenes.

GRÁFICO 3 DISTRIBUCIÓN DE LOS ENCABEZAMIENTOS PRINCIPALES – COINCIDENCIA ENTRE DOS INDIZADORES



Es importante resaltar que en este trabajo hemos asignado encabezamientos basándonos solamente en los resúmenes, a diferencia de otros estudios en los que la indización se hizo sobre el artículo completo. A pesar de esto, el nivel de coincidencia que hemos obtenido es significativamente más alto que el conseguido generalmente en esos otros estudios sobre indización. Algunas razones podrían ser:

(1) Nosotros poseemos una buena formación teórica en Documentación, así como muchos años de práctica en el campo. (Dos de nosotros hemos trabajado en diversas áreas de la Documentación alrededor de 25 años y el otro ha trabajado con resúmenes e índices durante 12 años).

(2) Hemos tenido la ventaja de haber estado en contacto previamente con el material usado en este estudio. Debido a nuestra relación e interés por ISA, este es el segundo estudio en el que hemos trabajado con los resúmenes o con los artículos que éstos representan.

(3) A medida que trabajábamos con el material y surgían dudas sobre la indicación, fuimos resolviéndolas y aprendiendo de nuestra experiencia.

(4) Debido a que estábamos usando resúmenes escritos por profesionales, podíamos estar seguros de que cada resumen identificaba con claridad los conceptos principales del artículo que representaba. Ya que los resúmenes eran breves y concretos, la incertidumbre era menor y, en consecuencia, hubo menos desacuerdos que si hubiésemos usado los artículos completos.

DESARROLLO Y USO DE LA TAXONOMÍA FINAL

A medida que avanzábamos en la validación de la primera taxonomía, hacíamos revisiones para corregir las deficiencias que íbamos encontrando. Además de usar nuestra experiencia práctica en la materia y la adquirida durante el proceso de validación, también examinamos los conceptos propuestos por Hawkins en el artículo anterior y nos aseguramos de incorporarlos en la taxonomía. Una gran parte de este trabajo se hizo de forma práctica ya que ISA continuaba publicándose mientras se llevaba a cabo esta investigación. De este modo pudimos aprovecharnos de lo que los autores estaban publicando en ese momento. En la taxonomía “final” (la consideramos final sólo en lo que afecta a esta investigación), los 13 encabezamientos se convirtieron en 11, y se modificaron asimismo algunas subcategorías.

También llevamos a cabo un segundo test de validación en el que dos de los autores indizaron manualmente las referencias bibliográficas aparecidas en ISA durante los meses de mayo, junio/julio y agosto del 2001, formadas por un total de 1.265 resúmenes. Los resúmenes se copiaron de la base de datos de ISA a una base de datos de Microsoft Access y la taxonomía también se convirtió en una base de datos de Access. El gráfico 4 es un ejemplo de un registro que muestra los campos de título y resumen y un enlace con la taxonomía para facilitar la búsqueda de los encabezamientos y la introducción del código del tema apropiado. Este segundo test dio como resultado la taxonomía presentada en el Apéndice, convirtiéndose en la “Taxonomía de ISA” a comienzos del 2002. Hasta hoy, se ha usado en la confección de seis números de la base de datos ISA con un total de 2.692 resúmenes. A cada uno de estos resúmenes se le debía asignar al menos a una clasificación temática y como máximo tres y el resultado final fue de 3.504 asignaciones. El gráfico 5 muestra cómo quedó la distribución.

GRÁFICO 4 ENTORNO PARA LA ASIGNACIÓN DE ENCABEZAMIENTOS A LOS RESÚMENES

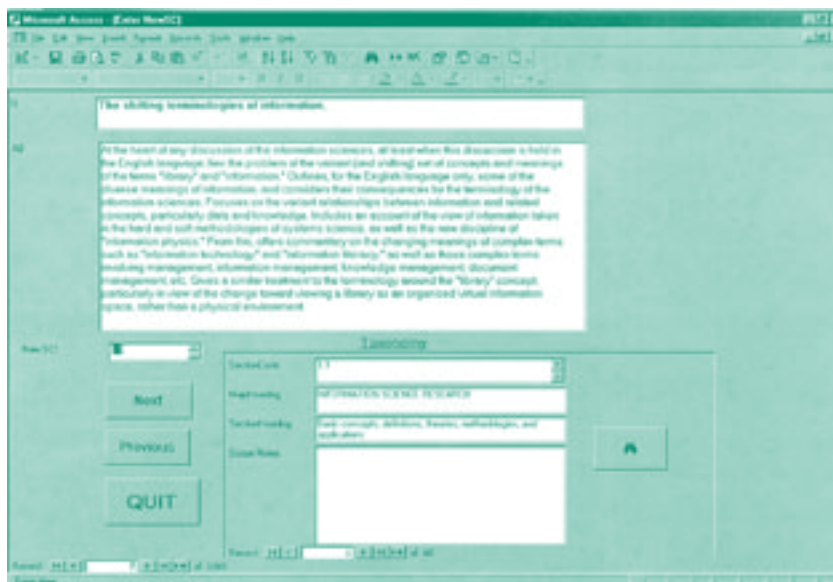


GRÁFICO 5 DISTRIBUCIÓN DE LA CLASIFICACIÓN USANDO LA NUEVA TAXONOMÍA



Este gráfico muestra que los resúmenes se distribuyen en las 11 secciones principales y subsecciones de la taxonomía y que algunas se utilizan con mucha frecuencia y otras muy poco. Examinamos cada una de las subsecciones menos usadas y llegamos a la conclusión de que deberían mantenerse como subsecciones independientes (tabla 2).

TABLA 2
SUBSECCIONES MENOS USADAS

1.8	Investigaciones operativas/matemáticas
6.3	Documentos secundarios
8.4	Sistemas de información geográfica
9.6	Noticias
11.3	Contratos y licencias
11.4	Cuestiones relacionadas con la responsabilidad legal
11.6	Estudios y políticas de información

La mayoría de estas subsecciones representan áreas importantes o en desarrollo de la Documentación o son importantes porque son específicas de esta disciplina, aunque no se publiquen regularmente artículos que traten de ellas. Debemos recordar que la muestra utilizada en este test recoge sólo las referencias aparecidas en los últimos tres meses en ISA. En uno o dos casos, nuestros años de experiencia en el campo influyeron en las decisiones que se tomaron.

Además, estudiamos las posibilidades de subdividir las tres subsecciones más utilizadas (tabla 3).

TABLA 3
SUBSECCIONES MÁS UTILIZADAS

1.4	Investigación en recuperación de la información
1.5	Comportamiento de los usuarios y uso de los sistemas de información
10.3	Automatización, procesos y planificación estratégica de bibliotecas

Apenas nos sorprende que estas subsecciones sean muy utilizadas ya que al representar la esencia de la disciplina contienen muchos artículos que tratan de temas estrechamente relacionados y, por lo tanto, decidimos no subdividir las.

CONCLUSIONES

Los autores de esta investigación son un editor, un bibliotecario de referencia y un resumidor/indizador y, por lo tanto, representan tres importantes categorías

profesionales dentro del campo de la Documentación. Creemos firmemente que la participación de una representación de cada una de estas categorías fue decisiva para el éxito de esta investigación ya que cada uno utilizó la taxonomía desde un punto de vista diferente.

1. Un editor ve la taxonomía como una herramienta para organizar un conjunto de información y dilucidar sus relaciones jerárquicas, yendo de lo general a lo específico.

2. El resumidor/indizador usa la taxonomía para ayudarse a decidir cómo clasificar los artículos en la base de datos. Los indizadores son los que realmente tienen un contacto directo con el sistema de clasificación y pueden opinar con conocimiento de causa sobre la facilidad de su uso, incoherencias, solapamientos y/o lagunas existentes. Son los que, en realidad, poseen la habilidad de identificar y resumir de forma rápida los puntos más relevantes y pertinentes tratados en el artículo, trabajando generalmente contrarreloj.

3. Los bibliotecarios de referencia, las personas que realizan búsquedas en línea y los investigadores ven la taxonomía como puntos de entrada a las bases de datos y no suelen depender completamente de ellos cuando buscan información. Ya que ellos son los usuarios finales de la información o actúan como intermediarios del usuario, no confían en que ni el editor o ni el indizador entiendan generalmente tan bien como ellos los conceptos o los requisitos especiales de sus ecuaciones de búsqueda. Por lo tanto, ven la taxonomía sólo como un componente más de su estrategia de búsqueda e incorporan además otros sinónimos y relaciones que se pueden utilizar en la búsqueda de información. Pese a la posible disparidad de los puntos de vista de los usuarios, hay que tenerlos muy en cuenta durante las fases de diseño y de evaluación de las herramientas que se ponen a su disposición. Este principio es importante sin importar lo que se esté desarrollando y aquellos que lo ignoren corren el riesgo de que sus productos no alcancen el éxito deseado en el mercado.

La taxonomía de Documentación que hemos desarrollado representa una importante mejora sobre la anterior. Refleja la definición moderna de la Documentación, tiene una orientación práctica y facilitará el uso de ISA tanto en su versión impresa como en la electrónica. Es flexible y fácil de mantener y a medida que avance y cambie la tecnología, se pueden añadir fácilmente nuevas secciones y subsecciones además de eliminar las que hayan quedado obsoletas.

Además de usar la taxonomía obtenida con este trabajo en la producción de ISA, sus aplicaciones son muy variadas. Pensamos que, al ser una taxonomía sobre la Documentación, se puede aplicar a *todos* los aspectos de esta disciplina (que como todos sabemos está constantemente cambiando y renovándose y se espera que lo siga haciendo en un futuro próximo), ya que está basada en la práctica real tal y como las publicaciones la reflejan. Nos gustaría hacer hincapié de nuevo en la opinión de Hawkins:

Las publicaciones de los servicios de indización y resumen [...] no sólo mantienen un registro de la historia de la disciplina [...] sino que además pueden servir como “mecanismo de seguimiento” de los cambios. (Hawkins, 2001)

La taxonomía que se presenta aquí usa los conceptos, la definición y el mapa previamente desarrollado de la Documentación para presentarlos de una manera organizada jerárquicamente. Existen otras clasificaciones más limitadas de la Documentación que hemos usado como punto de partida para esta investigación, pero, según nuestro entender, la taxonomía que se presenta aquí es la única que trata directamente de esta disciplina, por lo que representa una contribución importante.

BIBLIOGRAFÍA

- HAWKINS, D.T. (2001). *Information Science Abstracts: Tracking the literature of information science. Part 1: Definition and map. Journal of the American Society for Information Science and Technology*, 52(1), 44-53.
- HAWKINS, D.T., & WAGERS, R. (1982). “Online bibliographic search strategy development”. *Online* 6(3), 12-19.
- LEININGER, K. (2000). “Interindexer consistency in PsycINFO”. *Journal of Librarianship and Information Science*, 32(1), 4-8.
- MILSTEAD, J.L. (Ed.). (1998). “ASIS thesaurus of librarianship and information science”. *Medford, NJ: Information Today, Inc.*
- REICH, P., & BIEVER, E.J. (1991). Indexing consistency: The input/output function of thesauri. *College & Research Libraries*, 52(4), 336-342.
- SIEVERT, M.C., & ANDREWS, M.J. (1991). “Indexing consistency in Information Science Abstracts”. *Journal of the American Society for Information Science*, 42(1), 1-6.

APÉNDICE: TAXONOMÍA DE LA BIBLIOTECONOMÍA

1.	INVESTIGACIÓN EN DOCUMENTACIÓN
1.1	Conceptos básicos, definiciones, teorías, metodologías y aplicaciones
1.2	Propiedades, necesidades, calidad y valor de la información
1.3	Estadísticas, medición Bibliometría, análisis de citas, cientometría, informe tría
1.4	Investigación en recuperación de la información Técnicas de búsqueda (lógica booleana, lógica difusa, lenguaje natural), el proceso de búsqueda, precisión/relevancia, jerarquización de los resultados/exhaustividad de la recuperación, modelos de búsqueda, formulación de la búsqueda, ficheros invertidos, actualizaciones, estructuras de las bases de datos
1.5	Comportamiento de los usuarios y uso de los sistemas de información Tácticas de los usuarios, exceso de información, estudios de usuarios, estudios de usabilidad
1.6	Interfaz hombre-máquina Factores humanos, ergonomía, cuestiones de diseño
1.7	Comunicación Edición, redacción, lingüística, principios del diseño y composición de páginas web
1.8	Investigaciones operativas/matemáticas Creación de modelos, lógica booleana, codificación, análisis de sistemas, algoritmos, compresión de datos
1.9	Historia de la documentación, biografías
2.	ORGANIZACIÓN DEL CONOCIMIENTO
2.1	Tesauros, listas de autoridades Taxonomías, ontologías, redes semánticas, nomenclaturas, terminologías, vocabularios
2.2	Catalogación y clasificación Etiquetado, metaetiquetas, el Dublin Core, los DOIs, los OPACs, el formato MARC, las AACR2, mapas temáticos, procesos y teorías de la catalogación
2.3	Resúmenes, indización, reseñas bibliográficas Resúmenes e indización automatizada
2.4	Normas y protocolos NISO, Z39.5, XML, SGML, HTML, Open Archives Initiative (OAI), Encoded Archival Description (EAD), OpenURL, portable document format (PDF)
3.	LA PROFESIÓN DEL DOCUMENTALISTA
3.1	Los documentalistas Intermediarios, buscadores de información, bibliotecarios de referencia, brokers de información, traductores, educadores, los bibliotecarios y la biblioteconomía, asesoramiento, salidas profesionales, futuro de la profesión, ética profesional, conocimientos y destrezas.
3.2	Organizaciones y sociedades

4.	CUESTIONES SOCIALES
4.1	Ética de la información, el plagio, la credibilidad
4.2	Alfabetización informacional, formación permanente
4.3	La Sociedad de la Información Acceso y accesibilidad universal, impactos tecnológicos y socioeconómicos de la información, predicciones tecnológicas, flujos de la información, previsiones de futuro, preservación
5.	LA INDUSTRIA DE LA INFORMACIÓN
5.1	Gestión del conocimiento y la información Transferencia del conocimiento en las organizaciones, estrategias comerciales
5.2	Mercados y actores Perfiles y entrevistas con proveedores, tendencias
5.3	Aspectos económicos y precios Modelos comerciales, valores añadidos
5.4	Marketing, comercio electrónico
6.	EDICIÓN Y DISTRIBUCIÓN
6.1	Material impreso
6.2	Material electrónico Revistas electrónicas, libros electrónicos
6.3	Documentos secundarios Servicios de indización y resumen, directorios
6.4	Comunicación científica Evaluación por pares, futuro de las revistas, tesis, literatura gris
7.	TECNOLOGÍAS DE LA INFORMACIÓN
7.1	Internet World Wide Web, la Web Invisible, la Deep Web, buscadores, programas cliente, hipermedia, listas de correo, tableros electrónicos, portales, directorios, localizadores
7.2	Intranets, teleconferencia a través de la web
7.3	Software Lenguajes de programación, sistemas operativos, plataformas
7.4	Hardware
7.5	Multimedia
7.6	Gestión de documentos Gestión de imágenes, escaneo, recuperación de textos, digitalización, gestión de documentación administrativa, creación de favoritos, sistemas hipertexto, tecnologías de preservación, enlaces y referencias cruzadas electrónicas, almacenamiento, gestión de los derechos digitales
7.7	IA, sistemas expertos, agentes inteligentes Cybernética, visualización y mapeo, minería de datos, reconocimiento de patrones y caracteres, agentes y robots de búsqueda

7.8	Telecomunicaciones Redes, transmisión de la información por satélite u otros medios inalámbricos, Ordenadores de mano y otros PDA. Las redes LAN y WAN
7.9	Seguridad, control del acceso, autenticación, encriptación Marcas de agua digitales
7.10	Otros
8.	SISTEMAS Y SERVICIOS DE INFORMACIÓN ELECTRÓNICA
8.1	Sistemas y servicios de búsqueda y recuperación de información Bases de datos bibliográficas, numéricas y de imágenes, descripciones de los servicios en línea
8.2	Sistemas de información personalizados, servicios de alerta, servicios de actualización permanente
8.3	Sistemas y servicios de distribución de documentos El préstamo interbibliotecario, recursos compartidos
8.4	Sistemas de información geográfica
9.	FUENTES Y APLICACIONES DE MATERIAS ESPECÍFICAS
	Ciencias físicas Química, física, ingeniería, ciencias de la tierra, ciencias de la computación, energía, matemáticas
9.2	Ciencias de la vida Medicina, ciencias de la vida, agricultura, medio ambiente
9.3	Ciencias sociales, humanidades, historia, lingüística
9.4	Negocios Gestión, económicas, empresas
9.5	Derecho, ciencias políticas, gobierno Patentes y marcas registradas, propiedad intelectual, jurisprudencia
9.6	Noticias
9.7	Educación, biblioteconomía y documentación, referencia
9.8	Otros/multidisciplinar Bases de datos biográficas y genealógicas, enciclopedias, bases de datos de tesis
10.	BIBLIOTECAS Y SERVICIOS BIBLIOTECARIOS
10.1	Tipos y descripciones de bibliotecas Bibliotecas especializadas, del gobierno, académicas y públicas, archivos, museos, bibliotecas nacionales y estatales/regionales, bibliotecas de depósito legal
10.2	Servicios bibliotecarios
10.3	Automatización, procesos y planificación estratégica de bibliotecas
10.4	Consortios y redes de bibliotecas, coaliciones, cooperativas
10.5	Bibliotecas digitales y virtuales, bibliotecas híbridas
10.6	Educación y formación Educación a distancia, formación continua, formación de usuarios, facultades, cursos y planes de estudios de biblioteconomía y documentación

11.	CUESTIONES E INFORMACIÓN RELACIONADAS CON EL DERECHO Y EL GOBIERNO
11.1	Protección de la propiedad intelectual Cuestiones e implicaciones del derecho de autor, el uso razonable, las marcas registradas, leyes sobre patentes
11.2	Legislación, leyes, y normas (excepto el derecho de autor)
11.3	Contratos y licencias
11.4	Cuestiones relacionadas con la responsabilidad legal Filtros, censura, privacidad
11.5	Fuentes de información pública
11.6	Estudios y políticas de información Seguridad, encriptación, privacidad, libertad de información, censura, políticas de información nacionales y de otro tipo.
11.7	Sistemas e infraestructuras Transferencia de la tecnología