

EL ANÁLISIS DOCUMENTAL: INDIZACIÓN Y RESUMEN EN BASES DE DATOS ESPECIALIZADAS

María Cruz Rubio Liniers
CINDOC-CSIC

I. EL ANÁLISIS DOCUMENTAL. DEFINICIÓN Y CONCEPTO

El análisis documental es un trabajo mediante el cual por un proceso intelectual extraemos unas nociones del documento para representarlo y facilitar el acceso a los originales. Analizar, por tanto, es derivar de un documento el conjunto de palabras y símbolos que le sirvan de representación.

En este amplio concepto, el análisis cubre desde la identificación externa o descripción física del documento a través de sus elementos formales como autor, título, editorial, nombre de revista, año de publicación, etc., hasta la descripción conceptual de su contenido o temática, realizada a través de los lenguajes de indización, como palabras clave o descriptores del tesoro. El concepto de indización se identifica con el análisis del contenido en la medida que dichos lenguajes se utilizan para elaborar los índices temáticos por los que se recupera la información.

Muchos y muy diversos pueden ser los métodos de análisis utilizados para representar el documento. La mayor o menor profundidad del análisis dependerá en gran medida de los siguientes aspectos:

- ?? Tipo de usuarios y necesidades de información más o menos especializadas
- ?? Tipos de servicios que se ofrecen. Bibliotecas generales, empresas, Centros de Documentación, etc.
- ?? Documentos a analizar: Libros, artículos de revistas, literatura gris, prensa, legislación.
- ?? Bases de datos y sistemas de recuperación utilizados y sus posibilidades
- ?? Finalidad técnica del análisis: catálogación o recuperación especializada.

Podremos hacer una descripción física basada en las normas ISBD y un análisis de contenido superficial a partir de una clasificación universal o encabezamientos de materia, propio de los sistemas bibliotecarios de catalogación y cuya finalidad primordial es la creación de un inventario de documentos.

Por el contrario, un sistema documental especializado realiza el análisis de contenido basado en la significación o contexto. Es un auténtico análisis en profundidad que implica las técnicas documentales de indización y resumen.

II. LA INDIZACIÓN Y LOS LENGUAJES DOCUMENTALES

Un lenguaje documental es un sistema convencional de signos que permite representar el contenido de los documentos con el fin de encontrar aquellos pertinentes en respuesta a preguntas sobre un tema. Para la mayoría de los especialistas el concepto de lenguaje documental es sinónimo de lenguaje controlado, frente al lenguaje natural que es el que utiliza el discurso científico, técnico o literario. Dicho discurso está cargado de metáforas, sinonimias y hominimias y necesita ser organizado y normalizado para permitir una recuperación eficaz del documento a salvo de los dos grandes problemas documentales: el ruido, -exceso de información encontrada no pertinente,- y el silencio,- ausencia de información pertinente que existe en la base de datos y no es recuperada.

Un título como “ *Las sombras del pasado* ” difícilmente nos orienta sobre el qué, el quién, el cuándo o el dónde. Incluso en otros títulos más precisos la recuperación por unitérminos a partir de las palabras de los títulos crea muchos problemas derivados de la indización con lenguaje natural.

Los principales problemas del lenguaje natural son:

?? **Normalización de singulares y plurales.** Para dicho tema la indización debe ajustarse en lo posible a la Norma UNE 50-106-90. *Normas para el desarrollo de tesauros monolingües.*

Además el significado de muchos términos varía según su número (*Derecho; derechos/Obligación ;Obligaciones*).

?? **Ambigüedad y falta de especificidad.** La utilización de unitérminos, propio del lenguaje natural¹ crea además ambigüedad y falsas combinaciones en la recuperación. Según el ejemplo anterior un buen sistema controlado debería recoger formas como : *Obligación de declarar;Obligaciones del Estado.*

Muchos términos están perfectamente claros y definidos en su utilización. Así, las notas de alcance de un tesoro nos permiten aclarar aquellos conceptos que ayuden a su uso en el contexto exacto. (*Restauración N.A. Reinado de Alfonso XII*). Las posibles homonimias de este concepto pueden solucionarse apellidando el homónimo (*Restauración artística*) Además al apellidar un unitérmino, convirtiéndolo en palabra clave evitamos la ambigüedad y le damos mayor especificidad (*Restauración de la comunidad conyugal*)

?? **Sinonimias.** La utilización de sinónimos, que aporta valor y riqueza desde el punto de vista literario es el mayor problema en la recuperación en texto libre en las bases de datos. No olvidemos que el principio “ideal” de la indización está en el uso del concepto frente a la palabra. Así cada forma debe tener un único significado y cada significado debe ir asociado a una única forma. El tesoro controla las sinonimias mediante los reenvíos:

¹ Mortimer Taube idea en 1958 el sistema de indización por unitérminos en lenguaje natural extraídos de los títulos. Siguiendo este camino Lhun en 1959 crea los índices KWIC y plantea la elaboración de bases de datos a partir de títulos y texto como el CITATION INDEX.

Enfermedad laboral
Usado por Enfermedad profesional

Enfermedad profesional
Use **Enfermedad laboral**

?? **Homonimias o polisemias.** El distinto significado de una misma palabra es una característica de la terminología científica e incluso en las Ciencias Sociales es indicio de escuelas historiográficas o geográficas. (*Edad Moderna (Europa), Encomiendas; Ingenios (Historia de América Latina)*). En muchos casos la precoordinación de términos soluciona las homonimias (*Restauración artística*), en otros es necesaria una nota de alcance e incluso la adscripción de un término a una familia o grupo semántico en el tesoro, le confiere una clara conceptualización.

Un eficaz análisis implica la utilización de vocabularios controlados y procedimientos sintácticos convencionales que permitan crear un vínculo entre el que emite el mensaje y el que lo recibe. Un metalenguaje que armonice pregunta y respuesta.

III. LOS LENGUAJES CONTROLADOS Y LA INDIZACIÓN

Los lenguajes y vocabularios controlados utilizados en la indización han evolucionado desde los primeros lenguajes clasificatorios hasta los actuales tesauros especializados.

Las clasificaciones universales, (CDU, LC, UNESCO) resultan poco pertinentes como único lenguaje de recuperación en bases de datos. Parten de una división del saber hoy superada por la especificidad actual y los “puntos de vista” particulares de las disciplinas científicas. Su estructura jerárquica impide la combinación de los múltiples aspectos de una investigación, y no permite recoger temas muy específicos o novedosos.

Las clasificaciones especializadas, elaboradas para una disciplina o un sistema documental específico, tienen sin embargo utilidad como complementarias a los descriptores. Permiten englobar en un marco amplio todos aquellos documentos de una base de datos a los efectos de una acotación posterior por descriptores. Posibilitan además organizar subproductos de la base de datos, como bibliografías impresas. Dice Maniez ² que para elegir la clasificación más idónea habría que analizar como recoge cada una de ellas un tema o aspecto disciplinar concreto y tratar prácticamente de ajustar un conjunto de documentos recientes a distintas clasificaciones.

Los encabezamientos de materia utilizados tradicionalmente en las bibliotecas como descriptores del contenido de los documentos tienen también una serie de problemas a la hora de la recuperación. Concebidos para la elaboración de catálogos impresos, los mayores inconvenientes en el uso de encabezamientos vienen dados por el exceso de precoordinación de los términos y las normas de presentación en los índices. Dichas normas, excesivamente complejas, da como resultado una falta de criterios

² Maniez, J. *Los lenguajes documentales y de clasificación: concepción, construcción y utilización en los . sistemas documentales*. Madrid: Fundación Germán Sánchez Ruipérez, 1993

homogéneos a la hora de escoger los términos, convirtiendo este lenguaje en un sistema excesivamente “encorsetado” y rígido que impide las combinaciones en la recuperación.

El tesauro es el vocabulario controlado más eficaz para el análisis y la recuperación de la documentación especializada. Sus ventajas fundamentales son: la especificidad de sus términos y sus posibilidades combinatorias en la recuperación. En el encontraremos los términos preferentes de uso, así como los sinónimos rechazados; conceptos más amplios o más específicos de la misma familia semántica y aquellos otros que deben acompañarles en la indización por sus relaciones contextuales. Estos últimos dan un valor añadido a los tesauros frente a otro tipo de vocabularios de indización ya que permiten abordar el análisis del texto desde múltiples puntos de vista, incluso desde el plano contextual de las asociaciones de ideas que cobran sentido dentro de una disciplina o investigación.

La listas de autoridades o de identificadores son vocabularios alfabéticos independientes que permiten normalizar términos tales como nombres propios de personas, instituciones, títulos de obras, nombres geográficos o cualquier otro nombre necesario en la indización y que por múltiples razones no es posible incluir en el tesauro. No olvidemos que según la especialidad de la base de datos, necesitaremos diferentes campos complementarios de contenido.

IV. METODOLOGÍA DE INDIZACIÓN

1. Descripción del documento.

La primera fase de la indización implica la descripción física o análisis formal de los documentos. Se utiliza un lenguaje controlado. Mientras en la catalogación de bibliotecas se utilizan las normas ISBD o ISSD, en las bases de datos especializadas el panorama es muy heterogéneo, mediatizado por el diseño de las bases de datos. Sin embargo, los campos descriptivos fundamentales deberán ser todos aquellos necesarios para la identificación y obtención del documento.

Campos generales para todo tipo de documento

- ?? Número de registro
- ?? Autor/es
- ?? Lugar de trabajo de los autores
- ?? Título
- ?? Lengua de la publicación
- ?? Año de publicación, páginas

Campos específicos según tipo de documento

Monografías	Compilaciones	Artículos de revistas	Actas de Congresos
Lugar de edición	Título del artículo	Título del artículo	Título de la
Editorial	Título colectivo	Nombre de la revista	ponencia
Colección	Compilador o	Volúmen y/o	Nombre del
ISBN	Director	Número	Congreso
Signatura	(además del resto de campos de monografías)	ISSN	Lugar de celebración
			Año de celebración

1. Análisis de contenido: La indización

El análisis de contenido de la información implica la identificación y representación del contenido de los textos mediante dos técnicas fundamentales: la indización y el resumen.

Indizar es extraer una serie de conceptos que responden a los temas tratados en el documento, y que servirán como puntos de acceso para su recuperación.

El analista de la información que se enfrenta a un texto debe realizar un trabajo intelectual de lectura, abstracción, análisis y síntesis. Este ejercicio implica ³ tener en cuenta varios factores o elementos del texto:

- ?? Mensaje del documento o información que ofrece
- ?? La forma que adopta o como presenta la información
- ?? La estructura interna o sistema de relaciones semánticas
- ?? Las aportaciones del documento respecto al corpus científico general

Para un buen análisis de la información es necesario abarcar todos estos factores, desde el conocimiento de la materia, pasando por la comprensión de los principales elementos y relaciones que el autor plantea, hasta la capacidad de discernir la importancia de dicho texto para la comunidad de especialistas. A estos factores hay que añadir la práctica documental consistente en convertir el texto original en un documento secundario: el registro, que permita en cualquier sistema la transferencia de la información mediante acceso por los descriptores seleccionados.

³ García Gutiérrez, A.L. *Estructura lingüística de la documentación: teoría y método*. Murcia: Secretariado de Publicaciones, 1990

Selección de términos de indización.

La selección de términos de indización depende de las siguientes características:

- ?? Exhaustividad versus generalidad : Materia frente a tema.
- ?? Campos semánticos: De qué trata el documento?; ¿quiénes son sus protagonistas?; ¿a qué época se refiere? o ¿a qué lugar?. ¿Cómo se trata la información? y ¿qué tipo de estudio se realiza?.

Los principales tipos de términos de indización podrían ser:

- ☞ Objetos contables: *Libros*
- ☞ Procesos: *Emigración; Conservación*
- ☞ Acciones: *Contratación; Juicio*
- ☞ Conceptos históricos: *Guerra de Cuba; Ilustración*
- ☞ Colectivos: *Aliados; Ordenes religiosas*
- ☞ Documentos {
 - Literarios: *Martín Fierro*
 - Jurídicos: *Constitución de 1812*
 - Artísticos: *Las Meninas*
- ☞ Personajes {
 - Políticos: *Simón Bolívar*
 - Escritores: *Jorge Luis Borges*
 - Instituciones: *Congreso de los Diputados*
- ☞ Cronología {
 - Fechas: *14/2/84*
 - Años: *1898*
 - Siglos: *XX*
- ☞ Topónimo {
 - Países: *Ecuador*
 - U.Suprancionales: *Países Árabes*
 - Provincias: *Buenos Aires*
 - Ciudades: *Santiago de Cuba*
 - Físico-naturales: *Andes*
 - Históricos: *Nueva España*

Características y normas de la indización.

Los principales factores que afectan a una buena indización son:

Especificidad . Nos referimos al nivel de concreción que deben tener los descriptores en relación con el tema del documento. Así un trabajo sobre impuestos concretos como el IRPF o el IVA deberá ser analizado por dichos términos, siendo demasiado genérico la utilización del concepto “*sistema fiscal*”

¿Hasta que punto debe ser específica una indización?. Hasta el mismo que el tema del documento lo exija. Esto no significa que en algunos casos, y siempre que las necesidades de información de los usuarios así lo recomienden, podamos añadir un término genérico. Un trabajo referido a un ámbito geográfico muy concreto puede interesar a especialistas de su región, provincia o país. La investigación sobre la población de Ecija, debe llevar los términos de Sevilla (Provincia) e incluso Andalucía, para facilitar la recuperación. Sin embargo hay que procurar no caer en

la redundancia. Una buena especificidad se conseguirá además en la recuperación combinando los términos.

Exhaustividad. Relacionado con el concepto anterior, la exhaustividad implica sobre todo un número alto de descriptores. Es difícil valorar este aspecto ya que, aunque una gran cantidad de entradas puede facilitar la recuperación de un documento, muchas veces este factor entra en contradicción con la pertinencia del documento hallado.

Pertinencia. Es el elemento fundamental a tener en cuenta para una buena indización. La pertinencia de un descriptor representa su grado de adecuación al concepto expresado en el texto original. Un buen descriptor debe ser siempre un punto preciso y fiel de acceso y responder a las expectativas de búsqueda del usuario. El indizador debe siempre ponerse en el lugar del especialista que cuando busca por ese término que hemos utilizado va a encontrarse con el documento que se está analizando. Imaginemos un artículo que trata sobre el modo de vida de los indios araucas y en el que incluimos un descriptor como sociedad. El que busca información sobre la sociedad latinoamericana, ¿busca realmente este tipo de trabajos?. ¿O le interesan realmente otros más generales?.

Coherencia. A pesar del uso de un buen vocabulario controlado, la indización como todo trabajo intelectual tiene una alta dosis de subjetivismo. Dos o más personas pueden disentir de cuáles son los temas importantes de un trabajo o de los descriptores que deben representarlos. Incluso, una misma persona puede tener criterios diferentes en distintos momentos. La coherencia⁴ en la indización se refiere al grado de concordancia entre indizadores (coherencia interindizador) o de uno mismo (coherencia intraindizador). Una medida de coherencia puede hallarse a través del promedio de descriptores comunes. La coherencia depende del número de descriptores asignados, el vocabulario utilizado, la terminología del documento y en general a las capacidades de conceptualización de los indizadores

Factores que afectan a la calidad de la indización

A) Características del indizador.

- ?? Conocimiento de la materia y las peculiaridades de su discurso.
- ?? Conocimiento de la forma de búsqueda y necesidades de los usuarios.
- ?? Capacidad de lectura comprensiva.
- ?? Capacidad de análisis y síntesis.

B) Características del documento.

- ?? Complejidad.
- ?? Claridad, estructura y presentación del texto.
- ?? Existencia de elementos informativos:
 - . Título pertinente.

⁴ Lancaster, F.W. *Indización y resúmenes. Teoría y práctica*. Buenos Aires:EB Publicaciones, 1996

- . Capítulos o epígrafes.
- . Resumen.

C) Características del proceso de análisis:

?? Selección de descriptores.

- . ¿Generales o específicos?. Tan concretos como aparezcan en el documento original.
- . ¿Número de descriptores?. Tantos como sean pertinentes y aspectos toque el documento original.
- . ¿Cuáles?. Todos aquellos que sean pertinentes para el usuario. No solo los explícitos sino también los implícitos como el método utilizado, tipo de análisis efectuado, estado de la cuestión, etc.

?? Grado de precoordinación de los términos.

En este punto conviene aclarar algunos aspectos respecto a la forma de coordinación de los términos de indización. Pueden utilizarse como unitérminos aquellos descriptores cuya forma simple es clara y precisa (P.Ej. *Catedrales*). Los términos compuestos por necesidades de desambigüedad o especificidad podrán formarse de:

1. Sustantivo + adjetivo (*Plantas medicinales*)
2. Sustantivo + sintagma preposicional (*Hospitales para niños*)
3. Sintagma aposicional (*Relaciones Norte-Sur*)

Son formas gramaticales prohibidas los adjetivos, adverbios, verbos no sustantivados. En cualquier caso, la precoordinación de los términos en la indización o la postcoordinación en la recuperación dependerá siempre del tesoro utilizado.

?? Campos de indización de la base de datos.

La estructura y número de los campos de contenido de una base de datos mediatizan en gran medida el éxito de una buena recuperación, pudiendo ser tantos como los responsables del análisis y la recuperación consideren necesarios según la disciplina o especialidad que “vacíen”.

Entre otros podemos enumerar los siguientes:

- ?? **Descriptores.** Recoge todos los conceptos del tesoro
- ?? **Identificadores.** Recoge los nombres propios de personas, instituciones con sede, títulos de obras etc.
- ?? **Topónimos.** Nombres de lugares geográficos
- ?? **Legislación.** Para bases de datos jurídicas permite incluir leyes, decretos, órdenes etc.
- ?? **Período histórico.** Para bases de datos históricas. Irán en este campo fechas, años, siglos etc.

Un mayor número de campos de contenido necesitará usuarios más expertos en el manejo e implicará un mayor conocimiento de la estructura de la base de datos, por lo que en cada caso habrá siempre que sopesar las ventajas y las necesidades reales de recuperación.

Por último, un buen servicio de indización debe tener normas y manuales de uso para sus analistas que garanticen la homogeneidad en la indización. Estos instrumentos facilitarán la coherencia.

Ejemplo de indización de artículo de revista

Nº Registro	402005
Autores	Moisset de Españés, Luis
Título del artículo	La mora en el Código Civil de Paraguay de 1987
Lugar de trabajo	Acad. Nal. Derecho de Córdoba, Argentina
ISSN	0034-7922
Título de la revista	Revista de Derecho Privado
Datos fuente	1999, 83 (12): 827-847, 10 Ref.
Tipo de documento	AR
Lengua	ES
Clasificación	250270;850420 (Derecho de obligaciones);(Derecho de los Países latinoamericanos)
Descriptores	Mora;Incumplimiento contractual;Incumplimiento de obligaciones;Obligación a plazo;Mora del acreedor;Derecho comparado
Identificadores	García Goyena, Florencio
Topónimos	Paraguay
Legislación	Código Civil PAR. Art. 424
Período Histórico	1987

V . EL RESUMEN

V. 1. Concepto y finalidad

El resumen es una representación abreviada y precisa del contenido de un documento, sin interpretación crítica y sin mención del autor del documento (ISO).

Mientras que la indización tiene como finalidad crear índices para recuperar la información por cada uno de los conceptos de los que trata el documento, el resumen es una síntesis del tema o representación abreviada del texto. Es un trabajo intelectual de condensación. Así como la indización se realiza en lenguaje controlado el resumen se elabora en lenguaje natural.

V. 2. La finalidad del resumen

A) El resumen en el documento original

?? **Conocimiento del tema y de su interés para los usuarios.** La gran cantidad de información existente sobre cualquier tema o disciplina, unido a la cada vez mayor especialización de los usuarios, hace necesaria una labor previa de selección, antes de decidirse a solicitar el préstamo o fotocopia de un documento. Esto es aún más necesario en aquellos documentos que por su extensión o coste de localización y copia, (las tesis doctorales por ejemplo) obligan a estar seguros de su adecuación a nuestras necesidades. Acompañando a los documentos originales el resumen

permite identificar de forma rápida su tema y alcance, a fin de decidir sobre la conveniencia o no de su petición. Además contribuyen a superar las barreras lingüísticas y permite una puesta al día de los que desconocen el idioma del texto, así como la difusión de trabajos en lenguas minoritarias.

?? **Facilitar la indización.** Un buen resumen, tiene una gran cantidad de términos significativos que ayudarán al indizador a encontrar los descriptores más idóneos y representativos.

B) El resumen en las Bases de Datos

?? **Recuperación en bases de datos a partir del resumen.** La utilidad de este aspecto está siendo muy debatido entre los documentalistas y usuarios. No cabe duda que la recuperación en el campo de resumen en una base de datos, tiene los problemas de acceso en lenguaje libre, por lo que el “ruido” es altamente previsible. El resumen debe ir complementado por los descriptores para hacer más eficaz la búsqueda. Y en este caso existen distintas opiniones sobre si los términos del resumen deben repetir los descriptores, lo que sin duda parece redundante, o más bien aprovechar las ventajas de la recuperación en texto libre para introducir en el resumen aquellos sinónimos u otros términos no permitidos en el vocabulario controlado, para enriquecer y hacer más eficaz la recuperación. Nombres propios o aspectos más difíciles de representar mediante los descriptores (aspectos colaterales, método utilizado, consecuencias o conclusiones etc.)

?? **TIPOLOGIA DE LOS RESUMENES**

Según el autor del resumen:

Del autor del documento. En este sentido el resumen es también parte del documento primario como información sobre el tema. Las normas de elaboración de resúmenes (UNE 50-103) están dirigidas principalmente a los resúmenes de autor. Sus ventajas son el bajo costo y el conocimiento del tema pero muchas veces se desconocen las técnicas y normas documentales para su elaboración por lo que suelen ser los más ineficaces para la recuperación en bases de datos.

De un especialista en análisis documental. Los resúmenes elaborados en los servicios de indización por los documentalistas, tienen la ventaja de estar mejor estructurados y preparados para una mejor recuperación en las bases de datos.

Según el tamaño del resumen

Telegráfico. Es un título enriquecido, con palabras relevantes.

Indicativo. Suele tener una extensión entre 50 y 150 palabras. Y recoge el tema de forma global, la información esencial sin entrar en detalles.

Informativo. De mayor amplitud, entre 150 y 500 palabras, recoge todos los elementos informativos del documento original. Puede sustituir al original, como información.

Crítico. Este resumen se identifica más con la reseña, ya que contiene una valoración de la calidad y aportación del documento. Desde un punto de vista documental se desaconseja realizar este tipo de resúmenes.

METODOLOGIA DE ELABORACIÓN DE RESÚMENES

- 1) **Análisis formal.** En los servicios de elaboración de resúmenes, tanto para índices impresos como para bases de datos es necesario recoger en primer lugar la referencia bibliográfica según las normas ISO 690.
Los elementos a recoger serán : N° de identificación del documento, autor/es, título, y datos de publicación.

- 2) **Análisis de contenido.**
 - ?? Lectura comprensiva del texto para identificar las ideas principales, los elementos más significativos y la estructura del texto, fijándonos principalmente en aquellas partes
 - ?? Extraer aquellas ideas, conceptos más relevantes y con mayor contenido de información, deteniéndonos en las partes del trabajo como introducción, objetivos y conclusiones.

 - ?? Cuestionarse el tema del documento y cual fue la finalidad del autor al escribirlo.

 - ?? Redacción del texto a través de las frases más significativas encontradas. Aquí interviene los conocimientos de redacción y estructura gramatical así como la capacidad de síntesis del analista.

 - ?? Dar la mayor cantidad de información pertinente en el menor espacio es la principal finalidad de un buen resumen. Pero un conjunto de ideas claras no tiene que dar necesariamente un conjunto claro. La capacidad de síntesis consiste en crear un nuevo documento más conciso y claro que el original, con significado pleno en sí mismo y con estructura gramatical, buena sintaxis y coherencia semántica

 - ?? Es recomendable realizar un borrador previo antes de darle la forma final.

NORMAS PARA LA ELABORACION DE RESUMENES

PRESENTACIÓN

- ?? Un solo párrafo en resúmenes indicativos
- ?? Dos o tres párrafos en los resúmenes informativos
- ?? Puede iniciarse con un resumen del resumen: Una frase significativa del tema del documento
- ?? En resúmenes informativos conviene respetar el orden de la argumentación del autor

ESTILO

- ?? Máxima información con la mínima extensión. Ello implica un estilo con “nervio”, conciso y de oraciones cortas.
- ?? No debe incluirse ninguna palabra innecesaria y de deben evitar las ambigüedades.
- ?? Nunca deben repetirse palabras.
- ?? Evitar frases largas y complicadas, con incisos y perifrasias.
- ?? Conviene utilizar enumeraciones separadas por comas.
- ?? Uso de verbos en voz activa y tercera persona del presente de indicativo.
- ?? Debe emplearse una terminología adecuada al texto, huyendo de tecnicismos, acrónimos o siglas de difícil comprensión.

LOS RESUMENES COMO ELEMENTO DE RECUPERACIÓN

Los conceptos que se deben incluir en un resumen para mejorar la recuperación se pueden seleccionar en coordinación con los descriptores o se pueden añadir como elementos nuevos. En cualquiera de ambos casos los resumidores pueden utilizar listas de términos normalizados para nombres propios, siglas, etc. No hay que olvidar que la utilización de puntos entre siglas puede impedir una buena recuperación, que la utilización de la forma negativa puede dar resultados erróneos o que es conveniente incluir el nombre y adjetivo en cada una en las iteraciones o listas enumerativas (*documento primario, documento secundario y documento terciario*, mejor que: *documento primario, secundario y terciario*). Insistimos en la posibilidad de utilizar variedad de formas de un mismo concepto, siempre que se considere útil para la recuperación.

A la hora de elaborar resúmenes con vistas a un tratamiento automático, no debe olvidarse la importancia que tienen los programas de gestión y recuperación de las bases de datos. El control terminológico a partir de una lista de palabras vacías aplicables en cada caso a la información contenida o a la frecuencia de aparición y rango y la aplicación de los avances en lingüística computacional, harán más efectiva la recuperación a texto completo.

BIBLIOGRAFÍA

- ?? **AITCHISON, J.** Indexing languages: Classification schemes and thesauri. In: Anthony, L.J. (ed.). *Handbook of special librarianship and information work*. London: ASLIB, 1982.
- ?? **ASCHERO, B.** *Teoria y tecnica dell'indicizzazione per soggetto*. Milano: Editrice Bibliografica, 1993.
- ?? **AUSTIN, D.** *Precis: a manual of concept analysis and subject indexing*. London: British Library, 1984.
- ?? **BARDIN, L.:** *Análisis del contenido*. Madrid: AKAL, 1986.

- ?? **BELY, N.** *Procedures d'analyse semantique appliquées a la documentation scientifique*. París: Gauthier-Villars, 1970.
- ?? **BENEDITO CASTELLOTE, P.** Clasificación e indización en bibliotecas especiales. *Boletín de la ANABAD* (Madrid) 1994, vol. 44, nº 1, p. 69-80.
- ?? **BONVILLE, J.** *Manuel d'analyse et d'indexation*. Quebec: Universite Laval, 1983.
- ?? **BORKO, H.** *Indexing concepts and methods*. New York: Academic Press, 1978
- ?? **BROOKS, T.A.** Todos los descriptores son válidos: prueba de la estrategia de indización de un documento mediante la incorporación de los descriptores asignados por distintos indizadores. *Journal of the American Society for Information Science*, 1993, vol. 44, nº 3, p. 137-147.
- ?? **CARRERAS, C. y SEGUI, M.R.** Reflexions sobre la classificació i la indexació dels documents. *Item. Revista de Biblioteconomía i documentació*, 1987, nº 1, p. 37-43.
- ?? **CARRIÓN GUTIEZ, M** El catálogo alfabético de materias. En: *Manual de Bibliotecas*. Madrid: Fundación Germán Sánchez Ruipérez-Pirámide, 1987.
- ?? **CLEVELAND, D. B.** *Introduction to indexing and abstracting*. Colorado: Libraries Unlimited, 1983.
- ?? **CUNHA, I. M. R.** Análise documentaria. En: *Análise documentaria: a análise da síntese*. J. Smit (coord.). Brasilia: IBICT, 1987, p. 37-60.
- ?? **CUNHA, I.M. R.** Estructuração de vocabulário. En: *Análise documentaria: a análise da síntese*. J. Smit (coord.) Brasilia: IBICT, 1987, p. 61-85.
- ?? **CHAN, L. M.; RICHMOND, P. A. y SVENONIUS, E.** *Theory of subject analysis: a sourcebook*. Littleton (Colorado): Libraries Unlimited, 1985.
- ?? **CHAUMIER, J.** *L'analyse documentaire*. París: Enterprise Moderne d'Edition, 1977.
- ?? **CHAUMIER, J.** *Les langages documentaires*. París: Enterprise Moderne d'Edition, 1978.
- ?? **CHAUMIER, J.** *Análisis y lenguajes documentales. El tratamiento lingüístico de la información documental*. Barcelona: Mitre, 1986.
- ?? **FOSKETT, A.C.** *Classification and indexing in the social sciences*. London: Butterworth, 1974.
- ?? **FOSKETT, A.C.** *The subject approach to information*. London: Clive Bingley, 1982.
- ?? **GARCÍA GUTIÉRREZ, A.L.** *Lingüística documental*. Barcelona: Mitre, 1984.
- ?? **GARCÍA GUTIÉRREZ, A.L.** Suficiencia estructural y tipología de la omisión en análisis documental. *Documentación de las Ciencias de la Información*, 1990, vol. 13, p. 73-86.
- ?? **GARCÍA GUTIÉRREZ, A.L. y LUCAS FERNÁNDEZ, R.** *Documentación automatizada en los medios informativos*. Madrid: Paraninfo, 1987.

- ?? **GARDIN, J. C.** Document analysis and Linguistic Theory. *Journal of Documentation*, 1973, vol. 29 , nº 2, p. 137-168.
- ?? **GIL LEYVA, I.** La automatización de la indización de documentos. Gijón: Trea, 1999
- ?? **GIL URDICIAIN, B.** *Manual de lenguajes documentales*. Madrid: Noesis, 1996.
- ?? **LANCASTER, F.W.** *Indexing and abstracting in theory and practice*. London: Library Association, 1991.
- ?? **LANGRIDGE, D.W.** *Classification: its kinds, elements, systems and applications*. London: Bowker-Saur, 1992.
- ?? **LONG, B.** Linguistique et indexation. *Documentaliste. Sciences de L'Information*, 1980, vol. 17, nº 3, p. 99-106.
- ?? **MANIEZ, J.** *Los lenguajes documentales y de clasificación: concepción, construcción y utilización en los sistemas documentales*. Madrid: Fundación Germán Sánchez Ruipérez, 1993.
- ?? **NEET, H.E.** *A la recherche du mot clé: Analyse documentaire et Indexation alphabétique*. Geneve: Institut d'Etudes Sociales, 1989.
- ?? **OLMEDO, M.I. (et al.)** Bibliografía comentada sobre lenguajes de indización y automatización. *Revista Española de Documentación Científica*, 1987, vol. 10, nº 4, p. 453-471.
- ?? **PALMA VILLALON, Mª V.** Técnicas y métodos para mejorar la calidad de la indización y su recuperación en bases de datos documentales de CC.SS. y Humanidades. *Jornades Catalanes de Documentacio, biblioteques, centres de documentacio i serveis d'informacio*, Barcelona, 1995.
- ?? **PINTO MOLINA, M.** Fuentes para el estudio de la indización documental: Bibliografía selectiva anotada. *Revista Española de Documentación Científica*, 1988, vol. 11, nº 2, p. 177-239.
- ?? **PINTO MOLINA, M.** Introducción al análisis documental y sus niveles: el análisis de contenido. *Boletín de la ANABAD*, 1989, vol. 39, nº 2, p. 323-341.
- ?? **PINTO MOLINA, M.** *Análisis documental: fundamentos y procedimientos*. Madrid: Eudema, 1991.
- ?? **.RICHTER, N.** *Les langages documentaires encyclopediques: Guide pratique d'Indexation* . Marigné: Editions de la Queu du Chat, 1990.
- ?? **ROPER, V. de P.** *Análisis documental: nuevas tendencias*. Madrid: CINDOC, 1987
- ?? **ROWLEY, K.** *Abstracting and Indexing*. Londres: Clive Bingley, 1988.
- ?? **ROY, R.** *Introduction a l'indexation documentaire pour les candidats aux examens et concours des bibliotheques*. Le Mans: Biblioteque de l'Université du Maine, 1985.

- ?? **ROY, R.** *Classer and indexer: introduction a l' ndexation documentaire*. Le Mans: Bibliotheque de l'Université du Maine, 1987.
- ?? **RUIZ PÉREZ, R.** El análisis documental como disciplina docente en la formación de profesionales de la Documentación en España. *Ciencias de la Documentación*, 1990, vol. 1, p. 67-84.
- ?? **SLYPE, G. VAN.** Indexing languages. Classification schemes and thesauri. En: *Handbook of special librarianship and information work*, L.J. Anthony (ed). London: ASLIB, 1982, p. 207-261.
- ?? **SLYPE, G. VAN.** *Les languages d'indexation: conception, construction et utilisation dans les systemes documentaires*. París: Les Editions d'Organisation, 1987.
- ?? **VALE, E. A.** *Linguagens de indexação. Análise documentaria: a análise da síntese*. Brasilia: IBICT, 1987.
- ?? **VALLE BRACERO, A. y FERNÁNDEZ GARCÍA, J.A.** Automatización de la indización y coordinación de descriptores. *Revista Española de Documentación Científica*, 1983, vol. 6, nº 1, p. 9-16.

EL RESUMEN DOCUMENTAL

- ?? **CREMMINS, E.T.** El arte de resumir. Barcelona: Mitre, 1982. ISBN 84-86153-44-1
- ?? **MOREIRO GONZALEZ, J.A.** Los resúmenes de autor: su aplicación y su lugar de comunicación científica. *Cuadernos de Documentación de Cajas de Ahorros*, 1988, 3 (7-8), 33-56
- ?? **MOREIRO GONZALEZ, J.A.** Aplicación de las ciencias del texto al resumen documental . Madrid: Univ. Carlos III, 1993
- ?? **PINTO MOLINA, M.** El resumen documental: principios y métodos. Madrid: Fundación Germán Sánchez Ruipérez, 1992. ISBN. 84-86168-70-8
- ?? **RAYA, FIDEL.** Hacer resúmenes para la búsqueda en texto libre. *Boletín de la Asociación andaluza de Bibliotecarios*, 1990, (20), 25-35
- ?? **SANZ SANZ, I.** La creación de un resumen. Fases principales. *Boletín de la Asociación andaluza de Bibliotecarios* , 1991 (24), 51-55