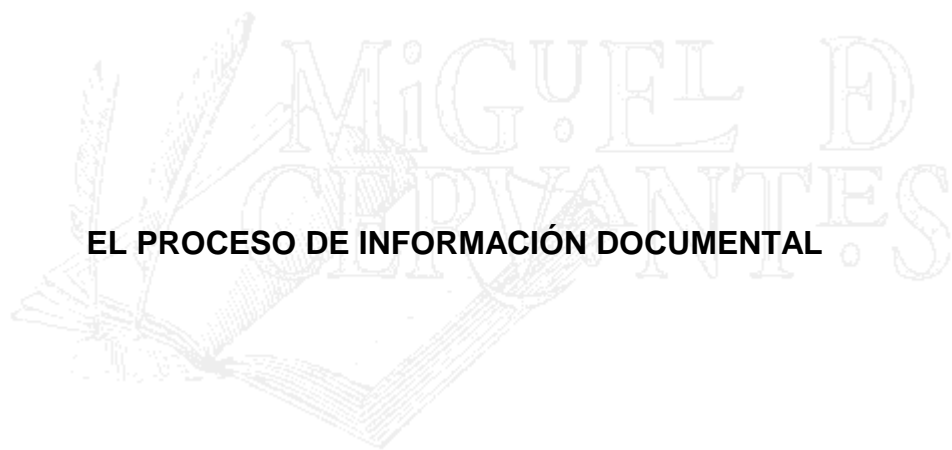


BIBLIOTECA VIRTUAL



EL PROCESO DE INFORMACIÓN DOCUMENTAL

EL PROCESO DE INFORMACIÓN DOCUMENTAL

2. INTRODUCCIÓN

La información contenida en el documento debe someterse a una serie de operaciones con el fin de que aquélla llegue a cuantos quieran utilizarla. En palabras de López Yepes⁸⁴ “el documento, aun siendo información vehiculada, es de suyo realidad estática si no se arbitran los recursos necesarios para su oportuna potenciación en la transmisión y difusión. Para que ello tenga lugar debe agilizarse un procedimiento, un proceso documental que posibilite la dinamización de la información guardada en el documento, por medio de su difusión”. Es decir, para permitir la difusión de la información documental y facilitar su accesibilidad óptima, el documento sufre un proceso compuesto por un conjunto de fases sucesivas, articuladas entre ellas. Este hecho ha motivado que algunos autores, entre ellos J. Chaumier y N. Amat, hayan denominado a este proceso *Cadena Documental*⁸⁵, porque dichas operaciones se relacionan entre sí de forma que las últimas dependen de las que preceden, como eslabones de una cadena.

Si partimos de la consideración teórica de que la Documentación es una ciencia informativa, pensamiento ya expresado por Otlet⁸⁶, el proceso documental se convierte en un proceso informativo cuyo esquema viene a coincidir con los elementos propios de la actividad documental (*emisor*: documentalista; *mensaje*: información documental;

⁸⁴ LÓPEZ YEPES, José et al. *Estudios de documentación general e informativa*. Madrid: Seminario Millares Carlo, 1981, p. 126.

⁸⁵ CHAUMIER, J. *Techniques documentaires*. París: PUF, 1979, p. 11 y AMAT, N. *Documentación científica y nuevas tecnologías de la información*. 2ª ed. Madrid: Pirámide: 1988, p. 21.

receptor: usuario; *medio*: documento). En este sentido, Martínez Comeche⁸⁷ sugiere que el análisis teórico de la actividad documental se sustente sobre tres rasgos principales:

- Los procesos documentales implican actos comunicativos, pues involucran sujetos -emisores y receptores- que intercambian mensajes entre sí.
- Puesto que la obtención de la información es el resultado esperable de la actividad documental, parece razonable que la acción que la origina constituye un proceso informativo.
- El tratamiento documental conlleva una puesta en forma especial de los mensajes que maneja, lo que implica que la actividad documental desarrolla no uno, sino varios procesos informativos.

Desde esta perspectiva el mensaje se constituye como el elemento más importante del proceso documental. Desantes Guanter⁸⁸ distingue varios tipos de mensajes documentarios en el proceso documental:

- a) *Mensaje documentado*: mensaje incorporado, analizado y conservado formando parte del documento y cuyo objetivo es su disposición para ser comunicado (difundido).
- b) *Mensaje referencial*: es el nuevo mensaje resultante de la referencia del mensaje documentado que se forma también en un nuevo documento.
- c) *Mensaje documental*: aquel mensaje documentado que pasa de ser difundible a difundido. Este mensaje viene así a cumplir la finalidad misma de la documentación: que la información llegue al usuario.

⁸⁶ OTLET, P. *Traité de Documentation*. Bruxelles: Mundaneum, 1934, pp. 216-217.

⁸⁷ MARTÍNEZ COMECHE, J.A. El proceso informativo-documental. En: *Manual de Información y Documentación*. Madrid: Pirámide, 1996, pp. 37-38.

⁸⁸ DESANTES GUANTER, José María. *Teoría y régimen jurídico de la documentación*. Madrid: Eudema, D.L. 1987, pp. 236-258.

d) *Mensaje marginal*: no puede tratarse como mensaje independiente desde el punto de vista de su tratamiento documental ya que está al servicio del mensaje documentado al que se refiere. Es un mensaje que presenta datos o anotaciones al mensaje documentado.

2.1. EL PROCESO DOCUMENTAL

La finalidad perseguida por todo sistema de información documental es el de conseguir que la información contenida en los documentos se difunda. Para cumplir este objetivo y lograr, pues, una correcta comunicación entre documentos y usuarios se hacen necesarias una tareas documentales, un *proceso documental*⁸⁹ que se concreta en las siguientes funciones:

- **Input o Entrada** de los documentos al sistema que permite crear y actualizar el fondo documental. Esta primera fase engloba las tareas siguientes:

Selección

Consiste en escoger, de entre toda la masa documental existente, aquellos documentos que se adecuen a las necesidades informativas de los clientes potenciales y de las disciplinas temáticas abarcadas por el sistema documental. Todo sistema de información debe elaborar una política de selección que garantice un alto grado de control de calidad y cantidad de los documentos que entran a formar parte del fondo documental. Así los criterios de selección vienen determinados por:

- Los objetivos del sistema

⁸⁹ PINTO MOLINA, M. *Análisis documental: fundamentos y procedimientos*. 2ª ed. Madrid: Eudema, 1993, p. 40.

- Las necesidades de los usuarios
- Y, los recursos económicos/humanos con los que cuenta el Servicio.

Adquisición

La adquisición permite constituir el fondo documental de un sistema de información. Las formas habituales de adquirir fondos son la compra, el intercambio o canje y la donación.

Registro

El registro es una tarea meramente administrativa encaminada a controlar el número total de documentos que constituyen el fondo de un Servicio de Documentación y, también, a fijar el orden de entrada de los mismos con una doble finalidad: controlar las pérdidas y servir de inventario económico.

- **Tratamiento.** De todas las fases del proceso documental, el *Tratamiento documental*, tal y como lo entiende María Pinto, ocupa un lugar destacado, el cual comprende dos operaciones interrelacionadas como son el *Análisis* y la *Búsqueda* o *Recuperación*. La importancia de esta fase reside en que la información contenida en los documentos es analizada y sintetizada a fin de posibilitar su control y, consecuentemente, su recuperación. Es por ello que ambas técnicas se configuran como dos eslabones de una misma cadena para la consecución de un fin: que el documento y el usuario se comuniquen.

La doble naturaleza del documento (soporte y mensaje) implica que el análisis documental se realice atendiendo a esta dicotomía. Así pues el análisis se desdobra en: *Análisis formal* o externo que se ocupa de los aspectos físicos y/o formales del documento, y *Análisis del*

contenido o interno, el cual se efectúa sobre el mensaje o contenido del documento.

- **Output o Salida.** El desenlace de todo proceso documental desemboca en la *difusión* de la información contenida en los documentos que se realizará mediante productos y servicios. Esta operación es la verdadera razón de ser de la Documentación: satisfacer las necesidades informativas de los usuarios.

2.2. TÉCNICAS DE ANÁLISIS DOCUMENTAL

El *Tratamiento Documental* ocupa un puesto relevante en el conjunto de actividades que se llevan a cabo en cualquier sistema de información, ya que la difusión se verá condicionada en gran medida por las dos fases que conforman esta función: el *Análisis Documental* y la *Recuperación*. Es por ello que ambas operaciones han sido y siguen siendo objetos de estudio por parte de los científicos de la información.

El *Análisis Documental* está constituido por una serie de operaciones que vienen a transformar el documento original, tanto en la forma como en el contenido, para facilitar y posibilitar la operación de recuperación de los documentos y, así, su difusión. A pesar de que algunos estudiosos -Gardin, Gayaud, Courier, Chaumier⁹⁰- conciben el análisis documental sólo en lo que respecta a las operaciones enfocadas a representar el contenido de los documentos (*indización y operación de resumir*), nosotros nos adherimos a los criterios propuestos por Foundin⁹¹, López Yepes⁹² y Pinto Molina⁹³, quienes

⁹⁰ Cit. por PINTO MOLINA, M. El análisis documental. En: *Fundamentos de Información y Documentación*. 2ª de. Madrid: Eudema, 1990, p. 273.

⁹¹ FOUNDIN, H. La structure et le vocabulaire de l'analyse documentaire. Contribution pour une mise au point. *Documentaliste*, 1977, vol. 14, nº 2, p. 12.

consideran que el análisis debe englobar el estudio de los dos elementos que conforman el documento: su forma (*Análisis Formal*) y su contenido (*Análisis de Contenido*).

Los objetivos esenciales que se persiguen con el Análisis Documental son, según Garrido Arilla⁹⁴, los siguientes:

- Posibilitar la elaboración de documentos secundarios a partir de los primarios.
- Posibilitar la Recuperación Documental.
- Es un instrumento de trabajo para el documentalista en favor del usuario, para lo cual ese profesional de Análisis Documental deberá elaborar un trabajo objetivo y normalizado.

Estos planteamientos del Análisis Documental se están viendo muy favorecidos con la aplicación de los medios informáticos y de telecomunicación que facilitan en gran medida la elaboración racionalizada de productos documentales secundarios, facilitando al usuario una rápida y selectiva recuperación de la información científica ante la ingente cantidad de información producida diariamente. De esta manera, el Análisis Documental actúa de filtro, desechando lo irrelevante y mostrando sólo aquellos datos precisos de identificación y de uso directo por parte del usuario.

2.2.1. EL ANÁLISIS FORMAL

⁹² LÓPEZ YEPES, J. *Nuevos estudios de documentación: el proceso documental en las ciencias de la comunicación social*. Madrid: Instituto Nacional de Publicidad, 1978, p. 112.

⁹³ PINTO MOLINA, M. *Op.Cit*, p. 82.

⁹⁴ GARRIDO ARILLA, M.R. Reflexiones sobre la problemática y estado actual del Análisis Documental. *Documentación de las Ciencias de la Información*, 1994, nº 17, pp. 73-74.

La operación de *Análisis Formal* tiene como objetivo la identificación física del documento. Este análisis se realiza previamente al análisis de contenido y se rige por una serie de instrucciones o normas convencionales, por países o universalmente aceptadas, para cada tipo de documento según el soporte físico que los sustenta. El campo de actuación del análisis de la forma se concreta en dos técnicas documentales: la *descripción bibliográfica* o *descripción documental*⁹⁵ y la *catalogación*.

2.2.1.1. LA DESCRIPCIÓN BIBLIOGRÁFICA

La descripción bibliográfica es la primera operación del análisis de la forma que permite la identificación unívoca del documento para que, posteriormente, pueda ser registrado, clasificado, almacenado y, consecuentemente, localizado dentro de una colección. El producto resultante de esta operación es la *referencia bibliográfica*, definida por la Organización Internacional de Normalización (ISO) como “un conjunto de indicaciones precisas y detalladas, suficientes para permitir la identificación de una publicación o parte de ella”⁹⁶. Esta referencia se plasmará físicamente en una ficha de cartón (cuyas dimensiones normalizadas atienden a las medidas de 12,5 x 7,5 cm) o en un soporte legible por ordenador.

⁹⁵ Otros autores abogan por descripción documental, término considerado más amplio ya que no se limitaría sólo a los documentos impresos, sino que comprendería también aquellos documentos procedentes de los nuevos avances tecnológicos (CD-ROMs, disquettes, vídeos, etc). No obstante, nosotros mantenemos el término clásico de descripción bibliográfica ya que los documentos que nos competen son documentos escritos o textuales.

⁹⁶ ISO. 690-1987. Documentation -Bibliographic References- Content, form and structure. En: *Documentation and Information*. Genève: ISO, 1988, pp. 436-447.

Según indican Guinchat y Menou⁹⁷, el procedimiento de la descripción bibliográfica consta de las siguientes etapas:

1. Conocimiento de la existencia del documento.
2. Determinación del tipo de documento y de las reglas aplicables al caso particular.
3. Determinación del nivel bibliográfico que conviene al tratamiento.
4. Identificación, para cada nivel bibliográfico, de los datos necesarios, siguiendo el orden de las áreas que indique la norma o el formato utilizado.
5. Transcripción de los datos anteriores según las reglas indicadas por la norma o e formato.
6. Verificación de la exactitud de la descripción y de su conformidad con las normas.
7. Entrega de la reseña para su elaboración definitiva o para su tratamiento posterior.

Estas fases corresponden al sistema manual. Pero en el caso que se utilicen los sistemas automatizados las decisiones se proceden generalmente en dos etapas. Se recoge primero en un formato de entrada y se graba luego en un soporte legible por ordenador.

2.2.1.2. LA CATALOGACIÓN

La operación de *catalogar* es más compleja que la descripción bibliográfica porque, además de describir y consecuentemente identificar, es ordenar, clasificar y concretar los puntos de acceso por

⁹⁷ GUINCHAT, Claire y MENOUE, Michael. *Introducción general a las ciencias y técnicas de la información y documentación*. 2ª de. corr. y aum. por Marie-France Blanquet. Madrid: CINDOC/UNESCO, D.L. 1992, p. 105.

medio de los cuales puede ser localizado un documento perteneciente a una determinada colección.

Catalogar significa transferir, conforme determinadas reglas, ciertos datos técnicos de un documento a un soporte documental, el catálogo. María Pinto distingue dos fases en este proceso⁹⁸: Una dedicada al estudio y observación de los datos del documento en cuestión, permitiendo determinar su punto de acceso y concretar la forma del encabezamiento y, otra relacionada con las tareas de confección del catálogo (reproducción de fichas secundarias, ordenación, mantenimiento, y actualización del catálogo).

El resultado de dicho proceso será el *catálogo* constituido por un conjunto de asientos bibliográficos, presentado en forma preestablecida en un soporte previsto, destinado a la consulta por el usuario de un determinado sistema de información, en ficheros manuales o automatizados.

Las nuevas tecnologías informáticas y de telecomunicación han propiciado la *catalogación compartida o cooperativa* que consiste en la creación simultánea por varios centros de una base común de datos a través de un sistema on-line de forma que la catalogación de un documento efectuada por un centro valga para el resto. El resultado de esta cooperación es un *catálogo colectivo en línea* que reúne los fondos de todos los sistemas de información participantes. El sistema de catalogación compartida también supone la utilización de unas normas de catalogación formal y por materias totalmente unificadas contribuyendo al programa de *Control Bibliográfico Universal (CBU)*,

⁹⁸ PINTO MOLINA, M. *Op.Cit*, p. 142.

propuesto por la IFLA⁹⁹, cuyos objetivos se centran en la organización, normalización y control de la producción bibliográfica mundial¹⁰⁰.

2.2.2. EL ANÁLISIS DE CONTENIDO

Para expresar el contenido de los documentos son necesarias dos operaciones: la *indización* y la *técnica de resumir*. Estas operaciones son extremadamente complejas pues se tratan de procesos intelectuales donde intervienen una serie de variables (analista, documento, usuario), que impiden que el análisis del contenido se convierta en algo mecánico o sistemático como ocurre en el análisis formal, cuyos procedimientos están altamente unificados y normalizados.

Según Guinchat y Menou¹⁰¹, otras finalidades se pretenden con estas técnicas de análisis de contenido como:

- Realizar cuando se requiera selecciones que permitan conservar o desechar un documento, determinar la forma y el nivel de tratamiento posterior y establecer las categorías de ordenación del documento.
- Colocar materialmente los documentos.
- Almacenar para recuperar, es decir, introducir en los ficheros las referencias de los documentos en apartados apropiados.

Para llevar a cabo las operaciones propias de Análisis de contenido, han de tener presente ciertos condicionantes objetivos, tales como:

⁹⁹ IFLA/FIAB: Federación Internacional de Asociaciones Bibliotecarias

¹⁰⁰ CARRIÓN GÚTIEZ, M. *Manual de bibliotecas*. Madrid: Fundación Germán Sánchez Ruipérez, 1988, pp. 315-317.

- Necesidades de los usuarios, que variará en función de su perfil profesional.
- La naturaleza del campo tratado influirá en la tipología documental que haya que analizarse.
- Los medios humanos y materiales y la organización de la unidad de información.
- Los productos y servicios que debe prestar la unidad de información.
- Y, por supuesto, la relación coste-beneficio.

2.2.2.1. LA INDIZACIÓN

A través de esta operación se define el contenido de un documento mediante un conjunto de conceptos que especifican el tema o temas de que aquél trata. Según advierte María Pinto la mayoría de los estudiosos sobre la indización se han centrado más sobre los productos de esta técnica (cuestiones de lenguajes, construcción de índices y su evaluación) que sobre el proceso y sus etapas. Si tomamos una definición emanada de una institución involucrada en la normalización documental como es la Unesco¹⁰², la operación de indizar es concebida desde un doble punto de vista:

- Como *proceso*, es la operación que consiste en describir y caracterizar un documento con la ayuda de representaciones de los conceptos contenidos en dicho documento.
- Desde el punto de vista de su *finalidad*, la indización va destinada a permitir una búsqueda eficaz de las informaciones contenidas en un fondo de documentos.

¹⁰¹ GUINCHAT, C. y MENOUE, M. *Ibíd.*, pp. 125-126.

¹⁰² UNESCO. *Principles d'indexation*. Paris: Unesco, 1975, p. 2.

De la misma manera se pronuncia la ISO, entidad que define la indización como una “operación destinada a representar los resultados del análisis de un documento o de una parte del mismo, a través de términos procedentes de un lenguaje natural o un lenguaje documental, para facilitar la recuperación”¹⁰³.

Por tanto, esta doble concepción expuesta, atiende a los dos períodos importantes del proceso documental: el almacenamiento y la recuperación. Es decir, el procedimiento que se haya empleado para caracterizar los documentos conceptualmente, bien sea el lenguaje natural de los documentos o un lenguaje documental previamente establecido por el sistema de información, tendrá una implicación directa en la recuperación de los mismos ya que el usuario habrá de realizar su petición de búsqueda documental en función del lenguaje utilizado en la indización.

2.2.2.1.1. PROCESO DE INDIZACIÓN

El proceso de indización lo conforman una serie de etapas que no tienen por qué ser secuenciales y dependiendo de los autores consultados, el número de éstas pueden oscilar entre las dos que propone Chaumier¹⁰⁴ y Lancaster¹⁰⁵ y las cinco que plantea Jones¹⁰⁶. Nosotros optamos por una postura ecléctica escogiendo las tres que ofrece una fuente autoritativa y normalizadora como es la ISO, con la

¹⁰³ ISO.5127/3a)-1981. Information and documentation. Vocabulary- Section 3a): acquisition, identification, and analysis fo documents and data. En: *Documentation and Information*. Genève: ISO, 1988, p. 80.

¹⁰⁴ CHAUMIER, J. *Análisis y lenguajes documentales*. Barcelona: Mitre, 1986, p. 28.

¹⁰⁵ LANCASTER, F.W. *Indexing and abstracting in theory and practice*. London: Library Association, 1991, pp. 8-9.

¹⁰⁶ JONES, K.P. Toward a theory of indexing. *Journal of Documentation*, 1976, vol. 32, nº 2, pp. 118-123.

norma ISO 5963-1985¹⁰⁷, la cual muestra una serie de pautas para el análisis de documentos, determinación de su contenido y selección de los términos de indización, recomendaciones dirigidas a la indización humana que utilice un lenguaje de indización controlado como tesauros o clasificaciones. Dicha norma, destinada a promover la utilización de una práctica normalizada en los centros y servicios de indización, desglosa el proceso de indización entre tres etapas, que tienden a solaparse en la práctica: a) *Examen del documento y determinación de su contenido*; b) *Identificación y selección de los conceptos principales del contenido*; y, 3) *Selección de los términos de indización*. Debemos advertir que la aplicación de elementos normalizadores a un proceso intelectual como es la indización, convierten a aquéllos en meras recomendaciones combinadas con el sentido común y talento del indizador.

a) *Examen del documento y determinación de su contenido*

En esta primera fase es recomendable que el analista tenga un buen conocimiento sobre el campo de que tratan los documentos para que le permita determinar con coherencia la naturaleza y objetivos de los mismos. Para Rowley¹⁰⁸ la familiaridad del indizador con la materia de que trata el documento objeto de indización constituye ser la primera etapa del proceso; además, el analista deberá apreciar los nuevos conocimientos que aporta a la materia el documento analizado.

¹⁰⁷ ISO 5963-1985. Documentation. Methods of examining documents, determining their subjects, and selecting indexing terms. En: Documentation and Information. Ginebra: ISO, 1988, pp. 579-584. Su equivalente española es UNE 50-121-91. Documentación. Métodos para el análisis de documentos, determinación de su contenido y selección de los términos de indización. Madrid: AENOR, D.L. 1991.

¹⁰⁸ ROWLEY, Jennifer. E. *Abstracting and indexing*. London: Clive Bingley, 1988, p. 51.

Según la norma ISO 5963-1985 al examinar los documentos, los elementos que deben tenerse en consideración en la indización son: título, resumen (si lo tiene), sumario o tabla de contenido, introducción, párrafos iniciales de los diferentes capítulos o apartados y las conclusiones, ilustraciones, diagramas, tablas y su leyenda o explicación, y palabras o frases que estén destacadas mediante una tipografía diferente o subrayada. En la indización a partir del *título* del documento se pronuncia Coll-Vinent¹⁰⁹ argumentando que es un sistema rápido y barato pues no exige ninguna pericia y aptitud especial en el indizador pero, por otro lado, matiza que no todos los títulos, sobre todo los pertenecientes a las ciencias sociales y humanísticas, expresan el contenido del texto que encabezan. En cuanto a la indización a partir del *resumen*, Guinchat y Menou¹¹⁰ exponen que si el resumen está bien hecho puede fundamentarse en la indización, representando una verdadera economía de tiempo. Sin embargo, estas autoras inciden en que el *documento completo* es el más procedente para realizar la indización y, por tanto, se apunta a la indización profunda, es decir, abarcando todos los temas que se tratan en el documento.

¹⁰⁹ COLL-VINENT, Roberto. *Curso de Documentación asistido por ordenador*. 2ª de. ampl. Madrid: Dossat, 1993, p. 141.

¹¹⁰ GUINCHAT, C y MENOUE, M., *ibíd.*, p. 180.

b) Identificación y selección de los conceptos

Una vez examinado el documento por parte del indizador éste procederá a la identificación de las nociones que describen su contenido. Para ello, el sistema documental fijará una serie de factores o criterios que considere importantes en el campo temático al que pertenecen los documentos con lo cual no todos los conceptos identificados van a ser utilizados como términos de indización. Con todo ello, el principal criterio de selección de conceptos debe ser, como indica la norma “su valor potencial como elemento de expresión del tema del documento para su recuperación”¹¹¹.

El indizador procederá, por tanto, a la lectura del documento (lectura superficial o profunda dependiendo de la naturaleza del documento), buscando frases o palabras significativas que indiquen el contenido del mismo. Puede servirse de dos indicadores formales: la frecuencia con la que aparece un término en el mismo documento, método utilizado en la indización automática, y el contexto lingüístico en el que aparece¹¹².

En cuanto al número de términos o descriptores que pueden asignarse a un documento (*exhaustividad o profundidad*) no debe limitarse de forma arbitraria ya que puede conducir a una pérdida de objetividad en la indización. Es preferible que el número de términos se determine enteramente por la cantidad de información contenida en el documento en relación con las supuestas necesidades de los usuarios a que va destinado el índice.

¹¹¹ ISO 5963-1985, *Ibíd.*, p. 582.

¹¹² CLEVELAND, D.B. y CLEVELAND, A.D. *Introduction to indexing and abstracting*. Englewood, Colorado: Libraries Unlimited, 1990, p. 107.

Otro aspecto relacionado con la selección de los términos de indización es la *especificidad* o *precisión* de los mismos, es decir “la exactitud con que un concepto particular que aparece en el documento está representado por un término de indización”¹¹³. La norma recomienda que ésta se aplique siempre que sea posible. No obstante, pueden utilizarse términos más genéricos cuando el indizador lo considere conveniente en los casos siguientes: 1) cuando pueda actuar de forma negativa sobre el sistema de indización, produciendo el fenómeno de silencio documental en la recuperación; y 2) cuando la idea representada por el término no esté lo suficientemente desarrollada.

Los valores de profundidad y precisión de la indización están íntimamente relacionados con la calidad de los sistemas de recuperación, valores de los que hablaremos en el apartado dedicado a la recuperación.

c) Selección de los términos de indización

Los términos candidatos elegidos en la etapa anterior se traducen a términos de indización. Si se utiliza una lista controlada y en ésta aparece el término, se retendrá como descriptor. Si no figura en la lista controlada habrá que localizar su equivalente, si lo hay, o bien un término más específico o genérico en su defecto.

En esta tercera etapa de descripción temática, la elección del lenguaje de indización cobra un interés excepcional pues éste se constituye como mediador entre el lenguaje de los documentos y el lenguaje utilizado por los usuarios cuando formulan sus demandas.

¹¹³ ISO 5963-1985, *Ibíd.*, p. 582.

Dada su importancia nos detenemos en conocer la tipología de lenguajes de indización.

2.2.2.1.2. LOS LENGUAJES DE INDIZACIÓN

Conviene precisar que no todos los autores utilizan la misma terminología para referirse a los lenguajes que enlazan el contenido de los documentos y las peticiones de los usuarios. Estos lenguajes reciben diversos calificativos como lenguajes documentales, lenguajes de recuperación, lenguajes de indización, etc... imprecisiones terminológicas motivadas, según aprecia María Pinto¹¹⁴, por el contexto profesional de la disciplina documental. Así en el ámbito anglosajón se prefieren los términos de lenguajes de recuperación o lenguajes de indización, frente al calificativo de lenguajes documentales utilizado en los países mediterráneos. No obstante, si partimos de que indizar es la operación destinada a representar el contenido de los documentos por medio de un lenguaje, ya sea documental o natural, y éste va a ser utilizado por el sistema de recuperación, la expresión *lenguaje de indización* parece la más adecuada a tal efecto al contemplar el uso del lenguaje natural o documental.

La tipificación de estos lenguajes atiende a los siguientes aspectos: *coordinación* de los términos, *estructura* que presenta el lenguaje y *control terminológico* de los elementos que lo conforman¹¹⁵, principios que exponemos a continuación.

a) Coordinación de los términos

¹¹⁴ PINTO MOLINA, M. *Op. Cit.*, p. 213.

¹¹⁵ AMAT, N. *Documentación científica y nuevas tecnologías de la información*. Madrid: Pirámide, 1988, p. 190.

El funcionamiento de un lenguaje de indización se caracteriza por el principio de *coordinación*. Este aspecto es contemplado en el esquema de Courier¹¹⁶ en el que distingue dos tipos de lenguajes: *los lenguajes precoordinados* que coordinan los diferentes conceptos que forman un tema antes del almacenamiento y por tanto, esta disposición metódica de los términos de indización provocará que se utilicen los mismos tanto en el momento de la indización como en la recuperación; y los *lenguajes postcoordinados* permiten la coordinación de conceptos en el momento de la recuperación; en este sentido, los conceptos y/o palabras utilizadas para representar el contenido de los documentos podrán combinarse en el momento del análisis.

BIBLIOTECA VIRTUAL

Son lenguajes precoordinados las *clasificaciones* y las *listas de encabezamientos de materia*; la utilización de éstos permiten pocos términos de indización por documento frente al empleo de múltiples puntos de acceso que facilitan los lenguajes postcoordinados. El nacimiento de los lenguajes postcoordinados (*unitérminos, listas de palabras clave, listas de descriptores y tesauros*) está íntimamente ligado al proceso de automatización de los sistemas documentales: las prestaciones de los sistemas informáticos aplicados a la automatización de la descripción de los documentos son múltiples permitiendo la coordinación de los términos en la operación de consulta o recuperación.

¹¹⁶ COURRIER, Y. Analyse et langage documentaires. *Documentaliste*, 1976, vol. 13, nº 5-6, p. 183.

b) Estructura de los lenguajes

Desde el punto de vista de su estructura, los lenguajes de indización se dividen en: *lenguajes jerárquicos*, *lenguajes asociativos* y *lenguajes de estructura sintáctica*.

- Los *lenguajes de estructura jerárquica* o *clasificatoria* siguen un principio de orden lógico que agrupa o aproxima los conceptos más sencillos o específicos dentro de los conceptos más generales. Ello conduce a establecer una clasificación sistemática lineal, en la cual los conceptos se encuentran ordenados siguiendo una jerarquía natural definida por el estado de los conocimientos en que ha sido elaborada. Cada concepto de estas estructuras jerárquicas se halla representado por un símbolo numérico, alfabético o alfanumérico que indica la situación de cada concepto.

- Los *lenguajes de estructura asociativa* o *combinatoria* presentan a sus términos ordenados alfabéticamente. Éstos pueden combinarse libremente sin quedar sujetos a una posición determinada del lenguaje. Esta estructura proporciona un acceso más inmediato a la información.

- Los *lenguajes de estructura sintáctica*, son lenguajes muy elaborados. Constan de un conjunto de descriptores más la sintaxis que excluye la ambigüedad en las relaciones entre conceptos.

c) Control terminológico

Este aspecto corresponde al dominio que ejerce el vocabulario asignado al documento en relación a la recuperación o no del mismo. Desde éste punto de vista, Van Slype¹¹⁷ distingue tres tipos:

- *Lenguajes libres*, constituidos *a posteriori*, sobre la base de indización en lenguaje natural de los documentos. Dentro de esta categoría se pueden diferenciar, a su vez, dos tipos de lenguajes combinatorios: *Las listas de palabras clave* compuestas por palabras significativas extraídas del documento original. Normalmente, éstas se recogen de forma automática a través del programa de gestión. Y las *listas de descriptores libres*, formadas por conceptos destacados del documento original. Los términos que designan a estos conceptos pueden no estar en los documentos .

- *Lenguajes controlados*, construidos *a priori*, antes de empezar a indizar los documentos. Equivalen a la concepción de lenguaje documental. Bajo esta tipología se encuentran las listas de autoridades y los tesauros. Conviene indicar que la elaboración de estos últimos es complicada y laboriosa; sin embargo, la recuperación puede ser más rápida y eficaz ya se dirigen a controlar los problemas lingüísticos de sinonimia, polisemia, etc que acontecen con la utilización de lenguajes libres.

- Los *lenguajes codificados* que utilizan un sistema de notación para la indización de documentos y demandas, muy utilizados en las bibliotecas.

¹¹⁷ VAN SLYPE, G. *Los lenguajes de indización*. Madrid: Fundación Germán Sánchez Ruipérez, 1991, pp. 22-23.

2.2.2.1.3. SISTEMAS DE INDIZACIÓN

Los sistemas de indización se conciben como "el conjunto de procedimientos prescritos dirigidos a organizar los contenidos de los registros con el fin de que puedan ser recuperados y difundidos"¹¹⁸. El tipo de lenguaje de indización empleado, ya sea natural o documental, va a determinar el sistema de indización. De esta manera, según se quiera expresar el contenido de los documentos a través de la jerarquía conceptual tema, palabra o concepto, se procederá a los respectivos sistemas de indización: por *materias*, por *unitérminos* y por *descriptores*.

- *Indización por materias*. La caracterización de los documentos, en función del tema, consiste en extraer la parte de la realidad a la que se refiere. Los encabezamientos de materia, *subject-headings* o *vedettes-matière*, serán los elementos utilizados para este tipo de indización. De esta manera, un documento podrá quedar descrito conceptualmente bajo el encabezamiento de la materia que predomine. El inconveniente que presenta este tipo de indización, a efectos de recuperación, es la falta de pertinencia, ya que los términos utilizados suelen ser amplios, además de la poca flexibilidad para realizar combinaciones entre ellos. Por este motivo, se ideó otro sistema con el afán de profundizar en la indización: la indización por *uniterms*.

- *Indización por unitérminos*. Con este sistema el documento queda caracterizado por las palabras significativas (generalmente sustantivos) tal y como aparecen en el mismo, en lenguaje natural. Así cada término (uniterm) representa una palabra clave caracterizada por una palabra sencilla y tomada del lenguaje natural del documento, teniendo la misma jerarquía en la lista, es decir, no existiendo uno

¹¹⁸ ROWLEY, Jennifer E. *Ibid.*, p. 53.

conceptualmente superior a otro. Este sistema, introducido por el norteamericano Mortimer Taube en 1953, supone un importante progreso en la indización pues, según Van Dijk¹¹⁹, el poder combinatorio de las palabras permite expresar un considerable número de conceptos y temas y, por tanto, la recuperación sería más pertinente que la obtenida a través de los encabezamientos de materia, pero puede ofrecer combinaciones falsas en la búsqueda documental, no adecuándose a la petición del usuario.

- *La indización por descriptores.* Este sistema se presenta en la actualidad como el más eficaz para fines recuperacionales, pues asegura la coincidencia entre el lenguaje utilizado por los emisores de información (analista) y el de quienes la buscan (usuario). El descriptor es un término o conjunto de términos extraídos de un vocabulario controlado que representan el contenido de un documento. El lenguaje de indización que presta un mayor control entre sus términos (descriptores y no-descriptores) es el *tesauro* que va a permitir pasar del lenguaje natural (de los documentos o de las demandas de los usuarios) a un lenguaje formado por términos controlados a través de las relaciones que se establecen entre ellos: relaciones de *equivalencia*, de *jerarquía* y de *asociación*. El objetivo de este lenguaje controlado es facilitar y unificar el análisis y la recuperación documentales.

¹¹⁹ VAN DIJK, M. y VAN SLYPE, G. *El servicio de documentación frente a la explosión de la información*. Buenos Aires: Consejo Nacional de Investigaciones Científicas y Técnicas, 1972, p. 51.

2.2.2.1.4. CONSIDERACIONES GENERALES SOBRE LA INDIZACIÓN AUTOMÁTICA

El incremento exponencial de la información científica a finales de los años cincuenta en todos los campos del saber científico y, especialmente en el de las ciencias experimentales, motivó a que investigadores de la información se cuestionaran que algunas tareas documentales repetitivas, entre ellas la indización, fueran realizadas por el ordenador al considerarse éste como una máquina exacta y precisa en las operaciones y, de esta manera, minimizar los errores en la selección de términos para la indización.

En la década de los sesenta, M.E. Stevens¹²⁰ revisó los criterios que aplicaban los ordenadores a la tarea de indizar y definía a la indización automática como el uso de máquinas para extraer o asignar términos de indización, sin intervención humana, una vez que se hayan establecido programas o normas relativas al procedimiento. Una definición más acorde con la finalidad de la indización (la recuperación) es la que ofrece Van Slype¹²¹ quien la define como la operación que consiste en que el ordenador reconozca los términos que figuran dentro del título, del resumen, o del texto completo, y que emplee estos términos tal cual, o traducidos a otros (ya sean equivalentes o conceptualmente próximos) con el fin de convertirlos en elementos de recuperación y queden disponibles en el fichero de búsqueda.

A pesar del futuro prometedor de esta técnica, desde su aplicación hasta la fecha todavía se sigue cuestionándose la preferencia o no de la indización automática frente a la indización

¹²⁰ STEVENS, M.E. *Automatic indexing: a state of the art report*. Washington, D.C.: National Bureau of Standards, 1965.

¹²¹ VAN SLYPE, Georges. *Op.Cit.*, p. 125.

humana. Gil Leiva y Rodríguez Muñoz¹²², en un estudio donde revisan la evolución de la indización automática y asistida por ordenador hasta nuestros días, aluden a los pros y contras de este método de indización. Entre sus *inconvenientes* se encuentran: la incapacidad de la máquina de realizar convenientemente la labor de indizar, ya que no puede llegar a captar todos los matices conceptuales como puede hacerlo el indizador humano; la validación del indizador humano de algunos sistemas de indización automática; y la relativa garantía de estos sistemas al restringirse a áreas de conocimiento muy concretas, donde la terminología está altamente normalizada. Sin embargo, los defensores de la indización automática alegan las siguientes *ventajas*: una mayor economía tanto de proceso como de presupuesto; una mayor objetividad, puesto que se aplicarían siempre los mismos parámetros; y una disminución de errores que son un claro inconveniente en el momento de la recuperación de la información en una base de datos.

Dejando al margen esta polémica, conviene indicar los métodos de indización automática o semi-automática que se utilizan en la actualidad. Básicamente estos se dividen en dos grupos: los *métodos no lingüísticos*, que agrupan a los estadísticos, la atribución de pesos, los probabilísticos y los de clustering, y por otro lado, los *métodos lingüísticos*, basados en el procesamiento del lenguaje natural (PLN). Chaumier y Dejean¹²³ clasifican los métodos de indización automática en tres grupos:

- *Métodos estadísticos*, basados en la frecuencia de aparición de los términos significativos, los cuales se retendrán para la indización. H.P.

¹²² LEIVA GIL, Isidoro y RODRÍGUEZ MUÑOZ, José Vicente. Tendencias en los sistemas de indización automática. Estudio evolutivo. *Revista Española de Investigación Científica*, 1996, vol. 19, nº 3, pp. 273-291.

Luhn¹²⁴ fue el primero en sugerir la utilidad de la frecuencia de aparición de los términos en una colección para la indización. Uno de sus inconvenientes es que no resuelve los problemas de homonimia y polisemia.

- *Métodos por asignación*, utilizando un tesoro preestablecido. Los términos significativos presentes en el texto son traducidos por descriptores del tesoro.

- *Métodos sintácticos*, introducen tratamientos morfológicos para llegar a un reconocimiento de las estructuras más significativas. La indización se realizaría en dos niveles: el primero consistiría en determinar las unidades léxicas del documentos en sus aspectos morfológicos y semánticos; el segundo nivel sería la interpretación sintáctica.

Actualmente, parece poco probable que la indización automática reemplace a la humana, pero M.L. Ward¹²⁵ aprecia que el número creciente de documentos primarios en formato electrónico supone ser una útil y productiva herramienta para su análisis automático ya que se presentan en toda su extensión.

En cuanto a los índices producidos por ordenador destacan:

- *Índices KWIC* (KeyWord In Context) o índices de palabras-clave en contexto. Reúnen cada palabra significativa del título del documento en una lista alfabética, permutando el título para que la palabra seleccionada aparezca siempre en la misma posición, precedida y

¹²³ CHAUMIER, Jacques y DEJEAN, Martine. L'indixation documentaire: de l'analyse conceptuelle humaine à l'analyse automatique morphosyntaxique. *Documentaliste*, 1990, vol 27, nº 6, p. 275-279.

¹²⁴ LUHN, H.P. A statical approach to mechanized encoding and searching of literary information. *IBM Journal of Research and Development*, 1975, vol. 1, nº 4, pp. 309-317.

sucedida por las que forman el contexto, que sirven para aclarar la correspondiente palabra clave.

- *Índices KWOC* (KeyWord Out of Context) o índices de palabras-clave fuera de contexto. Presentan las palabras significativas por orden alfabético fuera del título y presentadas y exhibidas como encabezamientos de los títulos completos.

- *Índices de Citas*. Son índices de autores, que citan junto a ellos sus correspondientes trabajos ordenados cronológica y alfabéticamente; añaden también a cada referencia la lista de autores y documentos que citaron esa referencia bibliográfica. Es muy apreciado para la recuperación de la información, pues permite localizar otros autores que han tratado las mismas materias y buscar documentos recientes sobre el mismo tema. Además, estos productos documentales adquieren gran importancia en la realización de estudios bibliométricos como es el cálculo del *Factor de Impacto* (FI) de una determinada publicación periódica.

2.2.2.2. LA OPERACIÓN DE RESUMIR Y EL RESUMEN

2.2.2.2.1. DEFINICIÓN

El segundo nivel del Análisis Documental de contenido lo constituye la *operación de resumir* que tiene por objeto la descripción sustancial del contenido de los documentos. El resultado de esta operación será el *resumen* que según la ISO, se define como “una representación abreviada y precisa del contenido del documento, sin

¹²⁵ WARD, Martin L. The future of the human indexer. *Journal of Librarianship and information Science*, 1996, vol. 28, nº 4, pp. 217-225.

interpretación crítica y sin mención expresa del autor del resumen”¹²⁶. En este sentido, el resumen, o *abstract* en su terminología inglesa, no debe confundirse con otras formas documentales que tienden a condensar el contenido de los documentos. Nos referimos a términos como la *anotación* (comentario o breve explicación acerca de un documento o de su contenido), el *extracto* (selección de una o más partes representativas de un documento) o *reseña* (interpretación de las principales ideas del documento del documento original analizado). El resumen, pues, se caracteriza por su objetividad en la descripción, rechazándose en su elaboración cualquier interpretación crítica por parte del analista/resumidor.

Una de las definiciones más acertadas de este producto documental es la que ofrece M. Pinto. Para esta autora el resumen es “la descripción reducida de la estructura cognitiva (contenido) del documento cuya función principal es la intermediación entre éste y el potencial usuario”¹²⁷. Esta definición recoge las aportaciones conceptuales de algunas disciplinas que intervienen en el Análisis Documental del contenido como son la Lingüística Textual, la Psicología Cognitiva, la Lógica Formal y la Teoría de la Comunicación.

La operación de resumir, a pesar de ser una técnica que cuenta con cierta antigüedad practicada ya por los archiveros de la cultura mesopotámica, se implanta socialmente con la aparición de las primeras revistas científicas en el siglo XVII, constituyéndose como el vehículo más idóneo de la actividad científico-informativa¹²⁸. Desde entonces hasta la fecha la revista científica no se ha desprendido de este provechoso instrumento de alerta.

¹²⁶ ISO 214-1976. Documentation. Abstracts for publications and documentation. En: *Documentation and Information*. Genève: ISO, 1988, pp. 420-439. Su equivalente española es AENOR. UNE 50-103-90. *Documentación. Preparación de resúmenes*. Madrid: AENOR, 1990.

¹²⁷ PINTO MOLINA, María, *Op.Cit.*, p. 174.

Las reflexiones teóricas en torno a esta técnica se han dirigido más hacia su producto (resumen), ofreciendo recomendaciones de preparación y presentación, que a la actividad misma de resumir. Sin embargo, las últimas aportaciones apuntan a establecer una teoría del resumen basándose en las variables que afectan en el proceso de resumir: el *documento*, el *resumidor/analista*, y los *usuarios*, elementos de los que hablaremos a continuación.

2.2.2.2. ELEMENTOS INVOLUCRADOS EN LA OPERACIÓN DE RESUMIR

En la operación de resumir intervienen, básicamente, tres elementos: el *documento original*, el *analista* y los *usuarios*.

a) *El documento original*

El objeto de la operación de resumir, el documento, puede ser, en principio, cualquier registro de información. Ahora bien, en la práctica, los documentos que más se someten a este tipo de análisis son los artículos científicos, tanto los producidos en el ámbito de las ciencias naturales (CN) como los derivados de las ciencias sociales y humanísticas (CS y H). Las aportaciones de la Lingüística Textual con su teoría de las estructuras retóricas (superestructuras) han ayudado en gran medida a diferenciar la tipología de textos. En este sentido, cabe indicar que la *estructura formal* de los textos pertenecientes al dominio de las ciencias naturales difiere en gran medida de la de los textos de las ciencias sociales y humanísticas; mientras que los textos de CN presentan, generalmente, una estructura formal dividida en

¹²⁸ LÓPEZ YEPES, José. *Op.Cit.*, p. 114.

cuatro *categorías* (Objetivos, Metodología, Resultados y Conclusiones) que facilitan más fácilmente la identificación las estructuras cognitivas, los textos de las CS y H presentan una estructura más versátil, menos normalizada y dependerá, en muchos casos, de la personalidad del autor¹²⁹. Además, en los textos humanísticos “el cómo decirlo” es también un fin, factor que viene a dificultar la técnica de resumir pues ya no se hacen tan identificables esas estructuras de las que hablábamos antes.

Pero los artículos científicos no son los únicos objeto de resumen. También se resumen otras tipologías documentales como las monografías, los informes técnicos, las tesis, las patentes, las conferencias, ponencias, comunicaciones y actas de congresos.

b) *El analista/resumidor*

El analista/resumidor será el encargado de ejecutar correctamente la operación de resumir. No obstante, existe la discusión de si debe ser el *autor* del documento original, familiarizado con la temática del documento a resumir, un *especialista* en la materia, o el *resumidor profesional*, conocedor de las técnicas documentales. A pesar de la posible idoneidad de los dos primeros, pues son conocedores del campo temático en el que se inserta la información a analizar, desconocen los procedimientos y técnicas de resumir y las necesidades de los usuarios a quienes van dirigidos. El autor del documento puede pecar de falta de objetividad; el experto carecer del adiestramiento previo para garantizar la adecuación y fiabilidad del producto. El resumidor profesional se perfila como el ejecutor más competente para llevar a cabo la técnica de resumir ya que reúne en su persona los conocimientos necesarios para tal fin como: experiencia en

¹²⁹ PINTO MOLINA, M. *Op.Cit*, pp. 65-67.

el tema objeto de resumen, conocimiento de lenguas extranjeras y conocimiento tanto de las técnicas de resumir en particular, como de las técnicas documentales en general.

c) Los usuarios

El usuario es el factor menos estudiado, a pesar de ser el destinatario de esta operación analítica. Es más, según apunta M. Pinto¹³⁰, el usuario debe ser el resorte en el que se cimiente todo el sistema de información como ya lo hacen las modernas teorías evaluadoras y de gestión de calidad, calidad dirigida hacia los servicios y los productos. El conocimiento de la tipología de usuarios a los que van destinados los resúmenes determinará una serie de toma de decisiones por parte de las unidades de información: nivel de descripción analítica, profundidad de los resultados, tipos de productos, formas de presentación, etc.

2.2.2.2.3. METODOLOGÍA DE LA OPERACIÓN DE RESUMIR

A pesar de la necesidad y vigencia de la operación de resumir, ésta no cuenta con unas reglas universales ni son aplicables a todo tipo de documento. Es decir, se trata de una operación donde intervienen un conjunto de variables (el texto, el analista y sus componentes), indicadas más arriba, que dificultan o hacen que la operación no sea una tarea mecánica sino una tarea intelectual, un arte como llaman algunos¹³¹. Otros autores, como es el caso de Endres-Niggemeyer¹³², realizan estudios empíricos con el objeto de desarrollar un modelo del proceso de resumir, recogiendo los principios organizativos de los

¹³⁰ PINTO, M. *Op.Cit.*, p. 317.

¹³¹ CREMMINS, E.T. *El arte de resumir*. Barcelona: Mitre, 1985, p. 13.

resumidores, las herramientas intelectuales que emplean y las estrategias utilizadas en cada paso del proceso. De acuerdo con M. Pinto¹³³ esta actividad se concibe como una unidad de actuación indivisible aunque, por razones metodológicas, pueden distinguirse tres fases: 1) *Lectura del documento*, 2) *Análisis* y 3) *Síntesis*. De la misma manera se expresa J. A. Moreiro González¹³⁴, para quien la actuación del analista sobre los documentos pueden considerarse en tres momentos: 1) *Reconocimiento* o análisis del documento para identificar su información más destacada; 2) *Reducción del contenido*, eliminando lo accesorio; y 3) *Representación* o redacción de la información fundamental del texto. Veamos pues desarrolladas estas tres fases.

a) *Lectura del documento*

Esta primera fase de lectura del documento original tiene como objetivos la identificación y comprensión de la macroestructura cognitiva del documento. Para ello, se llevarán a cabo dos lecturas. La primera debe realizarse de forma rápida y superficial con el objeto de identificar las características fundamentales del documento, como su forma, clase a la que pertenece y cómo estructura la información.

La segunda lectura debe ser más activa, más lenta y profunda, donde el resumidor trata de captar lo que el autor del documento original considera lo más importante. El resumidor prestará atención a las secciones de los epígrafes que caracterizan al documento científico como: Objetivos, Metodología, Resultados y Conclusiones.

¹³² ENDRES-NIGGEMEYER, B.A. Naturalistic model of abstractin. *Advances in knowledge organization*, 1994, vol. 4, pp. 181-187.

¹³³ PINTO MOLINA, M. Documentary abstracting: Toward a methological model. *Journal of the American Society for Information Science*, 1995, vol. 46, nº 3, pp. 225-234.

¹³⁴ MOREIRO GONZALEZ, J.A. La técnica del resumen científico. En: *Manual de Información y Documentación*. Madrid: Pirámide, 1996, pp. 376-380.

b) Análisis

Esta segunda fase se caracteriza por dos actividades complementarias: la *selección* y la *interpretación*. La *selección* consiste en eliminar aquellas unidades de información que se consideran irrelevantes para formar parte del resumen documental. Para M. Pinto¹³⁵ pueden identificarse tres grupos de unidades de información que pueden ser eliminadas: Unidades redundantes, unidades poco relevantes y unidades irrelevantes.

La *interpretación*, o asignación de un contenido al texto, es considerada como la actividad más subjetiva del proceso de resumir porque entran en juego factores extratextuales como el conocimiento del mundo del resumidor, el contexto y los objetivos documentales. El éxito o fracaso de este acto interpretativo estará condicionado por las competencias lingüística, situacional y discursiva del analista.

¹³⁵ PINTO, M. *Op. Cit.*, pp. 230-231.

c) Síntesis

Esta última fase es considerada como la más delicada de la operación de resumir, pues no existen técnicas totalmente válidas para todos los tipos de documentos y de analistas¹³⁶. La síntesis o descripción analítica, consiste en describir el contenido del documento derivado de su análisis. En esta descripción o reexpresión del contenido esencial del documento, las competencias y habilidades del analista toman su expresión, que según algunos autores, es considerada como el verdadero acto de resumir. De ahí se deduce que un gran número de estudiosos y practicantes ofrezcan recomendaciones o directrices que ayuden a la confección de este documento secundario, de las que hablaremos posteriormente en el apartado de normalización documental.

2.2.2.2.4. FUNCIONES DEL RESUMEN DOCUMENTAL

El resumen documental, puente entre el documento y el usuario potencial, deben atender a una serie de funciones, entre otras¹³⁷:

- Actualizar el conocimiento del especialista, ahorrándole tiempo y esfuerzo al evitar la lectura del documento original.
- Servir de anticipo del documento original. Tras su lectura permitirá al usuario excluir o seleccionar aquellos documentos idóneos para sus intereses.
- Actuar, a veces, como sustituto del documento original.
- Ayudar en las tareas de búsqueda retrospectiva y recuperación de información, pues cada vez va siendo más habitual que en las bases

¹³⁶ PINTO, M. *Op.Cit.*, pp. 180 y 182.

¹³⁷ MOREIRO GONZALEZ, J.A. *Ibid.*, pp. 375-376 y PINTO, M. *Op.Cit.*, pp. 183-184.

de datos referenciales cada registro bibliográfico venga acompañado de un resumen.

2.2.2.2.5. TIPOS DE RESÚMENES

Pueden distinguirse diversos tipos de resúmenes¹³⁸ en función de: su extensión, su autoría, la profundidad del análisis, el medio de difusión empleado, etc.

a) Según la *autoría* del resumen, se distinguen:

- de *autor*, realizados por el propio autor del documento original.
- de *especialistas*: confeccionados por expertos en un campo temático determinado.
- de *documentalistas*: confeccionados por expertos en técnicas documentales.

b) Según sus *objetivos* o *propósitos externos*:

- Orientados a disciplina: dirigidos a un área específica del conocimiento.
- Orientados a misión: para apoyar actividades operativas concretas.
- Parciales: dirigidos a los intereses de unos usuarios determinados.

c) Según la *tipología* de fuentes originales:

- De artículos de publicaciones periódicas
- De monografías
- De informes, tesis, patentes

¹³⁸ Varios teóricos de la documentación establecen una tipología. Véase N. Amat. *Documentación científica y nuevas tecnologías de la información*, pp. 179-182; J. Chaumier. *Análisis y lenguajes documentales*, pp. 22.28; M. Pinto. *El análisis documental*, pp. 183-188; J. A. Moreiro. La técnica del resumen científico. En : *Manual de información y documentación*, pp. 381-390.

- De documentos audiovisuales

d) Según el *modo de difusión*:

Difusión simultánea: documentos originales y resúmenes se publican a la misma vez.

- Revistas de resúmenes: resúmenes de artículos de revistas científicas que se publican conjuntamente.

- Bases de datos: que incluyen además de las referencias bibliográficas de las publicaciones originales sus resúmenes.

e) Según su *forma de presentación*:

- Telegráficos

- Discursivos

- Tabulares

Pero la categoría más estudiada es la que corresponde al nivel de *profundidad* del análisis. Atendiendo a este factor, el resumen puede clasificarse en:

- *Indicativo*. Es el tipo de resumen más breve con una extensión de aproximadamente 50 palabras. Redactado en lenguaje natural se aproxima al índice y describe la macroestructura del texto. La función principal del resumen indicativo es alertar al usuario de la existencia del documento.

- *Informativo*. Recoge todos los aspectos del documento primario, es decir, la idea central del documento junto con los temas tratados en la superestructura del mismo, especialmente en lo que atañe a los resultados y conclusiones del trabajo. De mayor extensión que el anterior (entre 100 y 300 palabras), la confección del mismo requiere de una alta especialización del tema de que trata el documento analizado.

- *Informativo-indicativo*. Es una combinación de los dos anteriores. Describe las ideas centrales expuestas a modo de resumen informativo, mientras que los aspectos de menor relevancia son tratados de forma indicativa.

- *Analítico*. Es el más extenso (500 palabras) pero también el más inusual en la práctica de los resúmenes. Representa con detalle todos los aspectos tratados en el documento primario y requiere un alto consumo de tiempo en su elaboración.

BIBLIOTECA VIRTUAL

2.2.2.2.6. PRESENTACIÓN Y ESTILO DEL RESUMEN

Existen unas reglas básicas para la presentación de los resúmenes, referidas, especialmente, a la objetividad, la precisión terminológica y la claridad expositiva del resumidor, y a la fidelidad al documento original. No obstante, al tratarse de una operación intelectual que se ve afectada por la pericia del resumidor, estas reglas básicas se convierten en meras recomendaciones o consejos. J. A. Moreiro¹³⁹ expone una serie de valores que deben considerarse en la redacción del resumen:

- *Entropía*, esto es, decir mucho con el mínimo de palabras.
- *Precisión* en la terminología empleada.
- *Pertinencia* en el origen de los contenidos, evitando juicios y valoraciones sobre los mismos.
- *Coherencia* en la exposición del asunto condensado, manteniendo el orden lógico que aparece en el original.

¹³⁹ MOREIRO, J.A. *Ibíd.*, pp. 378-380.

- *Corrección lingüística y gramatical.* El analista respetará la morfología, sintaxis y ortografía propias de la lengua.
- *Estilo.* Debe evitarse la expresión telegráfica y omitirse expresiones vagas y ambiguas, o redundantes. Se recomienda el uso de la tercera persona, el tiempo presente y formas activas de los verbos.
- *Informatividad.* El resumen documental está mediatizado siempre por su función informativa.

En cuanto a normativa específica para la presentación y forma del resumen, existen recomendaciones y pautas, no existiendo para la operación de resumir que, como hemos indicado anteriormente, dependerá de la personalidad y habilidades del analista. Así, la norma ISO 214-1976¹⁴⁰ recomienda que la referencia bibliográfica puede situarse antes o después del resumen, pero siempre en la misma página; y, en el cuerpo del resumen habrá que indicar los objetivos y alcance la investigación, la metodología empleada en la misma (técnicas y procedimientos), y los resultados y conclusiones descritos de forma concisa e informativa.

2.2.2.2.7. AUTOMATIZACIÓN DE LA OPERACIÓN DE RESUMIR

La automatización de esta técnica documental es una de las más complejas y difíciles debido a su dimensión intelectual y, en este sentido, el hombre sigue siendo el único capaz de realizar con eficacia esta labor. No obstante, la investigación y estudio sobre la mecanización de esta operación no se desecha y puede que, en un futuro próximo, sea la llave que abra el complejo mecanismo de la técnica humana. En este sentido, disciplinas como la Lingüística Computacional, la Traducción Automática, o la Organización e

¹⁴⁰ ISO 214-1976. *Op.Cit.*, pp. 420-430. Su equivalente española es AENOR. UNE 50-103-90. *Documentación. Preparación de resúmenes.* Madrid: AENOR, 1990.

Ingeniería del Conocimiento, trabajan conjuntamente. Para Mathis y Rush¹⁴¹ la secuencia operativa de un sistema automático de resúmenes es:

- Lectura del documento original.
- Análisis aplicando una serie de normas de selección y/o transformación para producir el resumen.
- Edición e impresión.

La etapa del análisis es la más compleja pues las normas de selección y/o transformación, basadas en técnicas estadísticas y lingüísticas, no están lo suficientemente desarrolladas, confeccionando productos de baja calidad. Sin embargo, las fases de edición e impresión es la que se ha visto más favorecida en la automatización, al mejorarse los sistemas recuperación.

2.2.2.3. LA CLASIFICACIÓN

2.2.2.3.1. DEFINICIÓN Y OBJETIVOS

Una tercera operación de análisis de contenido es la *clasificación*, entendida como la “operación intelectual por la cual el documentalista atribuye a una obra un índice correspondiente a una clase de materias, utilizando un lenguaje de clasificación”¹⁴². Se trata pues de indicar el tema principal al que pertenece el documento primario, configurándose como una operación sintética frente a la indización, de naturaleza analítica, que pretende detectar los conceptos

¹⁴¹ MATHIS, B. y RUSH, J. Abstracting. En: DYM, E. *Subject and Information Analysis*. Nueva York: M. Dekker, 1985, pp. 445-484.

¹⁴² MANIEZ, J. *Los lenguajes documentales y de clasificación*. Madrid: Fundación Germán Sánchez Ruipérez, 1993, p. 23.

claves representativos del documento. Para Guinchat y Menou¹⁴³ la clasificación tiene por objetivo permitir la ordenación de:

- Los documentos primarios en las estanterías correspondientes cuando los documentos se disponen sistemáticamente.
- Las reseñas bibliográficas en los ficheros sistemáticos
- Las reseñas o referencias bibliográficas en catálogos impresos u otros productos documentales como boletines bibliográficos o de resúmenes.

M. Pinto indica que el objetivo principal de la clasificación es “permitir el agrupamiento de materias a fin de poder almacenar y recuperar con posterioridad la información, al menos en los sistemas convencionales”. Pero también, continúa la autora, “afecta al contenido de las preguntas o demandas de los usuarios”¹⁴⁴, determinándose en qué clases o subclases de la clasificación se sitúa la información demandada.

2.2.2.3.2. ETAPAS DE LA CLASIFICACIÓN

El proceso de clasificación es muy similar al modelo seguido para la indización ya que con ambas técnicas se pretende dar a conocer el contenido de los documentos analizados. Las fases en las que se desglosa este proceso son¹⁴⁵:

- Determinación del tema principal de los documentos. Para ello no se requiere un conocimiento especializado del campo correspondiente por parte de analista. Se procederá a realizar una lectura superficial (título, índice de materias y un recorrido diagonal del texto).
- Determinación de la materia principal. Se trata de seleccionar el tema principal y desechar lo accesorio. Si el documento trata de una sola

¹⁴³ GUINCHAT, Claire y MENOU, Michael. *Ibíd.*, p. 169.

¹⁴⁴ PINTO, M. *Op.Cit.*, p. 107.

¹⁴⁵ MAINEZ, J. *Ibíd.*, p. 60.

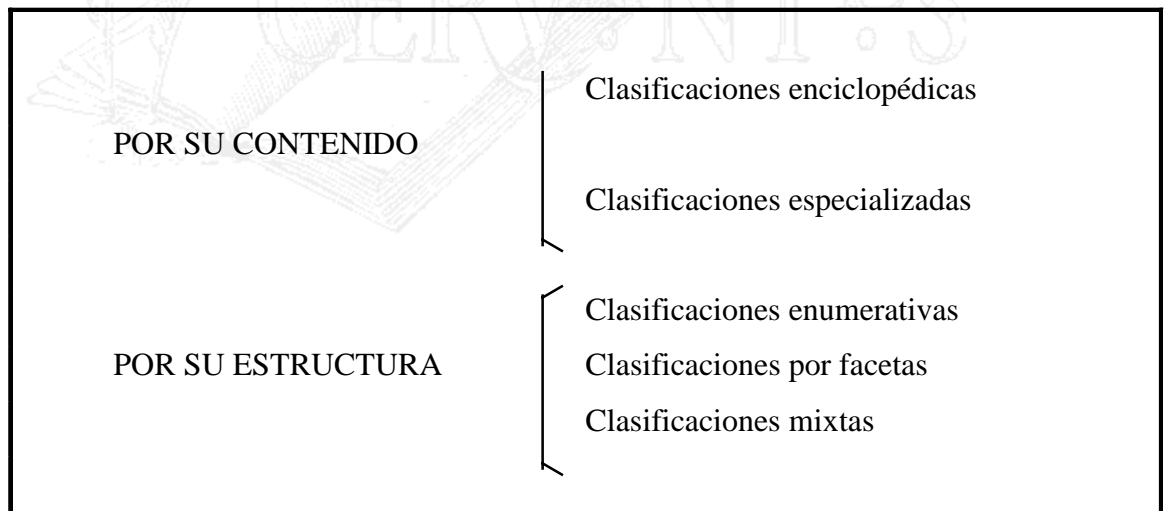
materia, la determinación de la misma se considera una tarea simple. El problema surge cuando el documento trata de más de una materia. En este caso, el analista tendrá que precisar la preferencia de una sobre la otra, en función de la importancia que le haya concedido el autor del documento original.

- Y, por último, la fase de codificación o traducción de ese contenido al lenguaje de clasificación seleccionado, lenguajes de los que hablaremos seguidamente.

2.2.2.3.3. LAS CLASIFICACIONES

Hasta ahora, sólo nos hemos referido a la clasificación como actividad. Pero, dado el carácter polisémico de este término, con él también podemos referirnos a un producto o sistemas de clases previamente establecido. Gil Urdicián¹⁴⁶ ofrece una tipificación de los lenguajes de clasificación de acuerdo con su *composición léxica* y su *estructura*.

¹⁴⁶ GIL URDICIÁN, B. Los lenguajes documentales (I). En: *Manual de información y documentación*. Madrid: Pirámide, 1996, p. 342.



Lenguajes de clasificación

Respecto al *contenido* las clasificaciones pueden ser *enciclopédicas*, abarcando todas las ramas del saber, o *especializadas*, cuando se centran en un campo temático determinado. Entre las primeras pueden mencionarse la Clasificación Decimal de Dewey

(CDD), la Clasificación Decimal Universal (CDU) y la Clasificación de la Biblioteca del Congreso (LCC). Entre las segundas destacan: la clasificación médica de la National Library of Medicine y la Clasificación de la OCDE, entre otras.

En relación con su *estructura* se distinguen: *jerárquicas*, *facetadas* y *mixtas* o híbridas. Las *clasificaciones jerárquicas* utilizan un modelo inspirado en la lógica clásica donde las clases o niveles principales se despliegan en subclases o subniveles, de manera que una clase -cualquiera que sea su nivel- se conecta con la raíz de la arborescencia siguiendo un sólo camino. Las *clasificaciones facetadas* permiten descomponer el campo científico o disciplina sometida a análisis en sus facetas correspondientes o puntos de vista fundamentales. Por último, las *clasificaciones mixtas* son básicamente jerárquicas pero hacen uso de facetas que permiten mayor precisión para la caracterización de los documentos.

2.3. RECUPERACIÓN DE LA INFORMACIÓN

La recuperación de la información es una fase del proceso documental que, como venimos insistiendo, se encuentra íntimamente ligada a la del análisis, en concreto, a la indización, pues la finalidad de ésta, según apuntaba la Unesco, va destinada a una recuperación y búsqueda eficaz de las informaciones contenidas en un fondo de documentos. El término Recuperación de la Información o *Information Retrieval* apareció en 1950 cuando Calvin Mooers lo utilizó para referirse a “una búsqueda de informaciones en un *stock* de documentos, efectuada a partir de especificaciones de los temas”¹⁴⁷. B.

¹⁴⁷ VICKERY, B.C. *Techniques modernes de Documentation*. Paris: Duod, 1962, p. ix. Cit. por LÓPEZ YEPES, J. *La Documentación como disciplina. Teoría e Historia*. 2ª de. act. y ampl. Pamplona: EUNSA, 1995, p.156.

C. Vickery, continuador de la teoría de Mooers en su libro *On retrieval systems theory*¹⁴⁸, expresa que “la búsqueda documental se limita a la operación por la cual los documentos son escogidos en el *stock* a petición del usuario. La búsqueda documental se basa, pues, esencialmente, en la estructura y la utilización de los medios de selección de las informaciones documentarias contenidas en un *stock* para responder a preguntas”¹⁴⁹. De estas afirmaciones puede deducirse que la búsqueda o recuperación de la información es el conjunto de *métodos* y *operaciones* encaminadas a satisfacer las necesidades específicas de información de un determinado usuario. Estos métodos y operaciones de búsqueda estarán mediatizados por el *sistema*. Así, en función de éste, los procedimientos podrán ser *manuales*, si se consultan directamente las fuentes de información impresas (directorios, catálogos, bibliografías, boletines de resúmenes, etc.) o bien *automatizados*, es decir, cuando se utilizan las bases de datos, lo que implica el uso del ordenador y, consecuentemente, el previo almacenamiento de la información en forma automatizada. En este sentido, para Mercedes Caridad¹⁵⁰ opina que el origen de esta aplicación informática en el entorno documental hay que buscarla en la operación de búsqueda y recuperación de la información, aunque hoy en día pueden encontrarse sistemas que automatizan total o parcialmente el proceso documental.

Son precisamente los años cincuenta cuando aparecen los sistemas de información automatizados, denominados *no convencionales* por “emplear o incorporar nuevos principios para la organización de materias o en la utilización de equipos informáticos

¹⁴⁸ VICKERY, B.C. *On retrieval systems theory*. 2nd. de. London: Butterworths; New York: Sohe String Press, 1965.

¹⁴⁹ Idem, p. 2-3. Cit. por LÓPEZ YEPES, J. *La Documentación como disciplina. Teoría e Historia*. 2ª de. act. y ampl. Pamplona: EUNSA, 1995, p. 157.

¹⁵⁰ CARIDAD, Mercedes. Introducción al estudio de la Teledocumentación. En: *Fundamentos de Información y Documentación*. 2ª de. Madrid: Eudema, 1990, p. 433.

para el almacenamiento y la búsqueda”¹⁵¹. Tefko Saracevic¹⁵² resume en ocho las características y objetivos básicos de la Information Retrieval y estos nuevos sistemas de información: 1) surgen para hacer frente a la “explosión de información”; 2) sus objetivos básicos son la relevancia y la orientación al usuario; 3) utilización de la indización coordinada para la representación y recuperación de la información; 4) el lenguaje natural como base para la recuperación de la información; 5) la tecnología de la información es un aspecto integral de los sistemas de recuperación de la información; 6) el enfoque sistémico se utiliza como método para el análisis, diseño y actuación de los sistemas de recuperación de la información; 7) las reglas de la oferta y la demanda sirven de base para su justificación económica; 8) se consideran esenciales una provisión de servicios activa, una continua exploración de nuevos servicios y productos, y un marketing agresivo.

2.3.1. ELEMENTOS DE UN SISTEMA AUTOMATIZADO DE RECUPERACIÓN

Un sistema automatizado de recuperación posee cuatro elementos, como cualquier otro sistema automatizado: el *hardware* o componentes físicos, el *software* o programas que permiten controlar el proceso de datos, *personal especializado* que mantiene y hace uso del sistema y, por supuesto, los *datos* procesados, objeto del sistema.

a) Hardware

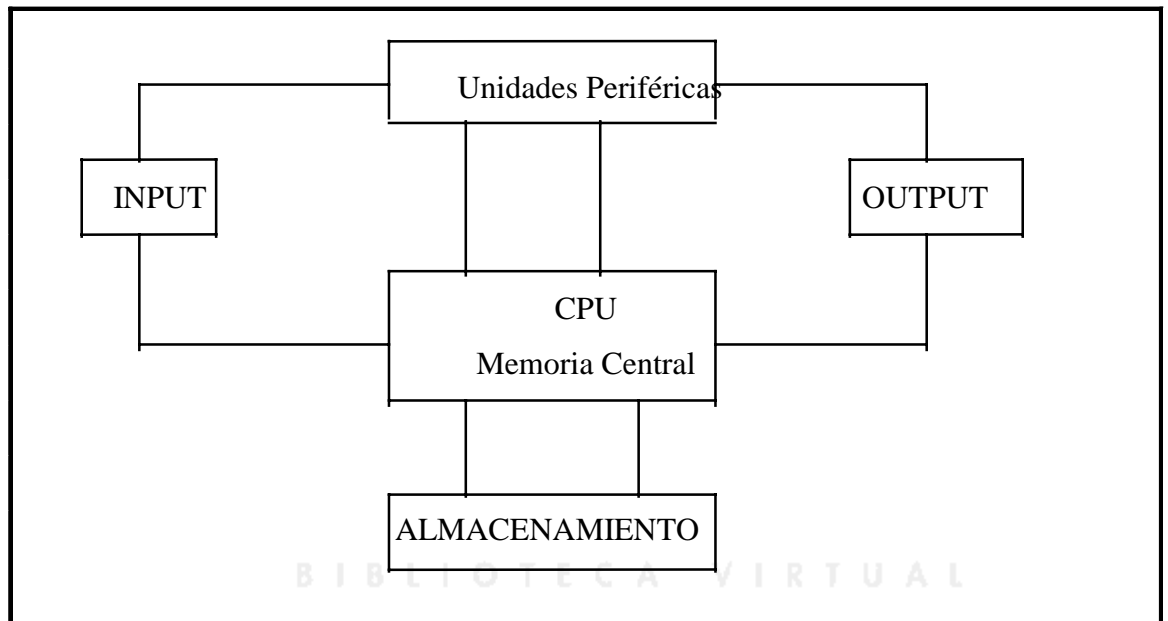
¹⁵¹ LILLEY, D.B. y TRICE, R.W. *A history of information science 1945-1985*. San Diego: Academic Press, 1989, p. 39.

¹⁵² SARACEVIC, Tefko. An essay on the past and future (?) of information science education.- I Historical overview. *Information processing and management*, 1979, vol. 15, nº 1, p. 8.

El ordenador procesa la información que se introduce en él, la almacena y la gestiona con el objeto de poder ser recuperada según las necesidades de los usuarios finales. Básicamente, el equipo físico de tratamiento electrónico de datos lo componen las siguientes unidades¹⁵³:

- Unidad Central de Proceso (CPU) donde se realiza el procesamiento de la información previamente introducida.
- Circuitos de transmisión de datos entre la CPU y los demás elementos del sistema.
- Memoria interna. Almacena los programas que se están ejecutando y los datos que se están procesando.
- Memorias auxiliares o externas. Sirven para almacenar datos antes y después de su tratamiento. Los soportes que la almacenan pueden ser discos duros, disquetes magnéticos, discos ópticos, etc.
- Unidades periféricas, como el teclado, la pantalla, la impresora o lector óptico, que permiten la entrada (INPUT) y/o la salida (OUTPUT) de los datos.

¹⁵³ CARIDAD, M. Introducción al estudio de la Teledocumentación. En: *Fundamentos de información y documentación*. Madrid: Eudema, 1990, pp. 433-435.



Equipo físico de tratamiento electrónico de datos

b) Software

Permite al ordenador ejecutar las tareas de la gestión documental y, por ende, las operaciones específicas encaminadas a la búsqueda y recuperación de la información. Entre estos programas se encuentran los *Sistemas de Gestión de Bases de Datos* (también conocidos con el acrónimo SGBD), configurándose como las aplicaciones más eficaces para las tareas específicas de recuperación de la información: la indización automática, creación de ficheros y las formas de consulta.

c) Usuarios del sistema

El conjunto de personas destinadas a la explotación y uso del sistema de recuperación de la información lo conforman, principalmente, profesionales de la documentación, bibliotecarios referencistas, etc. Aunque, hoy en día, las tendencias se dirigen a que sea el usuario final el que haga uso del sistema.

d) Los datos

El conjunto de datos y/o información objeto de recuperación por parte del usuario y correspondiente a un fondo documental, tema o problema determinado.

2.3.2. LAS BASES DE DATOS

2.3.2.1. CONCEPTO Y DEFINICIÓN

En la actualidad, las bases de datos constituyen la tecnología más avanzada para tratar grandes volúmenes de información, facilitando en gran medida las tareas propias del proceso documental y, en particular, el almacenamiento y recuperación de la información.

La expresión base de datos aparece por primera vez a mediados de los años sesenta en una conferencia organizada por el System Development Corporation en Santa Monica (California) para referirse a “un almacén de todos los datos de interés y valor para los usuarios del sistema de información”¹⁵⁴. Desde esta fecha el concepto ha evolucionado a medida que se han producido nuevos adelantos en la tecnología y nuevas exigencias en los usuarios. Según el Comité de Terminología de la FID, una base de datos es “un conjunto de datos

¹⁵⁴ COLL-VINENT, Roberto. *Información y poder: el futuro de las bases de datos documentales*. Barcelona: Herder, 1988, p. 75.

homogéneos, ordenados de una forma determinada, que se presenta normalmente en forma legible por ordenador y se refieren a una organización, materia o problema determinado”¹⁵⁵. Desde el punto de vista de la teoría general de sistemas de información, Lluís Codina concibe una base de datos como “un conjunto de informaciones relativas a una materia determinada del conocimiento, registradas en soporte legible por ordenador y articuladas en unidades lógicas determinadas *registros*”¹⁵⁶. Cada registro se encuentra estructurado según unos criterios precisos de manera que representa un objeto del mundo real (*entidad*). Desde esta perspectiva, una base de datos se entiende como un sistema de información estructurado que intenta modelar una parte de la realidad. Para este autor la ventaja que ofrece esta concepción frente a otras es que de ella se desprende con toda naturalidad el principio fundamental del diseño de una base de datos: saber que una base de datos ha de representar con la máxima fidelidad posible aquella parte del mundo que intenta modelar, según el modelo de sistemas mediante ordenadores de Walker¹⁵⁷.

Por otro lado, conviene precisar que el término base de datos suele utilizarse indistintamente tanto para referirse al conjunto de información que comparte características comunes como al Sistema de Gestión de Bases de Datos, aplicación que permite crear, modificar y mantener ese conjunto de informaciones al que nos referíamos antes. Se tratan de dos cosas conceptualmente diferentes que no deben confundirse: la primera acepción correspondería a la base de datos propiamente dicha, mientras que la segunda se referirá al programa informático (SGBD) que gestiona y manipula dicha base de datos. Una vez aclarado el concepto, nos detendremos, en primer lugar, en la

¹⁵⁵ Cit. por PÉREZ ÁLVAREZ-OSSORIO, José Ramón. *Introducción a la información y documentación científica*. Madrid: Alhambra, 1990, p. 40.

¹⁵⁶ CODINA, Lluís. *Sistemas d'información documental*. Barcelona: Editorial Pòrtic, 1993, pp. 69-71.

clasificación de bases de datos y, seguidamente, en los modelos básicos de Sistemas de Gestión de Bases de Datos, ya que no todos ofrecen las mismas ventajas para la recuperación documental.

2.3.2.2. CLASIFICACIÓN DE BASES DE DATOS

La naturaleza de la información va a determinar la organización de los datos que se van a almacenar y procesar, así como las formas en las que éstos se van a poder recuperar. Y, todo ello condicionado por las características y necesidades de los usuarios a las que van destinadas las bases de datos.

Según el *contenido*, las bases de datos pueden clasificarse en¹⁵⁸:

a) Referenciales. Son aquellas que contienen datos que hacen referencia a información primaria. las bases de datos referenciales suelen dividirse a su vez en:

Bibliográficas si contienen datos relativos a la descripción de documentos primarios, tales como libros, artículos de revista, patentes, informes, discos, documentos audiovisuales, etc.

Directorios, cuando describen información que hacen referencia a personas, entidades u organizaciones.

b) *Factuales o Fuente* denominadas así porque contienen los datos originales o información primaria. Este grupo se subdivide en:

¹⁵⁷ WALKER, D.W. *Sistemas de información basados en ordenador*. Barcelona: Marcombo, 1991.

Textuales cuando contienen los documentos íntegros en toda su extensión, tal y como han sido publicados.

Numéricas. Almacenan datos de naturaleza numérica como estadísticas, censos o encuestas.

Alfanuméricas, o también llamadas textual-numéricas, contienen tanto información textual como datos numéricos.

Icónicas. Presentan la información en forma de imágenes, ya sean gráficos, mapas, fotografías, etc

Multimedia, donde la información se recoge en forma de imagen y/o sonido.

En función del *tipo de público* al que vayan destinadas, pueden distinguirse:

a) *Bases de datos generales dirigidas al gran público* que almacenan información de tipo general.

b) *Bases de datos especializadas* dirigidas a usuarios de áreas desconocimiento muy diversas como: científicas, tecnológicas, sociales, económicas, jurídicas, médicas, etc.

¹⁵⁸ MOSCOSO, Purificación. Bases de datos documentales. En: *Manual de información y documentación*. Madrid: Pirámide, 1996, pp. 393-394.



Bases de datos según el contenido

En cuanto al *ámbito territorial* se distinguen entre:

- a) *Bases de datos internacionales*
- b) *Bases de datos de carácter nacional*, afectan a todo el territorio nacional
- c) *Bases de datos regionales*: su ámbito de aplicación está dirigida a un municipio o provincia

Por último, según la posibilidad de acceso o difusión, las bases de datos pueden ser:

- a) *Públicas*, comercializadas por productores o distribuidores. Son de libre acceso mediante pago de una tarifa.

b) *Privadas*, cuya difusión queda restringida al personal de la entidad privada o administrativa que las ha creado.



2.3.2.3. SISTEMAS DE GESTIÓN DE BASES DE DATOS (SGBD)

Los datos estructurados y almacenados en soporte legible por ordenador necesitan de una aplicación informática que permita crear, mantener y explotar esos datos. Esta aplicación informática es el denominado *Sistema de Gestión de Bases de Datos* o conjunto coordinado de programas y procedimientos que suministran los medios necesarios para describir y manejar los datos integrados en las bases. Básicamente, los programas de gestión de bases de datos realizan las siguientes *funciones*¹⁵⁹ de:

- *Administración* de bases de datos que comprende todas aquellas operaciones relacionadas con el diseño y creación de una base de datos, tareas como: definición del modelo de registro, determinación de la política de indización, etc.
- *Mantenimiento*. Encierra operaciones relacionadas con la alimentación de las bases de datos, bien por adición, modificación o supresión de registros.
- *Indización y recuperación*. Se ocupa de las tareas de creación de índices automáticos y de las modalidades de consulta.
- *Difusión de la información*, que abarca tanto los aspectos relacionados con la salida de registros como las operaciones que se ocupan de la incorporación y adaptación de ficheros externos de registros.

Ahora bien, dependiendo de la naturaleza de la información que se quiera gestionar, es decir, estructurada o no, se suelen distinguir dos modelos de sistemas de gestión de bases de datos: el *modelo relacional* y el *modelo documental* o textual.

¹⁵⁹ CODINA, L. *Op. Cit.*, p. 82.

a) Modelo Relacional

El modelo relacional es el más utilizado en el contexto empresarial y administrativo ya que la información que se gestiona en estos centros está altamente estructurada. Quiere decir esto que la naturaleza de la información y la extensión de los datos es previsible y estadísticamente regular, frente a la información documental, de naturaleza, extensión y formato imprevisible.

Los elementos esenciales que caracterizan un modelo relacional son: la estructura única de datos, llamada *tabla*, y las *operaciones relacionales* que afectan a las tablas. Las tablas, unidades bidimensionales, están formadas por *filas* (tuplas) y *columnas* homogéneas. Las filas representan las entidades (registros) y las columnas los atributos (campos). Los datos almacenados en las tablas se relacionarán a través de operaciones basadas en el álgebra y en el cálculo relacional, que implican operaciones lógicas entre conjuntos. Básicamente, estas operaciones son tres: la *selección* crea un subconjunto de las filas de las tablas que cumple con una o varias condiciones dadas, creándose una nueva tabla, normalmente temporal. La *proyección* es una operación similar a la selección que en lugar de actuar sobre las filas lo hace sobre las tablas. Y, por último, la *unión* crea una nueva tabla a partir de otras dos o más tablas¹⁶⁰.

El modelo relacional se configura como el modelo más sólido y seguro que existe actualmente para gestionar información estructurada. Sin embargo, este modelo no sirve para la gestión de información documental, precisamente, según indica L. Codina¹⁶¹, por la no

¹⁶⁰ GONZÁLEZ, Javier. ¿Qué es una base de datos?. *PC World*, marzo 1992, pp. 92-101.

¹⁶¹ CODINA, L. *Ibíd.*, 1993, pp. 84-87.

estructuración de la información que contienen los documentos que impide que pueda ser representada mediante campos homogéneos.

b) El Modelo Documental

El modelo documental o textual, modelo también conocido como *Sistema de Recuperación de Información (SRI)*, se configura para que sea capaz de recuperar documentos selectivamente a partir de cualquiera de los aspectos de su contenido. El modelo documental se basa en tres elementos fundamentales¹⁶²:

- Un modelo de registro que admite estructuras no homogéneas, es decir, con campos de longitud variable (autor, título, lugar de publicación, etc.)
- Un motor de recuperación de información que se caracteriza por los siguientes elementos: 1) capaz para crear una estructura de índices o ficheros inversos que reflejan las propiedades estadísticas de los textos de los documentos y que permiten que cada palabra del registro sirva de punto de acceso al documento; y 2) una lista de palabras vacías o términos no indexados.
- Un lenguaje de interrogación basado en operadores booleanos y operadores de proximidad y que eventualmente, es capaz de ordenar los documentos recuperados por su grado de probabilidad de satisfacer una necesidad de información concreta.

¹⁶² CODINA, L. *Ibid.*, 1993, pp. 87-88.

2.3.2.4. ESTRUCTURA DE UNA BASE DE DATOS DOCUMENTAL

Según se ha indicado anteriormente, una base de datos es un conjunto de datos almacenados en soporte informático, estructurados y organizados de tal manera que éstos puedan ser recuperados según las necesidades del usuario. La estructura y organización de la base de datos no se lleva a cabo de manera caprichosa sino que intenta dar respuesta a las necesidades de los potenciales usuarios. Es por ello, que creemos conveniente detenernos en los elementos que la componen. Básicamente, los elementos que integran una base de datos son los siguientes:

- *El registro* es la unidad básica de la que se compone una base de datos y es la representación de un documento para su almacenamiento y subsiguiente recuperación. Así, por ejemplo, un registro bibliográfico vendría a ser información bibliográfica estructurada referente a una monografía o artículo de revista. Esta información estructurada (autor, título, fecha de publicación, etc.) se la conoce por el nombre de *campos*.

- *Campo* es la subdivisión del registro e indica cada uno de los aspectos del registro que se quieran señalar. Los campos típicos y más comunes de un registro bibliográfico son: autor, título, lugar de publicación, editorial, año de publicación, y campos referentes al contenido como el de palabra clave o resumen. Cada tipología documental requerirá una serie de campos que no tienen por qué ser comunes a todos los tipos de documentos almacenados en la base de datos documental.

- *Formato*. Disposición de los datos en un soporte de almacenamiento, entrada o salida y el código o conjunto de instrucciones que dirigen esa

disposición¹⁶³. Existen una gran variedad de formatos, ya estén normalizados por las instituciones de estandarización (formato MARC¹⁶⁴) como diseñados por productores y/o distribuidores de base de datos. El diseño del formato requiere un estudio previo de las necesidades de los futuros usuarios. En una base de datos documental los campos pueden ser de longitud variable o fija, precisamente por la utilización de códigos o etiquetas que permiten que el sistema agrupe las diferentes unidades de acuerdo con el tipo de información que contienen y, en consecuencia, serán los elementos clave para la recuperación de registros¹⁶⁵.

2.3.2.5. MÓDULO FUNCIONAL DE RECUPERACIÓN DE LA INFORMACIÓN

Como ya hemos tenido ocasión de adelantar, los sistemas de gestión de bases de datos documentales realizan una serie de funciones (administración, mantenimiento, indización y recuperación, y difusión de las bases de datos). Pero es el módulo de recuperación de la información, junto con el sistema de indización, el que se configura como núcleo de todo Sistema de Gestión Documental (SGD) ya que la calidad y prestaciones que ofrezcan estas funciones dependerá la validez de toda la base de datos, cuya finalidad es que el usuario y el sistema se comuniquen.

¹⁶³ Glosario ALA de Bibliotecas y Ciencias de la Información. Citado por RODRÍGUEZ MUÑOZ, J.V. y ASENSI ARTIGA, V. Los formatos y su aplicación práctica. En: *Catalogación de documentos*. Madrid: Síntesis, 1994, p. 333.

¹⁶⁴ ISO 2709-1981.

¹⁶⁵ RODRÍGUEZ MUÑOZ, J.V. y ASENSI ARTIGA, V. *Ibíd.*, p. 333.

Generalmente, el SGD incorpora las siguientes formas de recuperación de información¹⁶⁶:

- *Consulta a través de índices* o ficheros invertidos que genera el propio sistema (indización automática). Un fichero invertido consiste en una lista ordenada de todas las palabras significativas (palabras clave o descriptores) existentes en la base de datos. De esta manera puede conocerse el número de veces que aparece ese término en la base, los registros en los que figura y el campo donde aparece. El número de índices variará en función del programa de gestión, pudiéndose crear un índice único o índices independientes para los distintos campos. Junto a estos ficheros índices se suele crear un *fichero de palabras vacías* (o antídicionario), que actúa como filtro y que recoge aquellos términos carentes de significado para la recuperación de la información como artículos, preposiciones, conjunciones, pronombres, etc.

- *Consulta asistida por menús*. Se trata de un sistema de búsqueda simplificada en el que el usuario no necesita conocer la sintaxis del lenguaje de interrogación de la base, sino tan sólo responder a las preguntas que el sistema le propone.

- *Consulta al fichero lineal o principal*. No se trata propiamente de una forma de búsqueda, y consiste en poder avanzar o retroceder por entre los registros de la base de datos, ofreciendo una visión general de su contenido y estructura. Los anglosajones utilizan el término *browse* para referirse a esta opción.

- *Consulta a través de un lenguaje de interrogación*. Casi todos los SGD disponen de un lenguaje de interrogación más o menos

¹⁶⁶ CODINA BONILLA, L. y ABADAL FALGUERAS, E. Gestió documental amb microordinadors: característiques, estructura y tecnología dels sistemes de gestió documental. *Item*, 1992, nº 11, pp. 87-90.

sofisticado, que sigue una sintaxis propia (no normalizada) y utiliza una serie de operadores par combinar los términos de búsqueda. A través de los operadores e instrucciones que incluye el lenguaje de interrogación, se ponen en contacto las consultas de los usuarios (formuladas en lenguaje natural) con la lógica interna del sistema. Así, para consultas complejas donde se tenga que combinar más de un concepto, se utilizarán los siguientes procedimientos:

- *Los operadores lógicos*, basados en el álgebra de Boole, que permiten la *intersección*, *unión* y *exclusión* de los términos utilizados como descriptores.

- *Intersección* (Y, AND, ET). Este operador permite obtener un conjunto de documentos (C) que contienen simultáneamente los términos A y B.

- *Unión* (O, OR, OU). Permite obtener un nuevo conjunto (C) formado por los documentos que contienen indistintamente el término A o B.

- *Negación* (NO, NOT, NON). Permite formar un conjunto con los documentos que contienen el término A pero no el B.

- *Operadores de proximidad*. Posibilitan poner un orden, una posición y/o una distancia relativa entre los términos de la ecuación de búsqueda.

- *Búsqueda limitada a un campo concreto*. Permite realizar la búsqueda en aquel(los) campo(s) donde aparezca el término.

- *Búsqueda por rango*. Permite buscar registros según criterios cuantitativos. Los principales rangos que se utilizan son “mayor que”, “menor que”, “igual que”.

- *Ponderación*. Consiste en destinar en la búsqueda un coeficiente o índice de ponderación para cada descriptor, coeficiente calculado según la importancia del término en el documento y/o en la consulta.
- *Truncamiento y máscaras* (*, \$, ?). Este método permite buscar las raíces de las palabras o bien la cadena de caracteres deseada no considerando los sufijos o prefijos, según se realice el truncamiento a la derecha o a la izquierda, respectivamente.
- *Búsqueda a partir de un tesoro en línea o con sinónimos*. Algunos programas permiten traducir el término de búsqueda a un término autorizado que se encuentra en un tesoro, o bien ampliar la búsqueda mediante el reenvío.

La efectividad de cada una de estas características funcionales del sistema de recuperación dependerá en cierta medida de la *estrategia de búsqueda*, función necesaria para obtener mejores resultados en la relación del sistema con el usuario.

2.3.2.5.1. LA ESTRATEGIA DE BÚSQUEDA DOCUMENTAL

El problema fundamental de un sistema de recuperación es la representación correcta de la petición expuesta por el usuario, es decir, que la pregunta formulada por éste, normalmente en lenguaje natural, sea entendida por el sistema de recuperación, que utiliza un lenguaje máquina, y, de esta manera, posibilitar la comunicación entre ambos. Para ello, se utilizará el lenguaje de interrogación o el conjunto de instrucciones tecnológicas características y propias del sistema, vistas anteriormente. Ahora bien, el uso de este lenguaje debe aplicarse mediante una serie de procedimientos y tácticas denominada *estrategia*

de búsqueda. Nuria Amat¹⁶⁷ diferencia cuatro etapas en la estrategia de búsqueda:

a) Entrevista con el usuario con el fin de poder definir la consulta demandada. Para ello, el documentalista se servirá de un cuestionario que reflejará las delimitaciones de la consulta como: la tipología documental deseada, el objeto de la consulta (investigación, actualización de conocimientos, etc.), idiomas de los documentos, etc.

b) A partir de esta entrevista, se elaborará un *perfil*: operación donde se identifican los conceptos y la expresión de las relaciones entre los términos¹⁶⁸.

c) Traducción de dichos conceptos al lenguaje documental de la base de datos interrogada. Para ello, se consultarán y utilizarán los descriptores específicos en el tesaurus, lista de palabras clave, etc.

d) Los términos de búsqueda se combinan formando la *ecuación de búsqueda*, mediante los métodos booleanos, sintácticos o de proximidad, de rango, etc.

¹⁶⁷ AMAT, Nuria. *Op.Cit.*, 1988, pp. 414-427.

¹⁶⁸ PÉREZ ÁLVAREZ-OSSORIO, J.R. *Op.Cit.*, p. 60.

2.3.2.5.2. VALORACIÓN DE LA BÚSQUEDA DOCUMENTAL

El propósito de un sistema de recuperación de información (SRI) no es proporcionar mucha información sobre un tema determinado, sino el de suministrar información relevante. Pero la relevancia no es una propiedad absoluta de la información sino una co-producción entre el documento y el usuario, aunque su medida se realizará siempre a partir de los documentos recuperados¹⁶⁹. Por tanto, para evaluar la eficacia de un SRI hay que determinar el grado en que éste satisface las necesidades informativas de sus usuarios.

Tradicionalmente, dos parámetros han sido utilizados para medir la relevancia de las recuperaciones¹⁷⁰: la *exhaustividad* o cociente entre el número de documentos relevantes recuperados y el número total de documentos relevantes contenidos en el sistema, y la *precisión* o cociente entre el número de documentos relevantes y el número total de documentos recuperados. Ambos conceptos son inversamente proporcionales. Si la exhaustividad es demasiado elevada la precisión de los términos disminuye y provocará el llamado *ruido documental*, es decir, documentos seleccionados por el sistema que no responden a la pregunta. Sin embargo, si la exhaustividad es demasiado reducida la precisión aumenta y aparecerá el fenómeno del *silencio documental* o conjunto de documentos no recuperados por el sistema que responden a la pregunta.

La medición de la relevancia en términos binarios, es decir, un documento es relevante (1) o no relevante (0) para el sistema, procede de la metodología ideada en la década de los cincuenta por el Cranfield

¹⁶⁹ BONILLA ÁLVAREZ, S. Información y relevancia. Una hipótesis acerca de cómo procesamos los seres humanos la información lingüística. *Revista Española de Documentación Científica*, 1996, vol. 19, nº 4, p. 392.

¹⁷⁰ LANCASTER, F.W. *El control del vocabulario en la recuperación de información*. València: Universitat, Servei de Publicacions, 1995, pp. 149-169.

Institute of Technology, método que ha asentado las bases en la investigación para la evaluación de los sistemas de recuperación llegando a ser considerado como el *paradigma físico* que evalúa los mecanismos operativos del sistema¹⁷¹. Sin embargo, una de las mayores objeciones que ha recibido el método Cranfield es la escasa consideración del papel del usuario final y el de sus necesidades informativas en la valoración de los resultados de la búsqueda documental¹⁷², aspecto que no puede ser ignorado en la evaluación ya que todo SRI trata de satisfacer las necesidades humanas de información. Efectivamente, las últimas tendencias investigadoras se dirigen a percibir el proceso de búsqueda de información en términos cognitivos, donde los procesos mentales del usuario se involucran con los mecanismos operativos del sistema. Desde esta perspectiva, las necesidades de información del usuario se conciben como construcciones mentales dinámicas que pueden variar durante el proceso de búsqueda. Siguiendo esta perspectiva cognitiva, el concepto de relevancia se relativiza y se tienen en cuenta, para su medición otros aspectos como el contexto situacional del usuario y el de sus necesidades de información. Es la denominada *relevancia situacional* o valoración que se refiere a la utilidad, provecho o beneficio de la información solicitada (output) en relación con una serie de elementos (los objetivos, intereses, contexto laboral o problema determinado) que determinaron esa búsqueda de información¹⁷³.

¹⁷¹ ELLIS, David. The physical and cognitive paradigms in information retrieval research. *Journal of Documentacion*, 1992, vol. 48, nº 1, pp. 45-54.

¹⁷² ROBERTSON, S.E. y BEAULIEU, M. Research and evaluation in Information Retrieval. *Journal of Documentacion*, 1997, vol. 53, nº 1, pp. 51-57.

¹⁷³ BORLUND, Pia e INGWERSEN, Peter. The development of a method for the evaluation of interactive informations Retrieval Systems. *Journal of Documentacion*, 1997, vol. 53, nº 3, pp 225-250.

2.4. DIFUSIÓN DE LA INFORMACIÓN

La *difusión* de la información consiste en transmitir al usuario la información que necesita, o bien, ofrecerle los medios para obtenerla. La difusión es el último eslabón de la cadena documental, el desenlace de todo el proceso documental, y la razón de ser la Documentación. Para F. del Valle Gastaminza la difusión “es el proceso por el cual el centro de documentación transmite al usuario la información o los documentos que la contienen o bien, mediante distintos productos documentales, le da la posibilidad de conocer la existencia de esos productos y le facilita su obtención”¹⁷⁴. Según esta definición se desprende que la difusión puede realizarse bien a través de *servicios documentales* ofrecidos por la unidad de información, o bien, mediante *productos documentales*.

a) *Servicios documentales*. Entre los servicios de información destinados a la difusión, se encuentran:

- *Búsquedas bibliográficas y búsquedas de información*. Se trata de responder a una demanda concreta de un usuario sobre un tema específico. La búsqueda bibliográfica retrospectiva (BR) podrá realizarse sobre productos secundarios internos o externos, es decir, generados o no por la unidad de información. Estas herramientas de recuperación y, consecuentemente, de difusión, son catálogos, bibliografías, índices, boletines de sumarios o de resúmenes, productos que bien pueden encontrarse de forma impresa o automatizada.

¹⁷⁴ VALLE GASTAMINZA, Félix del. Difusión de la información. Metodología y descripción de los instrumentos informativos. En: *Manual de información y documentación*. Madrid: Pirámide, 1996, p. 477.

En cuanto al *acceso del documento primario* se puede proceder de varias maneras:

- *Consultar la propia unidad de información* donde se solicita el documento. Si ésta es de libre acceso el propio usuario, previa identificación de los documentos mediante la consulta en los catálogos, buscará en los estantes. Si la unidad de información es de acceso controlado, el usuario consultará los catálogos y, tras cumplimentar la ficha de petición, el servicio le entregará los documentos solicitados.

- *El préstamo*. Este servicio ofrece la posibilidad de que los documentos puedan ser consultados fuera de la unidad de información durante un período determinado. Si el centro no dispone del documento solicitado por el usuario, aquél podrá pedirlo prestado a otro centro. Este tipo de préstamo es el que se conoce como *préstamo interbibliotecario*.

- *Obtención de documentos externos*. Existen en el mercado varios centros especializados en suministrar información original a través de fotocopias o copias en soporte informático. Algunos de estos centros son: el Document Supply Centre de la British Library en Gran Bretaña, el Institut de l'Information Scientifique et Technique en Francia o el Centro de Información y Documentación (CINDOC) del Consejo Superior de Investigaciones Científicas en España.

- Los *servicios de traducción* son un medio esencial de difusión de los documentos primarios. El *Centro Internacional de Traducciones* (ITC) se constituye para tal fin, esto es, recopila, procesa y difunde información sobre las traducciones científicas y

técnicas existentes de cualquier idioma original a los idiomas occidentales.

- *Servicio de Difusión Selectiva de la Información (DSI)*. Es un procedimiento que permite suministrar a cada usuario o grupo de usuarios información bibliográfica de forma periódica y que corresponde a un perfil o necesidad informativa individualizada.

- *Acceso a bases de datos y catálogos externos*. La unidad de información posibilita al usuario el manejo de bases de datos, bien editadas en CD-ROM o accesibles a través de las redes de telecomunicación. En este último caso, la red Internet, entre sus múltiples posibilidades, ofrece acceso a OPAC's (Online Public Access Catalogue).

b) Productos documentales

Los productos documentales con carácter de difusión son confeccionados en la unidad de información. La variedad de éstos estará condicionadas por el tipo de usuarios del sistema. Los más característicos son los de carácter secundario: bibliografías, catálogos, índices, etc. Pueden adoptar formas muy diferentes en cuanto a su contenido, presentación, periodicidad y objetivos.

Estos servicios y productos documentales de difusión no son comunes a todas las unidades de información. Según apunta C. Guinchat y M. Menou¹⁷⁵ existen una serie de limitaciones y dificultades de la difusión. Entre ellas, pueden mencionarse las relacionadas con la

¹⁷⁵ GUINCHAT Y MENO, *Op.Cit.*, p 360-361.

variedad de necesidades de usuarios, las limitaciones financieras de las unidades de información o las limitaciones técnicas.

Otros problemas relacionados con la difusión inciden en ciertas cuestiones jurídicas, en concreto, con los derechos de autor. Este instrumento legal protege a los autores de obras, y aunque la reproducción de éstas pueda hacerse siempre que se cumplan ciertas condiciones¹⁷⁶, puede constituir una limitación para la circulación y difusión de documentos.

2.5. NORMALIZACIÓN DOCUMENTAL

La normalización en el campo de la Documentación se concibe como una actividad ineludible para la transferencia de la información. Su aplicación en el contexto documental permite la unificación y simplificación de métodos, criterios y técnicas de trabajo con el fin último de facilitar la comunicación científica. En este sentido es definida por A. García Gutiérrez, para quien la *normalización documental* es “la organización racional de los conocimientos y sus soportes y el tratamiento y dinamización del conjunto acumulado en ellos”¹⁷⁷. Y la función que presta es, según la Asociación Francesa de Normalización (AFNOR), establecer reglas que aseguren la interconexión de sistemas y faciliten el tratamiento y la transferencia de la información. Estas reglas son aplicables tanto a las operaciones como a los productos documentales¹⁷⁸.

¹⁷⁶ Según la ley de Propiedad Intelectual, Ley 11 de noviembre 1987, nº 22/1987, BOE, 17 de noviembre de 1987, nº 275, pp. 4755-4772. Entre las condiciones se encuentran: para uso individual, sin ánimo de lucro, con fines didácticos, de investigación y de estudio.

¹⁷⁷ GARCÍA GUTIÉRREZ, Antonio Luis. Normalización general y documental: concepto, historia e instituciones. *Documentación de Ciencias de la Información*, 1985, vol. 9, p. 94.

¹⁷⁸ AFNOR. Au Sicob 78. L'actualité des normes françaises de l'informatique et de l'organisation de bureau. *Documentaliste*, 1974, vol. 16, nº 1, p. 37.

Efectivamente, la normalización está presente en cada una de las fases del proceso documental (selección, tratamiento y difusión), distinguiéndose así, normas relativas al vocabulario, a los soportes, a la identificación de documentos y a los procedimientos. En este sentido, son diversos los organismos que se encuentran involucrados en la normalización documental. A nivel nacional pueden mencionarse: AFNOR (Association Française de Normalisation) en Francia, AENOR (Asociación Española de Normalización) en España, ANSI (American National Standards Institute) en Estados Unidos, BSI (British Standards Institute) en Gran Bretaña, o DIN (Deutsches Institut für Normung) en Alemania. En el ámbito internacional se encuentran la ISO (International Organization for Standardization), UNESCO (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura), la FID (Federación Internacional de Documentación e Información) e IFLA/IFAB (Federación Internacional de Asociaciones Bibliotecarias). Veamos pues, algunas normas más significativas del proceso documental.

a) Normalización de la terminología documental

La normalización de la terminología documental ejerce un papel de primer orden, permitiendo a los especialistas de la documentación de diversos países, expresar sus ideas y conceptualizar la realidad. La ISO, a través de su Comité Técnico 37 “Terminología: principios y coordinación” (CT37) tiene la responsabilidad de establecer principios y métodos de normalización terminológica. Entre sus normas publicadas sobre el vocabulario utilizado en documentación e información se encuentra la ISO 5127¹⁷⁹.

¹⁷⁹ La norma ISO 5127 recoge los vocabularios referentes a las siguientes parcelas de la documentación e información: 1) Conceptos básicos; 2) Documentos tradicionales; 3) Documentos icónicos; 4) Documentos de archivo; 5) Adquisición,

b) Normalización del análisis documental

b1) Normalización del Análisis de la Forma

En este contexto las operaciones destinadas al *Análisis de la Forma* se encuentran altamente unificadas debido al consenso de los acuerdos internacionales. Entre las más significativas se destacan:

- Formato internacional de fichas catalográficas, utilizado para el catálogo de fichas sueltas y cuyas dimensiones son 7,5 x 12,5 mm¹⁸⁰.
- Códigos para la identificación de documentos. Para realizar el control bibliográfico se han definido números y códigos especiales que identifican con precisión al documento. Para la identificación de publicaciones unitarias la ISO elaboró el código ISBN (International Standard Book Number)¹⁸¹. Sin embargo, para la identificación de publicaciones periódicas y seriadas se utiliza el ISSN (International Standard Serials Number)¹⁸². Las ventajas que presentan ambos códigos en la comunicación de la información son¹⁸³: 1) eliminar las barreras lingüísticas, facilitando la circulación de la información; 2) evitar el riesgo de equivocar la identificación de las publicaciones; y 3) permitir identificar una

identificación y análisis de los documentos y sus partes; 6) Lenguajes documentales; 7) Recuperación y difusión de la información; 8) Almacenamiento, conservación y reproducción de los documentos; 9) Administración y organización de archivos, bibliotecas, centros de documentación y museos; 10) Aspectos legales de la documentación; y 11) Documentos audiovisuales.

¹⁸⁰ DIN 1461

¹⁸¹ ISO 2108-1978. Documentation. International standard book numbering (ISBN); Su equivalente española es UNE 50-130-94. Información y Documentación. Numeración internacional de libros (ISBN).

¹⁸² ISO 3297-1986. Documentation. International standard serial numbering (ISSN); Su equivalente española es UNE 50-107-89. Documentación. Número internacional normalizado de publicaciones en serie (ISSN).

¹⁸³ CARRIZO SAINERO, Gloria. Normalización. En: *Manual de fuentes de información*. Madrid: CEGAL, 1994, p. 103.

publicación periódica o una serie, sea cual sea el país, la edición o la lengua.

- Normalización de descripción bibliográfica de materiales¹⁸⁴. Diversas instituciones ya nacionales o internacionales, como públicas o privadas, han emitido normas al respecto. Las ISBD¹⁸⁵ (International Standards Book Description), publicadas por la IFLA para su Control Bibliográfico Universal (CBU) fueron establecidas por su Comité de Catalogación para facilitar la comunicación internacional de información bibliográfica¹⁸⁶. También la ISO contribuye a la normalización del análisis formal del documento con la norma ISO 690-1987¹⁸⁷, cuya adaptación española corresponde a la norma expedida por AENOR UNE 50-104-1994¹⁸⁸, norma que hemos aplicado en la elaboración de la base de datos objeto de esta investigación.

- Formato normalizado de intercambio internacional de datos bibliográficos legibles por ordenador, conocido por MARC(*Machine Readable Catalogue*)¹⁸⁹. El formato MARC creado en 1965 por la Biblioteca del Congreso, fue promocionado por la IFLA dentro del marco de información mundial UNISIST¹⁹⁰ para posibilitar el intercambio de información bibliográfica a nivel internacional en soporte informático. Este formato confiere

¹⁸⁴ Véase anexo de Normas de Descripción Bibliográfica.

¹⁸⁵ ISBD (M): Descripción de monografías; ISBD (S): Descripción de publicaciones en serie; ISBD (CM): Descripción de materiales cartográficos; ISBD (NBM): Descripción de materiales no librarios; ISBD (A): Descripción de publicaciones antiguas; ISBD (PM): Descripción de partituras musicales.

¹⁸⁶ FIAB/IFLA. *General International Standard bibliographic Description: annotated text*. London: IFLA, 1977.

¹⁸⁷ ISO 690-1987. Documentation -Bibliographic references- Content, form and structure. En ISO: *Documentation and information*. Genève: ISO, 1988, pp. 436-447

¹⁸⁸ UNE 50-104-94. *Referencias bibliográficas. Contenido, forma y estructura*. Madrid: AENOR, 1994.

¹⁸⁹ Formato cuya estructura fue adoptado por la ISO. ISO 2709-1981. Format for bibliographic information interchange on magnetic tape.

básicamente la posibilidad de identificar un registro bibliográfico en formato informático y, dentro de él, todos aquellos elementos informativos que van a ser necesarios en el contexto bibliotecario. Sobre el esquema del formato MARC se han creado formatos nacionales adaptados a sus normas de catalogación. Así, se encuentran: IBERMARC en España, UKMARC en Gran Bretaña, USMARC en Estados Unidos, CANMARC en Canadá, INTERMARC en Francia, o ANNAMARC en Italia.

b.2) Normalización del Análisis del Contenido

La normalización del *Análisis del Contenido* afecta a las operaciones de indización y a la confección de resúmenes.

- Para el *proceso de indización* se encuentra la norma ISO 5963-1985¹⁹¹, aprobada con la intención de provocar la utilización de una práctica unificada entre diferentes servicios de indización o de organismos documentales. Su equivalente española es la norma UNE 50-121-91¹⁹², norma de la que ya hemos dado cuenta al hablar del proceso de indización.

- Para la *operación de resumir* existen pautas o directrices de preparación y presentación de resúmenes de los documentos, bien preparados por los autores de los documentos primarios o por personas distintas de los autores que trabajan en servicios secundarios. Entre las normas se encuentran: ISO 214-1976¹⁹³ y

¹⁹⁰ Programa de Control Bibliográfico Universal.

¹⁹¹ ISO 5963-1985. Documentation. Methods for examining documents, determining their subjects, and selecting indexing terms. En: *Documentation and Information*. Genève: ISO, 1988, pp. 579-585.

¹⁹² UNE 50-121-91. *Documentación. Métodos para el análisis de documentos, determinación de su contenido y selección de los términos de indización*. Madrid: AENOR, 1991.

¹⁹³ ISO 214-1976. Documentation. Abstracts for publication and documentation. En: *Documentation and Information*. Genève: ISO, 1988, pp. 420-430.

su equivalente española UNE 50-103-90¹⁹⁴. María Pinto aprecia que estas normas son más ricas en consejos sobre la redacción del análisis que en recomendaciones sobre el verdadero análisis del contenido (proceso de análisis y síntesis).

c) Normalización de los lenguajes de indización

El incremento de la terminología en los ámbitos temáticos especializados ha estimulado una normalización terminológica organizada para frenar la dispersión designativa, normalización que debe garantizar la precisión comunicativa entre los especialistas. A tal efecto, se han creado las siguientes normas para la creación de tesauros (monolingües como multilingües) o vocabularios controlados para la recuperación de información: norma ISO 2788-1986, ANSI Z39.19-1974, AFNOR con su norma NF Z47-90, y AENOR con sus normas UNE 50-106-90; UNE 50-125-96.

¹⁹⁴ UNE 50-103-90. *Documentación. Preparación de resúmenes*. Madrid: AENOR, 1990.

d) Normalización en el contexto telemático

La proliferación de redes telemáticas en los años setenta va a provocar un proceso de normalización ante la anarquía imperante en la época donde cada fabricante tiene sus propias normas de comunicación lo que imposibilita el entendimiento con otros ordenadores que no sean los propios. Como solución a este problema en 1974 la ISO desarrolló un modelo abierto (es decir, no dependiente de ningún fabricante) de nominado OSI (*Open Systems Interconnection*) o Sistemas de Interconexión Abiertos al objeto de que todos los fabricantes fuesen adaptando sus normas de comunicaciones a él, y, de esta manera, proporcionar una estructura común para que ordenadores de distinta arquitectura y diferentes programas de aplicación pudieran conectarse.

El modelo OSI normalizado divide los complejos procesos de una comunicación en *siete niveles* relacionados jerárquicamente. Cada nivel se sustenta sobre el inmediatamente inferior y se comunica con el mismo nivel del sistema o sistemas distantes por medio de unas normas lógicas denominadas protocolos. Así, el primer nivel especificado es el *nivel físico*, que engloba los sistemas de telecomunicaciones. El segundo nivel es *el nivel de enlace de datos*, que controla la sincronización y coordinación de la transmisión. El tercer nivel o *nivel de red* se encarga de evitar que un ordenador reciba información dirigida a otro. Estos tres niveles de carácter eminentemente físico, son los que determinan lo que se denomina en general como redes. El cuarto nivel o *nivel de transporte* es el responsable del mantenimiento de la continuidad de la transmisión. Por último, los niveles de *sesión, presentación y aplicación* se encargan de controlar la secuencia de la comunicación con el otro dispositivo, de convertir los datos en diferentes formatos según se precise en el

destino y de proporcionar los recursos que requiera un determinado programa.

El modelo de referencia OSI se configura como un modelo de arquitectura de comunicaciones y muy útil para el entendimiento de la diversidad de redes. A partir de este modelo las normas se desarrollan e implementan. Sin embargo, hoy en día existe una gran discusión acerca de las ventajas e inconvenientes entre las normas OSI y los protocolos Internet. No es nuestro cometido detenernos en este debate pero sí indicar que los protocolos Internet (TCP/IP)¹⁹⁵ están ganando mercado frente a los protocolos OSI. Una de las razones hay que encontrarla en que la implementación de los protocolos TCP/IP se realiza primero, después se desarrolla la norma y por último, se configura el modelo.

2.6. TERMINOLOGÍA Y DOCUMENTACIÓN

En este trabajo no podemos pasar por alto el papel que juega la Terminología en el campo de la Documentación. Es más, ambas disciplinas se conciben como indisociables¹⁹⁶ ya que para realizar un trabajo terminológico se necesita acceder y manejar abundante documentación; por el contrario, la Documentación se sirve de la Terminología para los procesos documentales de la información especializada.

Efectivamente, la Terminología, base para la estructuración del conocimiento dentro de los lenguajes de especialidad y canalizadora para la transferencia del conocimiento, también constituye la base para

¹⁹⁵ Transmission Control Protocol/Internet Protocol

¹⁹⁶ RONDEAU, Guy. Terminologie et documentation. *Meta*, 1980, vol. 25, nº 1, pp. 152-170.

algunas técnicas del proceso documental, en concreto, la descripción, almacenamiento y recuperación de la información especializada¹⁹⁷. La utilización de una terminología estandarizada en estos procesos confiere a la documentación una sistematicidad y univocidad difícilmente alcanzable con el uso de la lengua común. De esta manera, mediante el empleo de herramientas documentales como tesauros y clasificaciones temáticas, se intenta ejercer un control tanto en la descripción del contenido de los documentos como en la recuperación de la información, apuntando hacia la eficacia del sistema documental.

2.7. REFLEXIONES SOBRE EL PROCESO DE INFORMACIÓN DOCUMENTAL

Como hemos venido insistiendo a lo largo de todo el capítulo, todo proceso documental desemboca en la difusión de la información. Para ello, todas las actividades involucradas en él tienden a optimizarse mediante la investigación, estudio y aplicación de nuevas teorías y tecnologías que permitan una mejora en sus resultados, es decir, asistir al usuario en sus demandas de información de la manera más rápida y eficiente.

La sociedad que nos ha tocado vivir, donde la información se ha convertido en un recurso imprescindible para el desarrollo de todo tipo de actividades, está cuestionando la figura del profesional de la información ante el avance tecnológico que provee una aparición continua de fuentes, soportes y sistemas de información de difícil actualización. Además, en los últimos años, esta tecnología está llegando al hombre de la calle, que desde su hogar o lugar de trabajo puede acceder directamente a grandes cantidades de información sin tener que recurrir a un intermediario. Sin

¹⁹⁷ CABRE, M. Teresa. La terminología. Teoría, metodología, aplicaciones. Barcelona: Editorial Antártida/Ampúries, D.L. 1993, pp. 111-114.

embargo, estos argumentos no constituyen amenaza alguna a la profesión, sino que sirven de estímulo para desarrollar otras facetas. Efectivamente, las innovaciones tecnológicas están abriendo nuevas vías y posibilidades en el tratamiento y transferencia de la información, posibilitando que las tareas mecánicas se agilicen y que el documentalista dedique mayor tiempo a conocer y controlar las necesidades informativas procedentes de una amplia variedad de usuarios (investigadores, científicos, empresarios, etc.), contemplando los contextos situacionales de las demandas. En este sentido, una de las soluciones es la que plantea Jean Michel¹⁹⁸: el cambio de mentalidad tanto del profesional de la documentación como del usuario, a través de una mayor comunicabilidad entre ambos. La aplicación de las técnicas de Programación Neurolingüística (PNL) en la comunicación documentalista-usuario es una de las apuestas investigadoras en este campo para perfeccionar en este binomio comunicativo¹⁹⁹.

Por otro lado, que el usuario se haga de esta tecnología desde su entorno doméstico o profesional no garantiza la correcta transferencia de la información demandada por aquél. Actualmente, el volumen de información que se oferta a través de las redes telemáticas es de tal envergadura que las tareas de discriminación informativa se dificultan, llevando al no especialista a la dispersión. En este sentido, la figura del documentalista se hace más necesaria que nunca convirtiéndose en un “hiperseleccionador” de la información, un intermediario que ejerce de filtro entre la masa documental ofertada y los usuarios finales.

¹⁹⁸ MICHEL, Jean. De la créativité en Documentation. Autres perspectives pour la formation. *Bulletin des Bibliothèques de France*, 1993, vol. 35, n° 2, p. 193. Cit. por ESPINOSA TEMIÑO, Blanca N. Documentación y creatividad. Aportaciones de la PNL (Programación Neurolingüística) o la interacción documentalista/usuario de la Documentación. *Documentación de las Ciencias de la Información*, 1995, n° 18, pp. 43-57.

¹⁹⁹ ESPINOSA TEMIÑO, Blanca N. *Op. cit.* pp.43-57.

Así, en esta era de la información altamente tecnificada, el estudio de las pautas conductuales y comunicativas del ser humano en los procesos de transferencia de la información es uno de los campos de investigación más interesantes en los que puede involucrarse el documentalista.



TRADUCCIÓN Y DOCUMENTACIÓN

Así, en esta era de la información altamente tecnificada, el estudio de las pautas conductuales y comunicativas del ser humano en los procesos de transferencia de la información es uno de los campos de investigación más interesantes en los que puede involucrarse el documentalista.



TRADUCCIÓN Y DOCUMENTACIÓN