

LAT
649

SEMINARIO SOBRE LOS SERVICIOS DE INFORMACION DOCUMENTAL
EN LA BIBLIOTECA UNIVERSITARIA Y ESPECIALIZADA.

GUANAJUATO, GTO., 8 AL 11 DE SEPTIEMBRE DE 1982.

BIBLIOTECA



CENTRO UNIVERSITARIO
DE INVESTIGACIONES
BIBLIOTECOLOGICAS

ANALISIS DE LA INFORMACION

Lic. Enrique Molina León.

Antes de comenzar mi exposición yo quiero pedir a ustedes su indulgencia: voy a mencionar aquí conceptos y puntos de vista, algunos de ellos muy elementales, conocidos por la mayoría de los presentes; si de todas formas lo hago, es porque creo que nunca resulta finalmente inútil recordar ciertas cosas sencillas pero que no dejan de ser fundamentales.

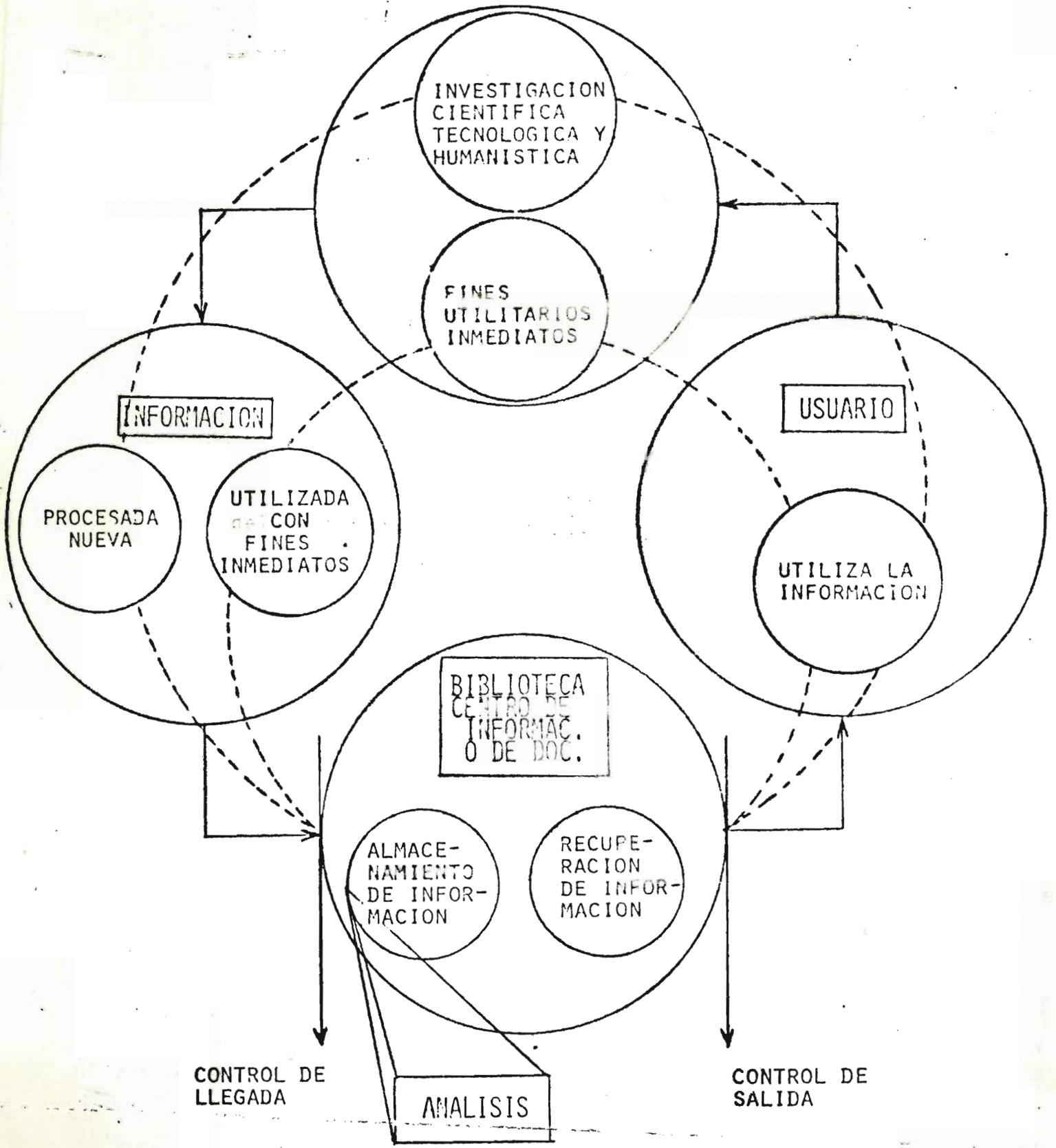
Una de las tareas más importantes dentro del quehacer bibliotecario es, quizá, la que tiene que ver con el análisis de la información. Es, al mismo tiempo, el paso inicial de todo servicio de información documental.

Con objeto de ubicar esta actividad dentro del proceso INFORMACION-BIBLIOTECA-USUARIO, y a manera de introducción, presento a continuación, esquemáticamente el CICLO DE LA INFORMACION.

Explicación del esquema: Los cuatro principales conjuntos muestran, sucesivamente, la INFORMACION, su llegada a la biblioteca o al centro de información o de documentación en donde es procesada: inicialmente analizada con miras a almacenarla y recuperarla ágilmente; el USUARIO, quien utiliza la información con fines utilitarios inmediatos o para realizar investigación y generar nueva información. El ciclo se cierra con el ingreso de la información al acervo de la biblioteca o del centro de

INFOBILA

EL CICLO DE LA INFORMACION



información o de documentación, ya sea para reintegrarse o para analizarse como información nueva e iniciar de nuevo el ciclo.

¿Qué se entiende propiamente por análisis de la información? Es toda operación o conjunto de operaciones que tienen por objeto representar un documento determinado bajo una forma diferente de la forma original, ya sea que se trate de traducirlo, resumirlo, indizarlo, etc., para facilitar su consulta o su localización por parte de los especialistas interesados.

"En la situación de un biblioteca convencional, nos dicen Becker y Hayes, el análisis consiste en la lectura del contenido de un artículo para asimilar las ideas expresadas en el texto y enseguida esquematizarlas por medio de encabezamientos escogidos dentro de listas especiales preparadas para este propósito." (1)

Se trata, pues, de un proceso intelectual de suma importancia, ya que está directamente relacionado con la eficiencia al momento de la recuperación de la información. Tanto más importantes se revelan las tareas del análisis de la información cuanto que la tendencia actual es de llegar a manejar unidades de información cada vez más pequeñas.

(1) BECKER, JOSEPH & ROBERT M. HAYES. Information Storage and Retrieval: tools, elements, theories. New York, Wiley, 1963. p. 110

En realidad la clasificación y la asignación de encabezamientos de materia, por ejemplo, implica mucho más que la simple habilidad para acomodar un libro dentro de un esquema preestablecido del conocimiento humano, o para encontrar un tema específico que lo represente.

Por otro lado, esta eficiencia en la recuperación de la información está también en relación directamente proporcional precisamente al grado de detalle en el análisis.

Tradicionalmente los libros han sido considerados como unidades independientes por la necesidad de clasificarlos - como tales; asimismo, para efectos de asignar encabezamientos de materia, se considera al libro como un todo. Sin embargo, las necesidades actuales de información nos obligan a analizar con mucho mayor detalle el material contenido en los libros.

Ranganathan anota con claridad esta idea al explicar el término "especializado" refiriéndose a la Bibliotecología - Especializada: es microscópica, más bien que macroscópica, - en cuanto que enfoca su atención a las partes específicas de los libros y de los artículos. (2)

(2) Comentado por: STRAUSS, LUCILLE J. en: Scientific and Technical Libraries. Their Organization and Administration. 2 ed. New York, Wiley, 1972. - p. 170

Si consideramos que el análisis de la información está encaminado, como ya se dijo, a facilitar el acceso de la misma al usuario, nos encontramos con que en la práctica no existe el análisis sin la clasificación, sin la indización, sin la asignación de encabezamientos de materia, etc., ya que de lo contrario no habría forma de aprovechar sistemáticamente el trabajo intelectual que implica el análisis.

Voy a referirme a continuación a algunas de las principales opciones que se presentan como actividades derivadas del análisis de la información.

En primer lugar, ¿es necesario establecer una versión resumida de cada documento? En caso afirmativo, ¿la operación debe ser confiada al hombre o a la máquina?

En segundo lugar, ¿se justifica la indización de cada documento analizado?; es decir, ¿es necesario representar brevemente el contenido por medio de descriptores o palabras claves, con fines diversos (almacenamiento en estantería, integración de catálogos, almacenamiento en un banco de datos automatizado, etc.)? En caso afirmativo, ¿esta indización debe realizarse partiendo del texto original o de su resumen? Además, ¿deberá llevarla a cabo el hombre o la máquina?

Como tercera opción, ¿qué utilidad existe en el empleo de un sistema simbólico de codificación para concretar la expresión del contenido de cada documento? En caso afirmativo, ¿a qué nivel de expresión debe llevarse a cabo la codificación (texto íntegro, resumen, indizado)? Y aquí también, ¿deberá realizarla el hombre o la máquina?

Finalmente, la eventual traducción del documento. Esta operación, en realidad, puede anexarse a cualquiera de las tres opciones ya mencionadas, ya sea que se quiera traducir a uno o varios idiomas determinados todo o parte de los textos originales, o solamente sus resúmenes o la formulación indizada o codificada de esos textos.

Cabe mencionar que gracias a los avances tecnológicos modernos, existe la posibilidad, como se acaba de mencionar, de una eventual automatización del análisis. A pesar de que actividades tales como la clasificación, la catalogación, la elaboración de resúmenes, etc., permanecen aún principalmente como actividades del intelecto humano, ya han sido nominadas candidatas a la automatización. (3)

Se encuentra uno finalmente en presencia de un buen número de caminos posibles para pasar de la presentación origini

(3) Cf. DOYLE, LAUREN B. Information Retrieval and Processing. Los Angeles, Ca., Melville, 1975. p. 247 ss.

nal de los documentos a una presentación distinta, mejor adaptada a las necesidades de la información actual.

1. El Resumen.

Numerosas son las discusiones sobre la utilidad del resumen como medio de información; sobre su necesidad, dentro de las perspectivas de la automatización; sobre su oportunidad, teniendo en cuenta su costo relativamente elevado.

a). En cuanto a la utilidad de los resúmenes se piensa generalmente que otros medios de información son más útiles y eficaces, tales como el intercambio personal de información entre especialistas de la misma área, o, en el otro extremo, la consulta de vastos bancos de datos por medios mecánicos o automatizados. Cualquiera de estos sustitutos propuestos tiene el inconveniente de estar a disposición de un número relativamente limitado de usuarios, mientras que las publicaciones de tipo bibliográfico tienen un acceso mucho más amplio.

b). En cuanto a la necesidad de los resúmenes, algunos especialistas de la automatización confían en la idea de que pronto será costeable guardar íntegramente la literatura científica y humanística en la memoria de las computadoras, y aplicar así los procedimientos de exploración automatizada sobre los textos originales, sin resumen previo. Sin descar-

tar esa posibilidad, no cabe duda de que se presentan aún en la actualidad problemas de orden teórico y práctico, tales como la duración y el costo del procesamiento computarizado. Aun si la etapa del resumen está destinada a desaparecer, parece por el momento que es imposible suprimirla, precisamente en vista de los límites actuales, particularmente de orden económico, que presentan el equipo y los métodos de automatización.

c). En cuanto a la oportunidad de esta etapa, desde el punto de vista económico, es difícil emitir una apreciación; si es cierto que la elaboración de resúmenes es una operación relativamente costosa, también es cierto que no se tiene una manera fácil de cuantificar el beneficio que se deriva para la difusión de la información y el progreso del conocimiento humano. Como quiera que sea, el número de publicaciones dedicadas a los resúmenes va en aumento; por otro lado, desde el punto de vista financiero, existen medios que pueden ayudar a abatir el costo, tales como la colaboración de los autores mismo de los textos.

Por lo que se refiere a la alternativa hombre/máquina en la elaboración de los resúmenes, es necesario precisar que el resumen automatizado, en cuanto a su utilidad práctica, estará siempre subordinado a la posibilidad de una captura total de los textos originales.

2. La Indización.

Por indización entiendo aquí, de manera muy amplia, toda caracterización de un documento determinado por medio de palabras o grupos de palabras, con miras a descubrir la fuente (autor, título, fecha, referencias, etc.) o el contenido. En el primer caso, el análisis y la indización permanecen en un nivel muy superficial. Al contrario, si, como en el segundo caso los términos utilizados deben constituir una representación auténtica y completa del contenido de cada documento, la indización se convierte en algo mucho más complejo. En un caso, por ejemplo, bastará con una caracterización por nombre del autor, temas principales, etc., mientras que en el otro se exige prácticamente una paráfrasis del texto (íntegro o resumido), despojándolo sólo de todas las imperfecciones lógico-lingüísticas (redundancias, ambigüedades, etc.), que puedan obstaculizar la búsqueda de la información, automatizada o no.

Consideremos, por otro lado, que si el progreso de la tecnología permite, tarde o temprano, almacenar en una memoria artificial el texto íntegro de los documentos, no por ello los problemas de la indización se verían resueltos de manera automática. El objetivo de la indización no es, en realidad, esencialmente, el de darle un carácter conciso a la información, -

sino más bien el de normalizar, en alguna forma, las expresiones que describen el contenido de los textos. Si esos textos se archivan en la memoria de la máquina bajo su forma literal original, sin normalización, es fácil darse cuenta de que la búsqueda de la información se entorpecería por problemas lingüísticos, por el hecho mismo de la diversidad que se presenta en la formulación de los contenidos documentales. (4).

3. La Codificación.

En la hipótesis admitida de un empleo tan amplio como sea posible, de métodos automatizados para el tratamiento de la información, debe tenerse en cuenta la necesidad de una codificación para la captura de datos. Raros son los casos en que se puede realizar la captura bajo la forma misma que revisten los datos en la literatura original, aún después del proceso del análisis. Se necesita una transposición a un sistema simbólico adaptado al equipo de computación que va a utilizarse. Cualquiera que sea el sistema, la operación de codificación es, a priori, enteramente susceptible de automatizarse, salvo la lectura inicial de los datos.

(4) Cf. DOYLE, LAUREN B. Information Retrieval and Processing. Los Angeles, Ca., Melville, 1975. p. 291

4. La Traducción.

Cabe distinguir dos categorías de traducciones: las que se hacen por medio de los especialistas y para ciertos textos solamente, y por otro lado, la traducción sistemática de todos los documentos. Por lo que se refiere a las primeras, no existe ningún problema aparente. En cuanto a las segundas se presenta el problema de decidir en que momento del análisis conviene llevarlas a cabo. Se presentan tres posibles soluciones:

a) Ya sea que se traduzcan los textos íntegros, con miras a someterlos posteriormente a una red de especialistas encargados de resumirlos.

so de los servicios de...
b) O bien, al contrario, realizar la traducción exclusivamente de los resúmenes, en la medida en que sean útiles para la indización o para la edición de publicaciones de tipo bibliográfico.

c) Finalmente, como tercera opción, manejar resúmenes en varios idiomas y efectuar la traducción en el momento de la indización, a un lenguaje normalizado y por consiguiente único.

La segunda solución es la que parece más viable; tiene, en efecto, la ventaja de facilitar los trabajos de indización, en la medida en que esta operación pueda llevarse a buen tér-

mino partiendo de los resúmenes, sin recurrir a los textos originales. Los especialistas encargados de esta tarea tienen que resolver entonces un solo problema de interpretación, del español, por ejemplo, al lenguaje artificial adoptado para la indización; la automatización eventual de la indización no exige en este caso más que un solo algoritmo, un solo diccionario, cualquiera que sea el idioma del documento original.

Conclusiones.

He presentado, de manera muy breve, las principales "actividades" del análisis de la información. Puedan estas líneas contribuir, en alguna forma, al mejoramiento y progreso de los servicios documentales en nuestro país. Si bien es cierto que vivimos actualmente una etapa de crisis que para nosotros se traduce, principalmente, en una escasez de recursos, reuniones como ésta deben animarnos a suplir esa escasez con una mayor voluntad de servicio por el bien de nuestros usuarios.