

Tema 2. Los Procesos de Organización del Conocimiento

El ámbito de Knowledge Organization (KO) u Organización del Conocimiento (OC) comprende el estudio de los principios, conceptos, técnicas y métodos que permiten formular enunciados cuya función es expresar los conceptos principales de un documento por medio de palabras, signos y códigos convencionales, con la intención de que éstos constituyan una representación que haga las veces del documento a fin de poder identificarlo y recuperarlo.

Los principales KOP/POC reconocidos serían los de indización y clasificación. La elaboración de otras representaciones documentales como los resúmenes es incluida también en ocasiones por Gnoli (2015), Souza, Tudhope y Almeida (2012), entre otros autores. Este mismo es el caso de la Catalogación (Hjørland, 2016 y 2017).

Centramos el contenido de este tema en la representación del mensaje del documento. En los procesos de acceso al contenido juegan un papel crucial los procesos de lectura y comprensión del texto que son expuestos por diversos estudiosos del ámbito de la Psicología y de la LIS/ByD como es el caso de Pinto (2005). El punto de partida es siempre la determinación de las estructuras semánticas de los contenidos de los diferentes tipos documentales. Se indaga, por tanto, en el contenido, en el mensaje de los documentos, y en su significado, aspectos en los que nuestra disciplina se ayuda de los conocimientos de la semiótica.

En el desarrollo de los contenidos de la asignatura queremos incidir en el reconocimiento y representación de los conceptos principales desarrollados en los documentos por medio de la elaboración de resúmenes y la indización. Se distingue entre los documentos textuales, los documentos en imagen fija, y aquellos que combinan códigos, dado que el acceso a los mensajes que contienen es diverso y ello condiciona su representación.

Los procesos de organización del conocimiento (KOP/POC) nacen como respuesta a los problemas que plantea el incremento incesante de la información documental, los métodos de identificación de los contenidos se presentan como el vehículo más idóneo para satisfacer las necesidades de los usuarios de acceder a los mensajes. La razón fundamental obedece a que las técnicas de análisis de contenido permiten la recuperación colectiva de documentos en respuesta a demandas por campos de conocimiento, materias, temas, palabras-claves, descriptores, etc.

Los principales objetivos de los KOP/POC son los siguientes: poner de manifiesto el contenido esencial de los documentos originales, seleccionar los principales elementos conceptuales de esos documentos y representarlos bajo formas diversas (palabras claves, encabezamientos de materia, descriptores, etc.) para facilitar el almacenamiento de la información. No obstante, su razón última estriba en contribuir a facilitar la recuperación de información, en consonancia con las demandas de los usuarios.

KOP/POC se entiende como el conjunto de operaciones dirigidas a representar y/o describir el tema o temas principales de que trata un documento. Esta tarea, eminentemente intelectual, puede ejecutarse a distintos niveles:

indizando mediante unos términos significativos o resumiendo en unos párrafos un documento original.

Así, tradicionalmente se ha distinguido, dentro de estos procesos la indización, orientada a poner de manifiesto el tema o temas del documento entresacando los elementos o aspectos característicos que lo representen para su posterior localización, y la construcción de resúmenes, cuyo fin es la elaboración de un resumen que permita el conocimiento en profundidad del contenido del documento analizado (Rodríguez-Bravo, 2011).

Conviene destacar, igualmente, el papel potenciador que la Sociedad de la Información ha otorgado a la Organización del Conocimiento mediante el uso de las clasificaciones. El estudio de las clasificaciones es abordado en la asignatura de Sistemas de Normalizados de Representación del Contenido, por lo que no desarrollaremos aquí sus contenidos.

Nos centramos en esta asignatura en el análisis de principios y procesos vinculados al contenido (KOP/POC) y en dos operaciones concretas: la que conduce a la representación de los documentos mediante un resumen y la de su representación mediante conceptos. Ambas basadas en el uso del lenguaje libre.

RESUMIR E INDIZAR: PRIMERAS NOCIONES

Resumir es una actividad compleja porque implica la necesidad de poner en juego diversas habilidades para abordar con éxito el doble proceso de análisis y síntesis que lleva aparejado. No sólo es preciso identificar la sustancia de los documentos sino que después esta debe ser expresada de forma que el lector pueda conocer las principales aportaciones de ese documento.

Moreiro (1993) apunta la necesidad de que el resumen se aborde desde una perspectiva interdisciplinar que tenga en cuenta factores psicolingüísticos, sociolingüísticos, gramaticales, estilísticos y literarios. De hecho, los avances en el estudio de los textos y de los procesos cognitivos de los resumidores están siendo de ayuda en los intentos de objetivar este proceso.

Respecto a la problemática de la objetivación, la ideología y la cultura del lector influyen excesivamente en los resultados de la comprensión lo que perjudica el sistema de recuperación. Sin embargo, el texto propone una información esencial que puede ser detectada por encima de esos filtros y que responde a la denominación de macroestructura (García Gutiérrez, 1992).

A la hora de proceder a la identificación de la macroestructura del documento original interesa conocer qué tipo de documento hay que resumir dado que cada uno tiene sus características propias, su estructura y su finalidad. Del mismo modo, la organización del discurso variará en función del ámbito científico de que se trate dado que los hábitos de comunicación varían en función de las disciplinas.

Tanto Moreiro (1993) como Pinto (2001, 2005) establecen estrategias metodológicas para identificar las estructuras del original basándose en el análisis lingüístico, sintáctico y semántico del texto, y estrategias para la reducción de este, a saber, segmentación de las unidades del texto, para

proceder posteriormente a su reducción y síntesis. En temas posteriores se profundizará en estas cuestiones.

Moreiro (2005) define los valores o criterios que presiden la construcción de resúmenes. Un buen resumen será aquel cohesionado a nivel sintáctico o superficial y coherente a nivel semántico o profundo. Ambas propiedades son la respuesta lógica a sus equivalentes en el texto original una vez sometido a las diferentes etapas del proceso resumidor.

Respecto al producto del proceso de resumir, Montesi (2006) recoge las principales definiciones del resumen en la literatura especializada y de las normas ISO/UNE. Realiza también una detallada distinción entre el resumen y otros productos que también tienen como objetivo la condensación del contenido como la anotación, el sumario, el extracto, la reseña, la sinopsis o el compendio. Asimismo, Simões (2015) realiza un completo recorrido por la evolución de este producto documental y por la historia de los servicios de resúmenes.

La indización, como la elaboración de resúmenes, está estrechamente ligada al contenido informativo de los documentos. En el caso de la indización cobra una relevancia más visible la teoría de los conceptos, dado que el objeto de la indización es la identificación, selección y representación de estos.

Hjørland (2001, 2017) ha analizado las relaciones semánticas de una serie de conceptos próximos con los que se vincula directamente la indización como son: *subject*, *topic*, *theme*, *domain*, *field* y *content*. La indización está conectada especialmente al concepto de *aboutness* que Hjørland considera sinónimo de *subject* y del que destaca la dificultad de su identificación en un documento. Señala que el *subject* de un documento es ese "algo" que la indización y la recuperación tienen que ser capaces de identificar. Se relaciona, por tanto, con las preguntas a que un documento tiene que responder. Dado que un documento puede, en principio, proporcionar respuestas a infinidad de cuestiones, la indización deberá establecer prioridades basadas en los grupos de usuarios a los que se dirige. La mejor indización será la que logre la más acertada prognosis del uso futuro del documento.

Gil Leiva (2008) ha reflexionado sobre el concepto de indización y ha sistematizado las definiciones de los principales estudiosos, al igual que las etapas del proceso de indización establecidas. Asimismo, las etapas del proceso de indización están recogidas en Rodríguez Bravo (2011) y se abordarán en los siguientes temas.

Existen diversos sistemas de indización puesto que el proceso puede estar fundamentado en las palabras (indización por unitérminos), en los conceptos (indización por descriptores) o en los temas (indización por materias). Además, los sistemas de indización pueden utilizar lenguaje libre (derivado del texto) o lenguaje controlado (asignado a partir de un vocabulario controlado o KOS/SOC, un tesoro o una lista de encabezamientos, por ejemplo). Igualmente, los KOS/SOC pueden ser postcoordinados (unitérminos y descriptores) o precoordinados (materias y clasificaciones). Los principales criterios para establecer la sistematización de los sistemas de indización aparecen recogidos en Gil Urdiciain (2004) o en Rodríguez Bravo (2011). Estos contenidos serán objeto también de desarrollo en temas

posteriores y en la asignatura Sistemas Normalizados de Representación del Contenido.

Referencias bibliográficas

García Gutiérrez, A. (1992). Análisis documental del discurso periodístico. Madrid: Centro de Tratamiento de la Documentación.

Gil Leiva, I. (2008). Manual de indización.teoría y práctica. Gijón: Trea.

Gil Urdiciain, B. (2004). Manual de lenguajes documentales. 2ª ed. Rev. Y amp. Gijón: Trea.

Gnoli, C. (2015). Knowledge Organization Systems (OCSs), En Golub, K. *Subject Access to Information: An Interdisciplinary Approach*. Santa Barbara, California: Libraries Unlimited, p. 43-65.

Hjørland, B. (2001). Towards a Theory of Aboutness, Subject, Topicality, Theme, Domain, Field, Content...and Relevance. *Journal of the American Society for Information Science and Technology*, 52 (8), p.774-778.

Hjørland, B. (2017b). Reviews of Concepts in Knowledge Organization: Subject (of Documents). *Knowledge Organization*, 44, 1, p. 55-64.

Montesi, M. (2006). *Métodos de evaluación y calidad de resúmenes documentales*. Gijón: Trea.

Moreiro González, J. A. (1993). *Aplicación de las ciencias del texto al resumen documental*. Madrid: Universidad Carlos III, BOE.

Moreiro González, J. A. (2005). La representación y recuperación de los contenidos digitales: de los tesauros conceptuales a las folksonomías. En Tramullas, J. (coord.). *Tendencias en documentación digital*. Gijón: Trea. P. 81-108.

Pinto Molina, M. (2001). *El resumen documental: paradigmas, modelos y métodos*. 2ª ed. Madrid: Fundación Germán Sánchez Ruipérez.

Pinto Molina, M; (2005). *Aprendiendo a resumir. Prontuario y resolución de casos*. Colaboradoras: M. Mitre, A.-V. Doucet, M. J. Sánchez Gijón: Trea.

Rodríguez Bravo, B. (2011). Apuntes sobre representación y organización de la información. Gijón: Trea.

Simões, M. da G. (2015). *Resumo documental: uma incursão na (des)construção da literatura científica*. Coimbra: Imprensa da Universidade.

Souza, R. R.; Tudhope, D.; Almeida, M. B. (2012).Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems. *Knowledge Organization*, 39, p. 179-192.