

## Metodologías *top-down* y *bottom-up* de análisis de dominio: Perspectiva desde la garantía literaria

Mario Barité  
Carlos Fernández-Molina

### 1 INTRODUCCIÓN

El análisis de dominio es el conjunto de aproximaciones teórico-metodológicas destinadas a representar el campo temático que comparte una comunidad de discurso. Se manifiesta habitualmente en el mapeo de una disciplina o especialidad (Hjørland 2002, Robinson 2009). En términos pragmáticos, para desarrollar un sistema de organización del conocimiento (SOC) destinado a la representación temática y la recuperación de información, se requiere un análisis de dominio que determine las fronteras, la extensión y los vínculos del dominio, a partir de la terminología y del sistema de relaciones conceptuales que utilice su comunidad de discurso. Pero ¿qué métodos se han usado para el desarrollo de un análisis de dominio para SOC's? Una revisión de la literatura científica del área permite agruparlos en dos diferentes enfoques: *top-down*, apoyado primordialmente en el método deductivo, y *bottom-up*, que apela al método inductivo.

La garantía literaria es uno de los principios más clásicos y controvertidos de la organización del conocimiento. Su punto de partida es que la terminología y las relaciones conceptuales vigentes en un dominio deben derivarse de la documentación, pues el volumen de literatura sobre un tema legitima su adscripción y situación dentro de una especialidad. A este respecto, la garantía literaria puede utilizarse como un principio, tanto teórico como metodológico, para fundamentar la selección de los tópicos a ser incorporados en cualquier SOC.

El objetivo de este trabajo es analizar los diferentes enfoques metodológicos para el desarrollo de análisis de dominio desde la perspectiva del principio de garantía literaria. Con ese propósito se realiza un breve repaso a las características básicas de los dos grandes enfoques metodológicos y a continuación se examina el papel que puede jugar en esta tarea el principio de garantía literaria.

### 2 ENFOQUES METODOLÓGICOS

El enfoque *top-down*, tradicionalmente impuesto en el desarrollo de sistemas de clasificación universales (Mills 2004; Gnoli & Mei 2006), sigue un proceso de división lógica del conocimiento, desde lo general a lo particular. Es el método que sugiere Ranganathan para desarrollar estructuras facetadas, partiendo de cinco categorías fundamentales (Personalidad, Materia, Energía, Espacio y Tiempo), y de la división de las clases principales en disciplinas (Ranganathan 1960). En los sistemas universales, como criterio habitual, se ubican en el primer nivel disciplinas, siguiendo una secuencia preestablecida (ejemplos: sistemas decimales como CDD y CDU, Colon Classification y Lista de Encabezamientos de Biblioteca del Congreso). En el segundo nivel se sitúan subdisciplinas o tópicos. En los siguientes niveles de división se ubican tópicos, es decir, los objetos de estudio más específicos de cada disciplina.

El enfoque disciplinario y la subordinación de tópicos a disciplinas están consolidados en la Organización del Conocimiento (Iyer 1995; Williamson 1998), ya

que las disciplinas son cortes tradicionales del conocimiento, avalados por universidades y especialistas. No obstante, se ha sugerido recurrentemente en la literatura que en el primer nivel de los SOC's se ubiquen tópicos que representen *phenomena* o conjuntos no disciplinarios de conocimientos, pues ello permitiría clasificaciones más adecuadas de campos interdisciplinarios como los Estudios de Género y otros (Iyer 1995; Gnoli, Bosch & Mazzocchi 2007).

Los SOC's desarrollados bajo el enfoque *top-down* son predominantemente jerárquicos. Cada término está subordinado a otro, salvo aquellos ubicados en el primer nivel de división. Cada subclase de elementos es un subconjunto de la clase inmediatamente superior (Cann 1997). Estas jerarquías reproducen taxonomías instauradas en las distintas disciplinas. De ellas se toman los sinónimos y las relaciones paradigmáticas que deben ser establecidas.

Por su parte, el enfoque *bottom-up* es básicamente inductivo, pues parte del análisis de los términos que se usan en la comunicación y la práctica cotidianas de una comunidad de discurso (Farradane 1950; Centelles 2005). En este enfoque, se soslayan las organizaciones previas del conocimiento y se privilegian los estudios de casos. Metodológicamente, se toma una definición tentativa del dominio de referencia y sus principales divisiones para establecer criterios de inclusión/exclusión de la terminología y el corpus de fuentes seleccionadas para el análisis.

La colecta de términos es exhaustiva o selectiva considerando los objetivos que se persigan. Una vez analizados y depurados los términos, estos se agrupan considerando sus atributos comunes más distintivos. Estas agrupaciones constituidas en torno a las afinidades conceptuales van configurando una estructura arbórea, que se construye también inductivamente. Si se pretende un sistema facetado, los términos se agrupan por facetas a partir de esos atributos, y se organizan siguiendo una secuencia que estará más o menos de acuerdo con las clasificaciones científicas o especializadas. Concomitantemente, se procede al establecimiento de relaciones de jerarquía, sinonimia y asociación entre ellos.

### 3 GARANTÍA LITERARIA

La garantía literaria es un principio teórico-metodológico formulado por Hulme en 1911 que, junto con otros de similar naturaleza, se utiliza para fundamentar la selección de los tópicos a ser incorporados en los SOC's de cualquier índole. El Sistema de Biblioteca del Congreso y la Clasificación Decimal Dewey son revisados conforme a este principio. La garantía literaria establece que la terminología y las relaciones conceptuales vigentes en un dominio deben derivarse de la documentación, antes que de clasificaciones científicas o filosóficas, o de la autoridad de los clasificacionistas, pues el volumen de literatura sobre un tópico legitima su adscripción y situación dentro de una especialidad (Yee 2001).

Los autores dividen opiniones respecto a situar a la garantía literaria en los dos enfoques mencionados. El enfoque *bottom-up* parece más claramente asociado con la garantía literaria (Cann 1997; Miwa & Kando 2007), en virtud de la pauta metodológica de identificación y extracción de los términos de un dominio de la literatura misma. Sin embargo, otros sugieren una estrecha vinculación entre garantía literaria y el enfoque *top-down*, pues asocian el principio de Hulme a la expresión de la opinión experta tal como se manifiesta en la documentación (Yiotis, Kovacevic & Harris 2004).

Ambos enfoques no son necesariamente excluyentes respecto a la garantía literaria. De hecho, otra línea de pensamiento sugiere comenzar un análisis de dominio por el método *bottom-up*, asegurando una recogida representativa de términos candidatos desde la documentación, y continuar por el enfoque *top-down*, basado en la opinión

experta tal como se expresa en la documentación, para ajustar la estructura conceptual y evitar lagunas en los esquemas (Cann 1997).

La falta de consenso respecto a la relación garantía literaria y enfoques *top-down* y *bottom-up* justifica la necesidad de más investigación y mayor precisión en el estudio de conceptos nucleares de la Organización del Conocimiento.

#### 4 CONCLUSIONES

La garantía literaria tiene un claro componente inductivo, pues extrae de la documentación los tópicos que se organizan en tablas de clasificación conforme a sus atributos comunes. Por tanto, en el proceso de extracción de terminología, la garantía literaria se inscribe en un enfoque *bottom-up*. También puede contribuir en el proceso de formación de la estructura conceptual a partir del análisis de los vínculos conceptuales surgidos de la documentación. No obstante, es posible construir la estructura de un dominio desde una perspectiva *top-down*, basada originalmente en la opinión experta, y justificar cada tópico en la garantía literaria. O inclusive, utilizar la opinión experta y la metodología *top-down* para diseñar un SOC, y luego utilizar el enfoque *bottom-up* en conjunción con la garantía literaria, para realizar las revisiones periódicas de los esquemas. En suma, la garantía literaria es una herramienta flexible que se ajusta a distintos métodos de diseño de SOC, que puede utilizarse en cualquier fase de un análisis de dominio.

#### REFERÊNCIAS

- Cann, J. 1997. *Principles of classification: suggestions for a procedure to be used by ICIS in developing international classification tables for the construction industry*. Disponible en: <http://www.icis.org/siteadmin/rtdocs/images/5.pdf> [consultado el 10 de julio de 2011]
- Centelles, M. 2005. Taxonomías para la categorización y la organización de la información en sitios web. *Hipertext.net*, 3. Disponible en: <http://www.hipertext.net/web/pag254.htm> [consultado el 10 de julio de 2011]
- Farradane, J.E.L. 1950. A scientific theory of classification and indexing. *Journal of Documentation*, 6: 83–99.
- Gnoli, C., Bosch, M. & Mazzocchi, F. 2007. A new relation for multidisciplinary knowledge organization systems: dependence. In: Rodríguez Bravo, B., Alvite, M.L., eds. *Interdisciplinarity and transdisciplinarity in the organization of scientific knowledge: proceedings of the 8th ISKO-Spain Conference*. León: Universidad de León, pp. 399-409.
- Gnoli, C. & Mei, H. 2006. Freely faceted classification for Web-based information retrieval. *New review of hypermedia & multimedia*, 12(1): 63-81.
- Hjørland, B. 2002. Domain analysis in Information Science: eleven approaches – traditional as well as innovative. *Journal of Documentation*, 58(4): 422-461.
- Iyer, H. 1995. *Classificatory structures: concepts, relations and representation*. Frankfurt/Main: INDEKS.
- Mills, J. 2004. Faceted classification and logical division in information retrieval. *Library Trends*, 52(3): 541-570.
- Miwa, M. & Kando, N. 2007. A naïve ontology for concepts of time and space for searching and learning. *Information Research*, 11(2). Disponible en: <http://informationr.net/ir/12-2/paper296.html> [consultado el 10 de julio de 2011]
- Ranganathan, S.R. 1960. *Colon Classification*. Bombay: Asia Publishing House.

- Robinson, Lyn. 2009. Information Science: communication chain and domain analysis. *Journal of Documentation*, 65(4): 578-591.
- Williamson, N. 1998. An interdisciplinary world and discipline based classification. In: El-Hadi, W. Mustafa, eds. *Structures and relations in knowledge organization*. Würzburg: Ergon-Verlang. pp. 116-123.
- Yee, M.M. 2001. Two Genre and Form Lists for Moving Image and Broadcast Materials: a Comparison. *Cataloging & Classification Quarterly*, 31(3/4): 237-295.
- Yiotis, K., Kovacevic, Z. & Harris, C. 2004. *Subject Headings for a Database: Developing the Controlled Vocabulary*. San Jose, Ca.: San Jose State University. Disponible en: <http://kristinhome.com/eportfolio/m/YiotisKovacevicHarris.pdf> [consultado el 10 de julio de 2011]

Sobre os autores:

Mario Barité  
Universidad de la República, Uruguay  
[mabarite@gmail.com](mailto:mabarite@gmail.com)

J. Carlos Fernández-Molina  
Universidad de Granada, España  
[jcfernand@ugr.es](mailto:jcfernand@ugr.es)