

# Organización del conocimiento y arquitectura de la información: tendencias y tensiones

Mario Barité<sup>1</sup>

## Resumen

Las nociones de “conocimiento” e “información” siguen siendo tan polisémicas y de naturaleza tan abierta, que cada vez que se refiere a ellas, debería hacerse explícita la significación particular que adquieren en un contexto determinado. Se mencionan nuevas dimensiones de la información, la que se presenta hoy organizada en distintos niveles o estratos -visibles, encubiertos, sucesivos, discontinuos- De entre los problemas nuevos que plantea Internet, se enfoca la necesidad de sistematizar los recursos de información y de determinar criterios para su inclusión entre los fondos disponibles en una biblioteca o centro de documentación. Se describe una categorización de documentos electrónicos que ofrece guías y orientaciones para su tratamiento bibliotecológico. Se explica por qué la Organización del Conocimiento como campo temático está en un proceso de expansión y de determinación más precisa de sus fronteras y de sus relaciones con otras áreas del saber. Se describen nuevas formas de estructuras conceptuales (taxonomías, ontologías, folksonomías) y se analizan las ediciones 2005 de las normas británica y estadounidense de creación y gestión de vocabularios controlados y estructurados. Entre otras conclusiones, se señala que el desarrollo de lenguajes de uso corporativo, la exigencia de atención a necesidades locales de información, las demandas particulares de comunidades de usuarios relativamente pequeñas en número pero fuertes en masa crítica o en el impacto social de su actividad, han resignificado la necesidad de centrar toda Organización del Conocimiento en una perspectiva de usuario. Se utiliza la obra de Salvador Dalí “La batalla de Tetuán” como disparador de reflexiones.

**Palabras clave:** Estructuras conceptuales. Lenguajes documentales. Dalí, Salvador.

---

<sup>1</sup>Profesor titular de “Clasificación” y “Terminología” Universidad de la República (Uruguay). Investigador en el área de Organización del Conocimiento. Diploma de Suficiencia Investigadora y Doctorando por la Universidad de Granada, España.

**Título:** Organização do conhecimento e arquitetura da informação: tendências e tensões

**Resumo**

As noções de “conhecimento” e “informação” permanecem tão polemicas e de natureza tão aberta, que cada vez que se refere a elas, deveria-se explicitar a significação particular que adquirem em um determinado contexto. São mencionadas novas dimensões da informação, a qual se apresenta hoje organizada em diferentes níveis ou estratos-visíveis, escondidos, sucessivos, descontínuos. Dentre os novos problemas que norteiam a Internet, enfoca-se a necessidade de sistematizar os recursos de informação e de determinar critérios para sua inclusão entre os fundos disponíveis em uma biblioteca ou centro de informação. Descreve-se uma categorização de documentos eletrônicos que oferecem guias e orientações para seu tratamento bibliotecário. Explica-se por que a Organização do Conhecimento como campo temático está em um processo de expansão e de determinação mais preciso de suas fronteiras e de suas relações com outras áreas do saber. Descrevem-se novas formas de estruturas conceituais (taxonomia, ontologias, folksonomías) e se analisam as edições de 2005 das normas britânica e estadunidense de criação e gestão de vocabulários controlados e estruturados. Entre outras conclusões, observou-se que o desenvolvimento das linguagens de uso corporativo, a exigência de atenção as necessidades locais de informação, as demandas particulares de comunidades de usuários relativamente pequenas em número, mas fortes em massa crítica ou em impacto social de sua atividade, tem aumentado a necessidade de centrar toda Organização do Conhecimento em uma perspectiva de usuário. Utilizando-se a obra de Salvador Dalí “A batalha de Tetuán” para despertar reflexões.

**Palavras chaves:** Estruturas conceituais. Linguagens documentais. Dalí, Salvador.

**1. La batalla de Tetuán: lecturas e incógnitas.**

La idea de acompañar este texto con una obra de Salvador Dalí tiene por objetivo hacer más visibles las líneas de cruce entre información y conocimiento, y la manera en cómo esas encrucijadas impactan tanto en la organización del conocimiento como en la arquitectura de información destinadas a la recuperación pertinente y relevante. Se eligió especialmente “La batalla de Tetuán” porque, siendo una de las obras mayores del artista español, no es de las más conocidas, ni es comprensible en toda su profundidad sin información previa. Está claro que su mera apreciación provocará un goce estético mayor o menor dependiendo de la

sensibilidad y de los gustos personales de cada observador. Más allá de los elementos oníricos y la mirada surrealista de la realidad, característicos de la obra de Dalí, este cuadro puede ofrecer sin embargo lecturas superpuestas, simultáneas o sucesivas, de índole semiótica, política, histórica, estética, cultural y también biográfica.



Figura 1: DALI, Salvador. A batalha de Tetuán.

“La batalla de Tetuán” fue pintada en 1961 por Salvador Dalí y no es en sí una obra “original”, pues se trata de una versión personal de una tela magnífica del gran artista catalán Mariano Fortuny pintada entre 1863 y 1873, y que su autor dejó inconclusa. La obra que reinventó Dalí es en sí misma un objeto de información, o más precisamente, *un objeto que irradia información*, que está disponible para distintas finalidades, y que por tanto aporta al acrecentamiento del saber o *conocimiento* de cada persona; puede ser utilizada, por ejemplo, para ilustrar las rémoras de colonialismo que perviven en ciertas partes del mundo (la bota que se planta en la tierra seca y firme de África); ofrece información sobre la indumentaria propia de la época, y sobre la cultura ecuestre de los árabes; invita a conocer la obra de Fortuny, es interesante para estudiantes de arte que están aprendiendo a dibujar caballos o a dominar el equilibrio de los ocre; nos brinda un retrato de Gala, la mujer y musa de Dalí, cuyo rostro aparece en medio del cuadro envuelta en un turbante, exhibiendo una sonrisa que no tiene sintonía aparente con la calidad dramática de la escena; ofrece la posibilidad de comparar los elementos originales del cuadro Fortuny con los que toma Dalí; motiva a saber más de la historia de esta batalla entre españoles y marroquíes, de la ciudad de Tetuán y de los

enclaves españoles que todavía existen en el norte africano, entre otras muchas posibilidades. Tiene también elementos ocultos, tan ocultos que están a la vista y sin embargo, si se hiciera un cuestionario acerca de cuáles podrían ser ellos, muy pocos los identificarían. Al final de esta comunicación se dirá cuáles son estos elementos que constituyen en sí un enigma de *información*, que sólo puede interpretarse, descifrarse, decodificarse con mayor *conocimiento*.

## **2. La tensión “conocimiento / información”.**

Las nociones de “conocimiento” y de “información” siguen siendo tan polisémicas y de naturaleza tan abierta, que cada vez que se hace referencia a una de ellas, debería hacerse explícita la significación particular que adquieren en un contexto determinado (Chalmers, 1999; Capurro & HjØrland, 2003). Como ese sería un ejercicio tedioso e insostenible, en un discurso cualquiera se suelen dar por sobreentendidas estas nociones, o al menos, se apela a las atribuciones más generales (y por ello, más vagas) que están presentes en ellas. Como apunta García Marco (1995), “los conceptos de información y conocimiento son conceptos abiertos, conceptos cruce, conceptos límite”, demasiado generales y ambiguos como para encontrar un punto de partida diferencial, que resulte satisfactorio para todo el mundo.

Con la intención de alcanzar definiciones operativas, pueden adelantarse algunos enunciados aceptables, como que la “información” se expresa siempre en conjuntos de datos mínimamente organizados y socialmente disponibles (más allá de eventuales restricciones debidas a la seguridad, al pago de un canon o a los derechos de las personas), con aptitud de ser recuperados, asociados, utilizados y reconfigurados por un usuario cualquiera para finalidades y objetivos particulares. La información tiene carácter objetivo y no tiene valor por sí misma, puesto que necesita ser interpretada y es “la materia prima del conocimiento.”

También en palabras de García Marco (1995) “el conocimiento es la información percibida y procesada por un sistema inteligente en virtud de su valor para ese sistema”. El sistema inteligente (por ejemplo, el hombre) utiliza la información con determinado propósito y por esa vía la asigna un significado particular, puesto que –siguiendo la cita- “el significado es el valor que posee una información para un sistema de información”. En la apropiación de información se dan como conocimiento tanto el proceso de negociación subjetiva entre dicha información y la inteligencia (el proceso cognitivo mismo), como su resultado, esto es, la efectiva integración de “nuevo saber” al universo conceptual propio de cada persona.

Es sabido que los seres humanos logran sostener y ponderar el conocimiento adquirido, a partir de estructuras de conceptos que van elaborando a lo largo de la vida, y que esas estructuras (siempre variables y dinámicas) constituyen una red de cristalización y asociación de ideas que es necesaria para situarse en la realidad, comprenderla, interpretarla y actuar en función de esa interpretación.

En todo caso, el conocimiento es una conquista de la inteligencia respecto a su comprensión de la realidad, y es un arma decisiva para la transformación de dicha realidad. Ante un problema nuevo (por ejemplo, una enfermedad desconocida) la investigación utiliza el conocimiento preexistente, ya devenido en información disponible, para obtener un cambio o un producto abstracto o material (por ejemplo, una vacuna o un medicamento; pero también una poesía, una pintura o una canción), el cual es registrado o patentado de alguna manera, y luego todo ello (la investigación, sus resultados, las consecuencias del descubrimiento o avance) es socializado a través de la documentación, una de las vías por las cuales el nuevo conocimiento es comunicado, apropiado socialmente y vuelto a reconvertirse (fragmentado en sus innumerables datos) en información.

Hubo momentos de la historia en los cuales se generaron grandes centros y depósitos de información, asimilables en cierto modo a lo que hoy es Internet y el conjunto de los contextos digitales e informacionales. En el siglo XVI, por ejemplo, los puertos por los que se canalizaba la comunicación con el Nuevo Continente, eran ámbitos fermentales, efervescentes, en los que el conocimiento práctico y el intercambio con las novedades que llegaban del otro lado del océano creaban un clima diferente, donde el informarse y el conocer tenían un tránsito fluido. Así, según Burke (2000), Sevilla por aquella época se había transformado en un emporio de información, pues era el lugar por donde entraba la plata que se traía desde América... y “con la plata llegaba también información sobre el Nuevo Mundo. La Casa de Contratación de Sevilla -por ejemplo- era un depósito de conocimientos sobre rutas marinas. Allí se guardaba una carta de navegación modelo (el llamado Padrón Real), que se actualizaba regularmente cuando los pilotos volvían de sus viajes.” Sólo de esa manera se explica que “el médico Nicolás Mornades fuera capaz de escribir sus famosos libros sobre las drogas de América sin dejar su Sevilla natal” (Burke, 2000).

El contexto científico-cultural, y en especial la creciente alfabetización informacional han exacerbado la percepción social de las ventajas de utilizar información para cualquier tipo de decisiones, asociar datos y utilizar “paquetes de información”.

Desde los ámbitos especializados, por su lado, se tiene cada vez más conciencia de que ninguna disciplina tiene capacidad de resolver bien los problemas que estudia por sí misma, se acepta la complejidad inherente a todo análisis y de este modo existe una tendencia a favorecer abordajes interdisciplinarios. La inversión en ciencia y tecnología sigue también esta tendencia. Se tiene claro que los problemas son dinámicos como que el conocimiento es provisorio y está en permanente revisión. Y si esto es válido en general, lo es en especial respecto a los problemas de la información.

Hoy puede hablarse de nuevas dimensiones de la información, modalidades en que se produce, se comparte, circula y se utiliza que difieren sustancialmente de las establecidas apenas quince años atrás. En la actualidad la información se presenta organizada en distintos niveles o estratos -visibles, encubiertos, sucesivos, discontinuos- (como en la pintura de Dalí), y su disponibilidad y circulación provoca dos fenómenos aparentemente contrapuestos: por una parte, es casi seguro que puede encontrarse información sobre cualquier asunto en cantidades, formatos, idiomas y presentaciones muy variados; por otra parte, es muy difícil acceder a una información sin “ruido”, que no esté inútilmente replicada decenas de veces, inclusive en versiones o traducciones distintas y de diversa calidad, o bajo la forma de resúmenes de escaso valor de referencia. El “ruido” se provoca muchas veces por razones lingüísticas (polisemia, homografía, variaciones y regionalismos, incluso errores ortográficos). El problema que se plantea actualmente, como nunca antes en la historia de la humanidad, es cómo lidiar contra la desmesura y la saturación de información, antes que con la carencia de datos.

El mundo de la información es cada vez más prolífico, ancho y abierto, está cada vez más disponible, pero a la vez, es un universo intrincado, frondoso, inabarcable, donde las estrategias deben apuntar a un refinamiento de la búsqueda de información, en cuyos procesos no pueden estar ausentes estrategias terminológicas. Por otra parte, el núcleo de una demanda de información hoy no tiene relación directa con un libro o documento, sin con un recurso de información cualquiera en sentido muy amplio: (imágenes aisladas, audiovisuales, música, datos concretos insertos en textos mayores, e inclusive las relaciones que se establecen entre un cuerpo de información y otro).

Si el conocimiento, para legitimarse y comunicarse, necesariamente debe alcanzar alguna (o muchas) forma(s) de registro documental, el desarrollo de una cultura electrónica alternativa y simultánea a la tradición libresca tenía que llevar –también necesariamente- a re-discutir y volver a significar la noción de documento. Briet (1951), Schamber (1996),

Buckland (1997), y en literatura española López Yepes (1997), han sabido explicar el tránsito complejo desde el libro como objeto socializador del saber por antonomasia, hasta los escurridizos y lábiles continentes de información digital (portales, sitios, documentos y libros electrónicos, etc.), HIPERvinculados por una red creciente de enlaces. Expresiones como “documento electrónico” o “documento digital” han logrado resistir (quien sabe por cuánto tiempo) la generalización más amplia que otorga el término “recurso de información”, como paraguas que puede contener –y ayudar a resolver- muchas de las cuestiones teóricas, con implicancias metodológicas y prácticas, que hoy enfrenta la Ciencia de la Información.

De entre los problemas nuevos que ha planteado Internet, se encuentra la necesidad de sistematizar los recursos de información presentes, y de encontrar criterios para su inclusión (o exclusión) entre los fondos disponibles en una biblioteca o centro de documentación. Gorman (2001), propuso una categorización de documentos electrónicos que ofrece guías y orientaciones. Así, identifica seis tipos de recursos de información en Internet:

- **Información de valor efímero.** Es la contenida en folletos, comunicados, boletines y otras, publicaciones de menor cuantía, que las bibliotecas tradicionalmente han desechado por su escaso valor como fuente de información, y que no tiene sentido que sea sometida ahora a control bibliográfico sólo porque se encuentra en Internet.
- **Información comercial.** Aquí, además de los sitios de empresas y de la publicación de catálogos de venta, Gorman incluye a la difusión de pornografía en este apartado, y afirma que “las bibliotecas nunca han recogido información comercial, ni - con pocas excepciones- pornografía”, por lo que tampoco se justifica un cambio de políticas a este respecto.
- **Recursos derivados del formato impreso.** La edición digital de obras impresas o conocidas previamente en papel, e incluso las obras que cuentan sólo con ediciones digitales no presentan mayores problemas de catalogación y control bibliográfico. Por otra parte, cuentan con mayor estabilidad que el material efímero o la información comercial.
- **Revistas electrónicas.** Las revistas especializadas, sea en su forma impresa o en edición electrónica, siguen siendo el soporte principal para la socialización del nuevo conocimiento y la evolución de las disciplinas. En los últimos años se ha dicho que las versiones electrónicas de las revistas sustituirán a las

impresas, pero ese proceso se va dando de forma gradual. Es cierto que un significativo número de revistas sólo se publican en versión digital, porque los costos son menores y su difusión mayor al disponerse de las colecciones siempre completas y disponibles en Internet. La tendencia creciente a la digitalización, y a la imposición de una filosofía de acceso libre al conocimiento científico y especializado (“*open access*”), fundada en que la investigación y los investigadores ya han sido financiados con anterioridad a la publicación, lleva a que cada vez existan más revistas en versión electrónica (exclusiva o compartida con la versión impresa), y más revistas disponibles sin pago previo de contraseña.

- **Archivos digitalizados** (textuales, sonoros y visuales). Como expresa Gorman (2001), “uno de los logros más importantes y valiosos de la era electrónica es la forma en la que los usuarios acceden a los grandes archivos, a los que, hasta ahora, sólo han podido acceder los investigadores con los medios y el tiempo para viajar al lugar donde se encuentra el archivo”. Hoy es posible acceder no sólo a una obra específica de Dalí, sino a un catálogo exhaustivo y crítico de todas sus obras, reunido o disperso en diversas colecciones particulares, con distintos niveles de información sobre cada obra en particular y sus obras en general, así como sobre su biografía.

- **Obras de creación original** (textuales, sonoros y visuales), surgidas y manifestadas en entornos digitales. Como afirma Gorman (2001), “se prevé que aparecerán nuevas generaciones de creadores en Internet, como por ejemplo, los escritores de hipertexto, los artistas digitales, los ciberpoetas y los músicos electrónicos”. Sus obras deberán ser catalogadas, indizadas y puestas a disposición tanto por las bibliotecas como por otros repositorios (museos y archivos digitales), y custodiadas también por productores de cine y vídeo y coleccionistas de arte.

Las estructuras de información que se desarrollen para la captura y el control de estos recursos, deben ser satisfactorias para cualquier usuario en cualquier circunstancia, pero también es esencial obtener una arquitectura de información de excelencia, que dé cumplimiento efectivo a preceptos de usabilidad, accesibilidad y amigabilidad, y que se apoye en sistemas lógicos y estructuras conceptuales adecuadas al tipo de información disponible. Detrás de toda arquitectura de información está el usuario, sus necesidades, sus problemas, incluso sus caprichos; y por ello toda arquitectura de información debe proporcionar hoy día



espacios interactivos para utilizar, reutilizar, comunicar y discutir conocimiento e información.

### **3. Estructuras conceptuales y Organización del Conocimiento.**

La Organización del Conocimiento como campo temático está en un proceso de expansión y de determinación más precisa de sus fronteras y de sus relaciones con otras áreas del saber. Si el hecho de que la revista *International Classification* cambiara su nombre por *Knowledge Organization* en 1974 fue suficientemente expresivo de la necesidad de ensanchar horizontes epistemológicos, no lo es menos el alcance que hoy día se da en dicha revista al campo, al señalar que las áreas que tienen contacto o están involucradas con la Teoría de la Clasificación son “filosofía y ciencia de la ciencia, organización de las ciencias y políticas científicas, matemáticas, estadísticas e informática, bibliotecología y ciencia de la información, archivística y museología, periodismo y ciencias de la comunicación, la disciplina relativa a los *commodities* y los productos industriales, la terminología, la lexicografía y la lingüística” (Knowledge Organization, 2006). Parece estar vigente la necesidad de someter a control bibliográfico las fuentes y las expresiones del conocimiento y los diferentes recursos de información, un control que precisa contar con herramientas mediadoras de representación del conocimiento, tales como los lenguajes documentales tradicionales: tesauros, listas de encabezamientos, listas de descriptores y sistemas de clasificación.

Una de las pruebas de la expansión del campo temático está dada por el proceso acelerado de diversificación y complejidad de las estructuras de conocimiento que se ha ido dando en los últimos quince años, proceso que no se ha dado caprichosamente, sino en función de necesidades y propósitos vinculados con el uso efectivo de terminología para el desarrollo de las ciencias, las especialidades, los entornos digitales y la comunicación corporativa. A las estructuras de conocimiento tradicionales se fueron sumando entonces, nuevas modalidades (no todas bibliotecológicas, aunque todas con repercusiones en sistemas de información), de las cuales se reseñarán algunas a continuación.

## **Taxonomías:**

Desde una perspectiva tradicional la taxonomía es la “ciencia que trata de los principios, métodos y fines de la clasificación. Se aplica en particular, dentro de la biología, para la ordenación jerarquizada y sistemática, con sus nombres, de los grupos de animales y de vegetales.” (Real Academia Española, 2001). Pero también se denomina taxonomía al producto de esas clasificaciones: la estructura conceptual misma, que permite organizar en forma sistemática, por ejemplo, a los vegetales, como resulta de la tradicional clasificación de Linneo. Las estructuras taxonómicas de la ciencia se presentan bajo forma de árbol, debido a que privilegian las relaciones jerárquicas entre seres u objetos, y por su vocación de internacionalidad, contribuyen a la formación de terminologías cerradas, unívocas, distintivas y monorreferenciales. La transferencia del término “taxonomía” a las áreas de Inteligencia Artificial, Ciencia de la Información e Informática se apoya en la noción central de jerarquía, en razón de que en todos los casos las estructuras resultantes se apoyan en niveles de términos superordenados y subordinados. No obstante, el nombre ha sido usado y abusado al punto de que Gilchrist (2003) identificó cinco sentidos diferentes del término “taxonomía” en la literatura de esas áreas: 1) como sinónimo de “directorio de buscador” o “directorio web”. 2) Para designar estructuras jerárquicas que soportan procesos de indización automática. 3) Taxonomías creadas mediante categorización automática. 4) Como “*front end filters*”, es decir, como estructuras de apoyo a la formulación de la búsqueda de información o a la navegación en Internet o Intranet (v. por ejemplo, [www.wordmap.co.uk](http://www.wordmap.co.uk)). 5) Taxonomía corporativas, como formas híbridas entre tesauros y ontologías de apoyo a la búsqueda y recuperación de la información, que se han vuelto comunes dentro de sitios institucionales.

## **Ontologías.**

Una **ontología** se concibe como una estructura funcional, que contiene entidades o elementos que se relacionan entre sí, para llevar a cabo determinados propósitos o para cumplir con ciertos objetivos, en un entorno habitualmente electrónico. Es funcional porque no pretende representar un segmento del conocimiento, sino desplegar una red de asuntos o acciones con sus relaciones, volviendo explícitos los circuitos que en su conjunto configuran un dominio. Más que una estructura de conocimiento, la ontología es ante todo un sistema relacional de acciones que persigue tanto una gestión corporativa de calidad como la satisfacción plena del usuario.

Lacroix, Marty y Roche (1998) distinguen tres niveles de conocimiento presentes en una ontología, cada uno respondiendo a una formalización lógica particular: el dominio; el conocimiento que describe situaciones del mundo real; y el conocimiento que resulta de deducciones e inferencias:

“1. El nivel del dominio describe el conocimiento que constituye el núcleo del dominio y nunca cambia. Esto incluye principalmente los conceptos y las relaciones permanentes entre ellos (por ejemplo, un cuadrado es un tipo de rectángulo, o una rodilla es una parte de la pierna).

“2. El nivel de la realidad del mundo expresa las relaciones entre objetos (por ejemplo, esta taza está sobre la mesa). Estas relaciones son aplicadas a objetos o casos que son reales en un determinado momento, y que pueden por tanto cambiar. Este nivel se basa en una lógica proposicional.”

“3. El nivel de inferencia se manifiesta como la expresión de relaciones de inferencias sobre conceptos u objetos (por ejemplo, si  $x$  es un triángulo, y si los lados de  $x$  tienen la misma longitud, entonces  $x$  es equilátero). La lógica que subyace es una lógica de razonamiento, como la lógica predicativa.”

Una ontología bien diseñada (por ejemplo, el sistema que sostiene el funcionamiento de los cajeros automáticos) toma en cuenta, antes que nada y sobre todo, los intereses y los comportamientos predecibles de los usuarios, así como los objetivos concretos que estos esperan cumplir dentro de la lógica de ese dominio.

Capítulo aparte merecen las **folksonomías** o clasificaciones sociales, que constituyen una metodología de recuperación de información en Internet basada en la asignación de etiquetas [*tags*, *labels*, palabras claves] generadas en forma asociativa por los mismos usuarios, para categorizar contenidos tales como recursos digitales de información, fotografías en línea o enlaces (Noruzi, 2006). Los sitios y sistemas de información web así organizados tienden a ser muy populares (por ejemplo, YouTube, [www.youtube.com](http://www.youtube.com)), y sólo recientemente se ha comenzado a estudiar la calidad de la recuperación de información a partir de folksonomías. Lo diferente de estas clasificaciones es que se van dando por acumulación, sin concierto previo y son tanto los emisores como los destinatarios de la información quienes participan en la “indización”, cuyo producto es una terminología inorgánica, desestructurada y sin posibilidades de un control posterior de vocabulario.

Otras estructuras conceptuales, tales como los **mapas conceptuales** y los **mapas de tópicos** (*Topic maps*), comienzan a ser utilizadas en los ámbitos relacionados con la información, por su capacidad para representar gráficamente nodos y relaciones de ideas y conceptos hacia adentro de un campo o dominio. Los manuales y las normas desarrollados inicialmente para la creación y gestión de tesauros han reconocido la necesidad de sistematizar directrices más flexibles que den cabida a estos nuevos tipos de estructuras de organización del conocimiento y recuperación de información.

#### **4. La norma británica.**

La última edición de las normas británicas BS 8723 (British Standards Institution, 2005) se denomina “*Structured vocabularies for information retrieval*”. Quiere decir que el término clave aquí es “vocabularios estructurados”, los que son definidos por esa norma como “los vocabularios utilizados para recuperar información a partir de recursos textuales, tales como bases de conocimiento, bases de datos bibliográficas o colecciones de documentos a texto completo, y también catálogos de imágenes o artefactos, como los que integran las colecciones de los museos”. Y agrega: “los vocabularios estructurados son utilizados en los sistemas de recuperación poscoordinada, directorios jerárquicos, índices precoordinados y sistemas de clasificación”. Excluye la cobertura de la “estructura de archivos de autoridad de nombres propios de personas, organizaciones y lugares, sin dejar de reconocer que ellos pueden algunas veces interactuar con vocabularios de términos temáticos”.

Esta tendencia a utilizar expresiones más abarcadoras que las de “tesauros”, “lenguajes documentales” o “vocabularios controlados” es un intento por cubrir razonablemente el universo posible de estructuras de conocimiento que pueden ser alcanzadas por normas comunes para su construcción, modificación y desarrollo; al tiempo que se enfatiza en metodologías de interfase y de comunicación entre esas estructuras, para potenciar sus posibilidades a la hora de clasificar, indizar, asignar etiquetas temáticas y facilitar la recuperación de información.

El peso específico que los tesauros tienen como herramientas de organización del conocimiento y recuperación de información se traduce en el hecho de que la parte 2 de la norma inglesa está destinada enteramente a ellos. El reconocimiento de la variedad se da en la parte 3, orientada a establecer reglas para la creación de vocabularios diferentes a los tesauros

(“*other than thesauri*”). La parte 4 refiere a a interoperatividad entre vocabularios, y la parte 5 a la interoperatividad entre vocabularios y otros componentes de los sistemas de almacenamiento y recuperación de información.

## 5. La norma estadounidense.

Ya desde el título, la edición 2005 de la norma estadounidense Z39.19 propone un nuevo horizonte de aplicación. En efecto, las normas de 1993 proporcionaban “*Guidelines for the construction, format and management of monolingual thesauri*”, en tanto que la nueva edición sustituye la expresión “*thesauri*” por la de “*controlled vocabularies*”, con orientaciones y reglas para cuatro tipos de herramientas utilizadas habitualmente para la representación de objetos de contenido en sistemas de organización del conocimiento: listas (*lists*), anillos de sinónimos (*synonym rings*), taxonomías (*taxonomies*) y tesauros. Estas normas interpretan que los cuatro tipos de lenguajes van de menor a mayor complejidad, en lo que hace a control de vocabulario. Así, las listas son vistas como la expresión estructural más sencilla, ya que se trata de organizaciones meramente alfabéticas o lógicas (como los planetas del sistema solar puestos en orden partiendo desde el sol), que se basan en la mutua exclusión de sus entidades, y por tanto apenas controlan la ambigüedad.

En un segundo nivel ubican a los anillos de sinónimos, herramientas que no pueden ser utilizadas para la indización por su misma naturaleza (en estos anillos todos los sinónimos tienen el mismo valor), pero sí auxilian en la recuperación, especialmente como una interfase en sistemas digitales de información en los cuales el acceso se da por lenguaje natural. Los anillos establecen en particular el control de la sinonimia y la cuasi sinonimia.

En el escalón siguiente, la norma norteamericana ubica a las taxonomías, limitando el alcance del concepto a los vocabularios controlados integrados por términos preferidos, conectados entre sí por vínculos jerárquicos o polijerárquicos. Las taxonomías controlan tanto la ambigüedad y las relaciones de identidad como las jerárquicas.

Finalmente, la norma Z39-19 establece directrices para los que considera vocabularios más elaborados y complejos, los tesauros, ya que en ellos el control de vocabulario abarca todas las posibles relaciones paradigmáticas existentes entre términos (de ambigüedad, identidad, jerarquía y asociativas). En la medida en que en el glosario de la norma se establece una noción amplia de tesoro, quedan compendidos por sus directrices todos los tipos de

tesauro que se han ido desarrollando y diversificando en los últimos años: tesauros facetados, tesauros de usuario, tesauros de software, tesauros de búsqueda, etc. (Gilchrist, 2003). Inclusive muchas de sus directrices son aplicables a instrumentos híbridos en evolución, como los tesauro / diccionarios (Buchan, 1989; Hudon, 1998).

## **6. Conclusiones.**

Lo que tienen en común estas conclusiones es la convicción de que los profesionales de la Ciencia de la Información deben adoptar y asumir una actitud especial ante los fenómenos –complejos, plásticos, intrincados, incesantes- del mundo de la información: la de apertura mental, sin prejuicios ni preconcepciones, para facilitar la comprensión de su rica realidad, entender la dinámica de sus procesos, la interacción cada vez más productiva entre herramientas informáticas y lógicas intelectuales y cognitivas que replican los procesos de buscar, asociar, conocer y producir. Si un investigador del área hubiera entrado en proceso de latencia a principios de los '90 y hubiera despertado hoy, se asombraría seguramente de encontrar una realidad tan diferente, y le costaría un buen tiempo reeducar su cabeza. Si bien no es posible dar respuestas definitivas a muchos “problemas de información”, lo cierto es que hoy estamos en condiciones de dar mejores respuestas, o por lo menos respuestas mejor fundamentadas a la creación y aprovechamiento de modalidades de organización, acceso y recuperación de información. De algún modo volvemos a vivir en la efervescencia social del conocimiento de la Sevilla del siglo XVI, ansiosos por conocer las novedades de los navegantes más audaces, preocupados por registrar los nuevos saberes y por dar con las herramientas más adecuadas para usar y reutilizar las fuentes documentales disponibles. Hoy hay que dar respuestas a la diversidad (de intereses, de visiones, de culturas), y parece volverse cada vez más esencial el trabajo interdisciplinario, especialmente hacia adentro del ámbito transdisciplinario de la información.

Una convicción que cabe reafirmar desde nuestra perspectiva disciplinaria, es que los inmensos caudales de información con los que nos toca lidiar requieren y seguirán requiriendo operaciones de organización documental, control de vocabulario y utilización de los lenguajes intermediarios de recuperación, porque más allá del debate interminable entre los defensores del lenguaje natural y de los vocabularios controlados, lo cierto es que los tipos de herramientas de lenguaje para diseño, ordenamiento y recuperación de datos no dejan de

crecer en número, de diversificarse, y de hibridarse entre sí. Los lenguajes tienen prestigio y siguen siendo ampliamente utilizados. Por debajo de esa incuestionable noticia de la realidad, subyace la idea de que las estructuras conceptuales sostienen necesaria y finalmente la organización temática de los sistemas de información, y favorecen su mejor aprovechamiento. Pocas profesiones como la nuestra están llamadas a contribuir a la mejor evolución de esas estructuras normalizadas y normalizadoras, en función de nuestro *expertise* en esa materia.

Por último, y no por ello menos importante, el desarrollo de lenguajes de uso corporativo, la exigencia de atención a necesidades locales de información, las demandas particulares de comunidades de usuarios relativamente pequeñas en número pero fuertes en masa crítica o en el impacto social de su actividad, han resignificado la necesidad de centrar toda Organización del Conocimiento en quien es nuestro mentor, nuestro abogado del diablo, nuestro esclavo y nuestro verdugo: el usuario. Un usuario acostumbrado a interactuar, navegante libre en el océano de información, que busca nuestro apoyo para ahorrar tiempo, encontrar las mejores rutas, los atajos más escondidos y la brújula para resolver sus problemas inmediatos –cada vez más inmediatos- de información. Sin usuario no hay Ciencia de la Información, excluyendo al usuario no hay organización del conocimiento que sea legítima y útil. “La batalla de Tetuán” acepta tantas miradas como usuarios tenga, y en la síntesis abierta que representa como obra de arte mayor, está implícito el llamado de la diversidad, de la verdad levantada como una construcción multi- e interdisciplinaria, y la capacidad de provocar demandas y respuestas múltiples.

## **7. Y finalmente, el enigma de “La Batalla de Tetuán”**

El enigma se resuelve en números.... números escondidos, números a la vista, agrupados, repetidos, desperdigados por toda la tela de Dalí, números que, como se dijo antes, sólo pueden “interpretarse, descifrarse, decodificarse con mayor *conocimiento*”.

### **Referencias bibliográficas**

BRIET, Susan (1951). **Qu'est-ce que la documentation ?** París : Edit, 1951.

BRITISH STANDARDS INSTITUTION (2005). **Structured vocabularies for information retrieval**. S.l.: BSI, 2005.

BUCKLAND, Michael (1997). What is a document? **Journal of the American Society for Information Science**, 48(9): p. 804-809.

BUCHAN, R.L. (1989). Intertwining thesauri and dictionaries. **Information Service & Use**, 9: p. 171-175 (1989).

BURKE, Peter (2002). **Historia social del conocimiento** : de Gutenberg a Diderot. Barcelona : Paidós, 2002.

CAPURRO, Rafael & Birger HJØRLAND (2003). Concept of Information. **Annual Review of Information Science and Technology**, 37: p. 343-411 (2003).

CHALMERS, A. F. (1999) **What is this thing called science**. 3rd. ed. Buckingham : OUP, 1999.

GARCÍA MARCO, Francisco Javier (1995). **Hacia un modelo de intervención en los procesos de transmisión del conocimiento**. *SCIRE*, 1(2): p. 105-138 (jul./dic. 1995).

GILCHRIST, Alan (2003). Thesauri, taxonomies and ontologies : an etymological note. **Journal of Documentation**, 59(1): p. 7-18 (2003).

GORMAN, Michael (2003). Control o caos bibliográfico: un programa para los servicios bibliográficos nacionales del siglo XXI. **Anales de Documentación** (6): p. 277-288 (2003).

HUDON, Michèle (1998). A preliminary investigation of the usefulness of semantic relations and of standardized definitions for the purpose of specifying meaning in a thesaurus. EN **Advances in Knowledge Organization**. V. 6 (1998). p. 139-145.

KNOWLEDGE ORGANIZATION : Official Quarterly **Journal of the International Society for Knowledge Organization**. Würzburg : ISKO. V. 33(3): 2006.

LACROIX, S; MARTY, J.C. & ROCHE, C. (1998) A Model of Ontologies by Differentiation. EN Structures and Relations in Knowledge Organization: Proceedings of the Fifth International ISKO Conference, 25-29 August 1998, Lille, France. Edited by Widad Mustafa el Hadi, Jacques Maniez and Steven A. Pollitt. **Advances in knowledge organization**. Vol. 6. Würzburg: Ergon Verlag, 1998. P.336-342.



LÓPEZ YEPES, José (1997). **Reflexiones sobre el concepto de documento ante la revolución de la información: ¿un nuevo profesional del documento?** *SCIRE*, 3(1):p. 11-29 (jul./dic. 1997).

NATIONAL INFORMATION STANDARDS ORGANIZATION (2005). **Guidelines for the construction, format and management of monolingual controlled vocabularies**. Bethesda: NISO Press, 2005. 172 p.

NATIONAL INFORMATION STANDARDS ORGANIZATION (1993). **Guidelines for the construction, format and management of monolingual thesauri**. Bethesda: NISO Press, 1993. 69 p.

NORUZI, Alireza (2006). Folksonomies: (un)controlled vocabulary? **Knowledge Organization**, 33(4): p. 199-2003 (2006).

REAL ACADEMIA ESPAÑOLA (2001). **Diccionario de la lengua española**. 22ª edición. Madrid: Espasa-Calpe, 2001. 2 v.

SCHAMBER, Linda (1996). What is a document? : Rethinking the concept in uneasy times. **Journal of the American Society for Information Science**, 47 (9): p. 669-671 (1996).