



ESTUDIOS / RESEARCH STUDIES

Retos y oportunidades en organización del conocimiento en la intersección con las tecnologías de la información

Daniel Martínez-Ávila*, Rosa San Segundo**, Francisco A. Zurian***

* Information Organization Research Group – FAPOI. São Paulo State University – UNESP, Brazil. Correo-e: dmartinezavila@gmail.com

**Universidad Carlos III de Madrid, España. Correo-e: rosa.sansegundo@uc3m.es

***Seminario Interuniversitario Permanente de Investigación Género, Estética y Cultura Audiovisual (GECA), Universidad Complutense de Madrid, Facultad de Ciencias de la Información, España . Correo-e: azurian@ucm.es

Recibido: 18-07-2013; 2ª versión: 11-02-2014. Aceptado: 13-02-2014

Cómo citar este artículo/Citation: Martínez-Ávila, D.; San Segundo, R.; Zurian, F. A. (2014). Retos y oportunidades en organización del conocimiento en la intersección con las tecnologías de la información. *Revista Española de Documentación Científica*, 37(3):e053. doi: <http://dx.doi.org/10.3989/redc.2014.3.1112>

Resumen: En este trabajo se estudia la intersección de la Organización del Conocimiento y las Tecnologías de la Información, haciendo especial incidencia en los retos y oportunidades que existen en este área para los expertos en Organización del Conocimiento, tanto a un nivel teórico como práctico. Primero se revisan algunos conceptos necesarios para la contextualización del trabajo. Después se analiza la historia de la Web, comenzando con Internet, y continuando con la World Wide Web, la Web Semántica, los problemas en Inteligencia Artificial, la Web 2.0, y Linked Data. Finalmente, se concluye el estudio con algunas aplicaciones de las tecnologías de la información para la Organización del Conocimiento en el campo bibliotecario, como FRBR, BIBFRAME e iniciativas de OCLC, así como algunos retos y oportunidades en los que la investigación y los expertos en Organización del Conocimiento podrían jugar un rol clave en relación con la Web Semántica.

Palabras clave: Organización del conocimiento; tecnologías de la información; web semántica; inteligencia artificial; linked data.

Challenges and opportunities for knowledge organization at the intersection with information technologies

Abstract: In this paper we study the intersection of Knowledge Organization with Information Technologies and the challenges and opportunities for Knowledge Organization experts that, in our view, are important to be studied and for them to be aware of. We start by giving some definitions necessary for providing the context for our work. Then we review the history of the Web, beginning with the Internet and continuing with the World Wide Web, the Semantic Web, problems of Artificial Intelligence, Web 2.0, and Linked Data. Finally, we conclude our paper with IT applications for Knowledge Organization in libraries, such as FRBR, BIBFRAME, and several OCLC initiatives, as well as with some of the challenges and opportunities in which Knowledge Organization experts and researchers might play a key role in relation to the Semantic Web.

Keywords: Knowledge organization; information technologies; semantic web; artificial intelligence; linked data.

Copyright: © 2014 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution-Non Commercial (by-nc) Spain 3.0.

1. INTRODUCCIÓN

Existe la percepción entre docentes e investigadores en Biblioteconomía y Documentación de que la investigación teórica y fundamental en Organización del Conocimiento es un área con cada vez menos relación con los campos de recuperación de la información en su vertiente más tecnológica. En el desarrollo de estas nuevas tecnologías la integración teórica no solo no ha sido siempre bien considerada, sino que las actividades tradicionalmente realizadas por personas han sido crecientemente desafiadas por técnicas de recuperación automática sin intervención humana (Hjørland, 2003; 2008). Nuevas tecnologías, como agentes inteligentes, aplicaciones asociadas a la Web Semántica, técnicas lingüísticas y estadísticas asociadas a reglas de inferencia o sistemas de autoaprendizaje, muchos de ellos relacionados del tipo que fuere con el campo de la Inteligencia Artificial y otros sistemas de procesamiento, a priori, no humano, parecen haber relegado las teorías y prácticas tradicionales de la Organización del Conocimiento a un segundo plano del panorama científico, en el que, cada vez más, el factor de formación teórica e intervención humana en Organización del Conocimiento parece no ser necesario para el uso y desarrollo de aplicaciones.

El objetivo del presente trabajo es estudiar el rol, retos y oportunidades de la intervención humana en estas nuevas tecnologías de la información, que fundamentan los nuevos sistemas de recuperación, inevitablemente cada vez más automáticos dada la naturaleza de los nuevos entornos digitales, para así reivindicar el potencial de la investigación teórica en Organización del Conocimiento, en su vertiente más humanística, ética y, si se quiere decir también, tradicional, que pueda enriquecer la vertiente tecnológica del área.

2. METODOLOGÍA

La metodología del presente trabajo corresponde con un Análisis del Discurso de Foucault aplicado al área de Biblioteconomía y Documentación/Ciencias de la Información (ver Frohmann, 1994; 2001; Budd y Raber, 1996; Budd, 2006; Martínez-Ávila, 2012a). En este sentido, el objeto de análisis son los discursos, genealogías y emergencias históricas de los principales conceptos relacionados con la intersección entre la Organización del Conocimiento y las Tecnologías de la Información, para probar así las formaciones discursivas, continuidades y estrategias de control que afectan a la construcción/mala representación de conceptos y concepciones y que, a un nivel fenomenológico, también pueden afectar, la construcción de percepciones sobre los conceptos (Martínez-Ávila y Smiraglia, 2013). Para ello, se han seleccionado algunos de los hitos considerados históricamente fundamentales en la posible discontinuidad de la dicotomía intervención humana/intervención automatizada (asumiendo/asignando el foucauldiano "otro" al actual estado

de la intervención humana) en la intersección de Organización del Conocimiento y Tecnologías de la Información, como son el nacimiento de Internet y la World Wide Web, los desarrollos históricos en Inteligencia Artificial, el nacimiento y discursos de la Web 2.0 y Linked (Open) Data, y se han estudiado las principales prácticas, formaciones discursivas y estrategias de estos momentos históricos tal como han sido reflejadas en la bibliografía. El criterio de selección bibliográfica ha sido el de relevancia histórica consensuada en el dominio científico que podría ejemplificarse con el consensuado reconocimiento histórico en la bibliografía de la relevancia de Tim-Berners Lee en el desarrollo de la Web, de la *Dartmouth Conference* en la emergencia y desarrollo del campo de la Inteligencia Artificial, etc.

3. DEFINICIONES PRELIMINARES Y APLICACIÓN DE LA ORGANIZACIÓN DEL CONOCIMIENTO

Partimos de que las Tecnologías de la Información son "aquellas herramientas y métodos empleados para recabar, retener, manipular o distribuir información. La tecnología de la información se encuentra generalmente asociada con las computadoras y las tecnologías afines aplicadas a la toma de decisiones" (Bologna y Walsh, 1997). Se desprende de esta definición que tanto en los procesos de obtención, retención, manipulación y distribución de información, la Organización del Conocimiento es parte implícita y necesaria para poder alcanzar el que comúnmente se considera fin último y parte indivisible del proceso: la recuperación de la información. Por otra parte, de esta definición también se desprende que la tecnología de la información se encuentra "generalmente" asociada a las computadoras, al menos en su desarrollo conceptual. El grado de asociación a las computadoras y el posible rol y grado de intervención humana será lo que se trate de dilucidar en el presente trabajo.

Respecto a Organización del Conocimiento serán necesarias algunas aclaraciones preliminares debido a la multiplicidad de textos existentes y a la variedad de contextos en los que puede ser aplicado. Por ejemplo, Arlene Taylor y Daniel Joudrey (2009), ponen de manifiesto la multiplicidad de contextos en los que hay una necesidad de organizar información, incluyendo bibliotecas de todo tipo, archivos, museos, Internet, bibliotecas digitales, arquitectura de la información, indización y resumen, "records management" y gestión del conocimiento ("knowledge management"). Aunque los términos información y conocimiento han sido identificados como conceptos diferentes (ver por ejemplo San Segundo, 2002), la investigación teórica y las prácticas docentes han venido solapándose en lo que conforma el dominio, quizás erróneamente bajo diversas denominaciones (Hjørland, 2012). En este sentido, Birger Hjørland, quien aboga por el uso del término Organización del Conocimiento para referir al dominio, reconoce que a pesar del título del

citado libro de Taylor y Joudrey (traducido por "La organización de la información"), Taylor "es una autora muy conocida de un libro sobre KO [Organización del Conocimiento]". En este sentido, en el presente trabajo también utilizaremos el término Organización del Conocimiento para referirnos al dominio de Organización del Conocimiento, centrándonos en el dominio "per se", y tratando de revelar como artificiales las exclusiones terminológicas que hay creadas tanto dentro como hacia él.

La definición de Organización del Conocimiento que utilizaremos ha sido recientemente revisada por Richard Smiraglia, y es: "el dominio en el que el orden del conocimiento es el principal paradigma para la investigación científica y la principal aplicación en el desarrollo de sistemas" (Smiraglia, 2012). Esta definición sigue también en gran medida una tradición más acorde con el discurso ISKO (*International Society for Knowledge Organization*) (ver Dahlberg, 2006), y, aunque incluye como principal aplicación desarrollo de sistemas, a priori, no hace ninguna distinción en la dicotomía del nivel tecnológico/humano de los mismos.

Siguiendo estas concepciones de Organización del Conocimiento y Tecnologías de la Información, su intersección, para el objeto de este trabajo, será reformulado como el estudio del orden del conocimiento e información para el diseño y evaluación de sistemas y herramientas que hacen posible recuperar y usar la información, tanto por computadoras como por humanos. Esta idea de Organización del Conocimiento debería convertir la investigación teórica del dominio en un aspecto clave de las Tecnologías de la Información. Sin embargo, quizás esa multiplicidad de contextos en la que la Organización del Conocimiento puede ser aplicada, así como su dispersión terminológica, sea lo que explica la relativa poca visibilidad y reconocimiento de la importancia de la Organización del Conocimiento, al menos en su concepción teórica, en muchas áreas de las Tecnologías de la Información; o quizás sea la ubicuidad del concepto e historia lo que hace que no se sea consciente de su presencia a pesar de su relevancia. Haciendo una analogía con el campo de la Inteligencia Artificial, puede explicarse la reticencia a usar el término en algunos contextos de las Tecnologías de la Información debido a su propio pasado, pasados contratiempos en el caso de la Inteligencia Artificial, y la reminiscencia a la vieja Biblioteconomía tradicional en papel en el caso de la Organización del Conocimiento. Sin embargo, y aunque se ha señalado (Markoff, 2005) que algunos informáticos e ingenieros del software han podido adoptar otros nombres y evitado deliberadamente el término "Inteligencia Artificial" por miedo a ser vistos como "soñadores locos", también es cierto que, tal como han señalado otros autores (Kurzweil, 2005), las bases de la Inteligencia Artificial subyacen en muchas de las tecnologías actuales incluso cuando su presencia no es reconocida. Tal como Nick Bostrom también apuntó "muchas de la

Inteligencia Artificial puntera se ha filtrado en aplicaciones más generales, muchas veces sin ser denominada Inteligencia Artificial, porque una vez se convierte en algo suficientemente útil y común ya no es etiquetado como Inteligencia Artificial" (Bostrom, 2006). Siguiendo con esta analogía, muchos de estos argumentos también podrían ser aplicados a la Organización del Conocimiento, una vez que los conceptos teóricos se han filtrado en las Tecnologías de la Información deja de tener como referente a la Organización del Conocimiento, o al menos se desvincula de sus fundamentos teóricos.

Por otra parte, cabe señalar que algunas de estas etiquetas y nombres relacionadas con la Organización del Conocimiento han adquirido suficiente relevancia como para convertirse en nuevos campos por sí mismos, como, por ejemplo, la informática, los sistemas de conocimiento (knowledge-based systems), los sistemas cognitivos, los agentes inteligentes (ver por ejemplo Mas, 2005), la arquitectura de la información, la gestión del conocimiento (knowledge management) y otros. Por ejemplo, Rowley y Harlton (2008), enumeran sistemas que emplean enfoques a la Organización del Conocimiento en el entorno digital y que, por lo tanto, se benefician de la investigación en Organización del Conocimiento, como por ejemplo: OPACs, que son catálogos en línea; bibliotecas digitales, que son conjuntos de recursos electrónicos y sus capacidades técnicas asociadas para crear, buscar y usar información; repositorios institucionales y de materia, que proveen acceso a objetos digitales que han sido depositados en ellos; motores de búsqueda y directorios, que facilitan la recuperación de la información en Internet; bases de datos de índices y resúmenes e índices de citas, que son bases de datos bibliográficas sobre materias -o con objetivos específicos- que registran las publicaciones en la disciplina y las hacen disponibles para futuras búsquedas; sistemas de gestión de contenidos (CMS), que facilitan a los usuarios el desarrollo y gestión de sitios web u otros medios de publicación, facilitando el manejo de texto, imágenes y objetos de información multimedia; sistemas de gestión de registros (records management systems), que se ocupan de la identificación, captura, almacenamiento, recuperación y, cuando es necesario, destrucción de registros; sistemas de recuperación de imágenes; y wikis y blogs. Todos estos sistemas emplean enfoques de la Organización del Conocimiento en conjunción con otros campos como la Recuperación de la Información (una obra clave para entender cómo funcionan los Sistemas de Recuperación de la Información en Internet y su relación con la Organización de Conocimiento, es Baeza-Yates y Ribeiro-Neto, 2011), la Arquitectura de la Información, la Archivística y más. Común a todos estos sistemas es que la aplicación de las Tecnologías de la Información, y especialmente Internet, ha proporcionado nuevas formas de acceder y organizar la información en las que, aunque pudiera parecer lo contrario, la intervención y colaboración humana es más necesaria que nunca.

4. REFERENCIAS HISTÓRICAS DE LA WEB Y SU FILOSOFÍA DE COOPERACIÓN

Los orígenes de Internet se remontan a la década de los años 60 con los desarrollos en conmutación de paquetes, una tecnología para comunicar bloques de datos, y ARPANET (Advanced Research Projects Agency Network), la primera red de conmutación de paquetes del mundo desarrollada por el Departamento de Defensa de los Estados Unidos. Aunque la creencia de que el objetivo de ARPANET era crear una red que fuera resistente a una guerra nuclear es bastante común, algunas de las personas involucradas en el desarrollo de Internet han manifestado que esa creencia es tan solo un "rumor falso" (Leiner y otros, 2000). De hecho, tal como Pekka Himanen señala, el director del proyecto Lawrence Roberts "visionó una red como un medio de avanzar en la cooperación de los informáticos teóricos" (Himanen, 2001; Roberts, 1967), y aquí el Network Working Group, un selecto grupo de hackers y estudiantes universitarios que lideraron el desarrollo de la red, operó bajo los principios de software libre de compartir y hacer disponible la información, es decir, todo el mundo podía contribuir con ideas que después eran desarrolladas colectivamente (Himanen, 2001). También es importante señalar, que aunque cuando se habla de Internet casi siempre se relaciona con el modelo TCP/IP, incluyendo el Protocolo de Control de Transmisión (TCP) y el Protocolo de Internet (IP), ésta tampoco era la única posibilidad. Los dos organismos de normalización más importantes en aquellos días, el CCITT y la OSI, también tenían sus propios estándares oficiales: el X.25 y la ISO. Sin embargo, según Janet Abbate (1999), una de las principales razones por la que los protocolos de estas organizaciones de estandarización no tuvieron éxito fue la naturaleza más cerrada de estos cuerpos, lo que pone de manifiesto y enfatiza el espíritu abierto y colaborativo de la red desde sus comienzos.

También durante los años 60, aunque independientemente del desarrollo de Internet, la Biblioteca del Congreso de Estados Unidos desarrolló e implementó el formato MARC (MACHINE READABLE CATALOGING) bajo los mismos principios de colaboración y compartir información electrónica que guiaron el desarrollo de Internet, cooperación de catalogadores en este caso. La filosofía de MARC consistía en que en lugar de que múltiples catalogadores repitieran la descripción de los mismos recursos, individualmente y de forma aislada, el esfuerzo de un solo catalogador podría ser compartido con muchos catalogadores si todos usaban un formato común. En la década de los años 70, el formato MARC se había convertido en el principal estándar internacional para la difusión de datos catalográficos bibliotecarios y las redes bibliográficas se habían organizado para fomentar la catalogación colectiva y compartir registros en forma de fichas informatizadas. Aunque, principalmente MARC se centraba en la manera en que se codificaban las descripciones de forma electrónica,

también hacía falta una manera en que esas descripciones fueran transmitidas, y, eventualmente, Internet y sus descendientes se convirtieron en algunos de los principales medios para almacenar, compartir y recuperar.

Durante las siguientes décadas, las ideas y tecnologías de Internet siguieron evolucionando de modo que a principios de los años 90 había varios candidatos para una red global, y la World Wide Web, diseñada por Tim Berners-Lee, era tan solo uno de ellos. Uno de los más serios competidores de la World Wide Web en aquellos días era el sistema de información Gopher, desarrollado por la Universidad de Minnesota. Sin embargo, tal como Tim Berners-Lee señala, el final de Gopher ocurrió cuando la Universidad anunció que se reservaba el derecho a cobrar por el uso de los protocolos de Gopher, algo que, aunque nunca hicieran, fue visto como un acto de traición entre las comunidades académicas y de Internet (Berners-Lee, 1999). Tim Berners-Lee, por el contrario, se aseguró de que el CERN, la institución para la que trabajaba entonces, le permitiera mantener el desarrollo de la Web completamente abierta. Algunos analistas han señalado que esta apertura sumada a su modelo social, como también fuera el caso de Internet, fue una de las principales ventajas y factores de éxito de la World Wide Web (Himanen, 2001).

De forma paralela, a principios de los años 90, OCLC también comenzó su proyecto de catalogación en Internet, en el que bibliotecarios, distintos tipos de bibliotecas, y de forma voluntaria, contribuyeran con registros MARC que habían creado para servidores Gopher, LISTSERVES, ftp y sitios Web. Es importante señalar aquí que, aunque estas redes bibliográficas habían estado funcionando y compartiendo registros desde los años 70, instituciones como OCLC o la Biblioteca del Congreso de Estados Unidos, supieron adaptarse y mantenerse alerta de los últimos avances en Tecnologías de la Información para aplicarlas a la Organización del Conocimiento de sus prácticas colaborativas.

Durante mediados de los años 90, una vez que la World Wide Web estuvo asentada y empezó a ser usada de forma masiva, Tim Berners-Lee comenzó a trabajar en lo que pensó que debía ser el futuro de la Web: la Web Semántica. En su "*Semantic Web Road map*" de 1998, Berners-Lee escribió: "Este documento es un plan para lograr un conjunto de aplicaciones para datos conectadas en la Web de manera que formen una web de datos lógica y consistente (una web semántica) [...] La Web Semántica es una web de datos, en algunas formas como una base de datos global" (Berners-Lee 1998a). En 2001, Tim Berners-Lee amplió esta definición afirmando que "la Web Semántica no es una Web separada sino una extensión de la Web actual, en la que a la información se la da un significado bien definido, ayudando mejor que los ordenadores y las personas trabajen en colaboración" (Berners-Lee y otros, 2001).

Las principales características de la Web Semántica son: Expresar significado, es decir darle estructura al contenido significativo de las páginas web, creando un entorno donde los agentes de software que recorren las páginas puedan realizar tareas que son sofisticadas para los usuarios sin discriminar entre tipos de información, lenguas, culturas, etc.; Representación del Conocimiento, para que la Web Semántica sea funcional, los ordenadores deben tener acceso a colecciones estructuradas de información y conjuntos de reglas de inferencia que puedan ser usadas para realizar razonamiento automático. Para conseguir esto, principalmente dos tecnologías fueron propuestas para el desarrollo de la Web Semántica: XML (eXtensible Markup Language, que permite a las personas crear sus propias etiquetas y añadir una estructura arbitraria a los documentos, y RDF (Resource Description Framework), para expresar significado y codificarlo en conjuntos de tripletes (sujeto, predicado y objeto), que pueden ser escritas usando etiquetas de XML; Ontologías, que son documentos o archivos que definen formalmente las relaciones entre términos, típicamente incluyendo una taxonomía y un conjunto de reglas de inferencia. Aquí, las ontologías pueden ser usadas para lidiar con problemas terminológicos relacionados con la ambigüedad; Agentes, que son programas que recogen contenido Web de diferentes fuentes, procesando la información, e intercambiando los resultados con otros programas. Para aumentar su efectividad, el contenido Web legible por ordenador y los servicios automáticos deberán incluir semántica; Evolución del conocimiento, es decir, asistir en la evolución del conocimiento humano como un todo, y aquí se dice que nombrando conceptos con URIs (Uniform Resource Identifiers, identificadores de entidades como personas, corporaciones, libros, conceptos abstractos y otros en la Web), cualquiera puede expresar/identificar nuevos conceptos (significados) que se inventen con un mínimo esfuerzo, y conectar esos conceptos en una Web universal en la que los significados no se vean confinados únicamente a comunidades o grupos pequeños.

Respecto al papel de las ontologías en la Web Semántica, éstas pueden ser utilizadas para minimizar la ambigüedad de los términos a través de relaciones que son establecidas entre los conceptos, entidades, términos y categorías de un determinado dominio. Las ontologías pueden ser utilizadas para explicitar el significado de los términos de forma que las máquinas puedan procesar de forma automática la información sin ambigüedades. La dependencia de esos significados y sus relaciones a un dominio específico (invalidando así la posibilidad de "ontologías universales") y la necesidad humana de definir esas relaciones en ese dominio específico, refuerzan su vínculo con la intervención humana y la Organización del Conocimiento en la Web Semántica. Mientras que en lo que respecta a la relación de las ontologías con la Inteligencia Artificial, Natalya Noy y Deborah McGuinness, de la Universidad de

Stanford, han señalado la multitud de definiciones de "ontología" en la bibliografía sobre Inteligencia Artificial, algunas de ellas contradictorias, y muchas de ellas mostrando también una fuerte relación con otros conceptos clásicos de la Organización del Conocimiento, como son la categorización y clasificación (Noy y McGuinness, 2000). En general, las ideas de representación del conocimiento expresadas en la Web Semántica tampoco son nuevas. Además de presentar similitudes con algunas de las prácticas tradicionales en Organización del Conocimiento y descripción de recursos, son reminiscentes de un concepto similar de Representación del Conocimiento que ha sido utilizado en el campo de la Inteligencia Artificial con resultados que no siempre han sido los esperados.

5. REFERENTES HISTÓRICOS Y PROBLEMAS CLÁSICOS EN EL CAMPO DE LA INTELIGENCIA ARTIFICIAL

El término Inteligencia Artificial fue utilizado por primera vez en 1955 por John McCarthy, en "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence" (McCarthy y otros, 1955), unos años antes al desarrollo de Internet. Sin embargo, el nacimiento del campo de la Inteligencia Artificial, generalmente, se establece en el verano de 1956, cuando la *Dartmouth Conference* tuvo lugar (Crevier, 1993; Russell y Norvig, 2003; McCorduck, 2004). Una definición de Inteligencia Artificial dada por McCarthy es "la ciencia e ingeniería de hacer máquinas inteligentes, especialmente programas de ordenador inteligentes. Está relacionada con la tarea similar de utilizar los ordenadores para entender la inteligencia humana, aunque la IA [Inteligencia Artificial] no tiene por qué confinarse a sí misma" (McCarthy, 2007). Una obra fundamental que puede consultarse sobre Inteligencia Artificial en español es Palma Méndez y Marín Morales (2008).

Aunque el campo de la Inteligencia Artificial vivió años de crecimiento, optimismo y abundancia de fondos durante las siguientes décadas, también ha vivido algunos contratiempos y decepciones debido a la imposibilidad de alcanzar algunos de los ambiciosos objetivos que se marcaron. Además de varios pequeños episodios que gradualmente minaron la confianza en el campo, la investigación en Inteligencia Artificial también tuvo dos crisis mayores, una en el periodo 1973-1980, debido a los recortes en investigación académica para la Inteligencia Artificial en los Estados Unidos y en el Reino Unido, y otra, en el periodo 1987-1993, con el colapso del mercado de la Máquina Lisp y la cancelación de nuevas inversiones en Inteligencia Artificial por el Strategic Computing Initiative del gobierno de los Estados Unidos. Estas grandes crisis en la historia de la Inteligencia Artificial han recibido el nombre de Inviernos IA (*AI Winters* en inglés), y dentro del campo de la Inteligencia Artificial la amenaza de un nuevo invierno junto a los proble-

mas teóricos clásicos han puesto casi cualquier nuevo gran proyecto de Inteligencia Artificial bajo sospecha. Por esta razón, algunos investigadores en Inteligencia Artificial han optado en sus trabajos por diferentes denominaciones incluso cuando las mismas tecnologías estaban siendo utilizadas en sus aplicaciones.

Respecto a los problemas clásicos en Inteligencia Artificial que también llevaron a esos Inviernos en Inteligencia Artificial, principalmente son aquellos relacionados con Conocimiento y Razonamiento de Sentido Común (Russell y Norvig 2003), estando muchos de estos problemas como los relacionados con la desambiguación y la traducción automática todavía abiertos a día de hoy. Otro gran problema en el campo de la Inteligencia Artificial es la Representación del Conocimiento, que se ocupa de representar el conocimiento en símbolos que faciliten la inferencia (razonamiento) desde los elementos de conocimiento y la creación de nuevos elementos de conocimiento. Desde el punto de vista de la Inteligencia Artificial, este concepto de Representación del Conocimiento sería el mismo que el expresado en la Web Semántica. Sin embargo, aunque en la Web Semántica se utilizan técnicas de Inteligencia Artificial para procesar la información, Tim Berners-Lee se apresuró a aclarar desde sus comienzos que la Web Semántica no es Inteligencia Artificial, a la vez que aclaró las diferencias entre su Web y algunas de las tecnologías clave de Inteligencia Artificial en el pasado como el CYC Representation Language (CycL) y el Knowledge Interchange Format (KIF): "Una Web Semántica no es Inteligencia Artificial. El concepto de documentos que puedan ser entendidos por máquinas no implica alguna inteligencia artificial mágica que permita a las máquinas comprender los balbuceos humanos. Tan solo indica una habilidad de la máquina para solucionar un problema bien definido a través de operaciones bien definidas sobre unos datos bien definidos. En lugar de solicitar a las máquinas que entiendan el lenguaje humano, incluye pedir a las personas que hagan ese esfuerzo extra" (Berners-Lee, 1998b). En definitiva, muchos de los problemas que afectaron a la representación del conocimiento en el pasado de la Inteligencia Artificial (como aquellos de centralización, escalabilidad, unión de bases de conocimiento separadas, la necesidad de usar los mismos términos para palabras comunes, etc.) se solucionarían con la participación de las personas, con el uso global y la descripción descentralizada de recursos utilizando las tecnologías adecuadas y aquellos estándares que permiten la interoperabilidad. Muchos de estos estándares también son desarrollados de un modo social bajo la coordinación del Consorcio Web (W3C). De hecho, esta participación social en la representación del conocimiento parece ser tan solo una extensión natural de lo que era la visión original de Tim Berners-Lee de la World Wide Web como una herramienta social: "La Web es más una creación social que una creación técnica. La diseñé

para que tuviera un efecto social -para ayudar a las personas a que trabajen juntas- y no como un juguete técnico. El fin último de la Web es apoyar y mejorar nuestra existencia en red en el mundo" (Berners-Lee, 1999).

Por otra parte, la importancia de los aspectos sociales y culturales en la creación y representación de conocimiento también se está empezando a reconocer en el campo de la Inteligencia Artificial. Por ejemplo, William Clocksin ha resumido algunos de los elementos para un nuevo marco conceptual en Inteligencia Artificial que ya habían sido introducidos por la observación de Weizenbaum en 1976: la inteligencia se manifiesta solo en relación a una matriz de contextos sociales y culturales (Clocksin, 2003).

6. DE UNA WEB SOCIAL A LA WEB SOCIAL (2.0): EL PODER DE LAS ETIQUETAS

Sin embargo, al hablar de una web social, o de la Web como una creación social, es inevitable no referirse a otro vocablo (o "etiqueta") que fue popularizado por Tim O'Reilly entre 2004 y 2005: el de Web 2.0. En su artículo llamado "*What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*," Tim O'Reilly revisaba algunos de los nuevos avances y desarrollos de la Web desde el estallido de la burbuja punto com de 2001, como por ejemplo Google, BitTorrent, Wikipedia y otros, afirmando que "El principio central detrás del éxito de los gigantes nacidos en la era Web 1.0 que han sobrevivido para liderar la era Web 2.0 parece ser éste, que han abrazado el poder de la web para emplear la inteligencia colectiva" (O'Reilly, 2005). Aunque algunas de las aplicaciones Web 2.0 usadas por O'Reilly pudieran estar ya superadas, destaca el alto grado de Organización del Conocimiento e intervención de los usuarios en ellas. Mientras que anteriormente se especulaba con nombres de dominios web, como por ejemplo registrando URLs que incluyeran nombres de instituciones o productos antes de que las verdaderas instituciones pudieran hacerlo, de forma que si querían estar accesibles en Internet debían comprarles el dominio a los especuladores, los algoritmos de posicionamiento en los motores de búsqueda como Google han hecho posible la recuperación de sitios web con nombres no relacionados en las URLs a través de la consideración del número de enlaces entrantes y los términos en el código fuente, teóricamente también considerando las palabras clave (meta keywords) utilizadas para describir el contenido de los sitios web.

Otra aplicación de Organización del Conocimiento en la Web 2.0, probablemente una de las más importantes, es el uso del etiquetado social en el que las personas categorizan recursos y el desarrollo de "folksonomías", lo que O'Reilly ha denominado "un estilo de categorización colaborativa de sitios utilizando palabras clave que son escogidas libremente, comúnmente referidas como etique-

tas" (O'Reilly, 2005). Taylor y Joudrey han comentado sobre esto que "la implicación es tal que si suficientes usuarios etiquetan suficientes recursos, suficientes datos pueden ser agregados para alcanzar estabilidad, fiabilidad y consenso. Para que estos datos sean útiles para mejorar los actuales enfoques de acceso por materias, sin embargo, una masa crítica de etiquetas debe ser acumulada. La idea es que muchas etiquetas aplicadas a recursos discretos por miríadas de individuos (que etiquetan por incontables razones) proporcionarán suficiente información para entender la naturaleza del recurso, y nos permitirán sacar partido tecnológico a una manera barata de organizar recursos de información accesibles a través de la web" (Taylor & Joudrey, 2009). En definitiva, la filosofía detrás del éxito del etiquetado social no es más que otra aplicación de lo que Eric Raymond bautizó como "modelo bazar" en su obra ya clásica *"The Cathedral and the Bazaar"* (Raymond 2001), en la descripción de la metodología de desarrollo del kernel Linux, resumida en "Dado un número suficiente de ojos, todos los errores son irrelevantes". Otro ejemplo de aplicación del modelo bazar en Organización del Conocimiento sería la evaluación del modelo del desarrollo de la Open Shelves Classification desarrollada por la comunidad de LibraryThing (Martínez-Ávila, 2012b). De hecho, tal como O'Reilly también ha señalado, gran parte del éxito del modelo colaborativo de la Web 2.0 le debe mucho a las tecnologías de software libre y a filosofías como las licencias GNU y Creative Commons.

Debido al sorprendente éxito de muchos proyectos colaborativos como Wikipedia, se ha pretendido ver en la Web 2.0 el triunfo de la organización "natural" humana sobre la Inteligencia Artificial de las máquinas, muchas veces representada por la Web Semántica a pesar de los esfuerzos de Tim Berners-Lee de distanciarse de ella. De hecho, la dicotomía entre Web 2.0 y Web Semántica podría ser vista también como la dicotomía entre publicación de datos no estructurados y datos estructurados, y, tal como la historia de la Web parecía estar demostrando, la caótica pero humana organización de contenidos estaba ganando la carrera.

Sin embargo, a pesar de todos los innegables logros de la Web 2.0, también debe reconocerse que muchas de las aplicaciones de Organización del Conocimiento en la Web 2.0 están lejos de ser perfectas. Tal como diversos autores han señalado, uno de los principales problemas de la Web 2.0 es que si cualquiera puede crear contenidos entonces cualquier también puede destruir contenidos, además de infringir aspectos de copyright y privacidad de un modo mucho más anónimo (Agarwal 2012). En un artículo mucho más crítico, Cory Doctorow listó las posibles causas de una pobre generación de metadatos por los usuarios: "la gente miente en un mundo competitivo, la gente común es vaga para hacer algo que no entienden, la gente se niega a tener cuidado y diligencia en sus creaciones de metadatos, la gente no se conoce a sí misma,

los schemas no son neutrales, las métricas influyen en los resultados, y hay más de una manera de describir algo" (Doctorow, 2001). Sin embargo al igual que en el caso de cualquier sistema social, los problemas de un comportamiento inadecuado con metadatos no solo estarían relacionados con el grado de civismo de cada individuo, sino también con el profesionalismo y formación teórica en Organización del Conocimiento. Incluso asumiendo que los usuarios pudieran estar generando metadatos con sus mejores intenciones, la diferencia entre buenos metadatos y malos metadatos podría yacer en el nivel de pericia en Organización del Conocimiento de aquellos que están haciendo el trabajo de organizar el conocimiento. Es decir, para que la información pueda ser útil, ésta no debe ser solo compartida de un modo preciso sino también de un modo estructurado, para que en un diferente escalón las tecnologías e infraestructuras globales fueran capaces de procesar y rentabilizar esas estructuraciones. Por otra parte, como solución a los problemas reconocidos de la Web y los malos usos de metadatos, a partir 2005 algunos autores también han abogado por una combinación explícita de las características de la Web 2.0 y la Web Semántica para así mejorar las características de la Web (ver por ejemplo Hope y otros, 2007; Gruber, 2007; El-Goarany y otros, 2008; Greaves y Mika, 2008; Jeschke y otros, 2008; Durville y Gandon, 2009; García Marco, 2009; Lemen, 2010), a veces incluyendo elementos de Inteligencia Artificial también.

7. LINKED (OPEN) DATA Y LAS NUEVAS OPORTUNIDADES PARA LA ORGANIZACIÓN DEL CONOCIMIENTO

Sin embargo, desde la perspectiva de Tim Berners-Lee, la idea de combinar la Web 2.0 y la Web Semántica también podría ser algo confuso, ya que nunca creyó que la Web 2.0 fuera algo diferente de la Web original, una web que fue diseñada como una herramienta social desde sus comienzos. En 2006, Berners-Lee dijo que "La Web 1.0 trataba de conectar a la gente. Era un espacio interactivo, y yo pienso que la Web 2.0 es, por supuesto, un pedazo de jerga, nadie sabe incluso lo que significa. Si la Web 2.0 es para ti blogs y wikis, entonces eso es personas a personas. Pero eso es lo que Web se suponía que era todo este tiempo. Y de hecho, ya sabes, la Web 2.0, cito textualmente, significa usar los estándares que han sido producidos por toda esa gente trabajando en la Web 1.0. Significa utilizar el document object model, significa para HTML y SVG, etcétera. Es usar HTTP, por lo tanto es hacer cosas usando los estándares Web, más JavaScript, por supuesto. Por lo tanto la Web 2.0, para algunas personas, significa desplazar algo del lado de la consideración del cliente para hacerlo más inmediato, pero la idea de la Web como interacción entre personas es realmente lo que la Web es. Eso era para lo que estaba diseñada, como un espacio colaborativo donde la gente pueda inte-

ractuar" (Laningham 2006). Y es también durante este tiempo cuando Tim Berners-Lee emergió con un nuevo concepto, el de Linked Data, que podría ser visto como el siguiente paso en la progresión de la Web Semántica, o simplemente como otro intento de aclarar lo que la World Wide Web realmente era: "La Web Semántica no trata solo de poner datos en la web. Tratar de hacer enlaces, de forma que una persona o máquina pueda explorar la web de datos. Con datos enlazados [linked data] cuando tienes algo de algo, puedes encontrar otros datos relacionados" (Berners-Lee, 2006). Para ello, Berners-Lee proporcionó cuatro reglas: 1. Utilizar URIs como nombres para las cosas; 2. Utilizar URIs HTTP para que las personas puedan buscar esos nombres, y aquí Berners-Lee enfatiza que los URIs HTTP son nombres (no direcciones); 3. Cuando alguien busca un URI, proporcionar información útil, utilizando estándares (RDF, SPARQL), es decir, hacerlos disponibles y de una forma que puedan ser entendidos y procesados por máquinas y otros; y 4. Incluir enlaces a otros URIs, de forma que se puedan descubrir más cosas. En 2010, Tim Berners-Lee actualizó esta definición añadiendo la necesidad de usar mecanismos legales como licencias derivadas de la GPL (como por ejemplo Creative Commons) para garantizar el uso libre de los datos, afirmando que "Linked Open Data (LOD) son Linked Data que son liberados bajo una licencia abierta, que no impide su reutilización con libertad".

Las definiciones de Linked Data y Linked Open Data también han servido para traer aspectos novedosos a la Web Semántica y sus tecnologías, como RDF, OWL y SKOS. Linked Open Data también ha inspirado algunos nuevos y prometedores proyectos como, por ejemplo, Linked Open Vocabularies (LOV), destinado a promocionar y promover el acceso a vocabularios en "la nube"; schema.org, un proyecto desarrollado por Google, Microsoft y Yahoo para proporcionar a los publicadores de la Web un vocabulario universal para describir sus páginas usando datos enlazados; HIVE (Helping Interdisciplinary Vocabulary Engineering), que se define tanto como un modelo como un sistema que soporta la generación automática de metadatos extrayendo descriptores de múltiples vocabularios controlados codificados con SKOS (Greenberg y otros, 2011); y también algunos nuevos enfoques a Linked Data desde la Inteligencia Artificial, como, por ejemplo, el AAAI Spring Symposium titulado "Linked Data Meets Artificial Intelligence" (<http://www.foaf-project.org/events/linkedai>), organizado por Dan Brickley entre otros, o las aplicaciones e investigación sobre Inteligencia Artificial en el Open Data Institute, lideradas por Nigel Shadbolt (Howard, 2013), dando también algo de espacio y circularidad a la, a veces contradictoria, relación entre la Inteligencia Artificial y la Web Semántica.

En el campo bibliotecario, con la llegada de la Web y las Tecnologías de la Información, las prácticas y formatos tradicionales como MARC también han visto la necesidad de renovarse para poder

compartir y reutilizar la información bibliográfica de una manera significativa y estandarizada, y ya no solo con otros catalogadores sino también con otras aplicaciones y agentes de la Web Semántica. Tal como la IFLA ya señalaba en 1997, algunos de los factores clave que han contribuido a este cambio incluyen "la introducción y desarrollo en curso de sistemas automatizados para la creación y procesamiento de datos bibliográficos, y el crecimiento de bases de datos a gran escala, tanto de ámbito nacional como internacional, que contienen registros que son proporcionados y usados por miles de bibliotecas que participan en programas para compartir registros (IFLA Study Group on the Functional Requirements for Bibliographic Records, 2009). Por otra parte, también se está sugiriendo que los nuevos retos y catalizadores para estos cambios están relacionados con la creciente necesidad de adaptar los códigos y prácticas de catalogación a la emergencia de nuevas formas de publicación electrónica, y a la aparición de acceso en red a los recursos de información. Estas ideas fueron aportadas en el "Final report on Functional Requirements for Bibliographic Records" (FRBR), que se trata de una propuesta de un modelo entidad-relación (E-R) para identificar las entidades, atributos y relaciones en las descripciones bibliográficas y procesos de uso. Y aquí también destaca que aunque los principios de la Web Semántica no son exactamente los mismos que los del modelo E-R, de hecho, tal como Tim Berners-Lee ideó su Web, sí que están estrechamente relacionados (Berners-Lee 1998b).

En este sentido otra iniciativa en el campo bibliográfico aunando estos principios es BIBFRAME, que es un intento de migrar de MARC al entorno Linked Data a la vez que se fomenta la creación entidades claramente identificadas y el uso de identificadores adecuados para un procesamiento automático que sirvan para la interpretación de esas entidades (Miller y otros, 2012). Finalmente, otra aplicación que también ha sido citada como ejemplo de combinación de recursos bibliográficos y Linked Data es la reciente actividad de OCLC en schema.org. Desde 2012 OCLC ha estado añadiendo descripciones basadas en marcado a schema.org así como otros enlaces entre sus descripciones bibliográficas, y otras iniciativas de datos enlazados con registros de autoridades (como VIAF), e iniciativas sobre materias y taxonomías (como FAST y la Clasificación Decimal de Dewey). Los resultados de esta iniciativa, en teoría, deberán hacer que los significados de la información almacenada en Worldcat.org sean más accesibles para los motores de búsqueda.

8. CONCLUSIONES FINALES: RETOS Y OPORTUNIDADES PARA LOS EXPERTOS EN ORGANIZACIÓN DEL CONOCIMIENTO EN EL ESCENARIO TECNOLÓGICO

Además de estas iniciativas, hay muchos otros aspectos en relación a la Web Semántica donde los

expertos en Organización del Conocimiento podrían ser útiles y necesarios. Algunos de estos aspectos parten de la premisa de que de que la inteligencia humana está todavía por delante de la Inteligencia Artificial y, pese a sus complejidades, también del componente social de la Web Semántica, el cual, ante sus características como inteligencia colectiva interconectada, está haciendo crecer también exponencialmente el crecimiento tecnológico (San Segundo y Martínez-Ávila, 2012). Algunos de estos retos y oportunidades para los expertos en Organización del Conocimiento en la intersección con las Tecnologías de la Información están resumidos en los siguientes puntos:

1. Asignación de términos: Las máquinas todavía no son tan buenas identificando la temática de los recursos de información sin algún tipo de intervención o descripción humana. La clasificación automática y los métodos de indización usando técnicas probabilísticas y lingüísticas, muchas veces incluyendo aspectos de Inteligencia Artificial, todavía no han probado ser tan precisos como para prescindir de una supervisión humana. Sin embargo, aquí Birger Hjørland también ha afirmado que la clasificación hecha por humanos podría ser tan infructuosa como la de las máquinas si usaran las mismas teorías de conocimiento (Hjørland, 2011), es decir, si no existe una correcta formación teórica en los individuos realizando el proceso de asignación. Una correcta formación de las teorías y paradigmas en indización y asignación de términos podría beneficiar tanto a los procesos de clasificación e indización humanos como de desarrollo de herramientas automáticas de asignación.

2. Unión de vocabularios: Aunque muchos de los problemas de unión de bases de datos y construcción de vocabularios interoperables han sido solucionados en parte con los estándares de la Web Semántica, todavía existen algunos problemas relacionados con el alineamiento y unión de vocabularios que son discutidos por la comunidad científica. Guus Schreiber ha señalado los siguientes aspectos, relacionados con la Organización del Conocimiento, en los que se hace necesaria una mayor investigación en este área: las bases de representación, necesidad de una base establecida de vocabularios publicados, problemas de alineación, y enriquecimiento de metadatos (Schreiber, 2011).

3. Análisis de Dominio: Aunque el concepto de análisis de dominio ha sido también utilizado en el campo de la Inteligencia Artificial, desde su introducción en el campo de la Documentación (Hjørland y Albrechtsen, 1995), el concepto de análisis de dominio se trata de un método comúnmente asociado al área de la Organización del Conocimiento y uno de los principales enfoques de investigación en Biblioteconomía y Documentación (López-Huertas, 2008). Respecto a su relación con la Inteligencia Artificial, Birger Hjørland indica que: "Los campos de la Inteligencia Artificial y las ciencias cognitivas han estado históricamente

relacionados y dominados por pensamientos individualistas en vez de sociales. Hay muchos tipos de investigaciones sobre como los investigadores, expertos y gente ordinaria piensa en y sobre diferentes dominios de conocimiento. Alguna de esta investigación está relacionada con el intento de construir sistemas expertos con inteligencia artificial. Esta investigación ha sido principalmente hecha con una visión mecánica del pensamiento humano, negando los aspectos históricos y culturales de la cognición humana. Un texto influyente con una presentación de los problemas es Sowa (1999). Dentro de la psicología, Hoffman (1992) introduce investigación sobre pericia humana" (Hjørland, 2002). Siendo la Inteligencia Artificial uno de los once enfoques propuestos por Hjørland para el análisis de dominio, el uso del análisis de dominio aplicado al campo de la Inteligencia Artificial podría proporcionar modelos mentales de un dominio o métodos para la extracción de conocimiento con la finalidad de producir sistemas expertos. Por otra parte, la técnica del análisis de dominio, tal como se ha venido enseñando en los cursos sobre Organización del Conocimiento en la Universidad de Washington, Seattle, por profesores como Jens-Erik Mai y Joe Tennis (Tennis, 2003), es utilizado por los expertos contratados por Amazon en Seattle para mejorar la organización y recuperación de recursos en el sistema.

4. Evaluación y diseño de sistemas: En la práctica, son usuales las quejas de bibliotecarios y usuarios sobre el uso de los sistemas y los problemas de comunicación y entendimiento con los vendedores respecto a las verdaderas necesidades en Organización del Conocimiento. Para una mejor comunicación con los vendedores, cuando esto sea posible, o incluso para el desarrollo de los sistemas, es necesario conocer las características específicas y deseables en Organización del Conocimiento que los sistemas deberían ser capaces de incluir, así como las posibilidades y limitaciones técnicas que existen para implementar esas características. Un mejor entendimiento entre usuarios y desarrolladores podría ser alcanzado con un mayor entendimiento e investigación en la intersección entre la Organización del Conocimiento y las Tecnologías de la Información. Aunque Taylor y Joudrey (2009) indican que el principal problema en el desarrollo de sistemas bibliotecarios es que los vendedores rehúyen seguir estrictamente los estándares, y otras recomendaciones, para poder distanciarse comercialmente de sus competidores, lo cierto es que muchos de los sistemas más innovadores (no solo utilizados en bibliotecas, como por ejemplo *Endeca*) explotan las posibilidades tecnológicas de investigaciones clásicas en Organización del Conocimiento como la búsqueda facetada.

5. Etiquetado social: Uno de los principales problemas del etiquetado social y de las folksonomías es la variable calidad de los metadatos, en ocasiones ocurre porque los etiquetadores pueden carecer de conocimiento sobre la materia y, en otros

casos, cuando es realizado por especialistas en la materia, porque estos carecen de formación en Organización del Conocimiento. Un corolario a esto es que si las Tecnologías de la Información permiten a los usuarios organizar la información de una manera simple, la formación y aspectos teóricos en Organización del Conocimiento podrían ser totalmente innecesarios y sustituidos por aplicaciones electrónicas sin fundamentos. En relación a esto, Birger Hjørland escribió en 2006 sobre las competencias profesionales en Biblioteconomía y Documentación en comparación con legos y especialistas en la materia: "Quizás la ideología dominante en la Biblioteconomía y Documentación es que los profesionales de la Biblioteconomía y Documentación saben cómo usar tecnologías avanzadas para organizar el conocimiento, y que esta es la principal diferencia entre legos y profesionales de la Biblioteconomía y Documentación. Hay por supuesto algo de verdad en esto, pero la aplicación de la tecnología no forma la base de la ciencia o el campo intelectual de estudio. La tecnología es tan solo una herramienta. La Organización del Conocimiento como campo de estudio está relacionada con teorías semánticas, teorías bibliográficas y similares. Si, por ejemplo, consideramos a un coleccionista de sellos, esa persona podría tener mayor conocimiento sobre fuentes de información relevantes y cómo organizarlas que un profesional de la Biblioteconomía y Documentación. Una idea sobre la diferencia entre legos y especialistas en la materia por una parte, y profesionales de la Biblioteconomía y Documentación por otra, es que legos y especialistas en la materia son expertos en un campo estrecho desde el que podrían ampliar sus competencias. En otras palabras: estás trabajando de abajo a arriba. Los profesionales de la Biblioteconomía y Documentación, por otra parte, están trabajando de arriba a abajo: ellos empiezan con un conocimiento sobre fuentes de información interdisciplinarias amplias, como por ejemplo bibliografías nacionales e índices de citas y desde este punto de partida se especializan tanto como pueden" (Hjørland 2006). Siguiendo con los argumentos expresados por Hjørland, haría falta algo más que solo conocimiento sobre la materia o entrenamiento en nuevas tecnologías para hacer un correcto trabajo en Organización del Conocimiento. Muchos de los actuales problemas del etiquetado social y otras prácticas colaborativas podrían evitarse con una buena combinación de conocimiento de nuevas tecnologías y pericia en Organización del Conocimiento.

6. Aspectos sociológicos y compromiso ético: Una de las visiones más pesimistas que se han dado sobre la Web Semántica, fue escrita por Robert Cairo y citada en la Third International Conference on Intelligent Systems Modelling and Simulation de 2012: "La Web Semántica nunca funcionará porque depende de negocios trabajando juntos, cooperando. Estamos hablando del grupo de gente más conservadora del mundo, gente que cree en éticas de negocio codiciosas y despiadadas, gente que roba-

ría la propiedad de otros si no estuvieran clavadas al suelo. La gente que diseñó la Web Semántica nunca leyó sus textos de epistemología (la parte de la filosofía que trata sobre el estudio de cómo conocemos cosas). Pero el gran problema es que creyeron que todo el mundo trabajaría junto: -se pondrían de acuerdo en estándares web -expondrían sus APIs confiadamente de manera que cualquiera pudiera usarlas" (Cairo citado por Agarwal, 2012). En relación a esto, es cierto que no se puede forzar a los demás a colaborar y trabajar juntos en la Web Semántica, pero lo que sí se puede hacer es contribuir en parte compartiendo y organizando datos, y, si se dispone de una buena formación en Organización del Conocimiento y Tecnologías de la Información, se podrá hacer de forma que la información pueda ser comprendida y procesada por otros. Por otra parte, el campo bibliotecario al que normalmente la Organización del Conocimiento es asociado, no siempre refleja esos codiciosos comportamientos de los intereses comerciales que son presentados en la cita. En el informe escrito por la Biblioteca del Congreso de Estados Unidos en 2012 sobre BIBFRAME, proyecto desarrollado y relacionado con la Organización del Conocimiento y MARC, se señala que "el enfoque de anotación de BIBFRAME es clave para crear un sistema guiado por la capacidad de conocimiento de las bibliotecas y no por los intereses comerciales que actualmente dominan la Web" (Miller y otros, 2012). Es por esta razón que el futuro de los expertos en Organización del Conocimiento pasará por además de una buena formación en Organización del Conocimiento y en Tecnologías de la Información, por un compromiso ético con aquellos aspectos que afectan a la profesión y al acceso a la información en sentido amplio. Y es en este punto de la intersección de la Organización del Conocimiento y la Tecnologías de la Información donde, además de guiar los principales éxitos y avances teóricos en el área, quizás también residan los principales retos y oportunidades en el futuro.

9. BIBLIOGRAFÍA

- Abbate, J. (1999). *Inventing the Internet*. Cambridge, Mass.: MIT Press, p. 274.
- Agarwal, P. R. (2012). Semantic Web in Comparison to Web 2.0. *Third International Conference on Intelligent Systems, Modelling and Simulation (ISMS)*, pp. 558-563. Ahmedabad, India: Jyoti Microsyst. Pvt. Ltd. <http://dx.doi.org/10.1109/ISMS.2012.49>
- Baeza-Yates, R.; Ribeiro-Neto, B. (2011). *Modern information retrieval: the concepts and technology behind search*. (2a ed.). New York: ACM, p. 913.
- Berners-Lee, T. (1998a.) *Semantic Web Road map*. Disponible <http://www.w3.org/DesignIssues/Semantic.html> [Consulta 19 de junio de 2013].
- Berners-Lee, T. (1998b). *Web design issues; What a semantic can represent*. Disponible <http://>

- www.w3.org/DesignIssues/RDFnot.html [Consulta 19 de junio de 2013].
- Berners-Lee, T. (1999). *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor*. New York: Harper-Collins, p. 256.
- Berners-Lee, T. (2006). *Linked Data*. Disponible <http://www.w3.org/DesignIssues/LinkedData.html> [Consulta 19 de junio de 2013].
- Berners-Lee, T; Hendler, J.; Lassila, O. (2001). The Semantic Web. *Scientific American*, May, 1-4.
- Bologna, J.; Walsh, A. M. (1997). *The Accountant's Handbook of Information Technology*. New York: John Wiley and Sons, p. 384.
- Bostrom, N. (2006). AI Set to Exceed Human Brain Power. *CNN* August 9. Disponible <http://edition.cnn.com/2006/TECH/science/07/24/ai.bostrom/> [Consulta 19 de junio de 2013].
- Budd, J. (2006). Discourse Analysis and the Study of Communication in LIS. *Library Trends*, vol. 55 (1), 65-82. <http://dx.doi.org/10.1353/lib.2006.0046> [Consulta 19 de junio de 2013].
- Budd, J.; Raber, D. (1996). Discourse Analysis: Method and Application in the Study of Information. *Information Processing & Management*, vol. 32, 217-226. [http://dx.doi.org/10.1016/S0306-4573\(96\)85007-2](http://dx.doi.org/10.1016/S0306-4573(96)85007-2)
- Clocksins, W. (2003). Artificial Intelligence and the Future. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, vol. 361 (1809), 1721-1748. <http://dx.doi.org/10.1098/rsta.2003.1232> [Consulta 19 de junio de 2013].
- Crevier, D. (1993). *AI: The Tumultuous Search for Artificial Intelligence*. New York: BasicBooks, p. 400.
- Dahlberg, I. (2006) Knowledge organization: a new science?. *Knowledge Organization*, vol. 33 (1), 11-19.
- Doctorow, C. (2001). Metacrap: Putting the Torch to Seven Straw-Men of the Meta-Utopia. *The Well*. Disponible <http://www.well.com/~doctorow/metacrap.htm> [Consulta 19 de junio de 2013].
- Durville, P.; Gandon, F. (2009). Filling the gap between Web 2.0 technologies and natural language processing pipelines with semantic web. *SEMAPRO '09. Third International Conference on Advances in Semantic Processing, 2009*, pp. 109-112. <http://dx.doi.org/10.1109/SEMAPRO.2009.23> [Consulta 19 de junio de 2013].
- El-Goarany, K.; Saleh, I.; Kulczycki, G. (2008). The Social Service Network - Web 2.0 Can Make Semantic Web Services Happen. *10th IEEE Conference on E-Commerce Technology and the Fifth IEEE Conference on Enterprise Computing, E-Commerce and E-Services*. pp. 419-423.
- Frohmann, B. (1994). Discourse Analysis as a Research Method in Library and Information Science. *Library and Information Science Research*, vol. 16, 119-138. [http://dx.doi.org/10.1016/0740-8188\(94\)90004-3](http://dx.doi.org/10.1016/0740-8188(94)90004-3) [Consulta 19 de junio de 2013].
- Frohmann, B. (2001). Discourse and Documentation: Some Implications for Pedagogy and Research. *Journal of Education for Library & Information Science*, vol. 42, 13-26. <http://dx.doi.org/10.2307/40324034> [Consulta 19 de junio de 2013].
- García Marco, F. J. (2009). Perspectivas sobre el uso de la Web Semántica en el tratamiento de información y documentación legislativa. *Scire*, vol. 15 (1), 59-85.
- Greaves, M.; Mika, P. (2008). Semantic Web and Web 2.0. *Journal of Web Semantics*, vol. 6 (1), 1-3. <http://dx.doi.org/10.1016/j.websem.2007.12.002> [Consulta 19 de junio de 2013].
- Greenberg, J.; Losee, R.; Pérez Agüera, J. R.; Scherle, R.; White, H.; Willis, C. (2011). HIVE: Helping Interdisciplinary Vocabulary Engineering. *Bulletin of the American Society for Information Science and Technology*, vol. 37 (4). Disponible http://www.asis.org/Bulletin/Apr-11/AprMay11_Greenberg_etAl.html [Consulta 19 de junio de 2013].
- Gruber, T. (2007). Collective Knowledge Systems: Where the Social Web meets the Semantic Web. *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 6 (1), 4-13. <http://dx.doi.org/10.1016/j.websem.2007.11.011>
- Himanen, P. (2001). *The Hacker Ethic and the Spirit of the Information Age*. New York: Random House Trade, p. 232.
- Hjørland, B. (2002). Domain analysis in information science: Eleven approaches - traditional as well as innovative. *Journal of Documentation*, vol. 58 (4), 422-462. <http://dx.doi.org/10.1108/00220410210431136>
- Hjørland, B. (2003). Fundamentals of knowledge organization. *Knowledge Organization*, vol. 30 (2), 87-111.
- Hjørland, B. (2006). Laymen as Knowledge Organizers. *Lifeboat for Knowledge Organization*. Disponible http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/laymen_as_knowledge_organizers.htm [Consulta 19 de junio de 2013].
- Hjørland, B. (2008). What is knowledge organization (KO)? *Knowledge Organization*, vol. 35 (2/3), 86-101.
- Hjørland, B. (2011). The importance of theories of knowledge: indexing and information retrieval as an example. *Journal of the American Society for Information Science and Technology*, vol. 62 (1), 72-77. <http://dx.doi.org/10.1002/asi.21451>
- Hjørland, B. (2012). Knowledge Organization = Information Organization? *Categories, Contexts and Relations in Knowledge Organization: Proceedings of the Twelfth International ISKO Conference 6-9 August 2012 Mysore, India*, pp. 206-211. Würzburg, Germany: Ergon.
- Hjørland, B.; Albrechtsen, H. (1995). Toward a new horizon in information science: Do-

- main-analysis. *Journal of the American Society for Information Science*, vol. 46 (6), 400-425. [http://dx.doi.org/10.1002/\(SICI\)1097-4571\(199507\)46:6<400::AID-ASIJ>3.0.CO;2-Y](http://dx.doi.org/10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASIJ>3.0.CO;2-Y)
- Hoffman, R. R. (1992). *The Psychology of Expertise: Cognitive Research and Empirical AI*. New York: Springer Press, p. 408. <http://dx.doi.org/10.1007/978-1-4613-9733-5>
- Hope, G.; Wang, T. (G.); Barkataki, S. (2007). Convergence of Web 2.0 and Semantic Web: A Semantic Tagging and Searching System for Creating and Searching Blogs. *International Conference on Semantic Computing, 2007. ICSC 2007*, pp. 201-208.
- Howard, A. (2013). Linking open data to augmented intelligence and the economy. *O'Reilly Radar* April 30. Disponible <http://radar.oreilly.com/2013/04/linking-open-data-to-augmented-intelligence-and-the-economy.html> [Consulta 19 de junio de 2013].
- IFLA Study Group on the Functional Requirements for Bibliographic Records. (2009). *Functional Requirements for Bibliographic Records: Final Report*. Disponible http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf [Consulta 19 de junio de 2013].
- Jeschke, S.; Wilke, M.; Natho, N.; Pfeiffer, O. (2008). KEA - a Mathematical Knowledge Management System combining Web 2.0 with Semantic Web Technologies. *4th International Conference on Innovations in Information Technology, 2007. IIT '07*, pp. 138-142.
- Kurzweil, R. (2005). *The Singularity is Near: When Humans Transcend Biology*. New York: Viking, p. 672.
- Laningham, S. (2006). *developerWorks Interviews: Tim Berners-Lee*. Disponible <http://www.ibm.com/developerworks/podcast/dwi/cm-int082206txt.html> [Consulta 19 de junio de 2013].
- Leiner, B. M.; Cerf, V. G.; Clark, D. D.; Kahn, R. E.; Kleinrock, L.; Lynch, D. C.; Postel, J.; Roberts, L. G.; Wolff, S. (2000). The History of the Internet. *Internet Society*. Reimpreso en Leiner, B. M.; Cerf, V. G.; Clark, D. D.; Kahn, R. E.; Kleinrock, L.; Lynch, D. C.; Postel, J.; Roberts, L. G.; Wolff, S. (2009). The History of the Internet. *ACM SIGCOMM Computer Communication Review* vol. 39 (5), 22-31. <http://dx.doi.org/10.1145/1629607.1629613>
- Lemen, C. (2010). The New Knowledge Processing: Integrating Semantic Web with Web 2.0. *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, pp. 314-317. <http://dx.doi.org/10.1109/ICCASM.2010.5620527>
- López-Huertas, M. J. (2008). Some current research questions in the field of knowledge organization. *Knowledge Organization*, vol. 35, 113-136.
- Markoff, J. (2005). Behind Artificial Intelligence, a Squadron of Bright Real People. *The New York Times* October 14. Disponible http://www.nytimes.com/2005/10/14/technology/14artificial.html?_r=0 [Consulta 19 de junio de 2013].
- Martínez-Ávila, D. (2012a). Problems and Characteristics of Foucauldian Discourse Analysis as a Research Method. En Smiraglia, R. P.; Lee, H-L. (editores) *Cultural Frames of Knowledge*. Ergon, Würzburg, Germany.
- Martínez-Ávila, D. (2012b). *DDC-BISAC switching a new case of reader-interest classification*. Tesis Doctoral. Universidad Carlos III de Madrid.
- Martínez-Ávila, D.; Smiraglia R. P. (2013). Revealing Perception: Discourse Analysis in a Phenomenological Framework. *Proceedings from North American Symposium on Knowledge Organization, Vol. 4. Milwaukee, Estados Unidos: University of Wisconsin-Milwaukee*. Disponible http://www.iskocus.org/NASKO2013proceedings/MartinezAvila_Smiraglia_Revealing_Perception.pdf [Consulta 10 de febrero de 2014]
- Mas, A. (2005). *Agentes software y sistemas multiagente: conceptos, arquitecturas y aplicaciones*. Madrid: Prentice Hall, p. 286.
- McCarthy, J. (2007). *What is Artificial Intelligence?* Disponible <http://www-formal.stanford.edu/jmc/whatisai/whatisai.html> [Consulta 19 de junio de 2013].
- McCarthy, J.; Minsky, M. L.; Rochester, N.; Shannon, C. E. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Disponible <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html> [Consulta 19 de junio de 2013].
- McCorduck, P. (2004). *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. (2.^a ed.) Natick, MA: A. K. Peters, p. 576.
- Miller, E.; Ogbuji, U.; Mueller, V.; MacDougall, K. (2012). *Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services*. Washington, DC: Library of Congress. Disponible <http://www.loc.gov/marc/transition/pdf/marcld-report-11-21-2012.pdf> [Consulta 19 de junio de 2013].
- Noy, N. F.; McGuinness, D. L. (2001). Ontology Development 101: A Guide to Creating Your First Ontology'. *Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001*. Disponible <http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.doc> [Consulta 10 de febrero de 2014].
- O'Reilly, T. (2005). *What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. Disponible <http://oreilly.com/web2/archive/what-is-web-20.html> [Consulta 19 de junio de 2013].
- Palma Méndez, J. T.; Marín Morales, R. (2008). *Inteligencia artificial: métodos, técnicas y aplicaciones*. Madrid: MacGraw-Hill, p. 1022.

- Raymond, E. S. (2001). *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary Revised and Expanded*. Cambridge, Mass.: O'Reilly, p. 241.
- Roberts, L. (1967). Multiple Computer Networks and Intercomputer Communication. *Proceedings of ACM Symposium on Operating System Principles*, Gatlinburg, Tenn. <http://dx.doi.org/10.1145/800001.811680> [Consulta 19 de junio de 2013].
- Rowley, J.; Hartley, R. (2008). *Organizing Knowledge: An Introduction to Managing Access to Information*. (4.ª ed.). Hampshire: Ashgate, p. 367.
- Russell, S. J.; Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*. (3.ª ed.) Upper Saddle River, New Jersey: Prentice Hall, p. 1152.
- San Segundo, R. (2002). A new concept of knowledge. *Online Information Review*, vol 26 (4), 239-245. <http://dx.doi.org/10.1108/14684520210438688> [Consulta 19 de junio de 2013].
- San Segundo, R.; Martínez-Ávila, D. (2012). New Conceptual Structures for the Digital Environment: From KOS to the Semantic Interconnection. En: Neelameghan, A.; Raghavan, K. S. (editores). *Categories, Contexts and Relations in Knowledge Organization: Proceedings of the Twelfth International ISKO Conference 6-9 August 2012 Mysore, India*. Würzburg, Germany: Ergon.
- Schreiber, G. (2011). Issues in Publishing and Aligning Web Vocabularies. Presentado en *Classification and Ontology: Formal Approaches and Access to Knowledge: UDC Seminar, 19-20 September, The Hague, The Netherlands*. Disponible <http://www.cs.vu.nl/~guus/talks/11-udc.pdf> [Consulta 19 de junio de 2013].
- Smiraglia, R. (2012). Knowledge Organization: Some Trends in an Emergent Domain. *El Profesional de la Información*, vol. 21, 225-227. <http://dx.doi.org/10.3145/epi.2012.may.01> [Consulta 19 de junio de 2013].
- Sowa, J. F. (1999). *Knowledge Representation: Logical, Philosophical and Computational Foundations*. Pacific Grove, CA: Brooks Cole Publishing Co., p. 594.
- Taylor, A. G.; Joudrey, D. N. (2009). *The Organization of Information*. (3ª ed.) Westport, Connecticut: Libraries Unlimited, p. 512.
- Tennis, J. T. (2003). Two Axes of Domains for Domain Analysis. *Knowledge Organization*. vol. 30 (3/4), 191-195.
- Weizenbaum, J. (1976). *Computer Power and Human Reason*. London: Penguin, p. 300.