

SCIRE

Representación y organización del conocimiento

SCIRE

Representación y organización del conocimiento

Vol. 21, n.º 2, julio-diciembre 2015

ISSN 1135-3716

Scire:

knowledge representation and organization
Vol. 21, n. 2, July-December 2015

Ibersid:

Red de Investigación
en Sistemas de Información
y Documentación

© 2015 Francisco Javier García Marco y los autores.

Reservados todos los derechos. Prohibida la reproducción por cualquier medio de la totalidad o parte de la presente edición sin permiso escrito de los titulares del copyright. Queda, sin embargo, autorizada expresamente la reproducción de los resúmenes y palabras clave de los artículos en inglés y español. También se permite la reproducción de sus textos con objetivos exclusivamente docentes para su uso ocasional en el aula, en cuyo caso se excluye el almacenamiento superior a un mes en un servidor.

Los artículos publicados están sometidos a evaluación científica y editorial en beneficio de la comunidad científica y los autores, pero dicha evaluación se realiza dentro de los límites del desarrollo actual de la disciplina, las limitaciones humanas de editores y revisores, y con una perspectiva ideológica, teórica y metodológica abierta, por lo que la dirección no siempre comparte las opiniones, métodos y procedimientos reflejados en la revista. Los autores se responsabilizan personalmente del cumplimiento de los códigos éticos y la legislación vigente, así como de la obtención de los permisos de copia de materiales ajenos.

ISSN: 1135-3716 = Scire (Zaragoza)
Depósito legal: Z. 1.790 — 1995

Edita: Ibersid® con la colaboración de Prensas Universitarias de Zaragoza

Imprime:
Servicio de Publicaciones. Universidad de Zaragoza.
Edificio de Ciencias Geológicas, C/ Pedro Cerbuna, 12.
50009 Zaragoza, España. Tel.: 976 761 330. Fax: 976 761 063.

Scire

representación y organización
del conocimiento

Alcance y objetivos

Scire: representación y Organización del Conocimiento es una publicación semestral de carácter interdisciplinar sobre la representación, normalización, tratamiento, recuperación y comunicación de la información y el conocimiento.

Difusión

Scire tiene difusión internacional. Agradecemos la inclusión en los siguientes servicios de referencia: Scopus, Information Science Abstracts, Information Services in Physics, Electronics and Computing, Library and Information Science Abstracts, Sociological Abstracts, ERIH Plus, Knowledge Organization Literature, Base de Datos ISOC y Catálogo Latindex.

Redacción, distribución y canje

Revista Scire
Departamento de Ciencias de la Documentación e Historia de la Ciencia
Facultad de Filosofía y Letras
Universidad de Zaragoza
C/ Pedro Cerbuna 12,
E-50.009 Zaragoza (Spain)

Tfno: int + 34 976 762239. Fax: 34 976761506.
E-mail: <mailto:scire@unizar.es>

Suscripciones y números sueltos

Suscripción anual / Annual subscription: 30 €.
Número suelto: 20 €.
(IVA incluido)

Scire

knowledge organization
and representation

Scope and aims

Scire: Representación y Organización del Conocimiento is an interdisciplinary journal published twice a year that is devoted to the representation, standardization, treatment, retrieval and communication of information and knowledge.

Dissemination

Scire has international distribution. We acknowledge its inclusion in the following reference services: Scopus, Information Science Abstracts, Information Services in Physics, Electronics and Computing, Library and Information Science Abstracts, Sociological Abstracts, ERIH Plus, Knowledge Organization Literature, Base de Datos ISOC and Catálogo Latindex.

Contact address

Revista Scire
Departamento de Ciencias de la Documentación e Historia de la Ciencia
Facultad de Filosofía y Letras
Universidad de Zaragoza
C/ Pedro Cerbuna 12,
E-50.009 Zaragoza (Spain)

Tel.: int + 34 976 762239. Fax: 34 976761506.
E-mail: scire@unizar.es

Subscriptions

Annual subscription: 30 €.
Issue: 20 €.
(VAT included)

Editor

Francisco Javier García Marco, Univ. de Zaragoza. E-mail: jgarcia@unizar.es

Consejo de redacción / Editorial council

Mario Guido Barité Roqueta, Universidad de Montevideo, Uruguay	João Batista Ernesto Moraes, Universidade Estadual Paulista, Brasil	Francisco Javier Martínez Mendez, Universidad de Murcia, España
José Augusto Chaves Guimarães, Universidade Estadual Paulista, Brasil	Francisco Javier García Marco, Universidad de Zaragoza, España	Álvaro Quijano Solís, Colegio de México, México

Consejo científico / Scientific council

Tomás Baiget, EPI S. A., España Mercedes Caridad Sebastián, Universidad Carlos III de Madrid, España	Alfonso López Yepes, Universidad Complutense de Madrid, España José López Yepes, Universidad Complutense de Madrid, España	Gloria Ponjuán Dante, Universidad de La Habana, Cuba José Vicente Rodríguez Muñoz, Universidad de Murcia, España
Alberto Carreras Gargallo, Universidad de Zaragoza, España	Pedro Marijuán Fernández, Universidad de Zaragoza, España	Adelaida Román Román, CINDOC (Madrid), España
Emilia Currás Puente, Universidad Autónoma de Madrid, España	Bonifacio Martín del Brío, Universidad de Zaragoza, España	Juan Ros García, Universidad de Murcia, España
Constança Espelt Busquets, Universidad de Barcelona, España	José Antonio Moreira González, Universidad Carlos III de Madrid, España	Francisco José Ruiz de Mendoza Ibáñez, Universidad de La Rioja, España
María Eulalia Fuentes Pujol, Universidad Autónoma de Barcelona, España	Purificación Moscoso Castro, Universidad de Alcalá, España	Félix Sagredo Fernández, Universidad Complutense de Madrid, España
Fernando Galindo Ayuda, Universidad de Zaragoza, España	Félix Moya Anegón, Universidad de Granada, España	Elías Sanz Casado, Universidad Carlos III de Madrid, España
Blanca Gil Urdicián, Universidad Complutense de Madrid, España	María del Carmen Negrete Gutiérrez, Universidad Autónoma de México	Carlos Serrano Cinca, Universidad de Zaragoza, España
Alan Gilchrist, Cura Consortium, Reino Unido	José Luis Otal, Universidad Jaume I de Castellón, España	Emilio Setién Quesada, Asociación Cubana de Bibliotecarios, Cuba
José María Izquierdo Arroyo, Universidad de Murcia, España	Manuel José Pedraza Gracia, Universidad de Zaragoza, España	Félix del Valle Gastaminza, Universidad Complutense de Madrid, España
María Pilar Lasala Calleja, Universidad de Zaragoza, España	María Pinto Molina, Universidad de Granada, España	

Revisores externos del número / External reviewers in this issue

Agradecemos la colaboración de los profesores Carlos Cândido de Almeida, José Luis Alonso Berrocal, María Luisa Alvite Díez, Montserrat García Alsina, Joaquín López del Ramo, Daniel Martínez Ávila, Michela Montesi, Luis Miguel Moreno Fernández, Juan Pastor Sánchez, Ana Ríos Hilarario, Antonia Salvador Benitez.

Candidaturas al consejo científico

Se aceptan candidaturas al consejo científico de especialistas del área de Biblioteconomía y Documentación y de otras disciplinas relacionadas (Informática, Ciencias Sociales, Lingüística, Filosofía, Psicología, etc.) con experiencia profesional e investigadora demostrada. En el sistema público de investigación español, suele ser equivalente al doctorado y dos sexenios de investigación o méritos equivalentes.

Scientific council membership policy

Candidatures of researchers from LIS and other related disciplines (Computer Science, Social Sciences, Linguistics, Philosophy, Psychology, etc.) with demonstrated professional and research experience are welcomed. In the Spanish public research system, for example, this usually means having a doctorate and two scientific productivity sexennia or equivalent outputs.

Tabla de contenidos en español

Table of contents in Spanish

<i>Tabla de contenidos en español</i>	9	<i>Procedimiento de evaluación</i>	83
<i>Tabla de contenidos en inglés</i>	11	<i>Instrucciones para la presentación de trabajos</i>	85
Artículos		<i>Índice de autores</i>	88
<i>La dimensión conceptual de la organización del conocimiento en el universo científico de la ISKO: un análisis de dominio de los congresos de ISKO-Brasil, ISKO-España, ISKO-Norteamérica e ISKO-Francia</i>		<i>Índice de materias en español</i>	88
<i>José Augusto Chaves GUIMARÃES, Daniela Fernanda de Oliveira MATOS, André Ynada dos SANTOS, Rodrigo de SALES</i>	13	<i>Índice de materias en inglés</i>	88
<i>Aplicación de la Teoría Crítica de Raza en la organización y representación del conocimiento</i>			
<i>Daniel MARTÍNEZ-ÁVILA, Marcio FERREIRA, José L. MAGRO</i>	27		
<i>Los datos bibliotecarios en la nube de datos: análisis de los datasets GLAM presentes en el LOD cloud diagram</i>			
<i>Ana B. RÍOS HILARIO, Blanca GIL URDICIÁN</i>	35		
<i>La investigación sobre políticas de indización en Brasil: avances y desafíos</i>			
<i>Paula Regina DAL'EVEDOVE, Mariângela Spotti Lopes FUJITA</i>	49		
<i>Edición de contenidos en un entorno colaborativo: el caso de la Wikipedia en español</i>			
<i>Ángel F. ZAZO RODRÍGUEZ, Carlos G. FIGUEROLA, José Luis ALONSO BERROCAL</i>	57		
<i>Archivo de Espacio P: propuesta de un cuadro de clasificación para su continuidad digital</i>			
<i>Federico CASTRO MORALES, Fátima GARCÍA LÓPEZ, Ana María MORALES GARCÍA, Karin OHLENSCHLÄGER OHLENSCHLÄGER</i>	69		
<i>Metateoría y diplomática: la construcción de un nuevo modelo de análisis para la identificación de los documentos de archivo</i>			
<i>Natália Bolfarini TOGNOLI</i>	75		

Tabla de contenidos en inglés

Table of contents in English

<i>Table of contents in Spanish</i>	9	<i>Evaluation process</i>	83
<i>Table of contents in English</i>	11	<i>Instructions for authors</i>	85
Articles		<i>Author index</i>	88
<i>The conceptual dimension of knowledge organization in the scientific universe of ISKO: a domain-analysis of ISKO-Brazil, ISKO-Spain, NASKO and ISKO France conferences</i>		<i>Subject index in Spanish</i>	88
<i>José Augusto Chaves GUIMARÃES, Daniela Fernanda de Oliveira MATOS, André Ynada dos SANTOS, Rodrigo de SALES</i>	13	<i>Subject index in English</i>	88
<i>Application of Critical Race Theory to knowledge organization and representation</i>			
<i>Daniel MARTÍNEZ-ÁVILA, Marcio FERREIRA, José L. MAGRO</i>	27		
<i>Library data in the data cloud: an analysis of the GLAM datasets in the LOD cloud diagram</i>			
<i>Ana B. RÍOS HILARIO, Blanca GIL URDICIAIN</i>	35		
<i>Research on indexing policy in Brazil: progress and challenges</i>			
<i>Paula Regina DAL'EVEDOVE, Mariângela Spotti Lopes FUJITA</i>	49		
<i>Editing content in a collaborative environment: the case of the Spanish Wikipedia</i>			
<i>Ángel F. ZAZO RODRÍGUEZ, Carlos G. FIGUEROLA, José Luis ALONSO BERROCAL</i>	57		
<i>A classification scheme for the "Espacio P" Archive: a proposal for its digital continuity</i>			
<i>Federico CASTRO MORALES, Fátima GARCÍA LÓPEZ, Ana María MORALES GARCÍA, Karin OHLENSCHLÄGER</i>	69		
<i>Metatheory and Diplomats: building a new model of analysis for the identification of archival records</i>			
<i>Natália Bolfarini TOGNOLI</i>	75		

A dimensão conceitual da organização do conhecimento no universo científico da ISKO: uma análise de domínio a partir dos congressos da ISKO-Brasil, ISKO-Espanha, ISKO-América Do Norte e ISKO-França

La dimensión conceptual de la organización del conocimiento en el universo científico de la ISKO: un análisis de dominio de los congresos de ISKO-Brasil, ISKO-España, ISKO-Norteamérica e ISKO-Francia

The conceptual dimension of knowledge organization in the scientific universe of ISKO: a domain-analysis of ISKO-Brazil, ISKO-Spain, ISKO-North America and ISKO France conferences

**José Augusto Chaves GUIMARÃES (1), Daniela Fernanda de Oliveira MATOS (2),
André Ynada dos SANTOS (3), Rodrigo de SALES (4)**

(1) (2) (3) Departamento de Ciência da Informação. Universidade Estadual Paulista - Av. Hygino Muzzi Filho, 737 – Marília –SP – BRASIL. Correio eletrônico: gpfapoi@gmail.com (4) Departamento de Ciência da Informação. Universidade Federal Fluminense – UFF. Rua Tiradentes, 148 - Ingá – Niterói - RJ – BRASIL. Correio eletrônico: rodrigo.sales.s@gmail.com

Resumen

Se analiza comparativamente la definición conceptual de la organización del conocimiento y los diferentes referentes teóricos que la sostienen en el ámbito de los congresos realizados por los capítulos brasileño, español, norteamericano y francés de ISKO en sus dos últimos eventos realizados (2011 y 2013). Para ello, se utiliza análisis de contenido y análisis de citas. Los resultados revelan que los cuatro capítulos analizados son bastante prolíficos en el abordaje de la cuestión conceptual del área, aunque ese aspecto todavía se encuentra en construcción. En particular, se detectan dos grandes preocupaciones: por un lado, abordar los aspectos relativos a la naturaleza y al objeto de la organización del conocimiento en su configuración epistemológica, y, por otro, una preocupación de naturaleza más pragmática y operacional, volcada a los procesos e instrumentos de la organización del conocimiento, muchas veces concebidos como objeto del campo. En este contexto, los referentes teóricos predominantes se centran en siete autores (Hjørland; Dahlberg; Gardin; Mai; Olson; Barité y Thelleffssen). Hjørland y Dahlberg asumen un fuerte liderazgo, lo que revela, respectivamente, el énfasis de los abordajes sociocognitivo y ontológico del área, siendo que la primera asume una mayor fuerza en la medida en que especialmente Mai y Thelleffssen, con la cuestión semiótica, y Olson, con los estudios culturales, refuerzan ese cuadro teórico.

Palabras clave: Organización del conocimiento. Concepto. Análisis de dominio. Análisis de contenido. Sociedad Internacional para la organización del conocimiento. ISKO. España. Brasil. Francia. América del Norte.

Abstract

We comparatively analyze the conceptual question of knowledge organization and the theoretical references that are fundamental in the last two Brazilian, Spanish, North American and French ISKO chapter meetings (of 2011 and 2013). From this, we conducted content analysis and citation analysis. The results show that the four analyzed chapters are quite fruitful in addressing conceptual issues, even when this aspect is still in development. This is verified by, on one hand, the concern with aspects such as the nature and object of knowledge organization in its epistemological configuration, and on the other hand, a more pragmatic and operational interest related to knowledge organization processes and tools, sometimes perceived as field objects. In this context, the main theoretical references involve seven authors (Hjørland, Dahlberg, Gardin, Mai, Olson, Barité and Thelleffssen). Hjørland and Dahlberg assume strong leaderships that show, respectively, the emphasis of social-cognitive and ontological approaches in the area. The emphasis of social-cognitive seems to take a greater force especially with Mai and Thelleffssen, with the semiotic question, and Olson, with cultural studies, that reinforce this theoretical framework.

Keywords: Knowledge organization. Concept. Domain analysis. Content analysis. International Society for Knowledge Organization. ISKO. Brazil. France. North America. Spain.

1. Introdução

A organização do conhecimento desempenha papel nuclear na Ciência da Informação, apresentando avanço teórico notadamente a partir da criação da ISKO, em 1989. No entanto, observa-se que esse campo ainda se encontra em fase de consolidação, em busca da construção e delimitação de sua própria base conceitual e seus limites, em virtude de distintas influências teóricas. Tal aspecto se revela, dentre outros, por uma diversidade no tocante ao seu aspecto conceitual, notadamente com relação a seu objeto, o que leva à necessidade de analisar de que forma o ambiente científico da ISKO vem construindo/delimitando essa dimensão conceitual, a partir de seu discurso científico oficial em âmbito internacional.

Em pesquisa anterior (Guimarães et al, 2014), relativa ao universo dos congressos internacionais da ISKO, foi possível observar que a organização do conhecimento constitui área ou campo de conhecimento e/ou de aplicação onde, relacionado à construção de discursos especializados. Sua natureza operacional encontra-se ora voltada à organização e representação de conceitos, ora voltada à busca de informação e ao acesso aos conhecimentos, em especial diálogo com a recuperação da informação. Nesse contexto, seu objeto reside na estrutura do conhecimento contida nos documentos (conhecimento registrado, socializado e publicado), com especial ênfase aos conceitos e à sua modelagem assim como às atividades discursivas em domínios científicos e práticas sociais e culturais específicos, a organização do conhecimento vale-se de instrumentos como sistemas de classificação, tesouros, vocabulário e outras linguagens de indexação. Como abordagens predominantes destacam-se as questões relativas a aspectos cognitivos, tecnológicos, lógico-conceituais, socioculturais e de gestão, o que sinaliza para os diálogos mais prementes na área.

Nesta oportunidade, e em continuidade à referida pesquisa, analisa-se comparativamente a questão conceitual da organização do conhecimento nos congressos realizados pelos capítulos brasileiro, espanhol, norte-americano e francês da ISKO nos últimos dois eventos realizados em cada (2011 e 2013) haja vista o fato de o capítulo brasileiro ser ainda recente.

Dessa forma, analisam-se especialmente os referentes teóricos que subsidiam essa dimensão conceitual do campo e até que medida existe diálogo entre essas quatro ambiências analisadas, a partir de seus referentes. Para tanto, vale-se de duas das onze vertentes de análise

de domínio propostas por Hjørland (2002): a epistemológica, a partir da análise de conteúdo das definições apresentadas, e bibliográfica, a partir das redes de citação que subsidiam essa produção.

2. A Organização do Conhecimento e a atuação científica da ISKO

Embora o termo organização do conhecimento remonte a Henry Evelyn Bliss, na década de 30 do século passado, sua constituição como campo de estudos interdisciplinar é bastante recente, podendo-se atribuir seu nascimento quatro décadas após, quando a partir da década de 70 quando Dagobert Soergel e Ingetraut Dalhberg buscaram conceituá-lo, aspecto que levou à criação, em 1974, na Alemanha, da Society for Classification e, em 1989, da International Society for Knowledge Organization-ISKO.

Para Dahlberg (2008), a Organização do Conhecimento possui natureza científica e destina-se a sistematizar unidades de conhecimento (conceitos) a partir dos elementos ou características que lhes são inerentes. A isso se alia a aplicação de conceitos e de classes de conceitos que possam permitir a transferência de conteúdos (assuntos). Para a autora (Dahlberg, 1993) duas concepções de conhecimento devem ser consideradas: enquanto processo individual, que não é transferível e aquele que pressupõe um consenso, a partir, de um registro, sendo, este último, objeto da organização do conhecimento. Assim, o termo Organização do Conhecimento, para a Ciência da Informação irá se dedicar às atividades de classificar, indexar e representar o conhecimento através de registros com o objetivo específico de suprir a necessidade de informação (Bufrem, 2008).

No âmbito da Ciência da Informação, a organização do conhecimento apresenta natureza mediadora, configurando-se em um conjunto de processos que estabelecem a intermediação entre um conhecimento que, uma vez produzido, foi materializado e socializado, de tal forma que esse conhecimento possa servir de base para a geração de um novo conhecimento. Este novo conhecimento, por sua vez, uma vez materializado e socializado, pode igualmente ser objeto de nova organização e representação caracterizando aquilo que se pode denominar como fluxo helicoidal da informação (Guimarães, 2008).

Tendo por objetivo promover o avanço teórico e aplicado da Organização do Conhecimento em diversos campos e em diferentes maneiras, congregando pesquisadores de distintos países, em áreas como Ciência da Informação, Linguís-

tica, Filosofia, e Ciência da Computação, entre outras. Desse modo, foram até então realizadas treze conferências internacionais (Darmstadt, 1990; New Delhi, 1992; Copenhagen, 1994; Washington, 1996; Lille, 1998; Toronto, 2000; Granada, 2002; Londres, 2004; Viena, 2006; Montréal, 2008; Roma, 2010; Mysore, 2012; e Cracóvia, 2014), além o que se tem toda uma atuação regional a partir do Germânico, Norte-americano, Francês, Espanhol, Italiano, Brasileiro, Polonês, Escandinavo, Maghreb e Indiano que, por sua vez, também realizam conferências nacionais e/ou regionais.

Nesse contexto, especial destaque merece quatro capítulos que vêm, ao longo do tempo, demonstrado forte produção científica na área bem como um alto nível de preocupação com as questões teóricas concernentes à organização do conhecimento: Espanha, França, América do Norte e Brasil.

A ISKO-Espanha, (hoje transformada em capítulo ibérico para abrigar também Portugal), foi criada em 1990, e possui já onze conferências realizadas (Madrid, 1993; Getafe, 1995 e 1997; Granada, 1999; Alcalá de Henares, 2001; Salamanca, 2003; Barcelona, 2005; León, 2007; Valencia, 2009; Ferrol, 2011; e Porto, 2013).

A ISKO-França A ISKO-França foi criada em 1998 e até o presente momento promoveu nove congressos: Lille, 1998; Lyon, 1999; Paris, 2001; Grenoble, 2003; Nancy, 2005; Toulouse, 2007; Lyon, 2009; Lille, 2011; e Paris, 2013.

A ISKO América do Norte – NASKO, criada em 2007, é formada pelo Canadá e pelos Estados Unidos, tendo realizado até então quatro congressos: Toronto, 2007; Syracuse, 2009; Toronto, 2011 e Milwaukee, 2013.

Por sua vez, o capítulo brasileiro surgiu a partir de um grupo de pesquisadores do Grupo de Trabalho 2: Organização e Representação do Conhecimento, da Associação Nacional de Pesquisa em Ciência da Informação (ANCIB), tendo sido criado em 2007. Até o momento foram realizados dois congressos: Brasília, 2011; e Rio de Janeiro, 2013.

3. Metodologia

A pesquisa desenvolveu-se a partir do conjunto de atas dos dois últimos congressos nacionais / regionais realizados (2011 e 2013) pelos capítulos brasileiro, espanhol, norte-americano e francês da ISKO, considerados capítulos de larga produção e que, ao longo do tempo, vêm mantendo certa dialogicidade científica.

A pesquisa centrou-se na abordagem da análise de domínio que, no âmbito internacional da Ciência da Informação, vem sendo tradicionalmente trabalhada, em termos teóricos e aplicados, por Hjørland & Albrechtsen (1995), Moya Anegón & Herrero Solana (2001), Hjørland (2002, 2004), Tennis (2003), e Smiraglia (2012), dentre outros, constitui uma importante abordagem para caracterização e avaliação da ciência, na medida em que permite identificar as condições pelas quais o conhecimento científico se constrói e se socializa. Para tanto, tem-se por domínio a concepção de Smiraglia (2012, p.114) enquanto

[...] um grupo que apresenta uma base ontológica que revela uma teleologia subjacente, um conjunto de hipóteses comuns, assim como um consenso epistemológico sobre as abordagens metodológicas, sociais e semânticas a serem utilizadas.

Desse modo, valeu-se de duas das onze abordagens de análise de domínio previstas por Hjørland (2002): a epistemológica e a bibliométrica. Para tanto, e tendo por base o percurso metodológico utilizado por Guimarães et al. (2014) partiu-se de um levantamento na coleção completa dos referidos Anais, tomando por base a incidência do termo "knowledge organization" (e seus correspondentes em francês, espanhol e português) suas variações em francês, nos títulos, palavras-chave, resumos e títulos de seções das comunicações publicadas. Em seguida, realizou-se a leitura das comunicações que revelaram tal incidência, selecionando-se aquelas que traziam conceitos, definições ou considerações teóricas acerca da natureza da organização do conhecimento.

A dimensão epistemológica foi abordada a partir do universo conceitual da literatura analisada. Dessa forma, no tocante às definições, procedeu-se à Análise de Conteúdo (Bardin, 2009), cuja escolha se justifica pelo fato de seus procedimentos possibilitarem uma análise com base em inferências extraídas de conteúdos de documentos – a partir de uma interpretação controlada por meio de variáveis ou indicadores, que proporcionam maior liberdade ao analista, sem que se perca a objetividade da investigação, e por se tratar de uma análise pautada em definições concebidas por autores/pesquisadores, ou seja, argumentos registrados textualmente. Dividida em três fases: pré-análise, exploração do material e, tratamento dos resultados (inferências e interpretações). Para tanto, a delimitação do corpus respeito as quatro regras básicas previstas por Bardin (2009): exaustividade (na medida em que todas as definições encontradas, quando da coleta de informações, foram consideradas materiais úteis à análise,

sem processo seletivo); representatividade (assegurada pela escolha do canal de comunicação científica, pois os trabalhos publicados nos Anais da ISKO são de autoria de pesquisadores e teóricos que representam uma parcela significativa da produção internacional em organização do conhecimento); homogeneidade (pois as definições analisadas são relativas a um mesmo tema, assim, são homogêneas tanto na estrutura textual quanto no assunto); e pertinência do material (as fontes são as próprias ideias explicitadas pelos autores/pesquisadores em suas definições).

As definições foram analisadas a partir das seguintes categorias: natureza, objeto, objetivo, instrumentos, processos, diálogos interdisciplinares e perspectivas ou abordagens.

Em seguida, procedeu-se à análise bibliométrica das fontes utilizadas nos referidos artigos de modo a identificar os referentes teóricos predominantes e as interlocuções estabelecidas entre os capítulos.

Para tanto, foram construídas e analisadas as redes de citação, valendo-se, para tanto do software PAJEK (2009).

4. Apresentação, análise e discussão dos resultados

Partiu-se de um conjunto de 85 trabalhos na ISKO-Brasil, 132 na ISKO-Espanha, 43 na NASKO e 45 na ISKO-França, totalizando 305 trabalhos. Desse conjunto, e a partir da incidência dos termos no título, chegou-se a um grupo de 33 trabalhos na ISKO-Brasil, 57 na ISKO-Espanha, 21 na NASKO e 9 na ISKO-França, totalizando 120 trabalhos. Em seguida, procedeu-se à leitura desse material, observando-se que nem todos traziam, efetivamente, conceituações, definições ou considerações teóricas sobre a organização do conhecimento. Dessa forma, chegou-se a um corpus final de 16 trabalhos na ISKO-Brasil, 17 na ISKO-Espanha, 10 na NASKO e 5 na ISKO-França, totalizando 48 comunicações.

Considerando o fato de que todos os autores das comunicações publicadas nos Anais da ISKO são pesquisadores e ou teóricos da área, cujas concepções contribuem efetivamente para a construção do universo epistemológico da organização do conhecimento em âmbito internacional, optou-se por nomeá-los em suas definições. Dessa forma, o universo dos autores das definições aqui levantadas é composto pelo seguinte conjunto de autores: a) NASKO: MacCall, 2011; Fox, 2011; Pattuelli, 2011; Adler & Tennis, 2013; Fox, Reece, 2013; Frické, 2013;

Jacob, Zhang, 2013; Marchese & Smiraglia, 2013; Olson & Howarth, 2013; Sales & Guimarães, 2013; Tognoli, Guimarães & Tennis, 2013; Schatz, Stennes-Spidahl & Loehrlein, 2013; b) ISKO-França: Arsenault & El Hadi, 2011; Couzinet, 2011; Gnoli, 2011; Sosinska-Kalata, 2011; c) ISKO-Brasil: Andrade, Berti Junior, Cervantes, Rodrigues, 2011; Miranda, Paranhos, Oliveira, Paes, 2011; Dodebei, 2011; Brascher, 2011, 2013; Ohly, 2011; Barite, 2011; Smiraglia, 2013; Alvarenga, 2013; Fujita, 2013; Varela & Barbosa, 2013; e d) ISKO-Espanha: Alfaya-Lamas, 2011; Sánchez, Blázquez, Rodríguez, 2011; Guimarães, Ferreira, Freitas, 2011; Tognoli, Milani, Barros, 2011; Café, Agustín-Lacruz, Barros, 2011; Ohly, 2011; Almeida, 2013; Alves Júnior, Saldanha, 2013; Dal'avedove, Almeida, Fujita, 2013; Martins, Moraes, 2013; Murguia, Sales, 2013; Semidão, Almeida, Moreira, 2013.

Aplicando-se a análise de conteúdo ao referido corpus, obtiveram-se os seguintes resultados:

4.1. ISKO-Brasil

Segundo a análise feita nos artigos publicados nos Anais da ISKO-Brasil, o objeto da O.C. está relacionado à classificação conceitual, à teoria, indexação e representação/terminologia de conceito (Andrade, Junior, Cervantes, Rodrigues, 2011; Ohly, 2011b; Fujita, 2013). Para alguns autores (Miranda, Paranhos, Oliveira, Paes, 2011) seu objeto são os fundamentos científicos e o desenvolvimento das técnicas de planejamento, construção, uso, gestão e avaliação das habilidades e ferramentas empregadas nos sistemas de informação para tratamento, armazenamento e recuperação de documentos. Guimarães (2001) e Barite (2009) entendem o conhecimento socializado ou registrado, o saber acumulado que se documenta por meio dos processos de publicação, edição, registro em patentes ou em formas de socialização do saberes similares como objeto da O.C. Tal definição está relacionada à informação como conhecimento (information-as-knowledge) e à informação como coisa (information-as-thing), duas aproximações à noção de “conhecimento materializado” e registrado (Barite, 2011). Segundo Smiraglia (2013) o objeto da O.C. é uma comunidade de discurso onde uma investigação rigorosa, autoconsciente acontece em relação àquilo que é conhecido, e suas várias ordenações e sequências, tanto as naturais ou heurísticas, e aquelas que são impostas. Para Brascher (2013) as representações de domínios do conhecimento são objeto de estudo da O.C.

Com relação aos objetivos os processos de representação, organização e recuperação ganham um destaque importante a partir das definições levantadas. Para Andrade, Junior, Cervantes, Rodrigues (2011) o objetivo da O.C é possibilitar a recuperação de objetos e conteúdos informacionais nos ambientes digitais, bem como estruturar e arranjar sistematicamente unidades do conhecimento (os conceitos) de acordo com os elementos de conhecimento (Andrade, Junior, Cervantes, Rodrigues, 2011; Ohly, 2011b). Ainda segundo Ohly (2011b) os objetivos da O.C são organizar: a) unidades de conceitos de conhecimento e b) todos os tipos de objetos (minerais, plantas, documentos, pinturas, objetos de museus, etc), relacionados a termos ou categorias particulares, para capturar o que é conhecido sobre o mundo em alguma forma ordenada permitindo seu futuro compartilhamento com outros. Ortega (2013) advoga a favor da promoção do acesso ao conhecimento, visando seu uso e nova produção bem como elaborar representações que possam ser significadas e manipuladas a favor de um certo público. Ainda segundo a autora a organização do Conhecimento deve ser capaz de fornecer respostas à sociedade para o problema do crescimento dos documentos, por meio de critérios científicos, como uma extensão da capacidade de organização e transferência daquilo que o ser humano sabe. Para Alvarenga (2013), a O.C deve prover uma estrutura ou esquema para armazenar e organizar dados, informação ou conhecimento sobre o mundo e sobre pensamentos, visando à compreensão, recuperação ou descoberta, ao raciocínio e a outros propósitos. Segundo Brascher (2013), delimitar o significado de termos no contexto de seus domínios e estabelecer relações conceituais que auxiliam a posicionar um conceito no sistema conceitual são os objetivos da Organização do Conhecimento. E, por último, Miranda, Paranhos, Oliveira, Paes (2011) definem a geração de um novo conhecimento como objetivo final da O.C.

Os processos apresentam já um maior consenso entre os autores pesquisados, uma vez que a descrição, indexação e classificação ou representação figuram entre os processos nucleares da O.C. Segundo Barité (2001) esses processos consistem na construção, desenvolvimento e gestão de sistemas de organização do conhecimento e outras estruturas conceituais, e nas atividades de classificação e indexação para a recuperação temática de documentos, recursos, dados e todo tipo de informação. Para Fujita (2013) os processos também são a descrição documentária, indexação e classificação, conceitos ligados à representação da informação apontados como processos por Varela e Barbo-

sa (2013). Ainda segundo os autores, a criação de recursos mediadores do acesso ao conhecimento também é entendida como um processo da O.C.

Como instrumentos, Andrade Junior, Cervantes, Rodrigues (2011) apontam a linguagem do sistema de informação (controlada).

No tocante à natureza da OC encontramos, ainda, mais divergência. Enquanto alguns a caracterizam como ciência (Andrade, Junior, Cervantes, Rodrigues, 2011; Ohly, 2011b) ou disciplina científica (Miranda, Paranhos, Oliveira, Paes, 2011), outros a definem como área temática (Ohly, 2011), domínio (Smiraglia, 2013), conjunto de procedimentos e ferramentas que ajudam os seres humanos a compreender e ordenar o caos cognitivo (Ortega, 2013), atividade-meio para o objetivo ou fim da CI (Ortega, 2013) e de caráter social, cognitivo e mediador (Varela, Barbosa, 2013).

Seus diálogos interdisciplinares apresentam as novas tecnologias (Brascher, 2011, Ortega, 2013), a fundamentação teórica à Indexação (Fujita, 2013) e as inter-relações de pesquisa com domínios além dos tradicionalmente contemplados (Brascher, 2011), ultrapassando barreiras disciplinares permitindo a fertilização de outras áreas do conhecimento (Dodebei, 2011).

4.2. ISKO - Espanha

Nos Anais da ISKO-Espanha é possível perceber uma divisão de autores que caracterizam a Organização do conhecimento, quanto a sua natureza, como disciplina científica aplicada que ainda carece de uma exploração mais efetiva de sua base teórica (Sánchez, Blázquez, Rodríguez, 2011; Guimarães, Ferreira, Freitas, 2011; Miranda, 2007; Ohly, 2011a), e outra divisão que a caracteriza enquanto ciência (Alfaya-Lamas, 2011; Ohly, 2011a) (Café, Agustín-Lacruz, Barros, 2011) (Semidão, Almeida, Moreira, 2013). Sua compreensão enquanto um campo de estudos relacionado à informação também se faz presente em algumas obras (Alfaya-Lamas, 2011; Dal'Evedore, Almeida, Fujita, 2013; Almeida, 2013). Outros autores possuem uma visão única e diferenciada da natureza da O.C. Para Café, Agustín-Lacruz, Barros (2011) é uma área de estudos interdisciplinar, enquanto para Semidão, Almeida e Moreira (2013) é uma área temática. Para Tognoli, Milani e Barros (2011) a O.C é um marco teórico-conceitual. Para Almeida (2013), ela é tanto uma atividade quanto um fenômeno. Martins e Moraes (2013) a consideram um processo de modelagem conceitual enquanto Murguia e

Sales (2013) acreditam que a O.C seja um fazer de natureza operacional.

O objeto de estudo da OC, segundo as definições encontradas nos anais da ISKO-Espanha, está ligado basicamente ao processo de organização. Assim, para Alfaya-Lamas (2011), Café, Agustín-Lacruz, Barros (2011), Ohly (2011a), Semidão, Almeida, Moreira (2013) o objeto da O.C é o arranjo estrutural e sistemático das unidades de conhecimento (conceitos) segundo elementos de conhecimento inerentes (características) e a aplicação de conceitos e clases de conceitos ordenadas para a atribuição dos conteúdos de referências de todos os tipos que valem a pena conhecer. Para Murguia e Sales (2013) é a organização de unidades de conhecimento ou organização de conceitos e para Guimarães, Ferreira, Freitas (2011) é a organização da informação em registros bibliográficos, incluindo índices de citação, texto completo e internet. Segundo Alfaya-Lamas (2011) seu objeto é a natureza e qualidade dos processos de organização do conhecimento e dos sistemas de organização do conhecimento. Para Tognoli, Milani, Barros (2011) é o conhecimento registrado e socializado. Para Almeida (2013) são os objetos existentes que impactam diretamente a vida de cientistas, das pessoas e das comunidades envolvidas e, por último, para Murguia e Sales (2013), o objeto de estudo da OC são os conceitos e estruturas conceituais formalizadas instrumentalmente em sistemas de organização do conhecimento, como sistemas de classificação, tesouros e ontologias.

O objetivo da O.C está fundamentalmente ligado às questões referentes à recuperação a acesso à informação registrada, como demonstram as definições a seguir. Jiménez (2011) e Alfaya-Lamas (2011) priorizam o acesso e a recuperação à informação, enquanto Semidão, Almeida, Moreira (2013) focam o armazenamento eficiente, recuperação e processamento da informação codificada em texto e em outros artefatos humanos. A organização e representação do conhecimento também são objetivos que estão presentes nas definições de Alfaya-Lamas (2011), Martins e Moares (2013).

No tocante aos diálogos interdisciplinares, as novas tecnologias e a inteligência artificial se fazem presente no texto de Sánchez, Blázquez, Rodríguez (2011). OS diálogos com a Biblioteconomia também estão presentes nos textos de Ohly (2011a) e Álvares Júnior e Saldanha (2013). A linguística e a teoria da classificação também são apresentadas por Garcia Marco como diálogos com a O.C. Bufrem, Breda, Silva e Fecchio (2013) destacam as relações com a gestão do conhecimento e inteligência competi-

viva. Almeida (2013) destaca que há uma necessidade do contato com outras disciplinas para a execução de seus produtos e processos.

Os processos estão fundamentalmente ligados à classificação, indexação, catalogação e representação do conhecimento, bem como à elaboração de resumos (Alfaya-Lamas, 2011; Guimarães, Ferreira, Freitas, 2011; e Tognoli, Milani, Barros, 2011), e ao tratamento temático da Informação (Murguia, Sales, 2013).

4.3. ISKO – França

Analisado os anais da ISKO-França, nota-se falta de consenso quanto à natureza da OC. Enquanto Gnoli (2011) a considera uma disciplina (em construção), Couzinet (2011) atenta para sua ligação à organização social do conhecimento, como um domínio que visa a constituir-se em uma ciência situada entre a indexação e a comunicação científica. No entanto, no tocante ao objeto, tem-se a abordagem do conhecimento (Gnoli, 2011; Arsenault & El Hadi, 2011) ou ainda o documento (Gnoli, 2011). Os objetivos da OC, por sua vez, estão ligados à atribuição de sentido, à recuperação, ao acesso e à apropriação da informação (Couzinet, 2011; Sosinska-Kalata, 2011).

Os processos da área encontram-se ligados a indexação, classificação e análise de conteúdo (Arsenault & El Hadi, 2011; Couzinet, 2011) sendo que os instrumentos de OC centram-se nas linguagens documentais (Couzinet, 2011) e as ferramentas dependentes da identificação e da representação de numerosas características formais e do conteúdo informacional dos recursos documentários (Sosinska-Kalata, 2011).

Os diálogos intersetoriais da área vão além da Ciência da Informação, encontrando um espaço de interlocução com a Informática, filosofia, linguística, as ciências cognitivas, as ciências da educação e mais geralmente as ciências humanas e sociais (Couzinet, 2011).

4.4. NASKO

Assim como nos demais capítulos analisados, não se verifica, nos anais da NASKO, um consenso no que tange à natureza da Organização do Conhecimento, pois enquanto Dousa (2009) e Smiraglia (2009) a consideram um subcampo da ciência e da Ciência da Informação e Biblioteconomia, Sales & Guimarães (2013) a consideram um domínio de natureza predominantemente operacional.

Não se observam abordagens relativas aos objetivos da OC, sendo que o objeto da área estaria nas estruturas conceituais que compõem

os Sistemas de Organização do Conhecimento, tais como os sistemas de classificação, tesouros e ontologias (Sales, Guimarães, 2013). Já os instrumentos da OC são caracterizados como os vocabulários controlados (Pattueli, 2011). Os processos, por sua vez, apresentam-se invariavelmente ligados à descrição, indexação e classificação de documentos de arquivo e bibliotecas. (Tognoli, Guimarães, Tennis, 2013).

Relativamente aos diálogos interdisciplinares, são citadas a Arquivística, a Ciência da Informação e a Biblioteconomia (Tognoli, Guimarães, Tennis, 2013) e Lógica (Frické, 2013).

5. Análise geral

De uma análise geral dos quatro capítulos analisados tem-se que a natureza da organização do conhecimento é ainda objeto de profícua discussão na área, na medida em que se verifica, por um lado, o reconhecimento de seu marco teórico-conceitual de modo a possibilitar sua classificação como ciência, disciplina científica (alguns destacando seu caráter aplicado e o fato de ainda se encontrar em construção), como domínio ou como área ou campo de estudos interdisciplinar ou ainda como subcampo da ciência e da Ciência da Informação e Biblioteconomia. Por outro lado, alguns autores destacam sua natureza prática, como atividade-meio, conjunto de procedimentos e ferramentas, fazer de natureza operacional ou processo de modelagem conceitual.

Relativamente ao objeto, abordado pelos quatro capítulos da ISKO analisados, preponderam as concepções relativas ao conhecimento, materializado, resgistrado ou socializado em documentos, assim como as unidades de conhecimento: conceitos e sistemas de conceitos. Considerando o campo em si, alguns autores centram o objeto nos fundamentos científicos, nas técnicas e na comunidade discursiva que o caracterizam. Por outro lado, e em uma dimensão mais pragmática, há autores que reconhecem o objeto do campo nos seus processos (organização, representação, indexação, classificação) ou nos seus instrumentos (sistemas de classificação, tesouros, ontologias).

No que tange aos objetivos, observa-se que apenas o capítulo norte-americano a eles não faz menção. Observa-se uma forte ênfase nos processos de representação, armazenamento, arranjo, recuperação, acesso e apropriação da informação, por meio do estabelecimento de relações conceituais que promovam a atribuição de sentido para fins de geração de novo conhecimento. Esses processos são efetivamente abordados nos quatro capítulos analisados,

encontrando-se diretamente ligados à questão do conteúdo ou tratamento temático. Nesse contexto, a indexação, a classificação, a descrição e a condensação são mencionadas como procesos de representação (da informação e do conhecimento), ao que se aliam as preocupações com a construção, desenvolvimento e gestão de sistemas de organização do conhecimento (SOS).

Os instrumentos / ferramentas para desenvolvimento dos processos são abordados pelos capítulos brasileiro, francês e norte-americano, sendo identificados como linguagens controladas dos sistemas de informação, vocabulários controlados ou linguagens documentais.

Os diálogos interdisciplinares, observados com maior ênfase nos capítulos francês e espanhol, são abordados no contexto da Ciência da Informação (notadamente com a arquivística e a biblioteconomia) e, no contexto geral das demais ciências humanas e sociais, tais como as ciências cognitivas, a educação, a filosofia, a gestão do conhecimento, a lingüística e a lógica. Destaca-se, ainda, a conexão com a dimensão tecnológica (informática, inteligência artificial e novas tecnologias em geral). Em seguida, e de modo a caracterizar o domínio, realizou-se uma análise bibliométrica, mais especificamente no que se refere à análise de citações, a partir das fontes (referentes teóricos) utilizadas pelos trabalhos selecionados, de modo a identificar diálogos, consonâncias correntes teóricas, etc.

Por meio da análise de citações torna-se possível analisar o impacto e a visibilidade de autores em um dado domínio científico, sendo especialmente útil para que se identifiquem escolas teóricas. Dessa forma, constitui especial ferramenta para que, a partir de análise qualitativas, se possa melhor compreender o universo epistemológico de um dado domínio, uma vez que a frequência de citação entre dois autores, além de evidenciar aquilo que Small (2004, p.71) denomina como “padrões de co-reconhecimento” determina a proximidade de conteúdos e de como a estrutura de conhecimento de uma área é percebida por seus pesquisadores. Assim, pode-se detectar tanto a similaridade quanto a contraposição de ideias (Vanz, Caregnato, 2003; Gmür, 2003).

A partir desses estudos torna-se possível a construção e visualização gráfica de redes sociais de citação e de cocitação, especialmente útil na visualização do comportamento de um dado domínio de conhecimento (Small, 2004; Guimarães et al., 2011).

Relativamente à rede do capítulo espanhol (e hoje ibérico) (Figura 2), da ISKO, mantém-se a proeminência de Dahlberg e Hjørland. Nesse contexto, observa-se a presença de um mais amplo número de referentes teóricos internacionais da área, como Albrechtsen, Barité, Capurro, Gardin, Olson, Peirce, Pombo, Rayward, Thellefssen e Zeng. A esse se aliam três refe-

rentes brasileiros (Fujita, Santaella e Souza). No âmbito espanhol destacam-se Izquierdo Arroyo e López Huertas.

A rede francesa (Figura 3), reitera o nome de Hjørland assim como o de Beghtol e acrescenta como referentes internacionais os nomes de Hartel, Svenonius e Szostak, assim como, no âmbito francês, vale-se de Meyriat.

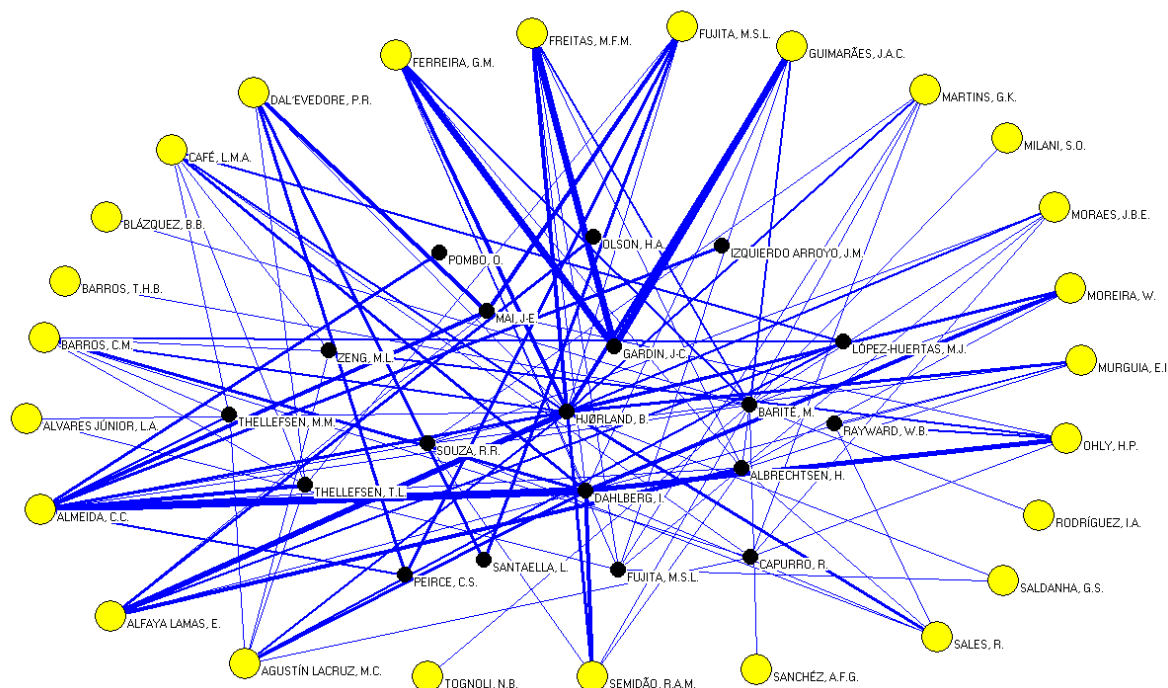


Figura 2. Rede de citação da ISKO-España (círculos amarelos: citantes; círculos pretos: citados)

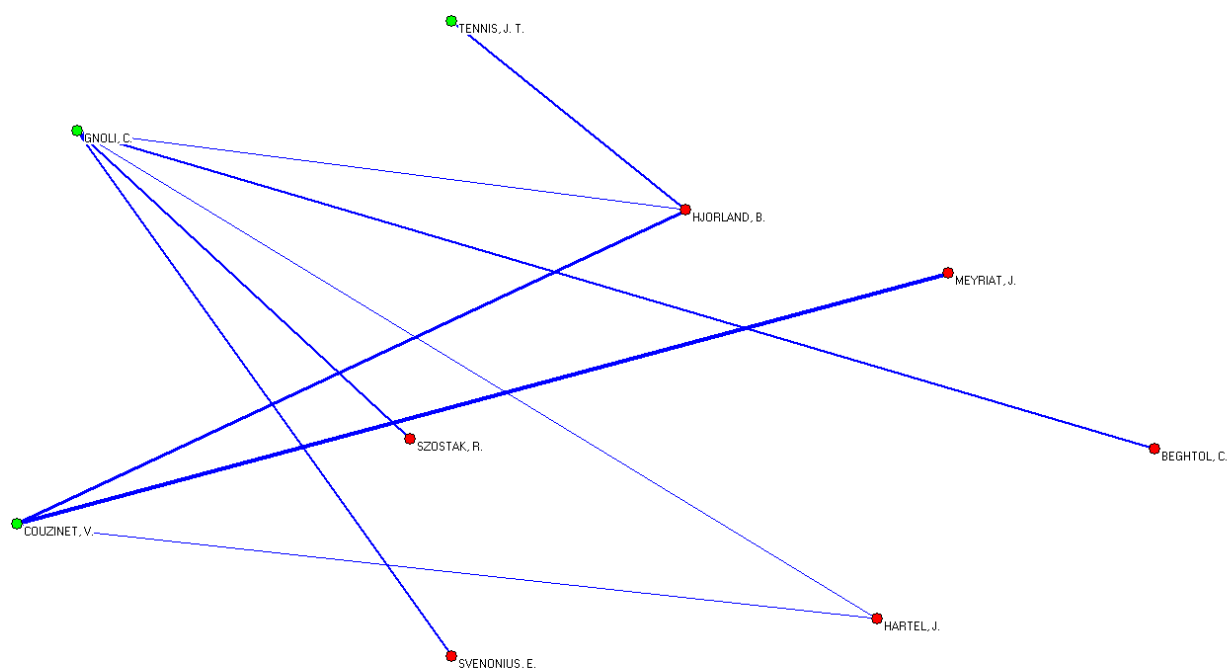


Figura 3. Rede de citação da ISKO-França (verde: autores citantes; vermelho: autores citados)

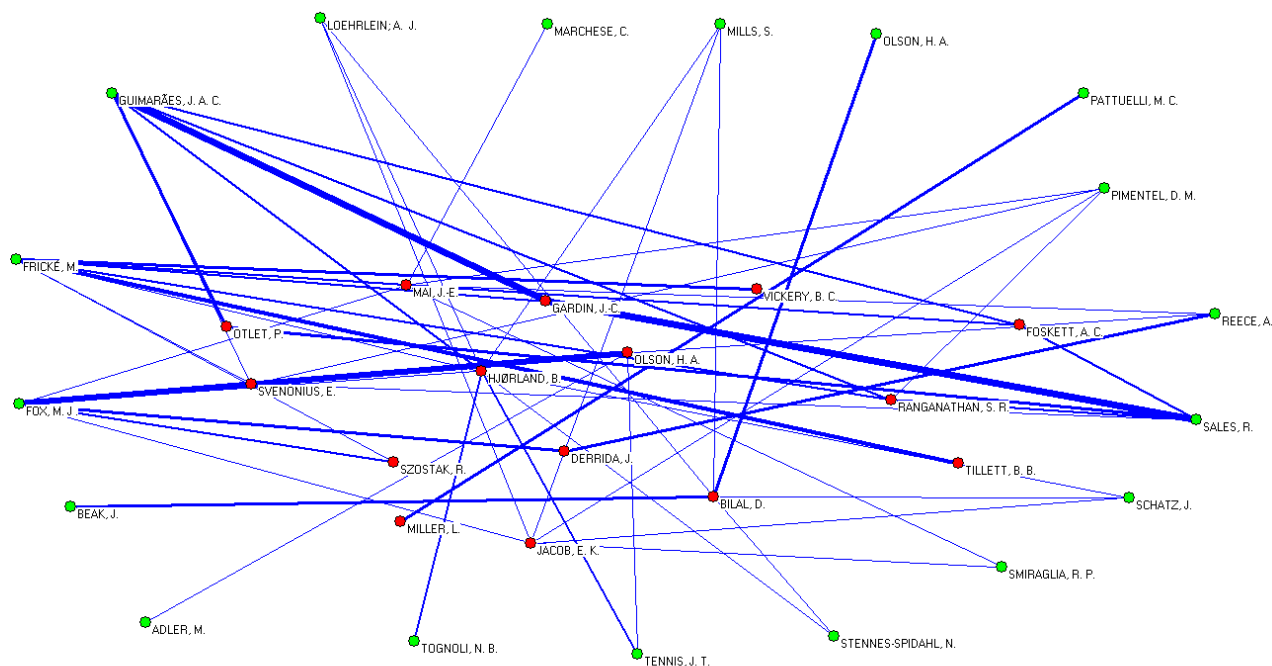


Figura 4. Rede de citação da NASKO (verde: autores citantes; vermelho: autores citados)

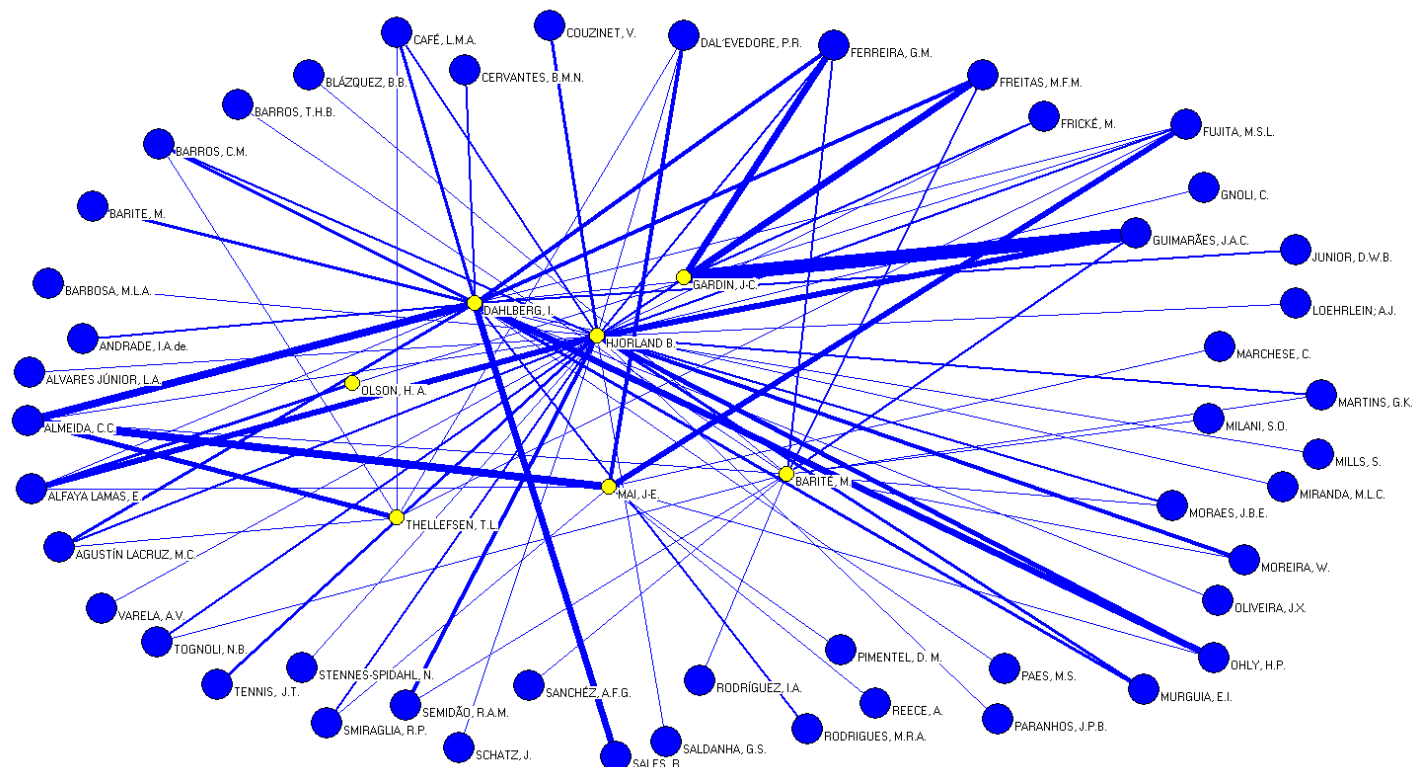


Figura 5. Rede de citação da ISKO: Espanha, Brasil, França e NASKO (círculos azules: citantes; círculos amarelos: citados)

No contexto norte-americano (Figura 4), verifica-se uma maior índice de citações de autores daquela região se comparados à realidade dos demais capítulos analisados (Bilal, Jacob, Miller,

Olson e Svenonius). No âmbito internacional, destacam-se os francófonos Derrida, Gardin e Otlét, os ingleses Foskett e Vickery, o dinamarquês Mai e o indiano Ranganathan. É intere-

sante observar que esse é o único capítulo em que a presença de Hjørland e de Dahlberg não se faz notar de maneira mais efetiva.

De uma análise geral, Hjørland desponta como o referente mais compartilhado no contexto analisado, na medida em que se encontra presente em três dos quatro capítulos (Brasil, Espanha e França). Em seguida, destacam-se seis outros autores que são citados em dois dos quatro capítulos analisados, a saber: Beghtol (Brasil e França), Dahlberg (Brasil e Espanha), Gardin (Espanha e NASKO), Mai (Brasil e NASKO), Olson (Espanha e NASKO) e Svenonius (França e NASKO). Considerando esse conjunto de autores mais citados, observa-se uma forte presença das abordagens sócio-cognitivas (Hjørland, Beghtol, Olson, Mai), o que parece sinalizar para uma tendência investigativa do campo de organização do conhecimento.

De modo a analisar a questão dos referentes teóricos a partir de uma dimensão global dos quatro capítulos analisados, chegou-se, a partir das redes individuais de cada capítulo, a um conjunto geral de 44 autores mais citados, o que, pela aplicação da Teoria de Elistimo de Price, levou a um núcleo de sete autores citados, com um mínimo de seis citações cada (Hjørland, com 45 citações; Dahlberg com 33; Gardin com 17; Mai com 15; Olson com 10; Barité com 8; e Thelleffsen com 6) que integram a rede geral (Ver Apêndice - Figura 5).

A análise da referida rede reafirma a centralidade de Hjørland no universo científico da organização do conhecimento, pois além de ser o autor mais citado, é aquele que recebe citações de uma maior diversidade de autores citantes. Em termos de vinculações teóricas, manifestadas pelas relações citante-citado, observa-se um forte nexos de Guimarães com as ideias de Gardin, de Ohly com as ideias de Dahlberg e de Almeida com as ideias de Mai, o que reflete, respectivamente, as dimensões temáticas relativas à análise documental, a teoria do conceito e aos estudos semióticos em organização do conhecimento.

6. Conclusão

A vista dos resultados obtidos, pode-se observar que os quatro capítulos da ISKO analisados (Brasil, Espanha, França e Norte-América) são bastante profícuos na abordagem da questão conceitual da área. No entanto, essa questão ainda se encontra em construção na medida em que se tem por um lado, os aspectos relativos à natureza e ao objeto da organização do conhecimento revelam uma preocupação com a sua configuração epistemológica, voltada para a sua

configuração como ciência, domínio ou campo centrado em um conhecimento produzido pela sociedade e por ela materializado para fins de permanência no tempo e portabilidade no espaço. Por outro lado, uma preocupação de natureza pragmática e operacional se coloca, voltada para os processos os instrumentos da organização do conhecimento, muitas vezes concebidos como seu próprio objeto.

No tocante aos referentes teóricos, observa-se uma forte diversidade – e complementaridade – entre os países, o que reforça a assertiva de uma diversidade de abordagens típica de um campo ainda em construção. Sete autores, nesse universo, revelam centralidade (Hjørland, Dahlberg, Gardin, Mai, Olson, Barité e Thelleffsen). Hjørland e Dahlberg assumem forte liderança, o que revela, respectivamente, duas abordagens teóricas da organização do conhecimento: a sócio-cognitiva e a ontológica, em que a primeira assume maior força na medida em que especialmente Mai e Thelleffsen, com a questão semiótica e Olson, com os estudos culturais, vêm se somar a esse quadro teórico.

Por fim, observa-se que a questão conceitual constitui tema relevante para o desenvolvimento da organização do conhecimento especialmente quando distintos capítulos da ISKO são criados e se consolidam, para que se possa pensar em traços distintivos mínimos que permitam a caracterização desse campo como tal.

Referências

- Adamic, L. (2008). *Introductory social network analysis with Pajek*. Ann Arbor (MI): School of Information, University of Michigan, 2008.
- Adler, Melissa; Tennis, Joe (2013). *Toward a Taxonomy*. // The Fourth North American Symposium on Knowledge Organization - NASKO, 2013: Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work, Milwaukee. *Proceedings from North American Symposium on Knowledge Organization*. Milwaukee: University of Wisconsin-Milwaukee, 2013, 1-19.
- Alfaya Lamas, E. (2011). *La asunción del género neutro en la teoría y práctica de la organización del conocimiento*. // Pérez Pais, María Carmem; Bonome, María G. (ed.). *Actas del X Congreso de ISKO-ESPAÑA - 20 years of ISKO Spanish Chapter*. Universidad da Coruña - Serviço de Publicaciones, 2011, 249-263.
- Almeida, C.C. (2013). *Notas sobre os conceitos Semióticos fundamentais à Organização do Conhecimento*. // Ribeiro, Fernanda; Cerveira, Maria Elisa. (ed.). *Informação e/ou Conhecimento: as duas faces de Jano*. Porto: Faculdade de Letras da Universidade do Porto/ CETAC.MEDIA, 2013, 68-81.
- Alvarenga, L. (2013). *Modelagem conceitual e sistemas de organização de conhecimento: uma reflexão baseada em experiências acadêmicas e profissionais no Brasil no domínio da Biblioteconomia e Ciência da Informação*. // Dobedei, Vera; Guimarães, José Augusto Chaves (ed.). *Complexidade e Organização do Conhecimento: desafio*

- os de nosso século . Rio de Janeiro: ISKO-Brasil ; Marília: FUNDEPE, 2013, 141-146.
- Alvares Júnior, L. S.; Saldanha, G.S. (2013). O mundo é o assunto – epistemologia da indexação no século XIX. // Ribeiro, Fernanda; Cerveira, Maria Elisa. (ed.). Informação e/ou Conhecimento: as duas faces de Jano. Porto: Faculdade de Letras da Universidade do Porto/ CETAC.MEDIA, 2013, 82-98.
- Andrade, I. A. de; Junior, D. W. B.; Cervantes, B. M. N.; Rodrigues, M. R. A. (2011). Organização e representação do conhecimento e os mapas conceituais. // Guimarães, José Augusto Chaves; Dobedei, Vera. (ed.). Desafios e Perspectivas Científicas para a organização e representação do conhecimento. Marília: ISKO-Brasil: FUNDEPE, 2011, 83-90.
- Arsenault, C.; Widad Mustafa El Hadi. (2012). Dynamism and stability in knowledge organization: from one conference to another Toronto 2000, Lille 2011. // Hadi, Widad Mustafa El (ed.). Knowledge Organization: dynamism et stabilité. Lille: Lavoiser, 2011, 21-31
- Bardin, L. (2009). *Análise de Conteúdo*. Lisboa, Portugal; Edições 70, LDA, 2009.
- Barite, M. (2011). La producción científica en Organización del Conocimiento vista através de la Sección Literatura de la revista Knowledge Organization: invisibilidad y transparencia. // Guimarães, José Augusto Chaves; Dobedei, Vera. (ed.). Desafios e Perspectivas Científicas para a organização e representação do conhecimento. Marília: ISKO-Brasil: FUNDEPE, 2011, 262-272.
- Brascher, M. (2011). Visibilidade científica e interlocução internacional em organização e representação do Conhecimento. // José Augusto Chaves Guimarães; Vera Dobedei (Org.). Desafios e Perspectivas Científicas para a organização e representação do conhecimento. 1ed. Marília: ISKO-Brasil: FUNDEPE, 2011, p.234-237.
- Bufrem, L. S. (2004). Levantando significações para significantes: da gestão do conhecimento a organização do saber. // *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*. 9, n. esp.(del), 1º sem. http://www.encontros-bibli.ufsc.br/bibesp/esp_01/1_bufrem.pdf.
- Café, L. M. A.; Agustín lacruz, M. C.; Barros, C. M. (2011). Organização do Conhecimento: análise conceitual. // Péres Pais, Maria Carmem; Bonome, María G. (ed.). Actas del X Congreso de ISKO-ESPANHA - 20 years of ISKO Spanish Chapter. Universidad da Coruña - Serviço de Publicaciones, 2011, 283-302.
- Couzinet, V. (2011). L'organisation des connaissances au regard des sciences de l'information et de la communication, une exception française?. // Hadi, Widad Mustafa El (ed.). Knowledge Organization: dynamism et stabilité. Lille: Lavoiser, 2011, 35-50.
- Dahlberg, I. (1993). Knowledge organization: its scope and possibilities. // *Knowledge Organization*. 20:4, 211-222.
- Dahlberg, I. (2008). The Information Coding Classification (ICC): A modern theory-based fully-faceted, universal system of knowledge field. // *Axiomathes*. 18:2, 161-176.
- Dal'evodore, P. R.; Almeida, C. C.; Fujita, M. S. L. (2013). A interpretação de conceitos no processo de indexação – aspectos e reflexões semióticas. // Ribeiro, Fernanda; Cerveira, Maria Elisa. (ed.). Informação e/ou Conhecimento: as duas faces de Jano. Porto: Faculdade de Letras da Universidade do Porto/ CETAC.MEDIA, 2013, 190-203.
- DODEBEI, V. Pesquisa em organização e representação do conhecimento no Brasil: Uma análise dos cursos de pós-graduação. // José Augusto Chaves Guimarães; Vera Dobedei. (Org.). Desafios e Perspectivas Científicas para a organização e representação do conhecimento. 1ed. Marília: ISKO-Brasil: FUNDEPE, 2011, p.165-170.
- Fox, M. J. (2011) Prototype theory: an alternative concept theory for categorizing sex and gender. // *Proceedings from North American Symposium on Knowledge Organization -NASKO*. Toronto, Canada, 2011, 13-22.
- Fox, M.; Reece, A. (2013). The impossible decision: social tagging and Derridas' s deconstructed hospitality. // *The Fourth North American Symposium on Knowledge Organization - NASKO*, 2013, Milwaukee. Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work, 2013, 58-74.
- Frické, Martin (2013). Logic and the organization of information: an introduction. // *The Fourth North American Symposium on Knowledge Organization - NASKO*, 2013, Milwaukee. Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work, 2013, 75-80.
- Fujita, M.S.L. (2013). A importância teórica e prática da indexação na fundamentação científica da organização e representação do conhecimento. // Dobedei, Vera; Guimarães, José Augusto Chaves (ed.). Complexidade e Organização do Conhecimento: desafios de nosso século. Rio de Janeiro: ISKO-Brasil ; Marília: FUNDEPE, 2013, 147-159.
- Gnoli, C. (2011). Des métadonnées représentant quoi?. // Hadi, Widad Mustafa El (ed.). Knowledge Organization: dynamism et stabilité. Lille: Lavoiser, 2011, 51-66.
- Guimarães, José Augusto Chaves. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). *Revista Ibero-Americana de Ciência da Informação*, - cidade -, v. 1, n. 1, 2008.
- Guimarães, J. A. C.; Sales, Rodrigo de; Martínez-Ávila, Daniel; Fernandes Alencar, Maira (2014). The conceptual dimension of knowledge organization in the ISKO proceedings domain: a Bardinian analysis. // Babik, Wieslaw. (ed.) Knowledge organization in the 21st century: between historical patterns and future perspectives. Wurzburg: Ergon Verlag, 2014, 101-106.
- Guimarães, J. A. C.; Ferreira, G. M.; Freitas, M. F. M. (2011). Correntes teóricas do tratamento temático da informação: uma análise de domínio da presença da catalogação de assunto e da indexação nos congressos de ISKO-Espanha. // Péres Pais, Maria Carmem; Bonome, María G. (ed.). Actas del X Congreso de ISKO-ESPANHA - 20 years of ISKO Spanish Chapter. Universidad da Coruña - Serviço de Publicaciones, 2011, 181-194.
- Hjørland, B. (2002). Domain analysis in information Science: eleven approaches traditional as well as innovative. // *Journal of Documentation*. 58:4, 422-462.
- Hjørland, B. (2004). Domain analysis: a socio-cognitive orientation for Information Science research. // *Bulletin of the American Society for Information Science and Technology*. 30:3. <http://www.asis.org/Bulletin/Feb-04/hjorland.html>.
- Hjørland, B.; Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. // *Journal of the American Society for Information Science*. 46:6, 400-425.
- Jacob, E. K.; Zhang, G. (2013). The role of virtual boundaries in knowledge and organization. // *The Fourth North American Symposium on Knowledge Organization - NASKO*, 2013: Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work, Milwaukee. Proceedings from North American Symposium on Knowledge Organization. Milwaukee: University of Wisconsin-Milwaukee, 2013, 127-135
- Maccall, S. L. (2011). Knowledge Organization under digital inversion: A theory for cooperative librarian organizing

- practices for online textual artifacts. // Proceedings from North American Symposium on Knowledge Organization -NASKO. Toronto, Canada, 2011, v. 3, p. 13-22.
- Marchese, C.; Smiraglia, R. P. (2013). Boundary objects: CWA, an HR firm, and emergent vocabulary. // The Fourth North American Symposium on Knowledge Organization - NASKO, 2013: Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work, Milwaukee. Proceedings from North American Symposium on Knowledge Organization. Milwaukee: University of Wisconsin-Milwaukee, 2013, 136-142
- Martins, G. K.; Moraes, J. B. E. (2013) Sistematização de conteúdos em Organização e Representação do Conhecimento – um estudos nos periódicos de Ciência da Informação no Brasil. // Ribeiro, Fernanda; Cerveira, Maria Elisa (ed.). Informação e/ou Conhecimento: as duas faces de Jano. Porto: Faculdade de Letras da Universidade do Porto/ CETAC.MEDIA, 2013, 421-436.
- Miranda, M. L. C.; Paranhos, J. P. B.; Oliveira, J. X.; Paes, M. S. (2011). Organização e representação do conhecimento em religiões yorubanas na Library of Congress Subject Headings. // Guimarães, José Augusto Chaves; Dobedei, Vera. (ed.). Desafios e Perspectivas Científicas para a organização e representação do conhecimento. Marília: ISKO-Brasil: FUNDEPE, 2011, 153-159.
- Moya-Anegón, F.; Herrero-Solana, V. (2001). Análisis de dominio de la revista mexicana de investigación bibliotecológica. // Información, cultura y sociedad, 5, 10-28.
- Murguia, E. I.; Sales, R. (2013). Práticas discursivas na Organização do Conhecimento – Otlet, Bliss e ISKO. // Ribeiro, Fernanda; Cerveira, Maria Elisa. (ed.). Informação e/ou Conhecimento: as duas faces de Jano. Porto: Faculdade de Letras da Universidade do Porto/ CETAC.MEDIA, 2013, 446-461.
- Ohly, H.P. (2011a). ISKO – Knowledge Organization in a Changing World. Challenges, Programs, and Mission. // Guimarães, José Augusto Chaves; Dobedei, Vera. (ed.). Desafios e Perspectivas Científicas para a organização e representação do conhecimento. Marília: ISKO-Brasil: FUNDEPE, 2011. 238-248
- Ohly, H. P. (2011b). Organization, Management and Engineering of Knowledge: Rivals or Complements? // Péres Pais, Maria Carmem; Bonome, Maria G. (ed.). Actas del X Congreso de ISKO-ESPANHA - 20 years of ISKO Spanish Chapter. Universidad da Coruña - Serviço de Publicaciones, 2011. 541-551.
- Olson, H. A.; Howarth, L. C. A (2013). Space of Transition: Rethinking Surrogates. // The Fourth North American Symposium on Knowledge Organization - NASKO, 2013: Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work, Milwaukee. Proceedings from North American Symposium on Knowledge Organization. Milwaukee: University of Wisconsin-Milwaukee, 2013, 161-166
- Ortega, C. D. (2013). Aspectos teóricos, procedimentos, normativos e pragmáticos como categorias para uma epistemologia da organização da informação. // Dobedei, Vera; Guimarães, José Augusto Chaves (ed.). Complexidade e Organização do Conhecimento: desafios de nosso século. Rio de Janeiro: ISKO-Brasil; Marília: FUNDEPE, 2013. 34-38.
- Pajek (2009). Introductory social network analysis with Pajek. School of Information- University of Michigan. http://ocw.mit.edu/courses/economics/14-15j-networks-fall-2009/assignments/MIT14_15JF09_pajek.pdf.
- Pattueli, M. C. (2011). Mapping subjectivity: performing people-centered vocabulary alignment. // Proceedings from North American Symposium on Knowledge Organization-NASKO. Toronto, Canada, 2011. 13-22
- Sales, R. Guimarães, J. A. C.. (2013). Contributions of Julius Otto Kaiser to the theoretical framework of Knowledge Organization: basis for the analytical-synthetic method. // The Fourth North American Symposium on Knowledge Organization - NASKO, 2013: Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work, Milwaukee. Proceedings from North American Symposium on Knowledge Organization. Milwaukee: University of Wisconsin-Milwaukee, 2013, 167-175.
- Sánchez, A. F. G.; Blázquez, B. B.; Rodríguez, I. A. (2011). Análisis bibliométricos de artículos relacionados con la Organización del Conocimiento en revistas españolas recogidas en la Web os Science. // Péres Pais, Maria Carmem; Bonome, Maria G. (ed.). Actas del X Congreso de ISKO-ESPANHA - 20 years of ISKO Spanish Chapter. Universidad da Coruña - Serviço de Publicaciones, 2011. 171-180.
- Schatz, J. S.-S. Nadia; Mills, S. Loehrlein; Aaron, J. (2013). Bibliographic Induction: How KO Systems Optimize Browsing by Supporting Library Users' Prior Knowledge. // The Fourth North American Symposium on Knowledge Organization - NASKO, 2013: Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work, Milwaukee. Proceedings from North American Symposium on Knowledge Organization. Milwaukee: University of Wisconsin-Milwaukee, 2013. 176-186
- Semidão, R. A. M.; Almeida, C. C.; Moreira, W. (2013). Diretrizes para análise conceitual – as perspectivas de Hjørland, Dahlberg e Lakoff. // Ribeiro, Fernanda; Cerveira, Maria Elisa. (ed.). Informação e/ou Conhecimento: as duas faces de Jano. Porto: Faculdade de Letras da Universidade do Porto/ CETAC.MEDIA, 2013, p.589-602.
- Smiraglia, R. P. (2012). Epistemology of Domain Analysis. // Lee, H-L.; Smiraglia, R.P. (ed.). Cultural frames of knowledge. Würzburg: Ergon, 2012.111-124
- Smiraglia, R. P. (2013). The Epistemological Dimension of Knowledge Organization. // Dobedei, Vera; Guimarães, José Augusto Chaves (ed.). Complexidade e Organização do Conhecimento: desafios de nosso século. Rio de Janeiro: ISKO-Brasil; Marília: FUNDEPE, 2013.17-25.
- Sosinska-Kalata, B. (2011) Evolution des systèmes d'organisation des connaissances et établissement de critères pour leur évaluation praxiologique?. // Hadi, Widad Mustafa El (ed.). Knowledge Organization: dynamism et stabilité. Lille: Lavoiser, 2011. 21-31
- Tennis, J. T. (2003). Two Axes of Domain Analysis. // Knowledge Organization. 30:3/4, 191-195.
- Tognoli, N. B.; Guimarães, J. A. C.; Tennis, J. T. (2013). Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work. // The Fourth North American Symposium on Knowledge Organization - NASKO, 2013: Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work, Milwaukee. Proceedings from North American Symposium on Knowledge Organization. Milwaukee: University of Wisconsin-Milwaukee, 2013. 217-225
- Tognoli, N. B.; Milani, S. O.; Barros, T. H. B. (2011). International Society for Knowledge Organization(ISKO): identificação de redes sociais a partir dos Anais de seus eventos (1991-2010). // Péres Pais, Maria Carmem; Bonome, Maria G. (ed.). Actas del X Congreso de ISKO-ESPANHA - 20 years of ISKO Spanish Chapter. Universidad da Coruña - Serviço de Publicaciones, 2011. 145-158.
- Urbizagástegui Alvarado, R. (2009). Elitismo na literatura sobre a produtividade dos autores. // Ciência da Informação, 38:2, <http://dx.doi.org/10.1590/S0100-19652009000200006>.

Varela, A. V.; Barbosa, M. L. A. (2013). O caráter técnico, social, e mediador da organização do conhecimento. // The Fourth North American Symposium on Knowledge Organization - NASKO, 2013: Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work, Milwaukee. Proceedings from North American Symposium on Knowledge Organization. Mil-

waukee: University of Wisconsin-Milwaukee, 2013. 277-282.

Enviado: 2015-04-16. Segunda versão: 2015-06-16.
Aceptado: 2015-07-02.

Aplicación de la Teoría Crítica de Raza en la organización y representación del conocimiento

Application of Critical Race Theory to knowledge organization and representation

Daniel MARTÍNEZ-ÁVILA (1), Marcio FERREIRA (1), José L. MAGRO (2)

(1) Department of Information Science, São Paulo State University – UNESP, Av. Higyno Muzzi Filho, 737, Marília (17525-900) – São Paulo – Brazil, dmartinezavila@marilia.unesp.br, soimers@marilia.unesp.br.

(2) The Graduate Center (CUNY), 365 5th Avenue, New York, NY, USA, jmagro@gradcenter.cuny.edu

Resumen

Se introducen las características y posibilidades metodológicas de la Teoría Crítica de Raza aplicada a la organización y representación del conocimiento, y más específicamente a los principales sistemas de clasificación bibliográficos utilizados en Brasil. Para ello, se identifican las clasificaciones bibliográficas como instituciones; y se esbozan los principales puntos de un proyecto de reemplazo de instancias de racismo por alternativas antirracistas.

Palabras clave: Teoría Crítica de Raza. Clasificaciones bibliográficas. Metodología. Organización del conocimiento. Aspectos éticos.

1. Introducción

Los sesgos y problemas éticos en las clasificaciones bibliográficas y otros sistemas de organización y representación del conocimiento son un aspecto que ha sido ampliamente expuesto en la literatura de las últimas décadas (por ejemplo, Berman, 1971; Foskett, 1971, 1984; Afolabi, 1992; Olson & Ward, 1997; Olson, 2000, 2001a, 2002; Olson & Schlegl, 2001; Milani et al., 2014). El paso de una hegemonía positivista hacia unas posiciones más pragmáticas ha permitido el rechazo cada vez más abierto de valores como la neutralidad y la objetividad en pos de la reivindicación de la explicitación de los puntos de vista en los procesos de organización y desarrollo de sistemas (por ejemplo, García Gutiérrez, 2002; 2007; 2011; Martínez-Ávila & Guimarães, 2013; García Gutiérrez & Martínez-Ávila, 2014). En términos más técnicos, Birger Hjørland lo expresa del siguiente modo (Hjørland, 2007, p. 8, traducción propia):

La teoría dominante en la biblioteconomía y ciencia de la información [LIS] ha sido que los vocabularios normativos representan soluciones 'neutrales', 'objetivas' que simplemente proporcionan sistemas de información más eficientes. Podemos denominar esta visión como 'visión positivista' y contrastarla con una 'visión pragmática' en la que cualquier vocabulario controlado tiende a favorecer al-

Abstract

In this paper we introduce the characteristics and methodological possibilities of applying the Critical Race Theory to knowledge organization and representation, and more specifically to the bibliographic classifications that are used in Brazil. We identify these classifications as institutions and outline the main points of a project of replacement of any instances of racism by anti-racist alternatives.

Keywords: Critical Race Theory. Bibliographic classifications. Methodology. Knowledge organization. Ethical aspects.

gunos tipos de consultas a la vez que hace otros tipos de consultas más difíciles de contestar.

Sin embargo, en la práctica, ese tipo de factores y sesgos que favorecen unos tipos de consultas sobre otros (perjudicando también a unos usuarios concretos sobre otros) están lejos de seguir criterios pragmáticos basados en el estudio de las características de las comunidades a las que sirven. Más bien, tal como indica Furner (2007, p. 154, traducción propia), "todo esquema de clasificación puede ser interpretado como sesgado en algún respecto, donde el sesgo es el inevitable reflejo de sus diseñadores", es decir, reflejando los sesgos, intereses y características de la comunidad dominante que los engendra, que en el caso de los principales sistemas de organización del conocimiento predominantes es blanca, étnicamente europea, burguesa, cristiana protestante, heterosexual, sin discapacidades físicas y masculina (1) (Olson, 2001b).

En este caso, los problemas prácticos de los usuarios pertenecientes a otros grupos diferentes de esta clase dominante, como por ejemplo las comunidades afrodescendientes, no se limitan únicamente a inofensivas consultas técnicamente más difíciles de contestar, sino que también pueden abarcar aspectos sociales mucho más sensibles como por ejemplo representaciones deformadas y la creación y continuación de estereotipos, omisión e invisibilización

de la comunidad o sus características, uso de lenguaje ofensivo para la comunidad, y agrupamiento o asimilación de la comunidad con otros grupos cuyo único nexo en común es compartir el estatus de “otro” en relación a la clase dominante.

Los casos de racismo en sistemas y otras instituciones de organización y representación del conocimiento han sido bien documentados en la citada bibliografía. En este sentido, tal como indican Bowker y Star (1999), las clasificaciones deben abandonar su estatus de herramientas neutrales y escondidas para pasar a ser elementos activos de trabajo ético y político. Así pues, adoptando una postura crítica, ética y emancipadora, los fundamentos y posibilidades de la Teoría Crítica de Raza (TCR) podrían suponer un perfecto cuadro de referencia teórico y metodológico para estudiar y eliminar estas instancias de racismo en sistemas utilizados por comunidades afrodescendientes.

En el presente trabajo se busca profundizar en las principales características y posibilidades metodológicas para la organización y representación del conocimiento de la TCR, y esbozar su posible aplicación en el contexto de dos de los principales sistemas de clasificación utilizados en Brasil: la Clasificación Decimal de Dewey (CDD) y la Clasificación Decimal Universal (CDU).

2. Uso de la Teoría Crítica de Raza

La TCR surgió a mitad de los años 1970s en los Estados Unidos, englobando principalmente abogados, activistas e investigadores en el área de Derecho que propusieron nuevas teorías y estrategias para combatir el racismo que había ganado terreno desde el parón en avances en derechos civiles de los años 1960s. (Delgado & Stefancic, 2001, p. 3-4). En la actualidad, tal como indican Delgado y Stefancic (pp. 2-5, traducción propia), “el movimiento de la Teoría Crítica de Raza es una colección de activistas e investigadores interesados en estudiar y transformar la relación entre raza, racismo y poder”, que también tiene relación con otros movimientos como los estudios legales, el feminismo radical, filósofos y teóricos europeos como Antonio Gramsci y Jacques Derrida, la tradición radical americana de Sojourner Truth, Frederick Douglass, W. E. B. Du Bois, Cesar Chavez, Martin Luther King, Jr., y los movimientos Black Power y Chicano de los 1960s y principios de los 1970s.

Derrick Bell, profesor de Derecho en la Universidad de Nueva York, es considerado el padre intelectual del movimiento, mientras que otras

figuras importantes incluyen a Alan Freeman, Kimberlé Crenshaw, Angela Harris, Charles Lawrence y Patricia Williams en su vertiente tradicional, Neil Gotanda, Eric Yamamoto y Mari Matsuda como autores asiáticos, Robert Williams como principal autor indio, y Richard Delgado, Kevin Johnson, Margaret Montoya, Juan Perea y Francisco Valdés como principales autores latinos.

En el contexto de la biblioteconomía y ciencias de la información, la aplicación de la TCR ya ha sido discutida en la archivística (Dunbar, 2006), Internet (Brock, 2009), servicios bibliotecarios (Kumasi, 2012), clasificaciones bibliográficas (Furner & Dunbar, 2004) y la versión en inglés de la CDD (Furner, 2007). Jonathan Furner puede ser considerado el principal precursor de la TCR en la organización y representación del conocimiento.

2.1. Características de la Teoría Crítica de Raza

No existen unas características uniformes y únicas en los autores que trabajan con TCR. Harris (1994) y Bell (1995) resumen estas diferencias en dos corrientes: la que se destaca por su compromiso con la crítica radical (normativamente deconstructivista), y la que se destaca por su compromiso con la emancipación radical (normativamente reconstruccionista). Común a toda la TCR es el objetivo de erradicar la injusticia racial, primero revelando su naturaleza institucional, y segundo especificando y conduciendo la acción radical necesaria para reemplazar las instituciones racistas por alternativas antirracistas (Furner, 2007, p. 144). En el caso del trabajo de Furner, una de las instituciones que son identificadas como racistas es la CDD, que en su caso es identificada como una institución de información. Enlazándolo con las teorías de justicia social, Furner (p. 155) también indica que los esquemas de clasificación pueden ser evaluados si son tratados como generadores de distribuciones de bienes conceptuales. Otros ejemplos de instituciones de información que también podrían ser objeto de estudio pueden ser cualquiera de los estándares, reglas y recomendaciones para catalogadores y bibliotecarios. En nuestro caso, hemos optado por continuar con la línea de trabajo de Furner en clasificaciones bibliográficas, identificando como instituciones racistas la aplicación de la CDD y CDU en Brasil, e introduciendo la TCR al discurso brasileño de organización y representación del conocimiento.

En ambos casos, las características comunes de esta teoría son compatibles con las visiones pragmatistas en organización y representación

del conocimiento en las que se rechaza la existencia y posibilidad de decisiones libres de valores en cualquier proceso o práctica profesional, que incluyen desde la toma de decisiones, elaboración de normas, leyes, etc. hasta el diseño de clasificaciones y otros sistemas. En este sentido, la TCR trabaja con la tesis de la “construcción social”, que afirma que las razas son únicamente productos de pensamiento y relaciones sociales. Las razas, en palabras de Delgado y Stefancic (2001, p. 7, traducción propia), “no son objetivas, inherentes o fijas, no corresponden con ninguna realidad biológica o genética; más bien, las razas son categorías que la sociedad inventa, manipula o retira cuando es conveniente”.

2.2. Sobre el concepto de raza y racismo

El concepto de “raza” no está demostrado como algo natural y absoluto desde un punto de vista biológico. Lo que es normalmente distinguido como “raza” es una agrupación artificial de elementos en función de criterios de semejanza fisiológica (y muy comúnmente también confundidos con atributos étnicos y culturales como lenguaje, religión, etc., e incluso geográficos) comúnmente utilizada con fines identitarios o discriminatorios.

No obstante, desde el punto de vista social constructivista, el hecho de que un concepto no sea natural no significa que no sea real. Es reconocido que la posición eliminativista —en la que el concepto de raza es eliminado o ignorado por no ser natural— no parece ser una solución efectiva mientras existan identidades que son construidas tanto activa como pasivamente por terceros con fines discriminatorios. El hecho de que no se pueda demostrar científicamente que existe una línea natural que distingue elementos de una raza de los de otra no ha prevenido que individuos hayan adoptado criterios artificiales para identificarse con un grupo y distinguirse de “otros” (los cuales estarían definidos, siguiendo a Derrida, por sus diferencias).

Un acto puede considerarse racista si trata injustamente a un individuo con motivo de ser percibido o identificado con una raza. En ese sentido, la TCR sí reconoce el derecho del individuo de auto-identificación con una raza, especialmente cuando es utilizado para explicar los actos discriminatorios motivados por los aspectos definitorios del grupo al que fue adscrito.

Dicho esto, la TCR no puede ser calificada ni de racista ni de ser equivalente al racismo que es ejercido contra los que son percibidos como no-blancos. David Ingram indica que “afirmar ‘orgullo negro’ no es equivalente a afirmar ‘orgullo

blanco’, ya que el primero —al contrario que el segundo— es una estrategia defensiva dirigida a rectificar un estereotipo negativo” (Ingram, 2004, p. 55, traducción propia). La TCR busca rectificar unas situaciones de injusto privilegio blanco a través de instituciones en sociedades, como por ejemplo la estadounidense o la brasileña, donde, a pesar de existir un porcentaje amplio de personas afrodescendientes, los sistemas no son diseñados para “favorecer sus tipos de consultas” o reflejar sus puntos de vista.

Tal como afirmaba A. C. Foskett, comentando el citado trabajo de Sanford Berman (1971), “la Lista de Encabezamientos de Materia de la Biblioteca del Congreso (LCSH) representa un punto de vista imperialista que puede ser probado como vergonzoso en una biblioteca africana” (Foskett, 1971, p. 121). Este punto de vista que menciona Foskett quizás podría ser pragmáticamente aceptable en un sistema diseñado exclusivamente para blancos imperialistas (aunque todavía podría plantear objeciones éticas dentro de su propia comunidad), pero no para comunidades con un considerable número de usuarios afros.

La TCR busca describir y explicar esas instancias de racismo y actuar desde un triple compromiso entre sus características: un compromiso epistemológico con la construcción social (de conceptos como la raza), un compromiso ético con la justicia social, y un compromiso metodológico con acciones radicales de tanto tipo físico como intelectual (Furner, 2007, p. 146).

2.3. Propuestas metodológicas de la Teoría Crítica de Raza para la ORC

Dentro de una postura epistemológica social constructivista, que podría coincidir con posiciones pragmatistas y postestructuralistas en el análisis y discusión de los datos, la TCR además presenta frecuentemente las siguientes técnicas en la construcción de sus textos: uso de la primera persona, storytelling, narrativa, contra-historias y contra-narrativas, alegorías, tratamiento interdisciplinar de la ley y todo tipo de usos creativos.

Como ejemplo de estas técnicas, Derrick Bell, en uno de sus relatos académicos más celebres, *The space traders*, presenta el hipotético escenario ético de una visita de extraterrestres a los Estados Unidos los cuales ofrecen cien años de prosperidad para la nación a cambio del sacrificio de toda la población negra, para lo cual el total de la población se somete a un referéndum (Bell, 1992). Explicando su método, Bell indica que prefiere (1995, p. 902, traducción propia):

Usar las historias como un medio de comunicar puntos de vista a aquellos que mantienen otros puntos de vista muy diferentes, en un tema tan emocionalmente cargado como es la raza. Las personas disfrutan con las historias y por eso muchas veces dejarán de lado sus creencias, escucharán las historias, y entonces compararán sus puntos de vista no con el mío, sino con el de aquellos que son expresados en la historia

Tal como indica Brown (1995, p. 514, traducción propia), siendo suscrito por Dunbar en su aplicación de la TCR al campo de la archivística (2006, p. 113), algunos de los puntos más importantes del uso de las contra-narrativas como metodología son los siguientes:

Los teóricos críticos de raza usan la narrativa en un esfuerzo de auto-consciencia para incluir las voces de la gente de color que tradicionalmente ha sido excluida de la ciencia legal convencional y 'apropiada'.

El uso de la narrativa reta el paradigma tradicional de la meritocracia en la academia al intentar subvertir lo que es visto como pretensiones de 'objetividad', 'neutralidad', 'meritocracia', y 'ceguera al color'.

Puede decirse, por tanto, que el uso de historias y contra-narrativas es un mecanismo que es utilizado para dar voz a grupos marginalizados y oprimidos cuyos discursos han sido tradicionalmente silenciados, especialmente en la academia. Por otra parte, también es una forma de construir realidades sociales alternativas, tan artificiales como las narrativas reales oficiales, y que sirven como llamada de atención sobre la injusticia social de éstas. Tal como resume Dunbar (2006, p. 115, traducción propia), Delgado (1989, pp. 2414–2416) también indica diversas razones por las que los grupos marginalizados utilizan contra-narrativas para construir una realidad más amplia:

- Desafían el status quo (de la dominación cultural).
- Muchos de ellos se centran en construir la comunidad (dentro de comunidades marginalizadas).
- Pueden abrir nuevas ventanas a la realidad (presentando la posibilidad de múltiples realidades o verdades, que comparten el mismo espacio social o filosófico).
- Pueden acelerar e involucrar conciencia (social).
- Pueden ayudarnos a entender cuándo es el momento de redistribuir el poder.
- Gran parte de la realidad es construida socialmente.

Por supuesto, estos heterodoxos métodos propios de la TCR, así como sus autores, han sido atacados en mayor o menor medida principalmente en la positivista área del Derecho donde surgió (para más detalles sobre el tema ver la amplia discusión sobre autores reaccionarios en Bell, 1995, p. 907-908). Sin embargo, tal como el propio Bell defiende las técnicas, utilizando su particular estilo (p. 908-910), el uso de la TCR podría ser comparado al de los cantos espirituales negros en tiempos de la esclavitud estadounidense: aunque los esclavistas blancos los escuchaban y se sentían entretenidos al pensar que eran inofensivas formas de entretenimiento sin seriedad alguna, en realidad se trataban de códigos y formas de comunicación secreta contra la opresión: avisaban sobre peligros, transmitían información sobre planes de fugas, o simplemente levantaban y fortificaban el ánimo.

En el campo de la organización y representación del conocimiento, los usos de la narrativa y primera persona también han sido utilizados para ilustrar las impresiones y efectos negativos en individuos de grupos marginalizados al utilizar sistemas de organización del conocimiento sesgados (por ejemplo, Grahn, 1984; Penelope 1989; Smiraglia, 2006; de la tierra, 2008). Quizás los ejemplos más ilustrativos sobre narrativas y primera persona, utilizados en relación a estos usos, son recogidos por Hope Olson en *The power to name* (2002, p. 1-3), incluyendo el relato específico en primera persona de Marielena Fina, publicado originalmente en la revista *Cataloging and Classification Quarterly*, al descubrir en 1972 que el encabezamiento utilizado para acceder a la información escrita por latinos/as era "bibliotecas y el discapacitado social" ("Libraries and the socially handicapped") después cambiado por "servicio bibliotecario para el discapacitado cultural" ("Library service to the culturally handicapped") (Fina, 1993, p. 269).

Estas técnicas y ejemplos sirven para ilustrar la primera parte de los objetivos de la TCR, la cual es exponer y analizar los casos de racismo sistémico en las instituciones. Para la segunda parte, la erradicación del racismo de estas instituciones y sustitución por alternativas antirracistas, esbozaremos algunos ejemplos de aplicación de la TCR en la CDD y CDU de Brasil.

3. Aplicación de la TCR a la organización y representación del conocimiento

Jonathan Furner, en su aplicación de la TCR a la CDD en inglés, indicaba algunas recomendaciones antirracistas para su aplicación a los esquemas de clasificación bibliográficos (Furner, 2007, p. 164, traducción propia). En la pre-

sente sección se discutirán estos principios para su aplicación a la CDD y CDU de Brasil, así como para ilustrar su extrapolación a otros sistemas de organización del conocimiento:

- Admisión por parte de los diseñadores de que el sesgo en los esquemas de clasificación existe, y de hecho es un inevitable resultado de los modos en los que están actualmente estructurados.
- Reconocimiento de que la adherencia a una política de neutralidad contribuirá poco a la erradicación del sesgo, y de hecho solo puede extender su vida.
- Construcción, recolección y análisis de las expresiones narrativas de los sentimientos, pensamientos y creencias de los usuarios de los sistemas de clasificación que se identifican con poblaciones particulares definidas racialmente.
- El desarrollo de esquemas de clasificación hechos a medida de las necesidades especiales de información de las múltiples y diversas comunidades, permitiendo a los usuarios buscar información sobre las vidas, intereses y actividades de la gente como ellos o con quien ellos se auto-identifican, usando términos de búsqueda con los que son familiares y se sienten cómodos.
- La implementación de programas diseñados para cambiar las actitudes de aquellos usuarios potenciales que tienen expectativas muy bajas en los esquemas de clasificación porque esos esquemas son percibidos como inútiles en el mejor de los casos (y fuera de lugar u ofensivos en el peor de los casos) dados sus intereses especiales.
- La adopción de procedimientos y criterios estandarizados para evaluar la utilidad de los esquemas de clasificación para los miembros de las comunidades locales.

Respecto al primer punto, no queda muy claro que los diseñadores de esos sistemas —OCLC en el caso de la CDD en los Estados Unidos, y el UDC Consortium en el caso de la CDU en La Haya— reconozcan la poca idoneidad de estos sistemas para grupos o contextos diversos como es el afrobrasileño. De hecho, la mera denominación de “universal” de la CDU, o las estrategias globales de marketing de la CDD (ver por ejemplo Martínez-Ávila et al., 2012a; 2012b) parecen indicar más bien todo lo contrario. Por otra parte, no deja de ser curioso que muchas de las reivindicaciones hechas por Furner (2007) fueran elaboradas a raíz de su experiencia como editor asistente de la CDD en OCLC, aunque aclarando que “los puntos de vista expresados en el trabajo son propios y no necesariamente representan los de los editores de la CDD o los de OCLC Online Computer Library Center, Inc.” (p. 144, traducción propia).

Para paliar esto, un gran avance sería el reconocimiento abierto por parte de los diseñadores

de los sistemas (así como su conocimiento y valoración por parte de las instituciones locales y nacionales encargadas de su adopción) de los peligros y problemas que puede suponer utilizar su sistema en países como Brasil para su comunidad afrodescendiente.

El reconocimiento de la no universalidad y no neutralidad de estos esquemas evitaría además la adherencia continuada a fallidos mecanismos de corrección y posturas positivistas que justifican la naturalidad del sesgo mayoritario y rechazan los puntos de vista de los grupos desfavorecidos. Sin embargo, estos grupos ven sistemáticamente silenciada su voz por culpa de unos sistemas de comunicación asimétricos en los que la influencia y exposición de los sistemas sobre los grupos no es proporcional a la capacidad de respuesta y participación de estos usuarios en sus mecanismos de revisión. La reforma o sustitución de estos mecanismos y políticas de participación tanto nacionales como internacionales propiciaría la participación de las comunidades afrodescendientes de Brasil (como el movimiento negro brasileiro y los grupos de investigación sobre temática africana), así como la identificación de estructuras, términos específicos, distorsiones, preconceptos, omisiones, etc., en las estructuras.

Implementando programas de participación más abiertos e inclusivos y sensibles a estos aspectos sociales, tanto en los procesos de revisión de la CDD en OCLC como en los Seminarios Internacionales de la CDU organizados por el UDC Consortium, se permeabilizarían las barreras idiomáticas, económicas y culturales que muchas veces impiden toda posibilidad de participación de estos colectivos.

Finalmente, las autoridades pertinentes en la implantación de los sistemas en Brasil deberán prever y disponibilizar mecanismos de evaluación locales para determinar los efectos de los sistemas sobre estas comunidades y facilitar la participación y amplificación de sus voces en los mecanismos de revisión de las instituciones.

4. Conclusiones y líneas de actuación futura

Derrick Bell afirma que “los estándares e instituciones creadas por y que fortalecen el poder blanco deben ser resistidas” (Bell, 1995, p. 901, traducción propia). Una vez identificados esas instancias de racismo en los estándares e instituciones bibliotecarias, el siguiente paso es resistirlas. En el presente trabajo hemos presentado algunas de las líneas teóricas y metodológicas de la TCR aplicadas a las clasificaciones bibliográficas en Brasil. En futuros trabajos de

investigación se espera presentar y proponer las alternativas específicas para eliminar el racismo y el sesgo en instancias específicas como la religión.

Notas

(1) Del inglés WEBCHAM: *white, ethnically European, bourgeois, Christian, heterosexual, able-bodied, male.*

Referencias

- Afolabi, Michael (1992). Spiritual Matters: Provision for Independent African Churches in General Classification Schemes. // *International Classification*. 19:4, 210-213.
- Bell, Derrick A. (1992). Faces at the bottom of the well: the permanence of racism. New York: BasicBooks, 1992.
- Bell, Derrick A. (1995). Who's afraid of Critical Race Theory? // *University of Illinois Law Review*, 893-910.
- Berman, Sanford (1971). Prejudices and antipathies: a tract of Library of Congress Subjects Headings concerning people. Metuchen, N. J.: Scarecrow Press, 1971.
- Bowker, Geoffrey C.; Star, Susan Leigh. (1999). Sorting things out: classification and its consequences. Cambridge, MA: MIT Press, 1999.
- Brock, Andre. (2009). Life on the wire. *Information*. // *Communication & Society*. 12:3 (Jun), 344-363.
- Brown, E.M. (1995). The tower of Babel: bridging the divide between critical race theory and "mainstream" civil rights scholarship. // *Yale Law Journal* 105:2, 513-547.
- De la tierra, Tatiana. (2008). Latina Lesbian Subject Headings: the Power of Naming. // Roberto, K. R., (ed.), *Radical Cataloging: Essays at the Front*. Jefferson, NC: McFarland & Company, Inc., 2008. 94-102.
- Delgado, Richard. (1989). The storytelling for oppositionist and others – a plea for narrative. // *Michigan Law Review* 87:8, 2411-2441.
- Delgado, Richard; Stefancic, Jean. (2001). *Critical Race Theory: An Introduction*. New York: NYU Press, 2001.
- Dunbar, Anthony W. (2006). Introducing critical race theory to archival discourse: getting the conversation started. // *Archival Science*. 6:1 (Mar) 2006, 109-129.
- Fina, Marielena. (1993). The Role of Subject Headings in Access to Information: the Experience of One Speaking-Patron. // *Cataloging & Classification Quarterly* 17:1/2, 267-274.
- Foskett, Anthony Charles (1971). Misogynists All: A Study in Critical Classification. // *Library Resources & Technical Services*. 15:2 (Spring), 117-121.
- Foskett, Anthony Charles (1984). Better Dead Than Read: Further Studies in Critical Classification. // *Library Resources & Technical Services*. 28 (November/December), 347-359.
- Furner, Jonathan. (2007). Dewey Deracialized: A Critical Race-Theoretic Perspective. // *Knowledge Organization*. 34:3, 144-168.
- Furner, Jonathan; Dunbar, Anthony W. (2004). The treatment of topics relating to people of mixed race in bibliographic classification schemes: A critical race-theoretic approach // McIlwaine, I. C., (ed.). *Knowledge organization and the global information society: Proceedings of the Eighth International ISKO Conference*. Würzburg, Germany: Ergon Verlag, 2004. 115-120.
- García Gutiérrez, Antonio. (2002). Knowledge organization from a "culture of the border": towards a transcultural ethics of mediation. // Lopez-Huertas, María José, (ed.). *Challenges in knowledge representations and organization for the 21st century: integration of knowledge across boundaries: proceedings of the seventh international ISKO conference*. Würzburg: Ergon Verlag, 2002. 516-522.
- García Gutiérrez, Antonio. (2007). Desclasificados: pluralismo lógico y violencia de la clasificación. Rubí, Barcelona: Anthropos, 2007.
- García Gutiérrez, Antonio. (2011). *Epistemología de la documentación*. Barcelona: Stonberg, 2011.
- García Gutiérrez, Antonio; Martínez-Ávila, Daniel. (2014). *Critical Organization of Knowledge in Mass Media Information Systems*. // *Knowledge Organization*. 41:3, 205-216.
- Grahn, Judy. (1984). *Another Mother Tongue: Gay Words, Gay Worlds*. Boston: Beacon Press, 1984.
- Harris, Angela P. (1994). Foreword: The Jurisprudence of Reconstruction // *California Law Review*. 82, 741-743.
- Hjørland, Birger (2007). Arguments for 'The Bibliographical Paradigm'. Some Thoughts Inspired by the New English Edition of the UDC. // *Information Research* 12:4 paper colis06. <http://InformationR.net/ir/12-4/colis06.html> (2015-03-05).
- Ingram, David (2004). *Rights, Democracy, and Fulfillment in the Era of Identity Politics: Principled Compromises in a Compromised World*. Lanham, MD: Rowman & Littlefield, 2004.
- Kumasi, Kafi. (2012). Roses in the concrete: a critical race perspective on urban youth and school libraries. // *Knowledge Quest*, 40:4 (May/Jun), 32-37.
- Martínez-Ávila, Daniel; Olson, Hope A.; Kipp, Margaret. (2012a). New Roles and Global Agents in Information Organization in Spanish Libraries. // *Knowledge Organization*. 39:2, 125-36.
- Martínez-Ávila, Daniel; Kipp, Margaret; Olson, Hope A. (2012b). DDC or BISAC: The Changing Balance Between Corporations and Public Institutions. // *Knowledge Organization*. 39:5, 309-19.
- Martínez Ávila, Daniel; Guimarães, José Augusto Chaves. (2013). Library classifications criticisms: universality, poststructuralism and ethics. // *Scire*. 19:2 (jul.-dic. 2013), 21-26.
- Milani, Suellen Oliveira; Guimarães, José Augusto Chaves; Olson, Hope A. Bias in subject representation: convergences and divergences in the international literature. // Babik, Wieslaw (ed.). *Knowledge organization in the 21st century: between historical patterns and future perspectives*. Würzburg: Ergon Verlag, 2014. 335-344.
- Olson, Hope A. (2000). Difference, Culture and Change: The Untapped Potential of LCSH. // *Cataloging & Classification Quarterly*. 29:1/2, 53-71.
- Olson, Hope A. (2001a). Sameness and Difference: a Cultural Foundation of Classification. // *Library Resources & Technical Services*. 45:3, 115-122.
- Olson, Hope A. (2001b). Patriarchal Structures of Subject Access and Subversive Techniques for Change. // *The Canadian Journal of Information and Library Science*. 26:2/3, 1-29.
- Olson, Hope A. (2002). *The Power to Name: Locating the Limits of Subject Representation in Libraries*. Dordrecht, Netherlands: Kluwer Academic, 2002.
- Olson, Hope A.; Schlegl, Rose (2001). Standardization, Objectivity, and User Focus: A Meta-Analysis of Subject Access Critiques. // *Cataloging & Classification Quarterly*. 32:2, 61-80.
- Olson, Hope A.; Ward, Dennis B. (1997). *Ghettos and Diaspora in Classification: Communicating Across the Limits*. // Frohmann, Bernd (ed.). *Communication and Information in Context: Society, Technology, and the Pro-*

fessions. Proceedings of the 25th Annual Conference/ Association canadienne des sciences de l'information: Travaux du 25e congrès annuel. Toronto, Ontario: Canadian Association for Information Science, 1997. 19-31.

Penelope, Julia. (1989). My Life as a Lesbian. // Penelope, Julia; Wolfe, Susan J. (eds.). The Original Coming Out Stories. Freedom, CA: Crossing Press. 43-54.

Smiraglia, Richard P. (2006). Curating and Virtual Shelves: an Editorial. // Knowledge Organization 33:4, 185-187.

Enviado: 2015-03-24. Segunda versión: 2015-05-22.
Aceptado: 2015-06-09.

Los datos bibliotecarios en la nube de datos: análisis de los datasets GLAM presentes en el LOD cloud diagram

Library data in the data cloud: an analysis of the GLAM datasets in the LOD cloud diagram

Ana B. Ríos HILARIO (1), Blanca GIL URDICIAIN (2)

(1) Departamento de Biblioteconomía y Documentación, Universidad de Salamanca, anarihi@usal.es

(2) Departamento de Biblioteconomía y Documentación, Universidad Complutense de Madrid, mbgil@pdi.ucm.es

Resumen

Se analiza la representación de los conjuntos de datos pertenecientes a galerías, bibliotecas, archivos y museos (GLAM) en la nube de los datos abiertos y vinculados. En primer lugar, se establecen las bases teóricas en las que se sustenta el análisis: por un lado, se define el concepto de linked open data (LOD); posteriormente, se delimitan las fuentes objeto de estudio, en concreto: el diagrama de la nube LOD en el que se exponen la totalidad de los conjuntos de datos; el catálogo de datos enlazados, en donde se obtiene información individualizada de cada caso específico y el informe titulado conjunto de datos, vocabularios de valores y conjuntos de elementos de metadatos, cuya tipología nos sirve para categorizar los datasets bibliotecarios. El análisis propiamente dicho se centra: primero, en el estudio de la muestra de los datos examinados sobre la totalidad de la nube; segundo, en la categorización de dichos conjuntos en función de la clasificación previamente establecida; finalmente, se describen los casos más representativos centrándose en la tecnología de datos vinculados empleada. Se concluye exponiendo los principales resultados obtenidos y estableciendo una serie de propuestas de mejora.

Palabras clave: Datos abiertos y vinculados (LOD). Galerías, bibliotecas, archivos y museos (GLAM). Diagrama de la nube LOD. Catálogo de datos enlazados Mannheim. Conjunto de datos, vocabularios de valores y conjuntos de elementos de metadatos (VO-CABDATASET).

1. Introducción

La web semántica permite pasar de una web de documentos a una web de datos en la cual unos datos se conectan o enlazan con otros. Se abre así la posibilidad de enlazar conjuntos de datos (*datasets*) con otros conjuntos y en última instancia datos con datos de acuerdo con una serie de principios y modelos de interrogación bien establecidos (*Linked data*, 2014). Estamos por lo tanto ante un nuevo fenómeno, el que hace referencia a los datos abiertos y vinculados,

Abstract

The representation of the data sets belonging to galleries, libraries, archives and museums (GLAM) cloud of linked open data is analyzed. First, theoretical basis on which the analysis is based are set: on one hand, the concept of open data linked (LOD) is defined; subsequently object of study sources are delimited, namely: the cloud diagram LOD in which all datasets are displayed; Mannheim Linked Data Catalog where individual information of each specific case is obtained and the report entitled Datasets, Value Vocabularies, and Metadata Element Set whose typology serves to categorize librarians datasets. The analysis itself focuses: first study sample of data examined of the entire cloud; second, in the categorization of such set according to the predetermined classification; finally, the most representative cases focusing on technology linked data used are described. It is concluded stating the main results and establishing a series of proposals for improvement.

Keywords: Linked Open Data (LOD). Galleries, libraries, archives and museum (GLAM). Cloud diagram LOD. Mannheim Linked Data Catalog. Datasets, Value Vocabularies, and Metadata Element Set (VO-CABDATASET).

más conocido por su acrónimo inglés LOD (*linked open data*).

Desde hace más de 20 años, en el entorno de las instituciones de memoria o de las entidades también conocidas con las siglas de LAM o GLAM (*galleries, libraries, archives and museum*) se están produciendo toda una serie de cambios, a menudo vertiginosos que han tenido como consecuencia un replanteamiento total en los modos de gestión y funcionamiento de dichas instituciones.

Centrándonos en el caso específico de las bibliotecas hemos contemplado como a finales de los 80 principios de los 90 se producía de modo masivo la automatización de dichos centros. El fin de siglo trajo consigo la aparición de nuevos modelos conceptuales, como el FRBR (*Functional Requirements for Bibliographic Records*), y la explosión del término metadato para referirse de modo específico a la descripción de los documentos digitales. En el nuevo milenio aparecen nuevos conceptos asociados a esa nueva tipología documental como son biblioteca digital, virtual o el más “novedoso y determinante” que es el término repositorio. Siguiendo con este rápido recorrido, llegamos al comienzo de la nueva década hablando de datos abiertos y enlazados.

Tomando como base las anteriores premisas el objetivo principal de este artículo se centra en estudiar los conjuntos de datos abiertos y vinculados pertenecientes al ámbito bibliotecario en el sentido más amplio del término. De este propósito general se establecen a su vez los siguientes objetivos específicos:

- Definir los conceptos claves objeto de estudio en lo referente a los datos abiertos y vinculados y delimitar las fuentes de información empleadas específicamente para la realización del análisis.
- Establecer una categorización de los distintos *datasets* GLAM presentes en el *LOD cloud diagram* (diagrama de la nube LOD) siguiendo la tipología establecida en el documento *Datasets, Value Vocabularies, and Metadata Element Sets (Conjunto de datos, vocabularios de valores y conjuntos de elementos de metadatos; VOCABDATASET)* realizado por el World Wide Web Consortium (W3C).
- Realizar un análisis estadístico que posibilite conocer la representación de los conjuntos bibliotecarios dentro de la nube en general y de modo más específico la representación de las categorías establecidas.
- Detallar los casos más representativos de cada categoría expuestos en el *Mannheim Linked Data Catalog (Catálogo de datos enlazados Mannheim)* centrándonos especialmente en la tecnología de datos vinculados empleada.

Dejando constancia de la imposibilidad de recoger y analizar la totalidad de *datasets* del ámbito GLAM publicados como datos vinculados en la actualidad, decidimos tomar como fuente de información principal el *LOD cloud diagram* debido a su representatividad y vigencia. El último

rastreo de la web *linked data* se efectuó en abril de 2014 y su difusión se hizo pública en agosto de ese mismo año. Para el estudio y comprensión de la “nube” nos ha sido de mucha utilidad la consulta del documento titulado *Adoption of the linked data best practices in different topical domains* (Schmachtenberg, Bizer y Paulheim, 2014a) y su versión abreviada disponible en el recurso *State of the LOD cloud 2014* (Schmachtenberg, Bizer y Paulheim, 2014b). Así mismo, para el análisis específico de los *datasets* acudimos al *Mannheim Linked Data Catalog* (2014). Otro instrumento clave para el desarrollo del estudio ha sido el *Datasets, Value Vocabularies, and Metadata Element Sets* (Isaac; Waites; Young; Zeng, 2011), fruto del informe final realizado por el *Library Linked Data Incubator Group (Grupo Incubador de Datos Vinculados de Bibliotecas)*.

Para cumplir con los objetivos propuestos hemos aplicado una triple metodología. En primer lugar, para sentar las bases teóricas se acudió a las propias fuentes anteriormente indicadas. Para la cuantificación de los *datasets* se llevó a cabo un análisis estadístico; y, finalmente, se aplicó el método descriptivo para definir los casos más representativos.

El artículo se estructura en dos partes. En la primera se define el concepto de *linked open data* y se explican las fuentes de información en las que se base el análisis. En la segunda se presenta el análisis propiamente dicho: tras la categorización y contabilidad de los conjuntos de datos bibliotecarios se muestran los casos más relevantes. Finalmente, en las conclusiones se exponen los principales resultados obtenidos y se realizan una serie de propuestas de mejora. El artículo incluye un apéndice final en el que figuran categorizados según las clases establecidas todos los conjuntos de datos bibliotecarios.

2. Conceptos claves entorno a la definición de linked open data (LOD)

Antes de pasar a realizar el estudio específico y objetivo general de este artículo, previamente debemos definir y delimitar una serie de conceptos claves sin los cuales sería difícil comprender el análisis propiamente dicho.

2.1. Linked open data (LOD)

En los últimos años, la palabra “data” aparece con frecuencia en artículos científicos y de divulgación, generalmente acompañado de otro término que delimita su significado como puede ser *open data*, *data mining* y el más reciente, *big data*. Cada vez es más amplio el número de

personas y organizaciones que están contribuyendo a este “diluvio de datos” al optar por compartir su información con los demás (Heath y Bizer, 2011). Sirvan de ejemplo instituciones tan importantes como Amazon, medios de comunicación como la BBC o el New York Times, los organismos públicos como el gobierno de los Estados Unidos y las iniciativas de investigación en diversas disciplinas científicas.

Por supuesto, en esta maraña de datos también están presentes los datos bibliotecarios cuya misión se centra en el intercambio de datos bibliotecarios y en el potencial para la creación de estos datos globalmente interconectados; el intercambio y utilización conjunta de datos con instituciones no bibliotecarias; la creciente confianza en el crecimiento de la web semántica, y en el mantenimiento de una gráfica cultural global de la información que sea fiable y persistente (Bauer y Kaltenböck, 2012). Como ejemplos más representativos de datos bibliotecarios podemos citar la presencia de bibliotecas como la Library of Congress, e iniciativas como Euro-peana y la Digital Public Library of America (DPLA).

El adjetivo “open” asociado a otros términos, ha dado lugar a los denominados “movimientos abiertos”. El citado movimiento está siendo objeto de atención por la comunidad bibliotecaria en los últimos años, ya que reporta numerosos beneficios ante el alto crecimiento de información. El concepto de *open* se incluye dentro de una expresión más amplia que es la de conocimiento abierto. Según el *Open Definition Advisory Council* (2014) este término haría referencia tanto al “contenido incluido en música, películas y libros; los datos de carácter científico, histórico, geográfico; o cualquier otro tipo información gubernamental y de otras administraciones públicas”.

Los datos enlazados es la forma que tiene la web semántica de vincular los distintos datos que están distribuidos en la Web, de tal manera que los datos se enlazan del mismo modo que lo hacen las páginas web. En la *Guía breve de Linked Data* (2015) se especifica que “la web semántica no se trata únicamente de la publicación de datos en la web, sino que éstos se pueden vincular a otros, de forma que las personas y las máquinas puedan explorar la web de los datos, pudiendo llegar a información relacionada a la que se hace referencia desde otros datos iniciales”. Si queremos una definición más precisa del término podemos acudir a la página oficial *linked data* en la que se define este concepto como “la utilización de la Web para conectar los datos relacionados que no estaban vinculados previamente, o el uso de la Web para

disminuir los obstáculos en la conexión de los datos actualmente vinculados mediante otros métodos”.

El término datos vinculados se refiere, por lo tanto, a un conjunto de buenas prácticas para la publicación y la interconexión de datos estructurados en la Web. Para hablar de LD los datos deben publicarse de acuerdo con los principios diseñados para facilitar los vínculos entre los conjuntos de datos, elementos y vocabularios controlados (Berners-Lee, 2006). Estas prácticas fueron presentadas en 2006 por Tim Berners-Lee y se han dado a conocer como los principios de *linked data* (LD). Tales principios son los siguientes:

1. Usar URIs (*uniform resource identifiers*) para identificar los recursos de forma unívoca.
2. Usar URIs http para que la gente pueda acceder a la información del recurso.
3. Ofrecer información sobre los recursos usando RDF.
4. Incluir enlaces a otros URIs, facilitando el vínculo entre los distintos datos distribuidos en la web.

Teniendo en cuenta las anteriores premisas podemos decir que, *linked open data* es un modo de publicar datos estructurados lo que permite que los metadatos se conecten y enriquezcan, de tal modo que diferentes representaciones de un mismo contenido puedan encontrarse y enlazarse entre recursos relacionados (Europeana Linked Open Data, 2015).

Según Mitchell (2013, p. 12) “LOD comprende dos conceptos distintos, el primero es que los datos publicados en la Web deberían conectarse fácilmente con información (“enlazados”) y al hacer esto debería ser accesible tanto para las computadoras como para los humanos (“datos”). El segundo concepto clave es que para que los datos se vinculen y se reutilicen, deben estar abiertos y libres de las restricciones de derechos y copyright (“abiertos”).

Aunque en algunos contextos se suele asimilar ya *linked data* con *linked open data*, no todos los datos enlazados son datos abiertos –y no tienen por qué serlo– sólo lo son si se liberan bajo acuerdos de licencia que permitan su libre reutilización. Por su parte, mientras *linked data* alude a la interoperabilidad técnica de los datos, *open data* hace referencia a su interoperabilidad legal.

2.2. La nube de los datos abiertos y enlazados

El *LOD cloud diagram* (*Diagrama de la nube LOD*) expone los conjuntos de datos que se han publicado en el formato de datos vinculados y que son recogidos por los colaboradores del proyecto Linking Open Data y de otra serie de personas y organizaciones (Figura 1). El propósito principal de este proyecto es “extender la Web como un bien común de datos mediante la publicación de varios conjuntos de datos abiertos como RDF en la Web y mediante el establecimiento de enlaces RDF entre elementos de datos de diferentes fuentes” (Linking Open Data, 2014). El diagrama se sustenta por los metadatos recopilados por los colaboradores del *Data Hub*, así como en los metadatos extraídos de un rastreo de la web *linked data* que se realizó en abril de 2014. El *Data Hub* a su vez podemos definirlo como como “un registro de datos en el que se puede compartir información sobre paquetes de datos de cualquier tipo y describirlos de forma colaborativa” (Isaac, Waites, Young y Zeng, 2011).

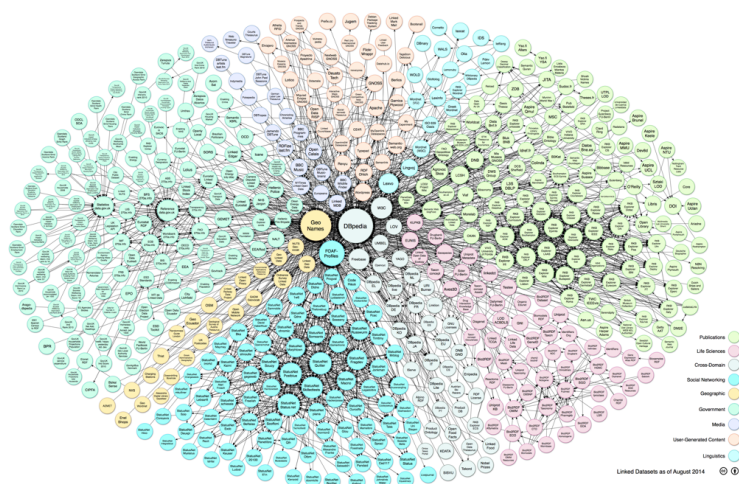


Figura 1. Diagrama de la nube LOD
Fuente: <http://lod-cloud.net/>

En dicho diagrama figuran dos documentos independientes que presentan información sobre los *datasets* en función del momento en el que se obtiene la imagen.

- El estado de la nube 2011: Proporciona información sobre los conjuntos de datos vinculados que fueron catalogadas en el *Data Hub* en septiembre de 2011.
- El estado de la nube en 2014: Presenta información sobre la estructura y el contenido del subconjunto rastreado de la nube LOD en abril de 2014.

La interpretación del diagrama es sencilla. La figura 1 muestra los conjuntos de datos publicados en formato *linked data* y que se entrelazan con otros conjuntos de datos de la nube. El tamaño de los círculos se corresponde al número de tripletas (sujeto, predicado, objeto) en cada conjunto de datos. Generalmente, los números son proporcionados por los editores de los conjuntos de datos y en ocasiones son cálculos aproximados.

Los conjuntos de datos referentes a la categoría *Publicaciones*, donde se incluyen los datos bibliotecarios en su más amplio sentido, se identifican por el color verde claro y se sitúan en el margen derecho del diagrama.

Las flechas muestran la existencia de al menos 50 enlaces entre dos conjuntos de datos. Un enlace, según el propósito de la nube, es una tripleta RDF donde los URIs del sujeto y objeto están en los espacios de nombres de diferentes conjuntos de datos.

La dirección de las flechas indica el conjunto de datos que contiene los enlaces, por ejemplo, una flecha de A a B significa que el conjunto A contiene tripletas RDF que utilizan los identificadores de B. Las flechas bidireccionales generalmente indican que los enlaces se reflejan en ambos conjuntos de datos. El espesor se corresponde con el número de enlaces.

2.2.1. El catálogo de datos enlazados

El *Mannheim Linked Data Catalog* (*Catálogo de datos enlazados Mannheim*) es un instrumento que proporciona información sobre los conjuntos de datos disponibles en la Web en el momento de la realización del último estudio del estado de la nube desarrollado en agosto de 2014. La creación de este catálogo, así como el diagrama de la nube LOD, han sido financiados por el proyecto de la Unión Europea *Planet Data*. El contenido de dicha herramienta se ha generado a partir de dos fuentes:

1. En abril de 2014 se realizó un rastreo de la web de LD y se analizaron los conjuntos de datos vinculados hallados que cumplieran con las prácticas de LD.
2. La comunidad *Linked data* recoge la meta-información de los conjuntos de datos disponibles en el catálogo *datahub.io*.

Dicho catálogo es un complemento al documento *Adoption of the linked data best practices in different topical domains* (Schmachtenberg; Bizer; Paulheim, 2014a) “permitiendo al lector desglosar y explorar la información sobre los conjuntos de datos detrás de cada resultado estadístico” (Schmachtenberg, Bizer, Paulheim, 2014b). Pre-

cisamente, en este trabajo nuestro objetivo es el estudio de los *datasets* bibliotecarios etiquetados bajo la categoría de *Publicaciones* y que figuran en la mencionada herramienta.

2.3. Conjuntos de datos, vocabularios de valores y metadatos

El W3C Library Linked Data Incubator Group (Grupo Incubador del W3C de Datos Vinculados de Bibliotecas; LLD-XG) se crea en el año 2010 con el objetivo de "contribuir a incrementar la interoperabilidad global de los datos de las bibliotecas en la Web, reuniendo a personas implicadas en actividades de la web semántica — centradas en los datos vinculados— en bibliotecas e instituciones afines, mediante el examen de las iniciativas en curso e identificando futuras vías de colaboración". Fruto de su trabajo durante un año fue la redacción de un informe final que se publica en octubre de 2011 bajo la denominación de *Library Linked Data Incubator Group Final Report (Informe Final del Grupo Incubador de Datos Vinculados de Bibliotecas)*.

Dicho informe se estructura en cuatro apartados:

1. Ámbito del informe.
2. Beneficios para la comunidad científica y académica, organizaciones, profesionales de la información y desarrolladores y firmas comerciales.
3. Estado de la cuestión en torno a problemas existentes con los datos bibliotecarios tradicionales, descripción de los LLD disponibles en la actualidad y las cuestiones relacionadas con los derechos.
4. Recomendaciones para directores de bibliotecas, organismos de normalización, ingenieros de datos y de sistemas y bibliotecarios y archiveros.

El informe final incluye a su vez dos documentos que lo complementan pero que se publican de modo independiente: *Use Cases (Casos de usos; USECASE)* y *Datasets, Value Vocabularies, and Metadata Element Sets (Conjunto de datos, vocabularios de valores y conjuntos de elementos de metadatos; VOCABDATASET)*.

El conjunto de casos de uso, estudios de casos, iniciativas e ideas se organizaron en el USECASE atendiendo a los ocho grupos que se citan a continuación: datos bibliográficos; datos de autoridad; vocabularios; archivos y datos heterogéneos; citas; objetos digitales; colecciones; y usos sociales

El USECASE presenta los casos de uso recopilados y ofrece un breve resumen de cada caso individual. Con el fin de proporcionar información más detallada para el lector interesado, las secciones de los grupos y los resúmenes de casos están relacionados con las páginas wiki que el Grupo Incubadora utilizó para recoger y conservar la información original.

Sin embargo, para el propósito de nuestro estudio el documento más relevante es el titulado *Conjunto de datos, vocabularios de valores y conjuntos de elementos de metadatos*. Su objetivo es la identificación de "un conjunto de recursos útiles para la creación o consumo de datos enlazados en el campo de las bibliotecas" (Isaac, Waites, Young y Zeng, 2011). La lectura de este informe permitirá, por un lado, "proporcionar a la comunidad de datos enlazados la oportunidad de comprender el punto de vista específico, los recursos y la terminología usada por parte del colectivo bibliotecario para sus datos"; y, al mismo tiempo, "servirá de ayuda a los profesionales de la información para comprender el modo en el que los conceptos relacionados con los datos enlazados encajan con sus propias tradiciones profesionales" (Isaac, Waites, Young y Zeng, 2011). Se concibe como un punto de partida para que los profesionales puedan encontrar, comprender y explorar algunos ejemplos de conjuntos de datos, vocabularios controlados y metadatos. El informe comienza por definir los conceptos claves que dan nombre al propio documento. Posteriormente se recopila un listado de los casos de uso citados en el documento precedente sobre cada uno de ellos.

Ya que estos mismos términos son los que vamos a emplear a la hora de categorizar los *datasets* bibliotecarios recogidos en el catálogo *Mannheim*, a continuación pasaremos a definirlos según se detalla en el propio VOCABDATASET (Isaac, Waites, Young y Zeng, 2011).

- *Conjunto de datos*: considerados como colecciones de metadatos estructurados, descripciones de cosas como los libros de una biblioteca. En el mundo bibliotecario el equivalente a un conjunto de datos es una colección de registros bibliográficos. Los registros bibliográficos son afirmaciones sobre cosas, en los que cada afirmación está formada por un elemento ("atributo" o "propiedad") de la entidad y un "valor" para ese elemento. Por ejemplo, un registro de un conjunto de datos sobre un libro determinado puede tener un elemento 'materia' tomado de *Dublin Core* y un valor para la materia tomado de los *Library of Congress Subject Headings* (LCSH).

- *Vocabulario de valores*: Un vocabulario de valores define recursos (como instancias de materias, estilos artísticos o autores) que se utilizan como valores de elementos en los registros de metadatos. Son componentes estructurales que sirven para llenar de contenido los registros de metadatos. Muchas bibliotecas usan obligatoriamente determinados vocabularios para seleccionar los posibles valores de un elemento de metadatos. Así, un vocabulario de valores es una lista de los valores permitidos para un elemento. Es el caso de los tesauros, listas de códigos, listas de términos, esquemas de clasificación, listas de encabezamientos de materia, taxonomías, ficheros de autoridades, nomencladores geográficos, esquemas de conceptos y otros tipos de sistemas de organización del conocimiento. Por ejemplo, el *Fichero Virtual de Autoridades (Virtual International Authority File; VIAF)* define autoridades de nombres (por ejemplo, Mark Twain).
- *Conjuntos de elementos de metadatos o conjuntos de elementos*: Los conjuntos de elementos de metadatos definen las clases y atributos utilizados para describir entidades de interés. En la terminología de datos vinculados, estos conjuntos de elementos se concretan generalmente por medio de RDF Schemas (esquemas RDF) o Web Ontology Language (Lenguaje de Ontologías Web; OWL), que con frecuencia se agrupan bajo el término "vocabulario RDF". Normalmente los conjuntos de elementos de metadatos no describen entidades bibliográficas, sino más bien proporcionan los elementos que se pueden utilizar para describir estas entidades. Términos equivalentes serían vocabularios RDF, esquemas RDF y ontologías (*Library terminology informally explained*, 2011). Por ejemplo, MARC21 define los elementos (campos) para describir registros bibliográficos y de autoridades.

3. Los datos bibliotecarios en el LOD cloud diagram

3.1. Análisis estadístico

Antes de comenzar el estudio pormenorizado de la categoría denominada *Publicaciones* es necesario obtener una visión global de la totalidad de conjuntos de datos presentes en el diagrama. Para ello, se ha tomado la información que se proporciona en el propio documento, en el que se clasifican por categorías temáticas los 1.014 conjuntos de datos que se identificaron en el último rastreo (Tabla I).

Categoría	Conj.	%
Gobierno	183	18,05%
Publicaciones	96	9,47%
Ciencias de la vida	83	8,19%
Contenido generado por el usuario	48	4,73%
Dominios cruzados	41	4,04%
Medios	22	2,17%
Geografía	21	2,07%
Web social	520	51,28%
Total	1.014	100%

Tabla I. Conjuntos de datos por categorías temáticas (Fuente: Mannheim, 2014)

Como puede observarse de las ocho materias identificadas, la categoría *Publicaciones* ocupa la tercera posición, con 96 conjuntos de datos identificados, lo que se corresponde con el 9,47% del espacio en la nube. Destacan por encima del resto, los conjuntos de datos relativos a la web social (520), que ostenta la primera posición.

Centrándonos en el análisis específico de la materia objeto de estudio, en primer lugar, tenemos que comenzar definiendo qué *datasets* se han etiquetado bajo dicha denominación.

Tal y como se indica en el propio documento, la categoría denominada *Publicaciones* comprende "conjuntos de datos de bibliotecas, información sobre publicaciones científicas y conferencias, listas de lectura de las universidades, y las bases de datos de citas".

El número de conjuntos que figura en el resultado final difiere del que se muestra en el catálogo Mannheim. En dicho catálogo aparecen 101 conjuntos de datos etiquetados dentro de la categoría *Publicaciones*. Tal diferencia viene dada porque los conjuntos pueden categorizarse en ocasiones en más de una categoría, pero a la hora de realizar el recuento final se ha optado por incluirlos en una única. De tal manera que del total de casos que figuran en el catálogo, un ejemplo aparece repetido y cuatro son contabilizados además en otra categoría. Es así como Uniprot-citation se incluye en el apartado de Ciencias de la vida, y DWS-Group, Data.dcs y Morelab se recogen en la tipología denominada Web Social.

Si hacemos una clasificación del total de los 96 casos contabilizados atendiendo a las subcate-

gorías establecidas en la anterior definición obtenemos los siguientes resultados (Tabla II).

Subcategoría	Número	Porcentaje
Bibliotecas (GLAM)	42	44%
Publicaciones científicas	23	24%
Listas de lectura	17	18%
Bases de datos	14	14%
Total	96	100%

Tabla II. Subcategorías de la categoría Publicaciones

Como puede observarse el mayor número de casos hace referencia a la subcategoría denominada *Bibliotecas* con un total de 42 ejemplos. Entiéndase el término bibliotecas en su más amplio sentido; es decir, abarcaría a todas las instituciones recogidas bajo las siglas GLAM. En un porcentaje similar se encuentran tanto la subcategoría que recoge las *Publicaciones científicas y conferencias* como las listas de lectura ofertadas por las instituciones académicas de enseñanza superior. La última subcategoría referente a las base de datos de citas presenta 15 casos recogidos en el mencionado catálogo.

Atendiendo a la clasificación que recoge el documento VOCABDATASET —que divide los datos vinculados bibliotecarios entre conjuntos de datos, vocabulario de valores y conjunto de elementos de metadatos—, de los 42 conjuntos de datos obtenidos en la anterior tabla como *datasets* bibliotecarios podemos obtener los datos de la Tabla III. Las colecciones bibliotecarias destacan ligeramente sobre los vocabularios de valores. En un porcentaje menor aparecen los conjuntos de metadatos.

Clases	Número	Porcentaje
Colecciones	20	48%
Vocabularios de valores	16	38%
Metadatos	6	14%
Total	42	100%

Tabla III. Clases según la clasificación establecida por el VOCABDATASET

Haciendo nuevamente un desglose del conjunto de casos que se han agrupado bajo la clase denominada *Colecciones*, aparecen en porcentajes similares los fondos de bibliotecas y archivos, bibliografías nacionales y catálogos (Tabla

IV). Aunque todos y cada uno de los casos referentes a esta categoría hacen referencia a “los conjuntos de datos considerados como colecciones de metadatos estructurados” (VOCABDATASET, 2011) los hemos diferenciado entre las diferentes colecciones que figuran en la tabla III, tan sólo hemos dejado un caso sin catalogar por la dificultad de clasificarlo en dichas categorías. Se trata del *Project Gutenberg* que consta de un conjunto de libros electrónicos gratuitos, tal y como figura en la propia definición del caso y que, aunque similar a los datos de una biblioteca, no está definido como tal.

Subclases	Número	Porcentaje
Bibliotecas	5	25%
Archivos	4	20%
Museos	2	10%
Bibliografías	4	20%
Colecciones	1	5%
Catálogos	4	20%
Total	20	100%

Tabla IV. Subclases para Colecciones

La siguiente clase definida en el documento elaborado por el LLD-XG hace referencia a los vocabularios de valores y vocabularios controlados. En esta ocasión, a la hora de subdividir esta clase, hemos empleado los mismos apartados que figuran recogidos en el propio documento. En este caso, aparecen con el mismo número tanto los sistemas de clasificación como los tesauros. Con dos ejemplos se presentan tanto los encabezamientos de materia como los datos de autoridad (Tabla V).

Subclases	Número	Porcentaje
Sistemas de clasificación	6	38%
Encabezamientos de materia	2	12%
Datos de autoridad	2	12%
Tesauros	6	38%
Total	16	100%

Tabla V. Subclases para Vocabulario de valores

Finalmente, si subdividimos la última clase que se presenta en el VOCABDATASET —denominada ‘conjunto de elementos de metadatos’—, es significativo que la mayoría de los

casos hacen referencias a ontologías. Tan sólo se da un único caso definido como metadatos, que es nuevamente el Project Gutenberg pero en este caso presentado como RDF (Tabla VI).

Subclase	Número	Porcentaje
Ontologías	5	83%
Metadatos	1	17%
Total	42	100%

Tabla VI. Subclases para Metadatos

3.2. Descripción de los casos más representativos

En este apartado pasaremos a describir brevemente los casos más representativos de cada una de las clases anteriormente identificadas.

Comenzando por los *datasets* incluidos dentro de la primera clase establecida —la referente al apartado *Colecciones*, y dentro de ella la subclase *Bibliotecas*—, destacan por un lado Europea, y los proyectos de datos enlazados de las bibliotecas nacionales de Francia y España.

the Cittie of London 31

Descripción:
This map has been attributed to Augustus Ryther, an engraver who prospered between 1572 and 1592, contributing to Saxton's Atlas of 1579. This plan was produced to satisfy a European market, and contains certain inaccuracies which a native Londoner would not have tolerated. The streets appear very much wider than they were in actuality. Houses are depicted as having large gardens, when these had, in fact, begun to disappear from London two centuries before. The map details the gap at the north end of London Bridge, caused by a fire in 1632.

Creador:
Cartographer: Ryther, Augustus

Cobertura geográfica:
531500, 181500; London, City of London

Fecha:
1633, Part of: Mid 17th century; From: Sat Jan 01 01:00:00 CET 1633 — To: Sat Dec 31 01:00:00 CET 1633

Fecha de creación:
1633

Tipo:
mapas

Formato:
jpeg, 368, 464, Milímetros, Engraving

Materia:
River Thames, London Bridge, London (England) — Maps, 912

Identificador:
00700000000001U00031000, Maps.Grace I

Es parte de:
Grace Collection Of Maps Of London, Collect Britain

Por el proveedor:
The British Library

Proporcionar Contenido:
The European Library

Providing country:
United Kingdom

Etiquetas generadas automáticamente *

Other items you may be interested in:

Figura 2. Ejemplo de Europea para Londres

Europeana, actualmente contiene metadatos abiertos sobre 2,4 millones de textos, imágenes, vídeos y sonidos recogidos por esta institución. Estos objetos proceden de proveedores de datos que han reaccionado pronto y positivamente a la iniciativa de Europea de promoción de datos más abiertos y nuevos acuerdos de intercambio de datos. Estas colecciones provienen de 8 proveedores directos de Europea que

abarca más de 200 instituciones culturales de 15 países. Cubren una gran variedad de objetos del patrimonio. Para el modelado de los datos siguen un esquema propio: *Europeana Data Model* (EDM). La adopción de este modelo basado en los estándares desarrollados por el W3C, ha permitido a Europea la compatibilidad con el paradigma de la web semántica. EDM está constituido por la reutilización de namespaces como RDF (*Resource Description Framework*), RDFs (*Resource Description Framework Schema*), OAI-ORE (*Open Archives Initiative Object Reuse and Exchange*), SKOS (*Simple Knowledge Organization System*) y DCMI Terms (*Dublin Core Metadata Initiative Terms*) (Figura 2).

Data.bnf.fr, recoge los datos de las diferentes bases de datos de la Biblioteca Nacional de Francia, con el fin de crear páginas web sobre obras, autores, materias y lugares, junto con una vista RDF de los datos extraídos. Fue uno de los proyectos pioneros que comenzó su actividad en julio de 2011. Como modelo conceptual utilizan el esquema FRBR de la IFLA y los vocabularios SKOS, FOAF (*Friend Of A Friend*), DC y RDA (*Resource Description and Access*).

DATOS.BNE.es

INICIO AUTORES OBRAS TEMAS Ayuda

Buscar un autor, obra o tema

Convenio Susseita, Miguel de / Don Quijote de la Mancha

Don Quijote de la Mancha

En la BNE

Disponibles 86 Versiones

Tema en 440 Obras

Versión destacada (Español) (ver las 863 ediciones)

Vida, y hechos del ingenioso caballero D. Quijote de la Mancha compuesta por Miguel de Cervantes Saavedra; parte II.

Título: Vida, y hechos del ingenioso caballero D. Quijote de la Mancha compuesta por Miguel de Cervantes Saavedra; parte II.

Lugar de publicación: En Madrid

Editorial: escuela de Santa Maria de Arnedo, escuela de Juan Antonio Bonet, mercader de libros, en France de S. Phelipe por Roque Rico de Miranda, impresor de libros.

Fecha de publicación: 1616

Tipo de recurso: Libro

Figura 3. La obra Don Quijote de la Mancha en la BNE

Datos.bne.es es el proyecto de datos bibliográficos vinculados de la Biblioteca Nacional de España. Incluye 4 millones de registros de autoridad y 2.4 millones de registros bibliográficos que dan lugar a más de 58 millones tripletes. Los datos españoles se modelan mediante vocabularios IFLA (FRBR, FRISAD, ISBD) y otros como RDA y DC. Para la migración y la vinculación se utilizó Marimba, una herramienta especial desarrollada por la propia institución y el Ontology Engineering Group y diseñada para ser utilizada por los bibliotecarios (Figura 3).

En el apartado denominado *Archivos* destaca por su importancia dentro de los datos vinculados la agrupación *Archives Hub Linked Data*. Este ejemplo recoge una muestra de datos de descripciones de fondos de archivos mantenidos en el Archives Hub, un agregador del Reino Unido, y emitidos como *linked data*. Ofrece una perspectiva sobre las personas, las organizaciones, materias y lugares relacionados con los archivos que se describen. Se proporcionan enlaces externos a otros conjuntos de datos, como el VIAF y la LCSH. Asimismo, proporciona una hoja de estilo para convertir los datos EAD (XML para archivos) en RDF XML.

Dentro de la subclase catalogada como *Museos* es relevante la colección del *British Museum*. Este servicio de datos vinculados y SPARQL (*Protocol and RDF Query Language*) proporciona acceso a la misma colección de datos disponibles a través de la web del Museo. El uso del estándar de datos abiertos RDF del W3C permite que los datos de la colección del museo se unan y se refieran a un creciente cuerpo de datos vinculados publicados por otras organizaciones de todo el mundo interesadas en la promoción de la accesibilidad y la colaboración. Los datos se han dispuesto mediante el empleo del CIDOC-CRM (modelo conceptual de referencia) fundamental para la armonización con otros datos del patrimonio cultural (Figura 4).

Predicate	Object
<code>rdf:type</code>	<code>skos:Concept</code>
<code>skos:P43_constitutes_of</code>	<code>http://www.oxfordjournals.org/doi/10.1093/ajph/117.12.1724</code>
<code>skos:P51_has_former_or_current_name</code>	<code>http://collection.britishmuseum.org/id/object/C0131318</code>
<code>skos:P52_has_current_name</code>	<code>http://collection.britishmuseum.org/id/object/C0131318</code>
<code>skos:P138_has_essential_uri</code>	<code>http://www.britishmuseum.org/collect/objects/AN00012400012840_001.jsp</code>

Figura 4. Hoa Hakananai'a del British Museum en rdf

El *Proyecto Gutenberg*, cuya dificultad en su clasificación ya hemos descrito anteriormente, es la primera y más grande colección de libros electrónicos gratuitos. Sobre este caso no se proporciona información técnica en el catálogo Mannheim desde el punto de vista de ejemplo de aplicación LOD. Su versión en RDF figura como un caso específico dentro de la categoría

metadatos, a la que posteriormente haremos referencia.

Si pasamos a la siguiente tipificación —*Vocabulario de valores*—, comenzando por la subclase *Sistemas de clasificación*, destaca la *Dewey Decimal Classification* (DDC). *Dewey.info* es un espacio experimental para los datos vinculados de la DDC. La intención de este prototipo es convertirse en una plataforma de los datos Dewey en la Web. Incluye como LD los sumarios DDC (los tres primeros niveles) de la vigésima segunda edición, en once idiomas, y todos los números asignables de la decimocuarta edición abreviada, en tres idiomas. La clasificación semántica se codifica en RDF utilizando SKOS y otros vocabularios. Cada clase tiene también una representación HTML (XHTML + RDFa) y varias serializaciones RDF —RDF/XML, Turtle (*Terse RDF Triple Language*) y JSON (*JavaScript Object Notation*)—.

Predicate	Object
<code>skos:Concept</code>	<code>http://id.loc.gov/authorities/0011738#concept</code>
<code>skos:P43_constitutes_of</code>	<code>http://id.loc.gov/authorities/0011738</code>
<code>skos:P51_has_former_or_current_name</code>	<code>http://id.loc.gov/authorities/0011738#concept</code>
<code>skos:P52_has_current_name</code>	<code>http://id.loc.gov/authorities/0011738#concept</code>
<code>skos:P138_has_essential_uri</code>	<code>http://id.loc.gov/authorities/0011738#concept</code>

Figura 5. La materia Libros en la LCSH

Si siguiendo con esta clase, en la subclase *Encabezamiento de materias* vamos a destacar por su interés los dos casos presentes: por un lado, la *Library of Congress Subject Headings* (LCSH) y la *Lista de Encabezamientos de Materia* (LEM). La LCSH es la lista de materias más utilizada a nivel internacional. Desde el año 1898 la Biblioteca del Congreso ha mantenido de forma activa esta lista para la indización de sus fondos. Los modelos utilizados para su implementación con LOD son los *Metadata Authority Description Schema* (MADS) —el esquema XML para los elementos de autoridad establecida por la propia institución— y SKOS (Figura 5).

Por su parte, mediante esta iniciativa la *Lista de Encabezamientos de Materia* (LEM) transforma su esquema a SKOS, y se adapta a *Europeana*

como uno de los *namespaces* constitutivos del *Europeana Data Model*. Ha establecido siempre que fue posible, un *skos:closeMatch* con los de la LCSH y se está progresando para proporcionar una funcionalidad análoga con RAMEAU (*Répertoire d'autorité-matière encyclopédique et alphabétique unifié*) o SWD (*Schlagwortnormdatei*). Incorpora además un fichero en zip de los registros MARC de autoridades de materia, es decir, aquellos basados en el campo 150 en los que se ha añadido las URI de las LEM y las de la LCSH (Figura 6).



Figura 6. El concepto Libros en la LEM

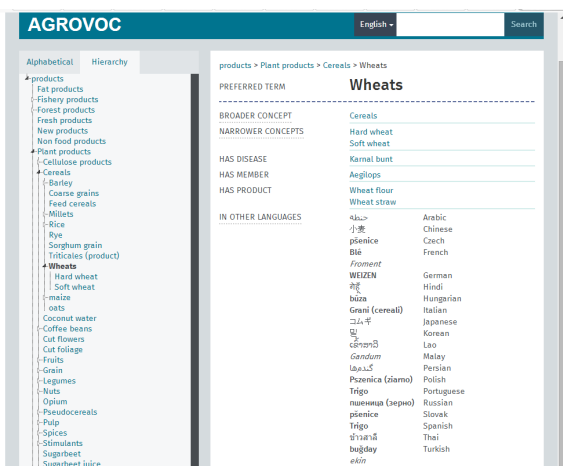


Figura 7. El concepto trigo en AGROVOC

Finalmente, en el último apartado de esta categoría destacaremos el tesoro AGROVOC publicado por la FAO (Food and Agriculture Organization). *AGROVOC Linked Open Data (LOD)* es un proyecto para convertir el tesoro AGROVOC en una columna vertebral terminológica multilingüe para artículos digitales agrícolas. Alojado por MIMOS, socio en la investigación, provee acceso web a los registros de datos estructurados sobre conceptos agrícolas y, más importante aún, enlaces sobre esos conceptos a

otros tesauros en línea. Actualmente describe más de 4.830.173 de recursos. La versión de LD del tesoro está en RDF/ SKOS-XL, y el conjunto de tripletas se almacena en Allegro-Graph (Figura 7).

Para terminar con este apartado descriptivo concluiremos haciendo referencia a la última clase establecida –*Metadatos*– y al último caso nombrado: *Project Gutenberg in RDF*. Esta iniciativa ofrece los metadatos sobre las obras de dominio público disponibles en el *Proyecto Gutenberg* al que anteriormente ya habíamos hecho referencia dentro de la subclase *Colección*. El proyecto ha sido realizado por el grupo de investigación Research Group Data and Web Science de la Universidad de Mannheim. Los datos estructurados se proporcionan de acuerdo a los principios de LD y se pueden acceder a los mismos a través del sitio web del grupo por medio de tres vías diferentes: HTML, RDF y SPARQL (Figura 8).



Figura 8. Moby Dick en el Proyecto Gutenberg

4. Conclusiones

Primeramente nos gustaría subrayar las dificultades a las que hemos tenido que hacer frente a la hora de realizar nuestro estudio y las anomalías encontradas tanto en la configuración del *LOD cloud diagram* como en la estructura del catálogo *Mannheim*, con el objetivo de que puedan ser solventadas en las próximas ediciones de ambos instrumentos repercutiendo en una susceptible mejora de los mismos. No obstante, ponemos de manifiesto nuevamente la relevancia y utilidad de este tipo de materiales siendo conscientes igualmente de la complejidad de su elaboración.

Primero, pensamos que la denominación de la categoría Publicaciones no es muy acertada; si bien es cierto que los autores de la nube han querido englobar en un número reducido las

categorías que comprende el diagrama y que esta denominación es un concepto muy amplio y en ella se pretende reunir la información científica y cultural. Nuestra propuesta sería o bien cambiar la denominación con un epígrafe similar al anteriormente indicado o utilizar dentro de esta categoría etiquetas más específicas que nos permitirían encontrar directamente las distintas subcategorías. En este sentido, para los conjuntos de datos GLAM, proponemos la misma clasificación que hemos empleado nosotros, cuya definición viene avalada por el W3C.

Por otro lado, la información individualizada de cada *dataset* que nos proporciona el catálogo *Mannheim* es muy dispar y en ocasiones muy escueta. De tal modo que en un número importante de casos hemos tenido que proceder a buscar el enlace del mismo teniendo en cuenta que no es igual la dirección de una biblioteca, por ejemplo, que la dirección al proyecto de datos enlazados de esa misma institución. Es decir, la URL de la Biblioteca Nacional es *bne.es* mientras que su proyecto de datos enlazados es *datos.bne.es*. Incluso hay casos en los que ni siquiera se proporciona información, por lo que hemos tenido que llevar a cabo toda una tarea de investigación para identificar el ejemplo en concreto y así poderlo categorizar. El responsable último de este hecho no es tanto el propio catálogo sino los responsables de los *datasets*. Por lo tanto, abogamos porque se proporcione la mayor información posible por parte de los gestores de los conjuntos de datos en aras a facilitar tanto el uso del propio *dataset* como el análisis del mismo.

Uno de los objetivos que nos planteábamos en el inicio del trabajo era la definición de los conceptos claves tanto en lo concerniente a la tecnología de los datos abiertos y vinculados como a la definición de conceptos más específicos de nuestro ámbito, en concreto, los referentes a los recursos bibliotecarios empleados: colecciones, vocabularios de valores y metadatos. A través de la consulta de la documentación especializada sobre el tema comprobamos que en ocasiones ambos conceptos se confunden y se está perdiendo precisión terminológica. Mediante la explicación de los términos en el apartado teórico y la sistematización de los ejemplos que figura en el apéndice esperamos haber cumplido con nuestro propósito.

Centrándonos ya en los resultados obtenidos en lo referente a los *datasets* culturales, la categoría *Publicaciones* en la que se incluyen los mismos ocupa el tercer puesto dentro de la nube pero su porcentaje no alcanza ni siquiera el 10 % de la totalidad. Sin embargo, dentro de esta categoría los datos de las instituciones GLAM

son casi la mitad del resto de clases que comprende esta división. De acuerdo a la clasificación del LLD-XG el número de los datos es muy parecido tanto para la clase *Colecciones* como para la de *Vocabulario de valores*, con un ligero predominio de los primeros. Por su parte, dentro de la subclase *Colecciones* se obtienen porcentajes similares dándose en igual número los conjuntos de *Bibliotecas*, *Bibliografías* y *Catálogos*. En la clase de *Vocabulario de valores* destacan por igual los *Sistemas de clasificación* y los *Tesauros*, mientras en la última división se da una supremacía de las *Ontologías* frente a los *Metadatos*.

En el último apartado hemos intentando describir los casos más representativos de las categorías previamente establecidas, centrándonos no sólo en su definición sino en la aplicación de las tecnologías de datos vinculados. Esperamos así conseguir que las instituciones culturales que estén interesadas en publicar sus datos como LOD puedan obtener con este informe una primera aproximación a la aplicación de esta tecnología, remitiendo a cada caso particular para obtener una información más exhaustiva.

Sobre este punto diremos que nos hubiera gustado profundizar en la explicación de los recursos empleados a la hora de disponer los datos como LOD. Sin embargo, este hecho no ha sido posible debido a la dificultad anteriormente comentada para encontrar la información referente a la utilización de las diferentes prácticas relativas a este aspecto.

Referencias

- Bauer, F.; Kaltentböck, M. (2012). *Linked Open Data: The Essentials: a Quick Start Guide for Decision Makers*. Viena, Austria: edition mono/monochrom, 2012.
- Bernes-Lee, Tim (2006). *Linked data – Design Issues*. [S.l.] : Bernes-Lee, 2006. <http://www.w3.org/DesignIssues/LinkedData.html> (2015-03-30).
- Europeana Linked Open Data (2015). European Union: Europeana Foundation: European Creative Project, 2015. <http://data.europeana.eu> (2015-03-30).
- Guía Breve de Linked Data (2015). W3C. Gijón: W3C Oficina España, 2015. <http://www.w3.es/Divulgacion/GuiasBreves/LinkedData> (2015-03-30).
- Heath, T.; Bizer, C. (2011). *Linked data: Evolving the Web into a Global Data Space*. Florida, USA: Morgan & Claypool Publishers, 2011.
- Isaac, Antoine; Waites, William; Young, Jeff; Zeng, Marcia. *Library Linked data Incubator Group: Datasets, Value Vocabularies, and Metadata Element Sets*. [S.l.]: W3C, 2011. <http://www.w3.org/2005/Incubator/ld/XGR-ld-vo-cabdataset-20111025> (2015-03-30)
- Library Linked Data Group Incubator (2011). *Linked Data Incubator Group Final Report*. [S.l.]: World Wide Web Consortium, 2011. <http://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/> (2015-03-30).
- Library Linked Data Group Incubator (2011). *Linked Data Incubator Group Use Case*. [S.l.]: World Wide Web Con-

- sortium, 2011. <http://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/#ref-USECASE> (2015-03-30).
- Library terminology informally explained (2011). [S.I.]: World Wide Web Consortium, 2011. http://www.w3.org/2001/sw/wiki/Library_terminology_informally_explained#Definitions (2015-03-30).
- Linked data: connect Distributed Data across the Web (2014). [S.I.]: [s.n.], 2014. <http://linkeddata.org> (2015-03-30).
- Linking Open Data (2014). [S.I.]: World Wide Web Consortium, 2014. <http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData> (2015-03-30).
- Mannheim Linked Data Catalog (2014). Mannheim: University of Mannheim, 2014. <http://linkeddatacatalog.dws.informatik.uni-mannheim.de> (2015-03-30).
- Mitchel, Erik (2013). Library Linked data: research and adoption. // *Library Technology Reports*. 5:49 (2013) 11-25.
- Open Definition Advisory Council (2014). Cambridge: Open knowledge, 2014. <http://opendefinition.org/> (2015-03-30).
- Schmachtenberg, Max; Bizer, Christian; Paulheim, Heiko (2014a). Adoption of the linked data best practices in different topical domains. Mannheim: University of Mannheim; Unión Europea: Planet Data, 2014. <http://dws.informatik.uni-mannheim.de/fileadmin/lehrstuehle/ki/pub/SchmachtenbergBizerPaulheim-AdoptionOfLinkedDataBestPractices.pdf> (2015-03-30).
- Schmachtenberg, Max; Bizer, Christian; Paulheim, Heiko (2014b). State of the LOD Cloud 2014. Mannheim: University of Mannheim; Unión Europea: Planet Data, 2014. <http://linkeddatacatalog.dws.informatik.uni-mannheim.de/state/#toc1> (2015-03-30).

Enviado: 2015-04-01. Segunda versión: 2015-05-11.
Aceptado: 2015-05-22.

Apéndice

	<i>Bibliotecas</i>	<i>Princeton Library FindingAids</i>	http://findingaids.princeton.edu/
		data.bnf.fr - Bibliothèque nationale de France	http://data.bnf.fr/
		datos.bne.es	http://datos.bne.es
		Europeana Linked Open Data	http://data.europeana.eu/
		Public Library of Veroia	http://libver.math.auth.gr
	<i>Archivos</i>	Airforcehistoryindex	http://airforcehistoryindex.org/
		Muninn World War I	http://rdf.muninn-project.org
		Archives Hub Linked Data	http://data.archiveshub.ac.uk/
		National Digital Data Archive of Hungary	http://lod.sztaki.hu/
<i>Colecciones</i>	<i>Museos</i>	British Museum Collection	http://collection.britishmuseum.org/
		Amsterdam Museum	http://semanticweb.cs.vu.nl/lod/am/
		Bibliografías Nacionales	
		British National Bibliography	http://bnb.data.bl.uk/
		Deutsche Biographie	http://www.deutsche-biographie.de/
		Deutsche Nationalbibliografie	http://www.dnb.de/EN/datendienste/linkedData
		LIBRIS	http://libris.kb.se
	<i>Colecciones</i>	Project Gutenberg	http://www.gutenberg.org/
	<i>Catálogos</i>	Sudoc bibliographic data	http://punktokomo.abes.fr/2011/07/04/le-sudoc-sur-le-web-de-donnees/
		Hungarian National Library (NSZL) catalog	http://nektar.oszk.hu/wiki/Semantic_web
	HeBIS	http://www.hebis.de/	
	85ZDB	http://www.zeitschriftendatenbank.de/services/schnittstellen/linked-data/	

Vocabularios de valores	Sistemas de clasificación	Mathematics Subject Classification	http://msc2010.org/mscwork/
		Faceted Application of Subject Terminology	http://experimental.worldcat.org/fast/
		Dewey Decimal Classification (DDC)	http://dewey.info
		JITA Classification System of Library and Information Science	http://www.destin-informatique.com/ASKOSI/Wiki.jsp?page=JITA%20Maintenance
		ICONCLASS - Multilingual Thematic Classification	http://www.iconclass.org/help/lod
		Glottolog	http://glottolog.org
	Encabezamientos de materia	Lista de Encabezamientos de Materia	http://id.sgcb.mcu.es
		Library of Congress Subject Headings	http://id.loc.gov/authorities/
	Datos de autoridad	IdRef: Sudoc authority data	http://punktokomo.abes.fr/2011/07/05/idref-des-pages-html-et-rdf-plus-riches/
		Gemeinsame Normdatei (GND)	http://d-nb.info/standards/elementset/gnd#
	Tesauros	yso-fi-allars	http://finto.fi/allars/en/page/Y46304?rdf+xml
		Unesconom	http://skos.um.es/unesco6/6101/rdfxml
		STW Thesaurus for Economics	http://zbw.eu/stw/versions/latest/about
		National Agricultural Library Thesaurus	http://agclass.nal.usda.gov/agt.shtml
		AGROVOC	http://aims.fao.org/vest-registry/vocabularies/agrovoc-multilingual-agricultural-thesaurus
		Mis Museos, índice semántico de museos, artistas y obras de arte (GNOSS)	http://mismuseos.net/comunidad/metamuseo
	Ontologías	Ontologías	yso-fi-ysa
		Bible Ontology	http://bibleontology.com/service/
		ASN:US	http://asn.jesandco.org
		Traditional Korean Medicine Ontology	http://tkm.kiom.re.kr
		FAO geopolitical ontology	http://www.fao.org/countryprofiles/geoinfo/en/?lang=en
Metadatos	Project Gutenberg in RDF	http://www4.wiwiss.fu-berlin.de/gutendata/	

Tabla VII. Listado de los conjuntos de datos bibliotecarios del LOD cloud diagram

A pesquisa sobre política de indexação no Brasil: avanços e desafios

La investigación sobre políticas de indización en Brasil: avances y desafíos

Research on indexing policy in Brazil: progress and challenges

Paula Regina DAL'EVEDOVE (1), Mariângela Spotti Lopes FUJITA (2)

(1) Universidade Federal de Pernambuco/UFPE. Av. da Arquitetura, s/n, Campus Universitário, CEP 50740-550 – Recife/PE, p.dallevedove@gmail.com (2) Universidade Estadual Paulista, Av. Hygino Muz-zi Filho, 737, Campus Universitário, CEP. 17525900, Marília, SP, Brasil, mariangelaslf57@gmail.com

Resumen

El objetivo de este estudio es reflexionar sobre la investigación brasileña sobre políticas de indización en el área de Organización y Representación del Conocimiento. A través de la investigación exploratoria bibliográfica, hemos tratado de contextualizar los principales avances en la literatura sobre la política de indización y, sobre todo, facilitar la comprensión del estado actual de la investigación sobre el tema. Se observó que, a pesar de los avances y resultados consistentes obtenidos por la investigación brasileña, debe prestarse atención urgente a nuevos enfoques de investigación que se centran en las cuestiones que permanecen abiertas o que no se han trabajado a fondo hasta el momento.

Palabras clave: Indización. Políticas de indización. Organización del conocimiento. Representación del conocimiento. Brasil.

1. Introdução

As diretrizes e normas adotadas no processo de indexação servem para dar direcionamento às ações profissionais, visando potencializar a condição de êxito dos sistemas de informação no que tange à criação da informação documentária e sua posterior recuperação. A expressiva importância de uma política de indexação estruturada e registrada para o encaminhamento da atividade de indexação decorre, especificamente, por esta direcionar a prática profissional mediante embasamento teórico e por compilar um conjunto de diretrizes e critérios elaborados e adaptados à realidade das unidades de informação.

Inserida na esfera da representação e recuperação da informação, a política de indexação constitui-se na formalização dos processos, procedimentos, instrumentos e filosofias institucionais, podendo ser entendida como a iniciativa de materializar em guias ou manuais os procedimentos adotados para o desenvolvimento da atividade de indexação (Gil Leiva, 2008). Neste sentido, a adoção de políticas de indexação é

Abstract

The aim of this study is to reflect about the Brazilian research on indexing policy within the Organization and Knowledge Representation field. Through a bibliographical exploratory research, we sought to contextualize the main advances in the literature on indexing policy and, above all, provide an understanding of the current state of the art. It was observed that, despite the advances and consistent results obtained by the Brazilian researchers, it is urgent to focus on new research approaches that look at the issues that remain open or that have not been worked in depth up to now.

Keywords: Indexing. Indexing policy. Knowledge organization. Knowledge representation. Brazil.

uma ação necessária para o bom funcionamento do sistema de recuperação de informação, produzindo uma correspondência precisa entre as ações profissionais e a qualidade dos produtos e serviços informacionais resultantes do processo.

Ao versar sobre a política de indexação, verifica-se ser esta uma temática pouco trabalhada pela literatura especializada em Ciência da Informação. No Brasil, as discussões acerca das políticas de indexação são recentes, sendo encabeçadas no contexto científico com o estudo intitulado *Diretrizes para uma política de indexação* apresentado por Marília Vidigal Carneiro no ano de 1985. A partir de então, a temática passou a ser trabalhada paulatinamente no âmbito da área de Organização e Representação do Conhecimento.

Diante disso, o presente estudo comporta uma discussão reflexiva acerca da pesquisa brasileira sobre política de indexação. Esta comunicação convida-nos a pensar sobre os avanços alcançados pela literatura especializada e, sobretudo, propicia uma compreensão do atual

cenário investigativo sobre a temática. Para isso, realizou-se um estudo exploratório de revisão bibliográfica, a fim de demonstrar o estado da arte sobre a questão.

2. Políticas de indexação

A política de indexação é considerada um contributo efetivo para o bom desempenho do sistema de recuperação de informação, atrelando questões inerentes à prática profissional no processo de indexação e questões de gestão, as quais versam para o contexto administrativo da unidade de informação enquanto instituição.

A atuação estratégica das unidades de informação está diretamente atrelada ao bom encaminhamento de uma política de indexação. Isto porque, ao atuar como um recurso auxiliar de orientação ao bibliotecário indexador, o estabelecimento de uma política de indexação não está condicionado apenas aos aspectos que perpassam o fazer técnico, mas, sobretudo, às questões organizacionais presentes nos mais variados contextos informacionais.

A função da política de indexação é definir as variáveis que influenciam o desempenho da atividade de indexação, tendo por objetivos “[...] estabelecer princípios e critérios que servirão de guia na tomada de decisões para otimização do serviço, racionalização dos processos e consistência das operações nele contidas [...]” (Carneiro, 1985, p. 222). Na ótica de Almeida (2000), as políticas ou diretrizes administrativas figuram na qualidade de planos gerais de ação que orientam na tomada de decisão ao estabelecerem guias mestras, o que confere estabilidade à instituição. Desta forma, no que tange à política de indexação, a questão administrativa da unidade de informação torna-se um aspecto indispensável a ser considerado quando do estabelecimento destas diretrizes.

Contudo, o contributo destas diretrizes não está relacionado apenas às questões organizacionais. A política de indexação atua como um recurso valioso na busca pela qualidade dos produtos e serviços informacionais disponibilizados nos sistemas de recuperação de informação, uma vez que proporciona equilíbrio entre o saber científico e a prática da profissão no que tange ao tratamento temático da informação. Consequentemente, contribui para a motivação e autonomia durante a realização do processo de indexação por propiciar segurança ao bibliotecário indexador no decorrer da atividade, contribuindo para a superação de barreiras que possam existir entre o tripé composto por teoria, prática cotidiana da profissão e ca-

racterísticas e especificidades da unidade de informação.

Rubi (2004, 2008) entende que a política de indexação é uma aliada para que o bibliotecário indexador realize sua tarefa profissional de maneira “racional e objetiva”, sendo indispensável que os profissionais bibliotecários percebam a importância da indexação em todo o ciclo documental. Para tanto, a autora advoga sobre a necessidade da política de indexação ser considerada parte administrativa da unidade de informação, como uma etapa de reflexão propriamente dita. Isto é, estas diretrizes não devem ser vistas como uma lista de procedimentos, e sim uma filosofia que reflita os interesses e objetivos da organização, o que exige do bibliotecário indexador uma efetiva compreensão de que a atividade de indexação exige parâmetros que o guiem no momento de tomada de decisão, minimizando subjetividades e incertezas.

Complementando o exposto, tem-se o fato da unidade de informação ser composta por partes interligadas que versam para um objetivo comum, o de disponibilizar a informação de modo a satisfazer as necessidades informacionais dos usuários. Esta visão indica que a indexação e sua política são uma das partes que compõem os sistemas de recuperação de informação. Com efeito, precisam integrar o planejamento global de toda e qualquer unidade de informação, como um parâmetro de sua administração no contexto gerencial.

Outro ponto importante a ser considerado é o de que a política de indexação condiciona os resultados de uma estratégia de busca por parte do usuário, o que produz uma exata correspondência com o assunto pesquisado em índices.

A política decide não só sobre a consistência dos procedimentos de indexação em relação aos efeitos que se necessita obter na recuperação, mas, principalmente, sobre a delimitação de cobertura temática em níveis qualitativos e quantitativos tendo em vista os domínios de assuntos e as demandas dos usuários. Isso nos leva a pensar sobre a indexação do ponto de vista gerencial e estratégico no contexto de unidades de informação, haja vista ter efeitos na entrada e na saída de informações do sistema (Fujita, 2012a, p. 17).

Carneiro (1985) sinaliza que no momento de se formalizar os elementos básicos para a constituição de uma política de indexação, três requisitos são indispensáveis para o bom planejamento de um sistema de recuperação de informação, quais sejam: identificação da organização à qual está relacionado o sistema de

indexação; identificação dos usuários ao qual se destina o sistema; e por fim, mas não menos importante, figuram os recursos humanos, materiais e financeiros envolvidos no processo. Rubi (2008) compreende que estes fatores ou requisitos podem ser delimitados em:

1. *Contexto*: conhecer os objetivos e atividades da instituição é fundamental para determinar o tipo de serviço a ser implantado, pois, além do tipo de atividade da organização afetar a demanda de informação, determina qual sistema de indexação a ser utilizado e os níveis de exaustividade e especificidade exigidos.
2. *Destinatário*: a principal finalidade de um sistema de recuperação de informação é o fornecimento de informação precisa aos usuários, em que identificar a clientela é um pré-requisito fundamental para o planejamento de qualquer sistema de informação. A partir de um estudo de usuários serão obtidas informações sobre o conhecimento do alcance exigido pelo sistema quanto aos assuntos centrais e periféricos e quanto aos níveis de tratamento exigidos; o núcleo de um vocabulário que refletirá os interesses do trabalho e necessidades de informação da clientela do sistema; conhecimento do tipo de resposta do sistema; e conhecimento do nível de exaustividade exigido na indexação, o grau de especificidade necessária, a linguagem de indexação e o nível de sofisticação desejável do sistema.
3. *Infraestrutura*: no qual serão considerados todos os recursos humanos, materiais e financeiros envolvidos no processo.

Atrelando o ciclo documental à política de indexação, Kobashi (1994) indica que tais diretrizes estão relacionadas às características do sistema documental: necessidades do usuário; instituição onde se desenvolve; domínio do assunto tratado; recursos humanos, físicos e financeiros; produtos e serviços; e relação custo/desempenho.

Observa-se que tais indicações apontam a política de indexação como uma ação indispensável a ser desenvolvida e considerada no âmbito da administração de toda e qualquer unidade de informação, ao passo que reflete de forma integrada e interdependente os objetivos e peculiaridades das atividades de organização, disponibilização e uso da informação (Cesarino & Pinto, 1978). Além da interdependência de seus elementos, a implementação de uma política de indexação pressupõe uma constante avaliação, visto que suas diretrizes estão diretamente relacionadas a um conjunto de fatores que são, por natureza, suscetíveis de alterações, por razões

relativas à atualização do acervo, possíveis mudanças das áreas de interesse da unidade de informação, perfis dos usuários, linguagens de indexação, dentre outros.

Sendo assim, tem-se a necessidade da política de indexação estar registrada e compor o manual de indexação, pois algumas incoerências cometidas durante o processo de tratamento temático da informação são resultantes da ausência de políticas e manuais próprios para o contexto de bibliotecas universitárias que sirvam como diretrizes para o trabalho do profissional e, sobretudo, como instrumento para a sua formação em serviço (Rubi, 2008; Dal' Evedove, 2014).

Ao sistematizar a política de indexação para um sistema de armazenamento e recuperação de informação, Carneiro (1985) evidencia dois eixos ou planos integrantes (Figura 1):

- *Eixo horizontal*: contempla atividades de gestão onde estão os requisitos para o bom planejamento de um sistema de recuperação de informação; e
- *Eixo vertical*: contempla as atividades de organização e representação da informação que reúne um conjunto de elementos e variáveis que afetam a atividade de indexação.

No eixo horizontal, entende-se que as atividades de gestão englobam toda e qualquer prática administrativa que vise ao alcance dos objetivos institucionais. Isto implica considerar a gestão de processos organizacionais resultantes da atividade humana. Por sua vez, no eixo vertical, os principais elementos a serem contemplados na política de indexação dizem respeito à cobertura de assuntos, seleção e aquisição de documentos-fonte, processo de indexação, estratégia de busca, tempo de resposta do sistema, forma de saída e avaliação do sistema. Neste eixo, as variáveis correspondem aos aspectos que interferem no processo de indexação e que influem na recuperação da informação, sendo estas: exaustividade, especificidade, escolha da linguagem, revocação e precisão (CARNEIRO, 1985).

Pode-se inferir que as duas visões de política de indexação apresentadas pela referida autora são complementares, em que se observa uma visão mais ampla,

[...] que considera o contexto da gestão de sistemas de recuperação da informação sobre as atividades de indexação e a visão mais direcionada aos procedimentos e elementos de indexação circunscritos ao manual de indexação, sua operacionalização e avaliação (Fujita, 2012a, p. 22).

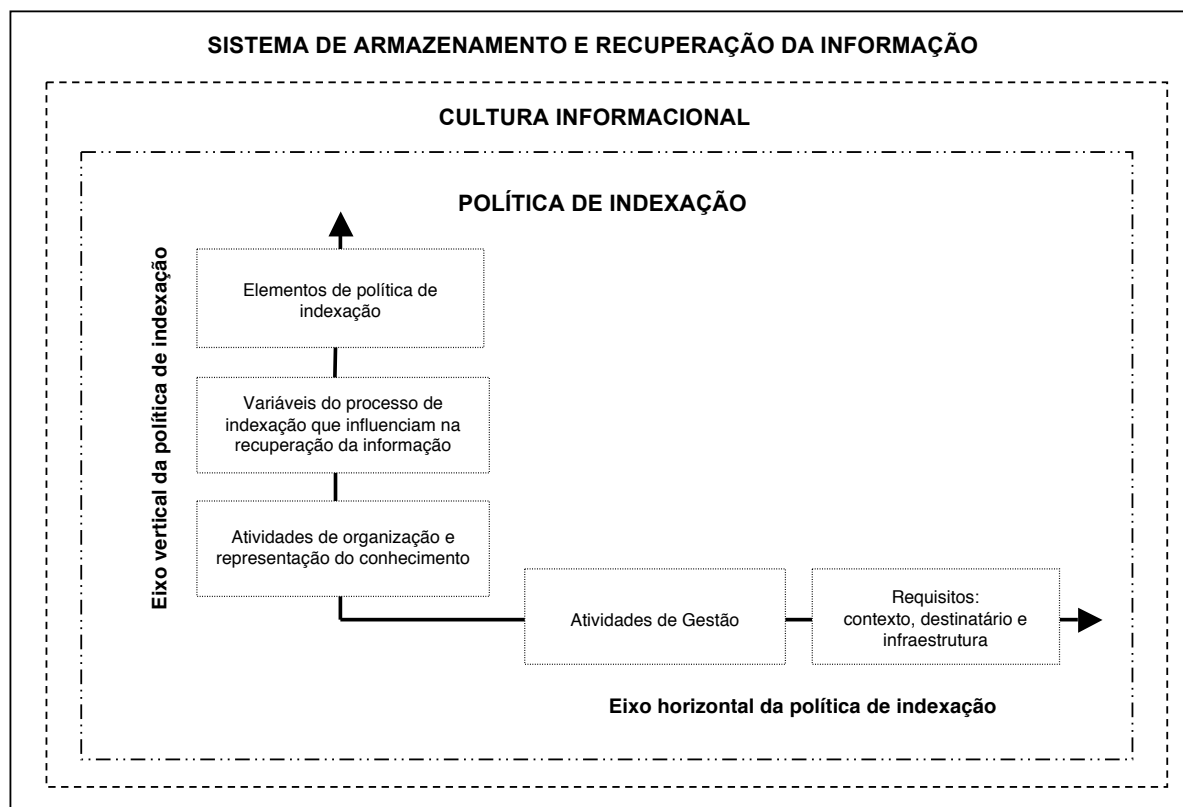


Figura 1: Eixos da política de indexação
(Fonte: Elaborado a partir de Fujita, 2012a)

Com base neste binômio inerente ao sistema de recuperação de informação que reflete uma dupla dimensão (Guimarães, 2004), a política de indexação é vista como um conjunto de procedimentos, materiais, normas e técnicas orientadas por decisões, as quais refletem os princípios subjacentes da cultura organizacional de uma unidade de informação.

3. A pesquisa brasileira na temática

Carneiro (1985) é quem projeta as primeiras diretrizes para o estabelecimento de uma política de indexação no cenário nacional, tomando-se como base os principais elementos envolvidos no planejamento de um sistema de recuperação de informação e a pouca literatura especializada que, segundo a própria autora era "bastante esparsa".

Tal cenário não mudou muito desde então e as pesquisas relacionadas ao tema ainda são limitadas, podendo-se citar em âmbito nacional os estudos de Albuquerque, Autran e Ramalho (2011); Albuquerque e Pereira (2007); Boccato (2012); Cervantes, Fujita e Rubi (2008); Dal'Evedove (2011 e 2014); Dal'Evedove e Fujita (2013); Fujita et. al. (2006a, 2006b e 2007);

Fujita (2012a, 2012b e 2012c); Fujita e Gil Leiva (2009, 2010 e 2012); Fujita e Rubi (2006a, 2006b, 2006c e 2007); Fujita, Rubi e Boccato (2009); Gonçalves, Fujita e Rubi (2006); Guimarães (2004); Kochani, Boccato e Rubi (2011); Lopes (2006); Lousada et al. (2011); Neves e Galvino (2010); Nunes (2004); Redigolo et al. (2012); Rubi (2000, 2004, 2008, 2010, 2012a e 2012b); Rubi et al. (2007); Rubi e Fujita (2003a; 2003b, 2006, 2007 e 2010); Rubi, Fujita e Boccato (2012a e 2012b) e Silva e Santos (2012), que em sua grande maioria fazem menção ao trabalho de Carneiro (1985) e buscam a sua maneira e dentro dos seus escopos investigativos contribuir com referenciais que sirvam de orientação para a definição de diretrizes que norteiem a atividade de indexação.

3.1. Caminhos percorridos

Debruçando-se na literatura brasileira especializada sobre política de indexação, observa-se que, apesar de não serem numerosos, os estudos dedicados ao tema apresentam resultados bastante significativos, sendo em sua maioria verticalizados no assunto.

Com base nos estudos de Carneiro (1985) e trabalhos subsequentes (Rubi, 2000, 2004; Rubi, Fujita, 2003 e Guimarães, 2004) sobre política de indexação, Rubi (2008) compila os principais elementos indicados para a elaboração destas diretrizes visando a sistematizar o conhecimento acerca do tema. Para tanto, apresenta uma proposta atualizada dos princípios de orientação para a elaboração de políticas de indexação. Neste estudo é apresentado à comunidade da área de Organização e Representação do Conhecimento um agrupamento dos elementos base a serem considerados na elaboração das políticas de indexação, no qual alia os elementos as suas fontes literárias (trabalhos em que tais indicações foram observadas).

O fruto deste estudo é uma proposta de diretrizes apresentada pela autora para a elaboração de política de indexação destinada à construção de catálogos cooperativos em bibliotecas universitárias, a partir da abordagem sociocognitiva com bibliotecários e usuários inseridos no contexto e da perspectiva conceitual dos aspectos teóricos que envolvem a indexação e a catalogação de assunto. O contributo desta compilação advém da indicação de um conjunto de elementos que servem de orientação para a elaboração de diretrizes de política de indexação destinadas à atividade de indexação em bibliotecas universitárias.

Como forma de avançar na questão, Dal'Evedove (2014) recomenda a observância dos aspectos cognitivos, culturais e sociais inerentes ao processo de tratamento temático da informação no delineamento das políticas de indexação. Neste estudo, foram apresentadas oito diretrizes complementares a serem consideradas no momento da elaboração da política de indexação para o contexto de bibliotecas universitárias brasileiras.

Traçando uma perspectiva inovadora, o estudo voltou-se para a perspectiva humana do processo de tratamento temático da informação a qual perpassa as questões relativas à política de indexação. O estudo apresentou novos subsídios a serem considerados no momento da elaboração de políticas de indexação para o contexto de bibliotecas universitárias, contribuindo sumariamente para tornar a política de indexação um recurso passível de uso por parte do profissional bibliotecário. Isto porque, foram consideradas as indicações e concepções teóricas; as especificidades da prática cotidiana da profissão relacionadas à atividade de indexação; e da realidade própria das bibliotecas universitárias, especificamente no que tange aos interesses informacionais dos usuários. O diferencial frente a outros estudos que tiveram

como pano de fundo a política de indexação assenta-se na observação dos modos sociais de pensamento dos atores testados, articulando suas opiniões, impressões, comportamentos e interesses pessoais e coletivos na busca por diretrizes que atendessem: os fundamentos teóricos e metodológicos presentes na área de Organização e Representação do Conhecimento; o contexto sociocognitivo do bibliotecário indexador; bem como as características e princípios culturais dos usuários (Dal' Evedove, 2014).

Tendo como preocupação central os aspectos organizacionais, Louzada et. al. (2011) voltaram-se para as relações existentes entre as políticas de indexação e a gestão do conhecimento em contextos organizacionais, como subsídio para a implementação de ações estratégicas. Como resultado, as autoras apontam as principais contribuições da gestão do conhecimento organizacional para as políticas de indexação, com vistas à construção do conhecimento. Ainda neste eixo investigativo, Rubi, Fujita e Boccato (2012) com a proposta de analisar os aspectos teóricos sobre política de indexação, sob a perspectiva do conhecimento organizacional, atestam a viabilidade da política de indexação ser elaborada a partir do conhecimento tácito do bibliotecário indexador produzido no contexto organizacional. Para tanto, o estudo é conduzida por meio da metodologia do Protocolo Verbal em Grupo (PVG), transformando o conhecimento tácito em conhecimento explícito na forma de manual de indexação.

Outro estudo bastante significativo na área foi o apresentado por Cervantes, Fujita e Rubi (2008) no qual são compiladas as terminologias relacionadas à política de indexação, mediante um estudo terminológico. Tendo um direcionamento mais pedagógico, Fujita e Gil Leiva (2012) oferecem uma rica contribuição à área de Organização e Representação do Conhecimento a partir da organização e publicação do livro intitulado "Política de Indexação" que aborda aspectos teóricos, metodológicos e pedagógicos sobre a temática.

3.2. Questões em aberto

Grande parte dos estudos desenvolvidos no âmbito nacional se restringe aos aspectos organizacionais, estruturais, metodológicos e pedagógicos que envolvem a política de indexação, comumente na perspectiva profissional e de usuários. Os elementos destinados à elaboração de políticas de indexação indicados pela literatura especializada são considerados princípios norteadores, aspecto que acentua a ne-

cessidade de estudos mais verticalizados sobre a questão.

Esforços estão sendo empreendidos para que o conhecimento sobre política de indexação na área de Organização e Representação do Conhecimento ganhe um delineamento sistematizado e, conseqüentemente, contribua para a inclusão de novos elementos que se mostrarem relevantes neste conjunto. Destaque especial seja dado às pesquisas conduzidas pelo Grupo de Pesquisa “Análise Documentária”, no qual a maior parte da literatura científica do Brasil acerca da política de indexação foi desenvolvida. Num primeiro momento, as pesquisas buscaram contribuir com fundamentação teórica e metodológica, baseada no conhecimento de Ciência da Informação e áreas interdisciplinares, de modo a sistematizar uma proposta teórica e metodológica de elaboração de política de indexação com e para bibliotecas universitárias. Atualmente, os esforços são direcionados para as questões relativas ao processo de elaboração da estrutura lógica e hierárquica de linguagens de indexação na perspectiva da política de indexação.

Apesar de a literatura especializada discorrer sobre a importância da adoção de tais diretrizes, nas quais são explicitadas as escolhas técnico-administrativas da unidade de informação pelas quais os bibliotecários indexadores precisam se orientar em suas rotinas profissionais, considerando como variáveis o usuário e o acervo, Fujita (2012) alerta sobre o desafio de aliá-las à prática cotidiana da profissão. Para a referida autora, esta associação é necessária para que haja conscientização sobre os elementos e as variáveis da política de indexação e os efeitos que estes causam na recuperação da informação. Logo, a política de indexação mostra-se imprescindível do ponto de vista gerencial e estratégico na medida em que estão subjacentes em suas diretrizes os princípios e critérios que servirão de guia na tomada de decisões, visando otimização do serviço e racionalização de processos.

Dessa forma, urge o encaminhamento de pesquisas que busquem atrelar teoria e prática profissional, considerando as questões inerentes e congruentes que envolvem estas duas esferas, ao passo que os resultados de tais investigações tendem a minimizar a barreira que separa as concepções teóricas e a prática cotidiana da profissão, propiciando o delineamento de políticas de indexação cada vez mais coerentes e passíveis de aplicação. Para tanto, entende-se que tais pesquisas precisam avançar no exame das questões de natureza cognitiva, cultural e social que permeiam a atividade

de indexação, mediante a observação das particularidades advindas da vertente humana que compõe a representação temática da informação nos mais variados contextos informacionais.

Na atual conjuntura, a elaboração de políticas de indexação deve considerar os fundamentos teóricos e metodológicos da área de Organização e Representação do Conhecimento, a prática cotidiana da profissão e, inevitavelmente, prever a realidade específica da unidade de informação para a qual se destinam. Esta indicação demanda, por exemplo, o encaminhamento de pesquisas que observem aspectos de natureza ética presentes na atividade de indexação, até então pouco trabalhados nos elementos de política de indexação.

Este posicionamento é sustentado, dentre outros, com a perspectiva de Rubi (2008) de que a indicação de diretrizes ou a própria elaboração de políticas de indexação deve ser representada por uma filosofia que traduza os objetivos da unidade de informação e sirva de guia para os profissionais na realização da atividade de indexação, auxiliando-os para a tomada de decisões. Adicionalmente, considera-se que qualquer instrumento de trabalho que sirva de orientação para a prática profissional necessita refletir as especificidades e particularidades do ambiente de trabalho e os interesses da comunidade usuária a qual se destina.

4. Algumas Considerações

Abordou-se neste estudo a política de indexação e o atual cenário das pesquisas brasileiras sobre a temática. Mediante uma exposição pontual sobre os avanços obtidos, tornou-se possível discorrer sobre os principais aspectos considerados pelos estudos e os desafios a serem vencidos para a verticalização da política de indexação no âmbito teórico-conceitual, com vistas à formulação de diretrizes de cunho prático-aplicadas nos mais variados contextos de informação.

É possível apontar que a política de indexação, na condição de objeto de estudo, apresenta uma sistematização teórica e metodológica consistente, fato que favorece o avanço da pesquisa brasileira na área de Organização e Representação do Conhecimento, uma vez que as bases investigativas sobre a temática foram construídas. Todavia, é preciso atender aos vários questionamentos secundários importantes que permeiam as políticas de indexação, a fim de proporcionar acréscimos às discussões presentes no universo teórico-conceitual e, sobretudo, contribuir com novos olhares e formas

de se pensar e elaborar diretrizes de política de indexação no âmbito prático-aplicado, considerando-se as características e particularidades dos diferentes contextos informacionais.

Diante das exposições realizadas no presente estudo, observa-se que a política de indexação configura-se um material de trabalho que demanda conhecimentos que perpassam o processo de indexação, cuja natureza complexa exige o encaminhamento de pesquisas contínuas. Nesse caminhar, registra-se a importância de novos olhares reflexivos sobre as políticas de indexação que considerem as problemáticas que envolvem a organização e representação temática da informação, bem como os fatores relacionados à gestão organizacional nas unidades de informação.

Referências

- Albuquerque, M. E. B. C.; Autran, M. M. M.; Ramalho, F. A. (2011). Política de indexação da Biblioteca Digital Paulo Freire. // Brennard, Edna Maria de Góes; Albuquerque, Maria Elizabeth Baltar Carneiro de (Org.). Paulo Freire: diálogos e redes digitais. João Pessoa: Editora universitária UFPB, 2011. 67-77.
- Albuquerque, M. E. B. C.; Pereira, A. R. (2007). Representação da informação de imagens paradas do Jornal da Paraíba: política de indexação. // Emeide Nóbrega Duarte; Alzira Karla Araújo da Silva. (Org.). Gestão de unidades de informação: teoria & prática. João Pessoa: Universitária UFPB, 2007. 499-512.
- Bocato, V. R. C. (2012). A linguagem documentária em catálogos on-line para política de indexação. // Gil Leiva, Isidoro; Fujita, Mariângela Spotti Lopes (Org.). Política de indexação. São Paulo: Cultura Acadêmica, 2012. 139-151.
- Carneiro, M. V. (1985). Diretrizes para uma política de indexação. // Revista da Escola de Biblioteconomia da UFMG, Belo Horizonte. 14:2 (set. 1985) 221-241.
- Cervantes, B. M. N.; Fujita, M. S. L.; Rubi, M. P. (2008). Terminologias em política de indexação. // García Marco, Francisco Javier (Org.). Avances y perspectivas en sistemas de información y documentación. Zaragoza: Universidad de Zaragoza, 2008. 1, 211-222.
- Cesarino, M. A. N.; Pinto, M. C. M. (1978). Cabeçalho de assunto como linguagem de Indexação. // Revista da Escola de Biblioteconomia da UFMG, Belo Horizonte. 7:2 (set. 1978) 268-288.
- Dal'Evedove, P. R. (2011). A Política de Tratamento da Informação na Percepção de Catalogadores de Assunto. // XXIV Congresso Brasileiro de Biblioteconomia e Documentação – CBBB. Anais: Sistemas de informação, multiculturalidade e inclusão social. Brasília: FEBAB, 2011.
- Dal'Evedove, P. R.; Fujita, M. S. L. (2013). O conhecimento profissional do catalogador de assunto sobre política de indexação em bibliotecas universitárias. // Revista Digital de Biblioteconomia e Ciência da Informação. 11 (Maio/Agosto 2013) 21-39.
- Dal'Evedove, P. R. (2014). O tratamento temático da informação em abordagem sociocultural: diretrizes de política de indexação em bibliotecas universitárias. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília. 266f.
- Fujita, M. S. L. et. al. (2007). Política de tratamento da informação documental na rede de bibliotecas universitárias da UNESP na perspectiva da Biblioteca Virtual. // XXII Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação. Anais do XXII Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação. Brasília, 2007.
- Fujita, M. S. L. (2012a). A política de indexação para representação e recuperação da informação. // Isidoro Gil Leiva; Mariângela Spotti Lopes Fujita (Org.). Política de indexação. São Paulo; Marília: Cultura Acadêmica; Oficina Universitária, 2012a. 17-30.
- Fujita, M. S. L. (2012b). Elaboração e avaliação da política de indexação na formação inicial do indexador. // Isidoro Gil Leiva; Mariângela Spotti Lopes Fujita (Org.). Política de indexação. 1ed. São Paulo; Marília: Cultura Acadêmica; Oficina Universitária, 2012b. 187-216.
- Fujita, M. S. L. et. al. (2006a). A política de tratamento de informações da rede de bibliotecas UNESP na perspectiva do catalogador: diagnóstico com uso de protocolo verbal individual. // XIV Seminário Nacional de Bibliotecas Universitárias. Anais do XIV Seminário Nacional de Bibliotecas Universitárias. Salvador, 2006a.
- Fujita, M. S. L. et. al. (2006b). O conhecimento organizacional sobre política de tratamento da informação documental da rede de bibliotecas da UNESP: o uso do protocolo verbal em grupo/leitura como evento social como abordagem qualitativa. // XIV Seminário Nacional de Bibliotecas Universitárias. Anais do XIV Seminário Nacional de Bibliotecas Universitárias. Salvador, 2006b.
- Fujita, M. S. L.; Gil Leiva, I. (2010). As linguagens de indexação em bibliotecas nacionais, arquivos nacionais e sistemas de informação na América Latina. // Seminário Nacional de Bibliotecas Universitárias. Onde estamos, para aonde vamos. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2010. 1-13.
- Fujita, M. S. L.; Gil Leiva, I. (2009). Política de indexação em latinoamérica. // García Marco, F. J. (Org.). Avances y perspectivas en sistemas de información y documentación en el entorno digital. Zaragoza: Universidad de Zaragoza. 2009. 1, 155-162.
- Fujita, M. S. L.; Gil Leiva, I. (2012). Política de indexação latino-americana. // Gil Leiva, Isidoro; Fujita, Mariângela Spotti Lopes (Org.). Política de indexação. São Paulo; Marília: Cultura Acadêmica; Oficina Universitária. 2012. 121-138.
- Fujita, M. S. L. (2012c). Perspectivas da política de indexação para a formação e atuação profissional de bibliotecários. // Gil Leiva, Isidoro; Fujita, Mariângela Spotti Lopes (Org.). Política de indexação. São Paulo; Marília: Cultura Acadêmica; Oficina Universitária. 2012c. 231-238.
- Fujita, M. S. L.; Rubi, M. P. (2006a). A política de indexação na perspectiva do conhecimento organizacional. // VII Encuentro asociación de educadores e investigadores de bibliotecología, archivología, ciencias de la información y documentación de Iberoamérica y el Caribe. Anais do VII Encuentro asociación de educadores e investigadores de bibliotecología, archivología, ciencias de la información y documentación de Iberoamérica y el Caribe. Marília: FFC/UNESP-PUBLICAÇÕES, 2006a. 638-654.
- Fujita, M. S. L.; Rubi, M. P.; Bocato, V. R. C. (2009). O contexto sociocognitivo do catalogador em bibliotecas universitárias: perspectivas para uma política de tratamento da informação documental. // Datagramazero. 10:2 (Abril 2009).
- Fujita, M. S. L.; Rubi, M. P. (2006b). Educação à distância para formação do bibliotecário de bibliotecas universitárias do consórcio CRUESP em política de indexação: perspectivas de conteúdo e aplicação de protocolo verbal em grupo. // VII Encuentro asociación de educadores

- e investigadores de bibliotecología, archivología, ciencias de la información y documentación de Iberoamérica y el Caribe. Anais do VII Encontro associação de educadores e investigadores de bibliotecología, archivología, ciencias de la información y documentación de Iberoamérica y el Caribe. Marília: FFC/UNESP-Publicações, 2006b. 176-187.
- Fujita, M. S. L.; Rubi, M. P. (2006c). Educação à distância para formação do bibliotecário de bibliotecas universitárias do consórcio CRUESP em Política de indexação: perspectivas de conteúdo e aplicação de protocolo verbal em grupo. // Seminário Nacional de Bibliotecas Universitárias. Anais do XIV Seminário Nacional de Bibliotecas Universitárias. Salvador, 2006c.
- Fujita, M. S. L.; Rubi, M. P. (2007). Protocolo verbal como metodologia sociocognitiva para coleta de dados e recurso pedagógico em sala de aula. // Machado, L. M.; Maia, G. Z. A.; Labegalini, A. C. F. B. (Org.). Pesquisa em Educação: passo a passo. Marília: Ed. M3T, 2007. 143-154.
- Gil Leiva, I.; Fujita, M. S. L. (Org.). (2012). Política de indexação. São Paulo: Cultura Acadêmica; Oficina Universitária. 2012.
- Gil Leiva, I. (2008). Manual de indexación: teoría y práctica. Gijón: Trea, 2008.
- Gonçalves, M. C.; Fujita, M. S. L.; Rubi, M. P. (2006). Política de indexação em sistemas de bibliotecas universitárias: levantamento de subsídios para o tratamento temático do acervo bibliográfico da UNESP. // XIV Seminário Nacional de Bibliotecas Universitárias. Anais do XIV Seminário Nacional de Bibliotecas Universitárias. Salvador, 2006.
- Guimarães, J. A. C. (2004). As políticas de indexação como elementos para a gestão do conhecimento nas organizações. // Vidotti, S. A. B. G. (Org.). Tecnologia e conteúdos informacionais. São Paulo: Polis, 2004. 43-52.
- Kobashi, N. Y. (1994). A elaboração de informações documentárias: em busca de uma metodologia. Tese (Doutorado em Ciências da Comunicação) - Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo. 195f.
- Kochani, A. P.; Boccato, V. R. C.; Rubi, M. P. (2011). Política de indexação para sistemas automatizados de coordenadorias de comunicação em ambientes universitários. // XXIV Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação. Maceió. Anais do XXIV Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação, 2011.
- Lopes, I. L. (2006). Diretrizes para uma política de indexação de fotografias. // Antonio Miranda; Elmira Simeão. (Org.). Alfabetização digital e acesso ao conhecimento. Brasília: UnB/CID, 2006.
- Lousada, M. et. al. Políticas de indexação no âmbito da gestão do conhecimento organizacional. // Informação & sociedade. 21:1 (2011) 191-202.
- Neves, D. A. B.; Galvino, C. C. T. (2010). Avaliação de política de indexação da Seção de Periódicos da Biblioteca Central da UFPB: pesquisa em andamento. // XI ENANCIB. Anais do XI ENANCIB. Rio de Janeiro: IBICT, 2010. 135-141.
- Nunes, C. O. (2004). Algumas considerações acerca da ausência de políticas de indexação em bibliotecas brasileiras. // Biblos. Rio Grande do Sul. 16, 55-61.
- Redígolo, F. M. et. al. (2012). Elementos de política de indexación para bibliotecas universitarias del área médica. // Scire (Zaragoza). 18, 75-86 (2012).
- Rubi, M. P. (2004). A política de indexação na perspectiva do conhecimento organizacional. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília. 135 f.
- Rubi, M. P. et. al. (2007). Política del tratamiento de la información documentaria en bibliotecas universitarias: estudio diagnóstico del contexto en la perspectiva del catalogador y del usuario. // Francisco Javier García Marco. (Org.). Revista de sistemas de información y documentación. Zaragoza: Ibersid, 2007. 71-80.
- Rubi, M. P.; Fujita, M. S. L.; Boccato, V. R. C. (2012a). Del conocimiento tácito al explícito: elaboración de un manual de política de indexación en bibliotecas universitarias. // Scire (Zaragoza). 18:2 (2012a), 67-74.
- Rubi, M. P.; Fujita, M. S. L.; Boccato, V. R. C. (2012b). Elaboração do manual de política de indexação na formação continuada do catalogador. // Gil Leiva, Isidoro; Fujita, Mariângela Spotti Lopes (Org.). Política de indexação. São Paulo: Cultura Acadêmica, 2012b. 217-227.
- Rubi, M. P.; Fujita, M. S. L. (2003a). Elementos de política de indexação em manuais de indexação de sistemas de. // Perspectivas em Ciência da Informação, Belo Horizonte. 8:8 (2003a) 57-66.
- Rubi, M. P.; Fujita, M. S. L. (2007). La política de indexación en la perspectiva del conocimiento organizacional. // Blanca Rodríguez Bravo; Maria Luisa Alvite Diez. (Org.). La interdisciplinariedad y la transdisciplinariedad en la organización del conocimiento científico. León: Universidad de León, 2007, 1, 451-458.
- Rubi, M. P.; Fujita, M. S. L. (2006). O ensino de procedimentos de política de indexação. // Perspectivas em Ciência da Informação, Belo Horizonte. 11:1 (2006) 48-66.
- Rubi, M. P.; Fujita, M. S. L. (2010). Política de indexação na catalogação de assunto em bibliotecas universitárias: a visão sociocognitiva da atuação profissional com protocolo verbal. // Revista Digital de Biblioteconomia e Ciência da Informação. 7 (2010) 168-200.
- Rubi, M. P.; Fujita, M. S. L. (2003b). Política de indexação na visão dos indexadores. // V Encontro Nacional de Pesquisa em Ciência da Informação. Anais do V Encontro Nacional de Pesquisa em Ciência da Informação. Belo Horizonte: ENANCIB, 2003b.
- Rubi, M. P. (2000). Os serviços de análise e sua política de indexação. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília. 80f.
- RUBI, M. P. (2012a). Política de indexação. // Gil Leiva, Isidoro; Fujita, Mariângela Spotti Lopes (Org.). Política de indexação. São Paulo: Cultura Acadêmica, 2012a. 107-120.
- Rubi, M. P. (2009). A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias. // Fujita, Mariângela Spotti Lopes (Org.). Os princípios da política de indexação na análise de assunto para catalogação: especificidade, exaustividade, revocação e precisão na perspectiva dos catalogadores e usuários. São Paulo: Cultura Acadêmica, 2009. 81-94.
- Rubi, M. P. (2008). Política de indexação para construção de catálogos coletivos em bibliotecas universitárias. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília. 169f.
- Rubi, M. P. Proposta para implantação de política de indexação em bibliotecas. // Gil Leiva, Isidoro; Fujita, Mariângela Spotti Lopes (Org.). Política de indexação. São Paulo: Cultura Acadêmica, 2012b. 171-183.
- Silva, L. C. de O.; Santos, C. A. C. M. (2012). Proposta metodológica para o estabelecimento de política de indexação para a biblioteca escolar. // CRB-8 Digital, 5 (2012) 173-182.

Enviado: 2015-04-01. Segunda versão: 2015-06-09.
Aceptado: 2015-06-16.

Edición de contenidos en un entorno colaborativo: el caso de la Wikipedia en español

Editing content in a collaborative environment: the case of the Spanish Wikipedia

Ángel F. ZAZO RODRÍGUEZ, Carlos G. FIGUEROLA, José Luis ALONSO BERROCAL

Instituto Universitario de Estudios de la Ciencia y la Tecnología, Edificio I+D+i - Universidad de Salamanca
{angelzazo | figue | berrocal}@usal.es

Resumen

Se analizan las características y la actividad que los usuarios editores de la Wikipedia en español realizan en el proceso de creación de contenidos. Tras volcar los datos de los artículos enciclopédicos, se han analizado aspectos cuantitativos de los artículos, como su longitud, enlaces entrantes y salientes entre ellos, y categorías a las que pueden ser asignados. En algunos casos, esas características siguen patrones similares a los encontrados en estudios webmétricos. El sistema de categorías, pese a que funcionalmente está bien constituido, no se utiliza de manera adecuada por los usuarios editores, lo cual menoscaba una buena forma de acceso al conocimiento. En cuanto a las ediciones que realizan los usuarios, se han obtenido patrones de actividad relacionados con la creación de artículos, la revisión de contenidos, días de actividad, reversiones, vandalismo y país de origen. Una parte importante del funcionamiento de Wikipedia recae en unos pocos usuarios que supervisan los nuevos contenidos, ayudados por robots que facilitan el proceso. En general, la creación de contenidos se lleva a cabo por dos grupos diferentes de usuarios: pequeñas contribuciones individuales de una gran legión de usuarios, y un gran número de contribuciones que realizan un reducido grupo de usuarios muy activos. Se ha obtenido el país de origen de muchos usuarios, lo cual ha permitido contabilizar las contribuciones realizadas desde cada uno de ellos.

Palabras clave: Wikipedia. Edición colaborativa. Estudios de usuarios. Organización del conocimiento.

1. Introducción

Wikipedia se autodefine como una enciclopedia libre, políglota y editada colaborativamente. En poco tiempo ha cosechado un enorme éxito, por lo que ha sido estudiada desde muy diferentes puntos de vista. Son muchos los artículos académicos y de investigación que la tienen como objeto de estudio, existiendo incluso portales web dedicados a recopilar la literatura gris que sobre ella se genera, como por ejemplo, Wikilit (wikilit.referata.com), WikiPapers (wikipapers.referata.com) o AcaWiki (akawiki.org). La propia Fundación Wikimedia dispone de servicios para la difusión de investigaciones, como el *Wikime-*

Abstract

This work uses the database backup dumps that collect content and history reviews of the encyclopaedic articles of Spanish Wikipedia since its creation, in order to characterize and understand the underlying activity of the editors in content creation. Some quantitative characteristics of articles are analyzed: length, assigned categories and in-links and out-links to other articles. Some characteristics have similar patterns to the ones found in webometric studies. The categories system, even though is functionally well built, is not used properly by the editors, which undermines the access to knowledge. We have also obtained patterns of the editors' activity related to article creation, content reviewing, activity days, reversiones, vandalism, and authors' countries of origin. We have found that an important part of Wikipedia lies on a few number of users who oversee the new content, aided by robots that facilitate the process. In general, content creation is performed by two different kinds of users: small individual contributions of a great legion of users and a large number of contributions made by a small group of extremely active users. For many users we have obtained the origin country, which has allowed us to know the contributions procedence.

Keywords: Wikipedia. Collaborative edition. User studies. Knowledge organization.

dia Research Newsletter. La Wikipedia ha sido muy estudiada desde el punto de vista de su contenido: amplitud, tamaño, evolución, calidad, actualidad, fiabilidad, etc. (Hu et al., 2007; Luyt et al., 2008; Magnus, 2009; Lewandowski y Spree, 2011). Otras veces se ha analizado como escenario colaborativo para la edición: participación, motivación, reputación, calidad de las contribuciones, redes de autoría, vandalismo, etc. (Wu et al. 2010; Leskovec et al. 2010). También como fuente de conocimiento para diversos entornos, como salud, educación, deporte, noticias, etc. (véase http://en.wikipedia.org/wiki/Wikipedia:Wikipedia_as_an_academic_source). Como gran corpus documental de ac-

ceso abierto que es, existen también muchos trabajos que la utilizan como fuente de datos en campos diversos, como la recuperación de información, el procesamiento de lenguaje natural, la construcción de ontologías y tesauros, etc. (Gabrilovich y Markovitch, 2009; Müller y Gurevych, 2009; Medelyan et al., 2009). Una buena recopilación de artículos sobre Wikipedia puede encontrarse en (Mesgari et al., 2015).

La inmediatez que ha supuesto Internet ha cambiado la forma en la que las personas intercambian conocimiento. Wikipedia es un ejemplo claro de este cambio. Al igual que *L'Encyclopédie* de Diderot y D'alembert supuso una revolución para su tiempo, al pretender recoger todo el conocimiento humano en una obra accesible a gran cantidad de personas, Wikipedia supone un ejemplo claro de colaboración voluntaria, que es consultada a diario por millones de personas.

La importancia de los procesos de organización y representación del conocimiento como intermediarios entre las personas y las fuentes de información, teniendo presente que la necesidad humana de informarse es consustancial al individuo, es lo que nos ha motivado a estudiar la Wikipedia desde el punto de vista de la creación de los contenidos que los usuarios de manera voluntaria realizan en esta enciclopedia.

Nuestro objetivo al estudiar la Wikipedia ha sido analizar de manera sistemática el proceso de creación y revisión de los artículos enciclopédicos. Se analizarán desde diferentes puntos de vista, desde su creación, su historial de revisión, los tipos de usuarios que editan y su origen, el uso de categorías y enlaces entre artículos (verdaderos elementos para representar el contenido), los procesos de reversión y vandalismo, etc. Previamente es importante conocer cómo Wikipedia organiza los contenidos y cómo los usuarios pueden acceder a ellos.

2. Funcionamiento de Wikipedia

La Wikipedia en español es la cuarta por número de páginas (páginas de todo tipo), es la décima por número de artículos enciclopédicos, y es la segunda en ratio de número de visitas por hora (véase <http://wikistats.wmflabs.org/>).

Wikipedia se basa en cinco pilares:

1. Solamente se aceptan contenidos enciclopédicos de carácter general, no se aceptan investigaciones originales ni ensayos, ni opiniones personales.
2. Se busca siempre el punto de vista neutral. Para ello deben ofrecerse todos los puntos de vista posibles, incluyendo fuentes autori-

zadas verificables. Las controversias deben solventarse en la página de discusión que poseen todos los artículos.

3. Cualquier persona puede incorporar, corregir o utilizar la información. El contenido está bajo licencia *Creative Commons* BY-SA 3.0.
4. Se exige respeto mutuo entre usuarios.
5. No hay otras normas aparte de las indicadas. Para colaborar en el proyecto solo es necesario guiarse por el sentido común, respetando en todo momento al resto de personas que en él colaboran.

El elemento fundamental de la Wikipedia es el artículo enciclopédico. Es el que recoge el conocimiento que buscan los lectores cuando consultan la Wikipedia. Todo artículo enciclopédico posee cinco elementos: el título, el contenido del artículo, la página de discusión en la que los usuarios editores realizan comentarios sobre el contenido, el apartado de edición que permite cambiar el contenido, y el historial de revisiones previas que hasta la fecha ha tenido el artículo.

Al igual que otras enciclopedias, Wikipedia incluye entradas que redirigen a otros artículos enciclopédicos. Tal es el caso de muchas siglas, acrónimos, seudónimos de personas y personajes, variantes léxicas del español y regionalismos, transliteraciones de nombres en otros idiomas, plurales, etc. A veces se utilizan para facilitar la navegación a usuarios que no puedan introducir determinados caracteres, como tildes, eñes, signos de puntuación invertidos o diéresis (por ejemplo, *Espana* redirige a *España*).

En el contenido de los artículos enciclopédicos se pueden incluir enlaces a otros artículos que permitan ampliar la información sobre dichos contenidos. Asimismo, cada artículo debe estar etiquetado en una o varias categorías. Su objetivo es poder encontrar grupos de artículos relacionados y navegar por ellos, ayudando a los lectores a conocer qué artículos existen sobre un determinado tema. Las categorías tienen a su vez subcategorías más específicas y supercategorías más generales, permitiendo navegar de lo general a lo concreto y viceversa. Al igual que los artículos, las categorías pueden ser editadas libremente. Existen categorías ocultas de mantenimiento que no se muestran a los usuarios.

Tanto los enlaces entre artículos como las categorías constituyen elementos de representación que permiten acceder a otros artículos que guardan relación con los contenidos (Pehcevski et al., 2010). Es importante, por consiguiente, estudiar cómo son utilizados por los usuarios editores. Hemos analizado la distribución del

número de artículos por categorías y del número de categorías por artículo. También hemos analizado la distribución del número de enlaces entrantes y salientes por artículo. En ambos casos el objetivo es determinar qué patrones de actuación siguen los editores al utilizar categorías y enlaces como mecanismos de acceso a otros contenidos, importantes desde el punto de vista del acceso a la información.

El segundo elemento importante de Wikipedia es el usuario editor (Sepehri et al., 2012). Los artículos pueden ser editados prácticamente por cualquier usuario, a excepción de algunos artículos protegidos contra el vandalismo y las guerras de ediciones. Todas las ediciones quedan identificadas por el usuario que las realiza, bien con el nombre de los usuarios registrados, bien con la dirección IP del ordenador desde donde editan los no registrados. Wikipedia no ofrece datos de los usuarios registrados, salvo la que los propios usuarios pueden incluir en sus páginas de usuario, si desean crearlas.

Hay varios tipos de usuarios, caracterizados por las diferentes acciones que pueden llevar a cabo. La página http://es.wikipedia.org/wiki/Wikipedia:Tipos_de_usuarios explica de manera detallada los diferentes tipos. Desde el punto de vista de la edición de los artículos, en nuestro análisis hemos utilizado cuatro categorías:

1. *Usuarios no registrados*: Son aquéllos que no disponen de cuenta o no inician sesión, se identifican por una dirección IP y pueden crear o editar páginas no protegidas.
2. *Usuarios registrados*: Aquéllos que inician sesión y pueden crear o editar páginas. Son los usuarios nuevos y los usuarios confirmados y autoconfirmados. Estos últimos son usuarios con una experiencia de al menos 50 ediciones, que pueden editar páginas semi-protegidas y trasladar o renombrar páginas.
3. *Usuarios administradores*: Hemos reunido bajo esta categoría a varios tipos de usuarios. En general son aquellos que pueden realizar acciones de administración de páginas o de usuarios, como eliminar o proteger páginas, editar páginas protegidas, verificar contenidos, revertir ediciones, bloquear o cambiar derechos a usuarios, etc.
4. *Usuarios bots*: Este último grupo lo forman programas automáticos (robots) que recorren los artículos de la Wikipedia realizando diferentes acciones, como revisiones ortográficas, verificación de enlaces externos e *interwikis*, asistir en la desambiguación, verificación y traslado de categorías, marcado o reversión automática del vandalismo, etc.

Existen otros elementos en la Wikipedia, como páginas sobre convenciones y políticas de uso, páginas de usuario, páginas de plantillas, páginas de categorías, páginas de anexos, y también *páginas de discusión* de los diferentes espacios de páginas. Aproximadamente la mitad de las páginas de la Wikipedia se corresponden con este tipo de páginas, en ellas se dirimen diversas cuestiones y sirven de medio de comunicación entre usuarios. En este sentido son especialmente importantes las páginas de discusión de los artículos enciclopédicos, donde se solventan las posibles controversias sobre su contenido y organización (Wu et al. 2011).

Cuando se crea una página nueva o se modifica el contenido de alguna existente se ponen en marcha una serie de mecanismos que tienen el propósito de garantizar, de acuerdo a los cinco pilares de la Wikipedia, tanto la calidad de los contenidos como la uniformidad en su presentación. Tales mecanismos podemos englobarlos dentro de procesos de organización del conocimiento (Fernandez-Molina y Guimarães, 2002).

Las páginas nuevas aparecen en la página especial de *PáginasNuevas*, con un indicador para determinar si la página ha sido verificada. Los artículos nuevos creados por usuarios que cuentan con el permiso de verificado o autoverificado aparecerán, por defecto, como verificados. Ello permite a los usuarios verificadores centrarse en aquellos artículos donde el riesgo de violación de las políticas es mayor. Los usuarios verificadores pueden dar paso a las nuevas páginas, marcarlas para su borrado, o asignarles una de las plantillas de seguimiento.

Cuando se modifica el contenido de un artículo los cambios aparecen en la página especial de *CambiosRecientes*. Cualquier usuario de Wikipedia puede patrullar los cambios recientes para detectar ediciones dañinas, esto es, ediciones que no siguen las reglas de la Wikipedia sobre el contenido, a menudo errores sin malicia de usuarios nuevos; otras veces son acciones deliberadas, entre las que se incluyen las ediciones vandálicas y las guerras de ediciones (reversiones mutuas entre uno o varios usuarios motivadas por el contenido del artículo).

Los usuarios reversores (entre los que se incluyen varios *bots*) son aquellos que pueden revertir ediciones rápidamente a una versión previa cuando detectan vandalismo o ediciones que no siguen los cinco pilares. Una gran parte de las ediciones de los administradores y de los bots tienen que ver con las acciones de verificación y reversión de ediciones. Son de especial interés las reversiones motivadas por vandalismo, pues es habitual su análisis para determinar la cali-

dad de las contribuciones de los usuarios editores (Adler et al., 2011b).

En nuestro estudio hemos diferenciado las actuaciones que cada uno de los tipos de usuarios ha realizado en cuanto a la creación y a la revisión de los artículos, incluida la reversión de ediciones; el objetivo es conocer en detalle sus características. También hemos analizado la cantidad de información que los diferentes tipos de usuarios ha generado o destruido a lo largo del tiempo.

Por último, dado que la Wikipedia en español se nutre de editores procedentes de muchos países hispanohablantes, hemos analizado el país de origen de los mismos.

3. Metodología

Wikipedia funciona gracias al software libre MediaWiki. Todas las acciones que se realizan en Wikipedia quedan registradas en bases de datos. Un aspecto destacable desde el punto de vista de la investigación, es que la Fundación Wikimedia mantiene el historial de todas las acciones que se realizan en sus proyectos, y este historial, a excepción de los datos personales de los usuarios registrados, es público.

El acceso a los datos se puede realizar en tiempo real utilizando la interfaz de aplicaciones de MediaWiki o a las herramientas *Wikimedia Tool Labs* (<http://wikitech.wikimedia.org>). Datos generales se ofrecen en las páginas de estadísticas que poseen todos los proyectos, que se recopilan en <http://stats.wikimedia.org/>. Otra forma es accediendo al volcado completo que se realiza de todos los proyectos de forma regular en <http://dumps.wikimedia.org>. Los datos suelen estar comprimidos en ficheros XML y SQL.

En nuestro estudio hemos utilizado el volcado de datos de la Wikipedia en español de fecha 5 de enero de 2015. El volumen de datos tratados ha sido de 1 TB de información. Los datos fueron procesados con diferentes programas desarrollados *ad hoc* por los autores, para, en primer lugar, extraer la información y almacenarla en el gestor de bases de datos MariaDB (en parte utilizando los volcados SQL), y, en segundo lugar, para obtener los resultados que se presentan más adelante.

De todas las páginas de la Wikipedia en español hemos limitado nuestro análisis a las páginas que son artículos enciclopédicos, dado que son el objetivo principal de las personas que visitan Wikipedia. De hecho, la mayoría de estudios sobre Wikipedia ponen el punto de mira en los artículos enciclopédicos y en las acciones que realizan los usuarios editores sobre ellos (Kim-

mons, 2011; Yasseri y Kertész, 2013). Un aspecto a tener en cuenta es que hemos unificado los datos de todas las redirecciones de las entradas existentes al mismo artículo enciclopédico. Eso ha supuesto reducir el número de artículos de 2,7 a 1,1 millones.

Las características que hemos analizado de los artículos enciclopédicos son las siguientes: número de artículos, longitud media y distribución de artículos por longitud, número de categorías, distribución del número de categorías por artículo, distribución del número de artículos por categoría, número de enlaces entrantes y salientes, y las distribuciones asociadas por artículo. Varias de estas medidas se corresponden con medidas típicas de estudios cibernéticos y web-métricos (Berrocal et al., 2003; Stuart, 2013).

Los usuarios editores de la Wikipedia pueden crear artículos nuevos o modificar el contenido de los ya existentes. Hemos analizado ambos procesos, teniendo en cuenta los distintos tipos de usuarios. Los aspectos que hemos analizado son los siguientes: creación de artículos, número de artículos editados, distribución de usuarios editores considerando periodos de actividad y número de artículos editados, y cantidad de información creada o destruida.

Es importante señalar que hemos unificado todas las revisiones que un mismo editor ha realizado de manera sucesiva sobre el mismo artículo, utilizando para ello también un programa desarrollado por los autores. Por un lado, es bien sabido que una medida de prestigio de los editores es el número de ediciones realizadas (Adler et al., 2011a; Francke y Sundin, 2010). Una sencilla artimaña para aumentar ese número es realizar pequeños cambios sucesivos del mismo artículo. Por otro lado, muchos usuarios a menudo van incorporando información poco a poco en cada revisión, comprobando en cada paso el resultado del artículo. Con ambas unificaciones, el número de revisiones de artículos enciclopédicos se reduce significativamente de 55 a 37 millones.

Para determinar el país de origen de los editores hemos utilizado, por una parte, la geolocalización IP de los usuarios no registrados (se ha utilizado un programa desarrollado por los autores que utiliza las bases de datos de MaxMind, www.maxmind.com), y por otra, la información de aquellos usuarios registrados que lo indican es sus páginas de usuario, al asignarse categorías del tipo “Wikipedistas de España”, “Wikipedistas de Argentina”, etc. Este proceso se complica pues muchos editores se asignan categorías muy locales (por ejemplo, “Wikipedistas de La Serena”), y ha sido necesario revi-

sarlas manualmente. Para usuarios muy activos se ha procedido también a revisar manualmente su página de usuario para determinar el país de origen. Los bots no tienen país de origen.

En cuanto a las reversiones, en nuestro estudio hemos detectado solamente las reversiones vandálicas, utilizando para ello información sobre el tipo de usuario que realiza la reversión y los comentarios que aparecen en el campo correspondiente de la base de datos.

4. Resultados y comentarios

Este apartado está organizado de acuerdo a los aspectos de interés que se han indicado previamente, para caracterizar los artículos enciclopédicos de la Wikipedia y los usuarios editores que los crean o los revisan. El primer apartado presenta datos cuantitativos de la Wikipedia en español a la fecha del volcado de datos. El segundo muestra resultados sobre la actividad de los diferentes tipos de usuarios editores, en lo referente a la creación de artículos, a la frecuencia de actuación y a la cantidad de información creada o destruida. El tercer apartado se centra en las revisiones de los artículos, ofreciendo resultados sobre número de artículos editados por editores humanos. También se incluyen en este apartado resultados relativos a las reversiones originadas por vandalismo. El último apartado ofrece resultados sobre el origen de los editores de la Wikipedia en español.

4.1. Artículos enciclopédicos

La Tabla I muestra el número de usuarios, el número de artículos creados y el número de revisiones (unificadas) por tipo de usuario, y el total, desde que la Wikipedia en español inicia su andadura en mayo de 2001 hasta la fecha del volcado de datos.

Tipos de usuarios-editores	Nº de usuarios	Nº de art. creados	Nº de ediciones
Registrados	557.970	422.874	7.980.795
Administradores	856	535.846	7.252.615
Bots	424	4.217	10.539.347
No registrados	5.793.501	148.863	11.179.884
TOTAL	6.352.751	1.111.800	36.952.641

Tabla I. Número y actividad de los usuarios de la Wikipedia en español

El número de artículos enciclopédicos es 1.111.800. Su longitud media es de 5.304 bytes, con una alta desviación de 9.616, lo cual indica

una alta variabilidad. La distribución de frecuencias de artículos por longitud puede verse en el Gráfico I. A partir de los 10 KB de longitud la línea de tendencia sigue una ley de potencias inversa (ley de Zipf) con parámetro 0,5. Este parámetro es menor que el obtenido para estudios webmétricos (Baeza-Yates et al., 2005; Baeza-Yates et al., 2006), que está en torno a 2,7. Los dos artículos más largos son *Alsacia en 1789* e *Historia de Euskaltel-Euskadi*, con 800 KB y 736 KB respectivamente.

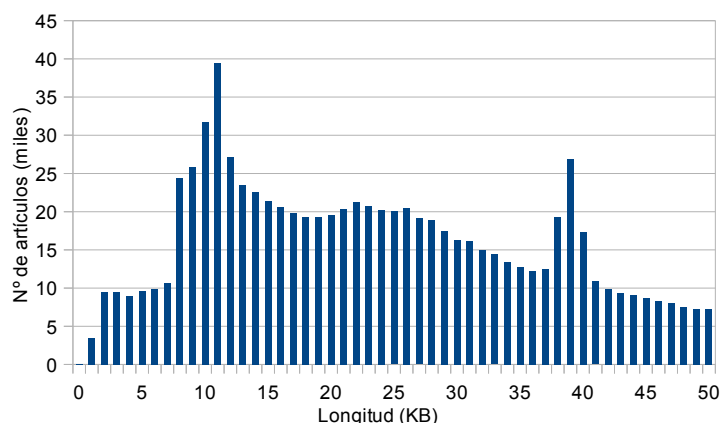


Gráfico I. Distribución de frecuencias de los artículos por longitud

El número de categorías no ocultas diferentes asignadas a los artículos es 227.152. La distribución de frecuencias de artículos por número de categorías puede verse en el Gráfico II. Hay unos pocos artículos sin categoría asignada. Se puede ver que la mayoría de ellos se etiquetan con una, dos o tres categorías. Los dos artículos asignados a un mayor número de categorías son *José María Pemán* y *Mario Vargas Llosa*, con 47 y 46 categorías, respectivamente.

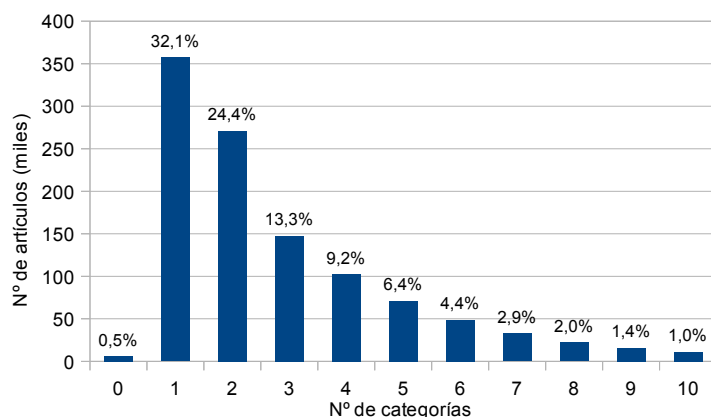


Gráfico II. Distribución de los artículos por categoría

Por otra parte, la inversa, esto es, la distribución de frecuencias de categorías por número de artículos se puede ver en el Gráfico III. Se aprecia que hay un 25% de categorías (57.156) que solamente aparecen en un artículo. Es decir, la asignación de categorías no siempre cumple el objetivo de que sirvan de navegación entre artículos. De acuerdo con este resultado podemos indicar que en la Wikipedia no se hace un uso adecuado de las categorías como elementos principales de representación del contenido para la obtención de nuevo conocimiento.

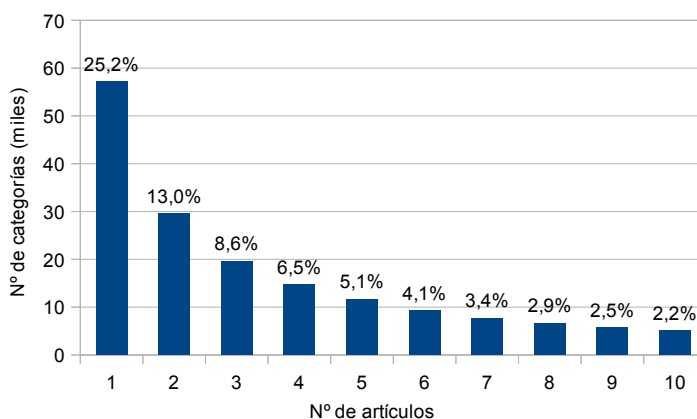


Gráfico III. Distribución de categorías por artículo

Se destaca que la categoría más frecuente es *Wikipedia:Desambiguación* con 46.243 artículos. Se trata de páginas con enlaces a artículos de títulos susceptibles de ambigüedad, en las que se ofrece una breve descripción de sus temáticas.

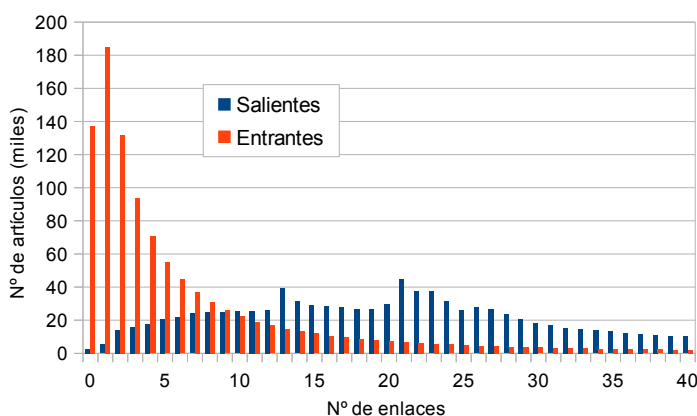


Gráfico IV. Distribución de frecuencias de enlaces entrantes y salientes por artículo

La distribución de frecuencias de enlaces entrantes y salientes, que permiten navegar directamente entre artículos, se muestra en el Gráfico IV. El número total de enlaces es de 33 mi-

llones. La distribución del número de enlaces entrantes es muy diferente a la de enlaces salientes. La distribución de enlaces entrantes tiene una distribución típica de larga cola o ley de potencias inversa, con parámetro 1,5. Este valor es ligeramente menor que para estudios cibernéticos sobre la Web (Baeza-Yates et al., 2005; Baeza-yates et al., 2006), que suelen estar en torno a 2,1. Es de señalar que hay un 12% de artículos que no reciben enlaces. Los dos artículos más citados son *Coordenadas geográficas* y *Estados Unidos*, con 235.971 y 197.008 enlaces entrantes, respectivamente.

La distribución del número de enlaces salientes, por el contrario, no sigue la típica distribución de potencias que sí se aprecia en estudios cibernéticos. Los usuarios editores suelen incorporar un elevado número de enlaces a otros artículos, con frecuencia a artículos de temática general o artículos relacionados solo tangencialmente con el contenido que están editando. Hemos comprobado que hay una correlación positiva de 0,76 entre la longitud de los artículos y el número de enlaces salientes a otros artículos. Los dos artículos que más enlaces salientes poseen son *Historia del Arte* y *Edad Contemporánea*, con 4.246 y 2.308 enlaces, respectivamente. Son dos artículos muy largos también.

4.2. Usuarios editores

El número total de usuarios aparecía en la Tabla I. De los 3,6 millones de usuarios registrados que ofrecen las estadísticas oficiales, en la práctica solamente han editado artículos 557.970. El número de bots es bastante elevado si lo comparamos con el de administradores. No obstante, la mayoría de ellos ha ejercido su labor durante unos pocos días y con muy pocas ediciones: suelen ser bots en pruebas (92 bots han realizado menos de 100 ediciones cada uno, estando activos 1 o 2 días). Por el contrario existen bots que están activos muchos días y realizan cientos de miles de ediciones: hemos comprobado que 31 bots han realizado más de 7,2 millones de ediciones, esto es, el 70% de todas las ediciones de los bots, el 20% de todas las ediciones (unificadas) de la Wikipedia. Si consideramos ediciones totales, sin unificar, el porcentaje de ediciones de todos los bots es del 19%. Este dato es específico para la Wikipedia en español, pues Wikipedias en otros idiomas ofrecen valores distintos (Mesgari et al., 2015).

En la misma tabla aparecen el número de artículos nuevos que cada uno de los tipos de usuarios ha creado, y las ediciones globales que ha realizado. Se destaca que el número de artículos que crean unos pocos administradores

es muy elevado; analizando más en profundidad los datos se comprueba que tan solo 28 usuarios (25 administradores y 3 usuarios registrados) han creado el 25% de los artículos. No obstante, no hay correlación significativa entre los usuarios que crean los artículos y los usuarios que los editan posteriormente. Estos resultados son similares para ediciones de Wikipedias de otros idiomas (Yasseri y Kertész, 2013; Gandica et al., 2015).

El Gráfico V muestra la evolución mensual en la creación de artículos para los diferentes tipos de usuarios. Desde el año 2006 hasta mediados del 2011 el ritmo de creación de artículos fue de unos 10.000 al mes. A partir de ese momento el ritmo ha sido de unos 7500 cada mes, siendo menor en el último año. Se aprecian meses de gran actividad: se debe a unos pocos editores que crean en muy poco tiempo una gran cantidad de artículos de la misma temática (astros, asteroides, plantas con ficha de taxón, coleópteros, arbustos, municipios, etc.). El número de artículos creados por usuarios administradores está muy por encima del resto. Vemos también que apenas existen artículos creados por bots, al contrario de lo que sucede en Wikipedias de otros idiomas, como la inglesa o la sueca (Okoli et al. 2012).

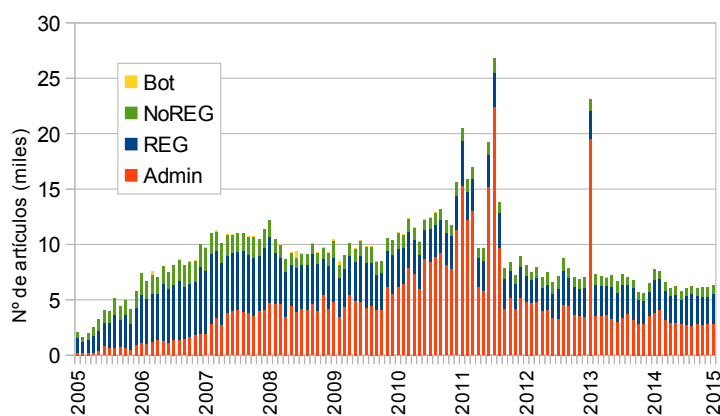


Gráfico V. Creación de artículos al mes

En el Gráfico VI se muestra la distribución de frecuencias relativas de editores por días de edición. La mayoría de usuarios anónimos y registrados están activos solamente un día, y editan un único artículo (Gráfico VII). Los usuarios administradores son los que más días dedican a labores de edición, y los que realizan más ediciones sobre un mayor número de páginas, excluidos, como era de esperar, los bots. El editor que más revisiones ha realizado es Diegusjaimes (156 mil), pese a dejar de editar artículos el 31-12-2012.

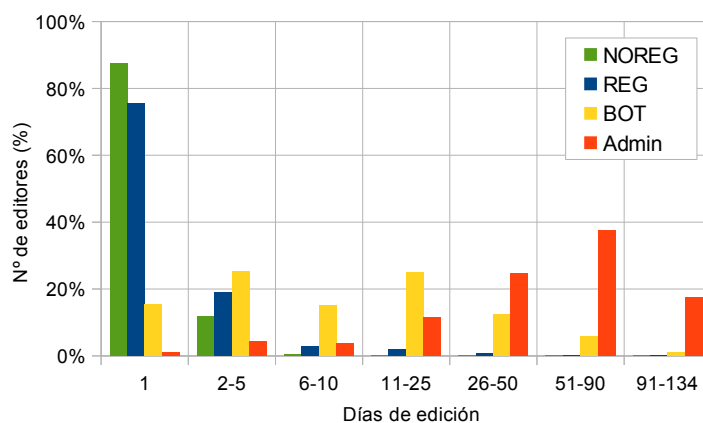


Gráfico VI. Distribución de frecuencias relativas de editores por días de edición

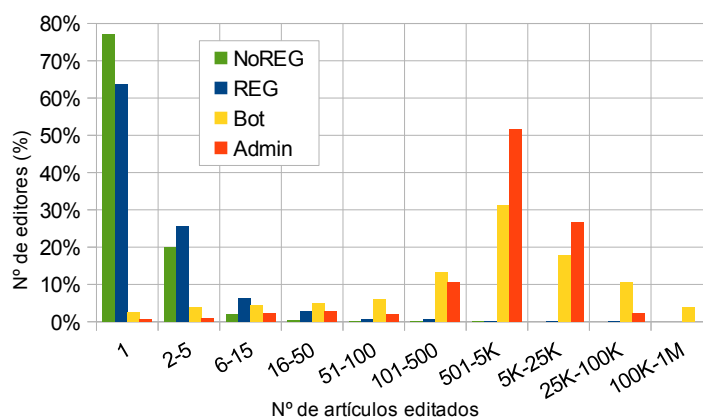


Gráfico VII. Distribución de frecuencias relativas de editores por número de artículos editados

Estas dos gráficas dan información muy valiosa sobre el protagonismo que tienen los diferentes tipos de usuarios en la Wikipedia. Podemos decir que existe un gran número de usuarios que realiza unas pocas contribuciones a la Wikipedia, y que una gran parte de las ediciones las realizan unos pocos usuarios administradores. Ello es general para todas las Wikipedias (Mesgari et al., 2015).

Hemos analizado también cuál es la actitud de los usuarios respecto de la creación de contenidos, teniendo en cuenta el número de bytes creados o destruidos en cada edición. En el Gráfico VIII se puede ver que desde el año 2007 se incorporan cada mes casi 50 MB nuevos a la Wikipedia. Hemos utilizado bytes y no palabras, como en estudios más completos (Adler et al., 2011a), porque hemos comprobado manualmente con varios cientos de artículos que los resultados son similares. La extraordinaria disminución de marzo de 2013 se debió al cambio de los enlaces *interwikis* entre Wikipedias de

diferentes idiomas al formato de WikiData, proceso que fue realizado por bots.

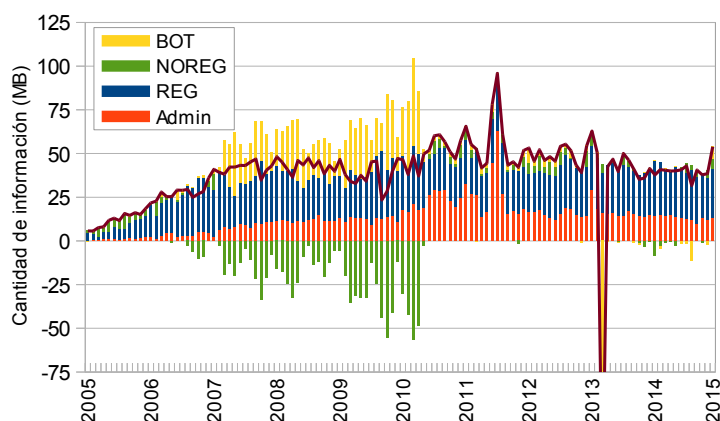


Gráfico VIII. Cantidad de información creada al mes

Se destaca también en el Gráfico VIII una actividad bastante negativa entre los años 2007 y 2010 de los usuarios no registrados. Este periodo coincidió con la publicación de la obra de Soliman y Gourdain (2008) que supuso una crítica tenaz contra Wikipedia, y muchos usuarios no registrados llevaron demasiado lejos algunas de sus propuestas de introducir pequeños cambios para determinar cuán rápidamente eran revertidos, realizando completos blanqueos de página. Podemos apreciar que en ese periodo de tiempo la actividad del resto de usuarios, especialmente de los bots, va casi pareja, pero en sentido contrario, a la de los usuarios no registrados, con lo que el volumen de información siguió incrementándose. Precisamente fue en este periodo cuando más se desarrollaron los bots antivandalismo.

4.3. Revisión de artículos

El Gráfico IX muestra el número de artículos que se revisan al mes una o más veces por editores humanos, no por bots, después de haber sido creados. Es un número bastante uniforme desde 2009, unos 100 mil al mes. Esto da idea de que el número de editores efectivos de Wikipedia no ha cambiado sustancialmente desde entonces. El número de revisiones sigue una tónica similar, aproximadamente desde 2009 cada mes se realizan unas 250 mil revisiones por editores humanos.

En el Gráfico X se muestra la distribución de frecuencias de artículos por número de editores humanos diferentes. Un 21% de los artículos (233.652) ha sido editado solamente por un editor; es decir, fue creado y no se ha vuelto a editar por nadie más. La mitad de los artículos han sido editados por cuatro o menos editores.

Los artículos más editados son, como cabía esperar, artículos muy antiguos, en los que han participado muchos editores, y que además han sufrido muchas reversiones. Los tres artículos más editados son *Venezuela*, *Club América* e *Idioma español*, creados todos ellos entre 2002 y 2005, con más de 6.200 revisiones cada uno de ellos, y con un porcentaje de reversiones en torno al 28% de todas sus revisiones.

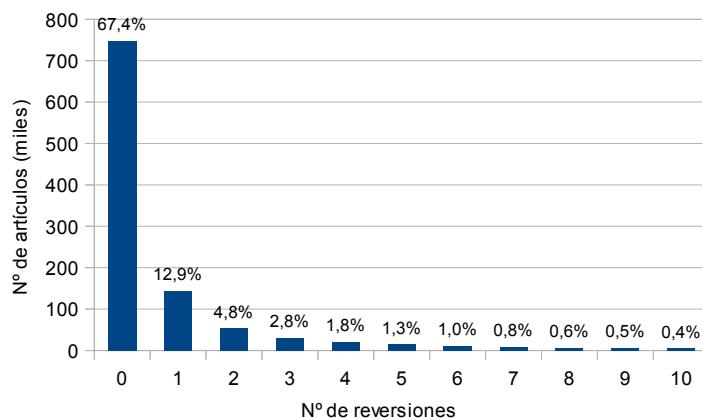


Gráfico IX. Artículos editados al mes por usuarios humanos, no por bots

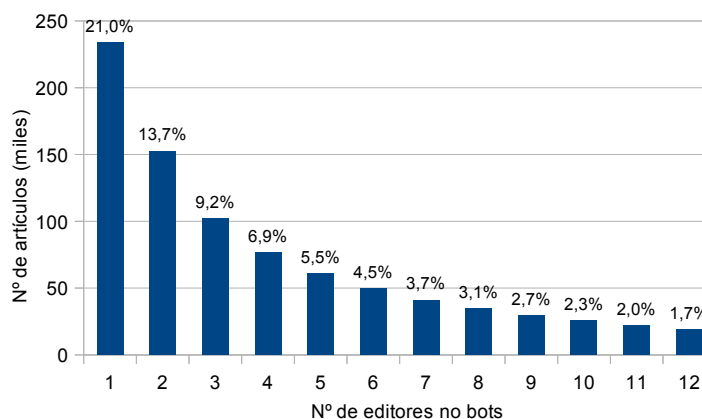


Gráfico X. Distribución de frecuencias de artículos por número de editores humanos diferentes

En el Gráfico XI se muestra la distribución de frecuencias de artículos por número de reversiones. El 67% de los artículos no se ha revertido nunca, y el 13% lo ha sido solamente una vez. Ello, junto con los resultados del gráfico anterior, nos indica que hay una buena disposición por parte de la mayoría de usuarios a la hora de incorporar contenidos a la Wikipedia. Los artículos más revertidos son *Club América*, *Dragon Ball*, *Baloncesto* y *Venezuela*. Son, como ya se indicó previamente, artículos muy antiguos.

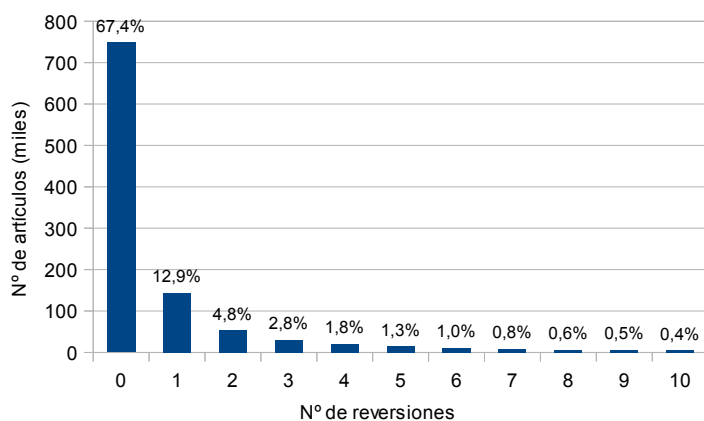


Gráfico XI. Distribución de frecuencias de artículos por número de reversiones

Los usuarios con atributos de reversor (administradores y algunos bots) son aquéllos que expresamente patrullan páginas para revertir vandalismos. Muchos usuarios registrados, sobre todo los más activos, también dedican mucho esfuerzo en esta labor. El número total de ediciones vandálicas revertidas por todos estos usuarios ha sido de 3,5 millones; es decir, que al menos otro número igual o mayor ha sido fruto del vandalismo. Al menos 3,5 millones de ediciones de un total de 19,2 millones de ediciones de usuarios de los tipos *registrado* y *no registrado* (los bots y los administradores no realizan ediciones vandálicas) son actos de vandalismo, lo cual supone un 18,3% de dichas ediciones. Es un porcentaje elevado, si bien, hemos comprobado que proceden de un número relativamente pequeño de usuarios. Los artículos más vandalizados son *Leyes de Newton*, *Romanticismo*, *Sistema operativo* y *Derechos humanos*.

Los reversiones humanos se ayudan de bots que marcan posibles ediciones vandálicas, a menudo teniendo en cuenta usuarios (o direcciones IP) con historiales de acciones vandálicas previas. Algunos autores (Geiger y Ribes, 2010; Priedhorsky et al., 2007) indican que estas herramientas permiten que “voluntarios medios”, que quizás tengan poco conocimiento del contenido del artículo en cuestión, puedan revertir ediciones vandálicas. Dicho proceso de revisión está en marcado contraste con las formas más tradicionales de producción de conocimiento profesional y académico de los expertos, que están en condiciones de contribuir a la Wikipedia en virtud de su conocimiento de un dominio determinado.

4.4. Origen de los editores

No debemos olvidar que el español es una de las lenguas más extendidas del mundo, y son muchas las contribuciones que se realizan desde muy diversos países a la Wikipedia en español. El Gráfico XII muestra el número de editores, creadores y ediciones realizadas en función de los países de origen de los editores para los que hemos podido obtener esa información (se indica el total en el gráfico). Las contribuciones latinoamericanas son muy importantes, por ello es frecuente encontrar variaciones dialectales, lo que en algunos casos ha llevado a algunas controversias.

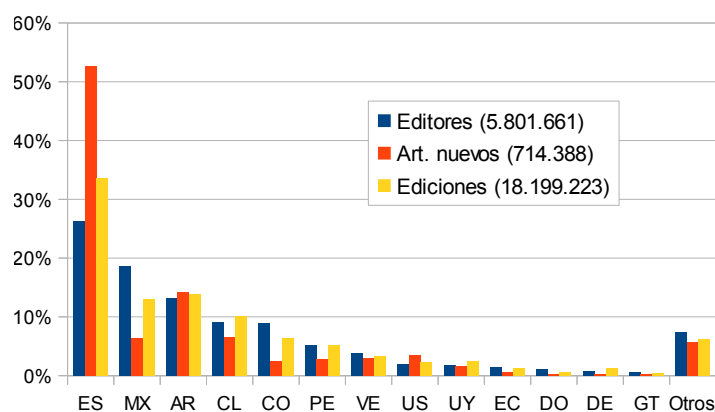


Gráfico XII. Origen de los editores por país (no bots)

5. Conclusiones

En este artículo hemos presentado el proceso de edición de la Wikipedia. Hemos estudiado los patrones de actuación que siguen los usuarios en sus contribuciones a esta herramienta de edición colaborativa. En cuanto a las características de los artículos enciclopédicos, hemos encontrado que para la longitud y los enlaces entrantes desde otros artículos se siguen distribuciones parecidas a las encontradas para el Web, si bien, con valores diferentes para los parámetros característicos. Para el caso de los enlaces salientes no existe tal analogía: ello se debe a que los usuarios editores suelen incluir gran cantidad de referencias a otros artículos enciclopédicos de temática general o relacionadas solo tangencialmente con el contenido que están editando.

En relación con el sistema de categorías que utiliza Wikipedia, cuyo objetivo es encontrar y navegar por grupos de artículos relacionados, hemos visto que no se utiliza adecuadamente como mecanismo de acceso a otros contenidos, tan importante en los sistemas de organización de la información. Es una de las deficiencias

más importantes que hemos encontrado, a pesar de que Wikipedia proporciona una buena infraestructura tecnológica al respecto y existen páginas con recomendaciones sobre el uso de categorías. Pensamos, por ello, que esta deficiencia se debe a una falta de sensibilidad o de formación de los usuarios editores a la hora de asignar categorías a los artículos.

Entre las conclusiones más importantes podemos destacar que la creación de contenidos se lleva a cabo por dos grupos diferentes de usuarios. Por un lado, pequeñas contribuciones individuales de una amplia legión de usuarios, en general usuarios registrados. Por otro lado, una gran cantidad de contribuciones que realizan un pequeño grupo de usuarios extraordinariamente activos, que poseen además atributos para la administración. La cantidad de información que se ha ido incorporando a la Wikipedia ha sido bastante uniforme en los últimos años, lo cual corrobora esta división de usuarios. La mitad de los artículos enciclopédicos han sido editados por cuatro o menos editores diferentes.

Destacamos también que una parte significativa del funcionamiento de Wikipedia recae en unos pocos usuarios que supervisan los contenidos nuevos que se van incorporando por usuarios anónimos o poco frecuentes, ayudados por bots que facilitan el proceso. Este resultado es general para Wikipedias de otros idiomas (Kimmons, 2011). Hay un gran número de revisiones que se realizan automáticamente por bots en busca de errores ortográficos y de formato, verificación de categorías y enlaces externos, etc., que ayudan al mantenimiento de la Wikipedia.

Una parte importante de las acciones de los usuarios administradores y de los usuarios registrados más activos es patrullar y revertir acciones vandálicas. La utilización de bots que marcan posibles ediciones vandálicas agiliza el proceso de reversión, no sólo por la rapidez, sino también porque hace que usuarios con conocimientos medios puedan realizar reversiones que de otro modo deberían ser realizadas por expertos del dominio. Es de señalar que el porcentaje de artículos que nunca ha sido revertido es muy elevado (67%), y otro 13% lo ha sido solamente una vez, lo cual indica que las principales acciones de vandalismo recaen sobre un número no muy elevado de artículos.

Para un porcentaje alto de usuarios se ha obtenido el país de origen, lo cual ha permitido contabilizar la contribución que se realiza desde cada uno de ellos. Este aporte es importante, pues son pocos los estudios específicos de sistemas de información en línea que obtienen información del origen de los contribuyentes, no

ya por la dirección IP (por otro lado bastante frecuentes en analizadores de visitas web), sino por el sistema de representación que tiene Wikipedia en lo referente a las categorías que pueden asignarse los propios usuarios para indicar su país o región de origen.

6. Agradecimientos

Este trabajo ha sido parcialmente financiado por la Fundación Memoria de D. Manuel Solórzano Barruso de la Universidad de Salamanca, proyecto de Ref. FS/5-2014.

Referencias

- Adler, B.T.; De Alfaro, L.; Kulshreshtha, A.; Pye, I.; (2011a). Reputation systems for open collaboration. // *Communications of the ACM*, 54:8, 81-87.
- Adler, B.T.; De Alfaro, L.; Mola-Velasco, S.M.; Rosso, P.; West, A.G. (2011b). Wikipedia Vandalism Detection: Combining Natural Language, Metadata, and Reputation Features, LNCS: Computational Linguistics and Intelligent Text Processing, 6609, 277-288.
- Baeza-Yates, R.; Castillo, C., Lopez, V. (2005): Characteristics of the Web of Spain. *Cybermetrics*, 9(1).
- Baeza-Yates, R.; Castillo, C., y Graells, E. (2006): Características de la Web Chilena. Universidad de Chile.
- Berrocal, J. L. A.; Figuerola, C. G.; Zazo, Á. F. (2004). *Cybermetría: nuevas técnicas de estudio aplicables al Web*. Ed. Trea.
- Fernández-Molina, J. C.; Guimarães, J. A. C. (2002). Ethical aspects of knowledge organization and representation in the digital environment: their articulation in professional codes of ethics. // *dvances in Knowledge Organization*, 8, 487-492.
- Francke, H.; Sundin, O. (2010). An inside view: credibility in Wikipedia from the perspective of editors. // *Information Research*, Special Supplement: Proceedings of the 7th International Conference on Conceptions of Library and Information Science, London. 15:3.
- Gabrilovich, E.; Markovitch, S. (2009). Wikipedia-based semantic interpretation for natural language processing. // *Journal of Artificial Intelligence Research*. 34, 443-498.
- Gandica, Y.; Carvalho, J.; dos Aidos, F. S. (2015). Wikipedia editing dynamics. // *Physical Review E*. 91:1, 012824.
- Geiger, R. S.; Ribes, D. (2010). The work of sustaining order in wikipedia: the banning of a vandal. // *Proceedings of the 2010 ACM conference on Computer Supported Cooperative Work*. 117-126.
- Hu, M.; Lim, E. P.; Sun, A.; Lauw, H. W.; Vuong, B. Q. (2007). Measuring article quality in Wikipedia: models and evaluation. // *Proceedings of the sixteenth ACM Conference on Information and Knowledge Management*. 243-252.
- Kimmons, R. M. (2011). Understanding collaboration in Wikipedia. *First Monday*. 16:12.
- Leskovec, J.; Huttenlocher, D.; Kleinberg, J. (2010) Governance in social media: A case study of the Wikipedia promotion process. // *Proceedings of the International Conference on Weblogs and Social Media, ICWSM'10*.
- Lewandowski, D.; Spree, U. (2011). Ranking of Wikipedia articles in search engines revisited: Fair ranking for reasonable quality? // *Journal of the American Society for Information Science and Technology*. 62:1, 117-132.
- Luyt, B.; Aaron, T. C. H.; Thian, L. H.; Hong, C. K. (2008): Improving Wikipedia's accuracy: Is edit age a solution? //

- Journal of the American Society for Information Science and Technology. 59:2, 318–330.
- Magnus, P. D. (2009). On trusting Wikipedia. // *Episteme: A Journal of Social Epistemology*. 6:1, 74-90.
- Medelyan, O.; Milne, D.; Legg, C.; Witten, I. H. (2009). Mining meaning from Wikipedia. // *International Journal of Human-Computer Studies*. 67:9, 716-754.
- Mesgari, M.; Okoli, C.; Mehdi, M.; Nielsen, F. Å.; Lanamäki, A. (2015). The sum of all human knowledge: A systematic review of scholarly research on the content of Wikipedia. // *Journal of the Association for Information Science and Technology*. 66, 219–245.
- Müller, C.; Gurevych, I. (2009). Using wikipedia and wiktio-nary in domain-specific information retrieval. // *Evaluating Systems for Multilingual and Multimodal Information Access*. Berlin, Heidelberg: Springer. 219-226.
- Okoli, C.; Mehdi, M.; Mesgari, M.; Nielsen, F. Å.; Lanamäki, A. (2012). The people's encyclopedia under the gaze of the sages: A systematic review of scholarly research on Wikipedia. // *Social Science Research Network*, 2021326.
- Pehcevski, J.; Thom, J. A.; Vercoustre, A. M.; Naumovski, V. (2010). Entity ranking in Wikipedia: utilising categories, links and topic difficulty prediction. // *Information Retrieval*, 13:5, 568-600.
- Priedhorsky, R.; Chen, J.; Lam, S. T. K.; Panciera, K.; Terveen, L.; Redl, J. (2007). Creating, destroying, and restoring value in Wikipedia. // *Proceedings of the 2007 International ACM Conference on Supporting Group Work* (pp. 259-268). ACM.
- Sepehri, H.; Makazhanov, A.; Rafiei, D.; Barbosa, D. (2012). Leveraging editor collaboration patterns in Wikipedia. // *Proceedings of the 23rd ACM Conference on Hypertext and Social Media, HT '12*. 13-22.
- Soliman, M.; Gourdain, P. (2008). *La revolución Wikipedia*. Madrid: Alianza Editorial, ISBN: 978-84-206-8236-5.
- Stuart, D. (2013). *Web metrics for library and information professionals*. London: Facet Publ.
- Wu, G., Harrigan, M., Cunningham, P. (2011) Characterizing Wikipedia pages using edit network motif profiles. // *Proceedings of the 3rd International Workshop on Search and Mining User-generated Contents*. 45-52.
- Wu, Q.; Irani, D.; Pu, C.; Ramaswamy, L. (2010). Elusive vandalism detection in Wikipedia: a text stability-based approach. // *Proceedings of the 19th ACM International Conference on Information and Knowledge Management, CIKM '10*. 1797-1800.
- Yasseri, T.; Kertész, J. (2013). Value production in a collaborative environment. // *Journal of Statistical Physics*. 151:3-4, 414-439.

Enviado: 2015-04-12. Segunda versión: 2015-08-10.
Aceptado: 2015-08-17.

Archivo de Espacio P: propuesta de un cuadro de clasificación para su continuidad digital

A classification scheme for the "Espacio P" Archive: a proposal for its digital continuity

**Federico CASTRO MORALES, Fátima GARCÍA LÓPEZ,
Ana María MORALES GARCÍA, Karin OHLENSCHLÄGER OHLENSCHLÄGER**

Instituto Agustín Millares de Documentación y Gestión de la Información, Universidad Carlos III de Madrid, Getafe, España
{castromofatimajlamorales}@hum.uc3m.es, k.ohlenschlaeger@gmail.com

Resumen

El objetivo del estudio es proponer un cuadro de clasificación del fondo generado por Espacio P, especializado en videoarte, como instrumento pionero en su campo que permita organizar y recuperar la documentación que conserva este archivo mueístico y otros semejantes.

Palabras clave: Espacio P. Cuadro de clasificación. Garhel, Pedro. Cultura digital. Arte electrónico. Performances. Archivos. Continuidad digital.

Abstract

A pioneer classification scheme is proposed for the Espacio P holdings, as a tool to organize and retrieve the documents preserved in this archive, a reference in videoart.

Keywords: Espacio P. Classification scheme. Garhel, Pedro. Digital culture. Digital arts. Performances. Archives. Digital continuity.

1. Introducción

Espacio P(erformance), también llamado Espacio P(olivalente) (1981-1997), fue una parte esencial del proyecto artístico de Pedro Garhel (Puerto de la Cruz, 1952 – La Guancha, 2005), un centro independiente, autogestionado y autofinanciado, pionero en el género del videoarte en España, capaz de generar recursos propios y crear redes de proyección nacional e internacional.

Espacio P diseñó festivales y programas de vídeo y coordinó exposiciones a través de la colaboración con instituciones, empresas del sector de las TICs y colegios profesionales, en el Museo Nacional de Escultura de Valladolid, el Museo Español de Arte Contemporáneo, el Círculo de Bellas Artes de Madrid, Injuve, o el Goethe Institut. Además, promocionó a videoartistas españoles a nivel nacional y europeo con su distribuidora Alliance Video Art (AVA), actividad que contribuyó al sostenimiento del proyecto (Ohlenschläger, 2011).

Espacio P generó sus propios cauces de producción y proyección, entroncando con algunos eventos de gran prestigio internacional, como Ars Electronica, la Documenta de Kassel o el programa Amsterdam Capital Cultural Europea (1987), donde participaron artistas vinculados a Espacio P.

A la iniciativa de Espacio P se debe la introducción en España de artistas hoy reconocidos —

como Ulay y Marina Abramovic, Dekay Pitch, Ulrike Rosenbach, Fabrizio Plessi, Bob Wilson o Peter Weibel—; y el desarrollo de líneas de investigación y producción a través de innovadores proyectos performativos, audiovisuales y multimedia dentro del ámbito de las artes visuales, escénicas y sonoras.

Espacio P, es a su vez una parte significativa del legado del artista canario Pedro Garhel; de ahí, que hayamos iniciado un gran proyecto de investigación en torno a las prácticas artísticas interdisciplinarias y la continuidad digital, liderado por el Instituto de Investigación Agustín Millares Carlo de Documentación y Gestión de la Información (Universidad Carlos III de Madrid) y financiado por el Ministerio de Economía y Competitividad (Convocatoria 2013 de Proyectos I+D+i, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2013-2016).

En este proyecto se pretende abordar el tratamiento archivístico de este fondo que contiene una información extraordinaria, en ocasiones fragmentaria. Para ello, y como resultado de la primera fase, se ha elaborado un cuadro de clasificación funcional, aplicado a dos secciones del Archivo "Espacio P".

Entre las dificultades para su concepción cabe señalar la ausencia de modelos similares, debido a que los cuadros de clasificación para fondos documentales del ámbito del arte contemporáneo se centran, sobre todo, en las funcio-

nes y actividades llevadas a cabo por las instituciones museísticas. En este sentido, es reseñable el trabajo realizado en el Museo Nacional de Arte de Cataluña (Blesa, 2008). Sin embargo, la conservación del patrimonio del arte experimental en el tránsito de la era analógica a la digital es bastante precaria (García Melero, 2005)

Paralelamente, el equipo solicitante, trabaja en el plan de difusión (web, editorial, expositivo y difusión digital: blog, redes y exposición permanente en Europeana), de interés para el Museo Nacional Centro de Arte Reina Sofía (MNCARS) y el futuro Espacio de Arte Experimental de Puerto de la Cruz (Tenerife).

2. Archivo de Espacio P: Propuesta metodológica para su continuidad digital

El Archivo de Espacio P y el archivo personal de Pedro Garhel, depositados en el Archivo Histórico Provincial de Santa Cruz de Tenerife, constituyen dos secciones del legado superpuestas y entrelazadas.

Espacio P generó un rico patrimonio documental que se suma al de su fondo videográfico, sonoro, y a las producciones audiovisuales emprendidas desde este centro autogestionado. Igualmente relevante es la documentación sobre sus actividades formativas y expositivas, así como el registro administrativo y contable (un volumen ingente de presupuestos, facturas, estados de cuentas bancarias...), una información especialmente reveladora en el contexto actual, marcado por una nueva consideración de la autofinanciación de la cultura y de reafirmación de la autogestión cultural y de incitación a la participación ciudadana en el sostenimiento de las iniciativas culturales.

El Archivo de Espacio P ocupa un total de 100 metros lineales de estantería en el Archivo Histórico Provincial de Tenerife. Hasta la fecha solo se ha podido inventariar el 60% de los fondos documentales, por tanto —a título indicativo y sin perjuicio de que puedan encontrarse más documentos de éstas y de otras tipologías—, se han registrado más de 1.600 cartas, 66 invitaciones, folletos y carteles de las actividades de Espacio P, 1.200 discos, así como 134 casetes de audio y CDs. También se han clasificado cerca de 300 carpetas colgantes con currícula, dossiers de artistas, festivales y distribuidoras de videoarte nacionales e internacionales, así como 403 audiovisuales de distintos formatos (Betacam SP, U-matic, VHS, Betamax y Super 8).

Con el Proyecto de investigación *Archivo de Espacio P: Propuesta Metodológica para su*

Continuidad Digital (2014-2017) se pretende concluir las tareas ya emprendidas; así como crear e implementar un Cuadro de clasificación, completar el inventario, desarrollar un tesoro específico y emprender la digitalización del fondo para hacerlo accesible, con el fin de poner en valor este importante episodio de la reciente historia del arte contemporáneo en los años iniciales de la Democracia.

La digitalización del Archivo Espacio P contribuirá así a la toma de conciencia acerca de la importancia que tiene el legado desde la perspectiva del patrimonio digital, y también desde el punto de vista del patrimonio artístico contemporáneo (Felipe y Ohlenschläger, 2011). Analizar las estrategias y vías de financiación que tuvo el proyecto, ayudará a comprender mejor el modelo de autogestión artística y a ponderar su vigencia y viabilidad en las circunstancias actuales.

Tratándose de una experiencia pionera en España, la digitalización del Archivo Espacio P supondrá una oportunidad de avance en la reflexión metodológica sobre las técnicas de creación, almacenamiento y preservación digital de archivos, registros y documentos electrónicos, y en el estudio transdisciplinar sobre los orígenes de la sociedad de la información y la sociedad red en nuestro país.

Comparten la investigación documentalistas, expertos en archivística y comunicación, restauradores de soportes audiovisuales, historiadores del arte y del uso creativo de los nuevos medios e ingenieros informáticos.

Por último, este proyecto de investigación tiene como finalidad contribuir a la difusión de los resultados de esta investigación a través de una exposición monográfica sobre el Archivo de Espacio P, que tendría su punto de partida en una institución cultural española, con itinerancia nacional e internacional, y propósito de permanencia como muestra virtual en Europeana.

3. Propuesta de cuadro de clasificación

De todas las actividades archivísticas, posiblemente, la clasificación sea la gran pieza angular que sustenta la organización y gestión del fondo archivístico, al ser “la operación intelectual que consiste en el establecimiento de las categorías y grupos que reflejan la estructura orgánica y/o funcional del fondo” (Diccionario Terminología, 1993, p. 25).

Desde la perspectiva Archivística, el concepto de clasificación ha estado vinculado al principio de procedencia y respeto al orden original, con-

siderados tradicionalmente como los dos pilares básicos que fundamentan la teoría y la práctica de esta disciplina. El concepto de clasificación ha sido objeto de revisión por parte de teóricos canadienses y anglosajones que la consideran una actividad fundamental, no sólo para la organización de los archivos, sino también como apoyo a la gestión de documentos en las instituciones públicas y privadas (Barros y Moraes, 2012)

Desde la perspectiva de la organización del conocimiento, la clasificación debe proporcionar una estructura lógica del fondo archivístico que permita sistematizar el contexto en el que la documentación se ha producido, además de configurarse como un instrumento para la representación y la posterior recuperación de los contenidos informativos (Ribeiro, 2014)

El primer problema que se planteó en la organización del fondo documental, fue seleccionar el criterio de clasificación más idóneo. Su legado, nos ofrecía una tipología documental muy variada, al tratarse de un proyecto artístico plurifuncional, con proyección nacional e internacional, aspecto que dificultó, en un principio, el establecimiento de clases jerárquicas. En esta tipología se fundían correspondencia privada y profesional, documentación bancaria personal y de Espacio P, recursos docentes, una biblioteca y una mediateca, e incluso algunos testigos de los instrumentos de descripción creados por el propio centro para intentar gestionar su fondo documental. A ello, se sumó otra dificultad añadida, la identificación entre arte y vida que marcó la existencia de Pedro Garhel, plasmada en la integración de la documentos familiares y escolares previos a la creación de Espacio P, así como posteriores, principalmente en su etapa final, cuando impulsa la creación de la “Fundación P” en Tenerife.

Asimismo, nos encontramos con un fondo desordenado, debido a los traslados y a su precario almacenamiento, donde no existía ni un inventario, ni un listado de la aduana, correspondiente al traslado del fondo de Espacio P a Tenerife.

La primera fase del proceso de clasificación fue, por tanto, una labor compleja, que exigió, en primer, lugar crear una serie preliminar “Fondo Personal Pedro Garhel” y posteriormente, establecer el corte cronológico de 1996, momento en el que el artista tiene un problema de salud, que le obliga a clausurar Espacio P.

Una vez concluido este reconocimiento físico, pudimos proceder a la elaboración del cuadro de clasificación. Este instrumento “supone una estructura jerárquica y lógica de los fondos, los

divide en grupos evidenciados por ser el reflejo de una acción, función o actividad [...] Además, es inestimable para la localización conceptual de las series documentales, pues orienta las búsquedas, según criterios persistentes y objetivos” (Mundet, 1994, p. 244).

Desde esta perspectiva, hemos considerado el Cuadro de Clasificación, como el instrumento que permita organizar jerárquicamente, mediante el establecimiento de clases, los documentos del fondo, respetando el principio de procedencia. En nuestro caso, además, pretendemos que sea de utilidad como sistema de información sobre las funciones desempeñadas por Espacio P. Siguiendo este criterio, el fondo documental del Archivo se ha estructurado en seis secciones:

- Sección 1: Espacio P (1981-1997)
- Sección 2: Fondo de Documentación Espacio P
- Sección 3: Fondo Personal Pedro Garhel
- Sección 4: Fundación P/Empaquetado (2000-2001)/IEHC (1998-2005)
- Sección 5: Actividades en torno a P (2006-)
- Sección 6: Materiales de Espacio P

Partiendo de los objetivos marcados en el proyecto, se ha elaborado un Cuadro de clasificación, aplicado a dos secciones del fondo documental del Archivo.

3.1. Sección 1: Espacio P (1981-1997)

Dada las características del fondo de Espacio P, se optó por un criterio funcional. Una vez seleccionado el criterio de clasificación, los siguientes pasos se orientaron a identificar y denominar las funciones y actividades y ordenarlas dentro de una estructura jerárquica de esta sección.

1.1. Agenda de actividades (time line)

1.2. Distribución: AVA

- 1.2.1. Anteproyecto
- 1.2.2. Diseños, papelería, carátulas...
- 1.2.3. Cartas circulares
- 1.2.4. CV
- 1.2.5. Directorios
- 1.2.6. Folletos
- 1.2.7. Prensa
- 1.2.8. Programas de videocreación
- 1.2.9. Propuestas de ciclos de videocreación

-
- 1.3. Plataforma de producción
 - 1.3.1. Proyectos P
 - 1.3.2. DD Records
 - 1.3.3. P Records
 - 1.3.4. Plató
 - 1.3.5. Revista Dimensión infinito
 - 1.4. Coproducciones y coordinación
 - 1.4.1. Norrköping'83. Seminar on Independent and Open Art
 - 1.4.2. Videoforum internacional
 - 1.4.3. Fabrizio Plessi
 - 1.4.4. Darío Corbeira
 - 1.4.5. Documenta
 - 1.4.6. K18
 - 1.5. Propuestas e iniciativas
 - 1.5.1. Proyectos propios
 - 1.5.2. Participación en proyectos ajenos
 - 1.5.3. Textos y proyectos de Karin Ohlenschläger
 - 1.5.4. Guiones
 - 1.5.5. Propuestas de colaboraciones
 - 1.6. Formación
 - 1.6.1. Talleres
 - 1.6.2. Cursos
 - 1.6.3. Conferencias
 - 1.6.4. Seminarios
 - 1.6.5. Ciclos
 - 1.6.6. Estancias en Espacio P
 - 1.7. Colección P (Obras de otros artistas)
 - 1.8. Difusión de espacio P
 - 1.8.1. Folletos y papelería
 - 1.8.2. Dossieres de prensa
 - 1.8.3. Notas y comunicados de prensa
 - 1.9. Gestión económica y administrativa
 - 1.9.1. Contratos (alquiler, compra-venta, escrituras)
 - 1.9.2. Documentación contable (facturas, recibos y extractos bancarios)
 - 1.9.3. Documentación legal (denuncias, juicios, procesos)
 - 1.9.4. Comunicaciones (informes sobre envíos y acuses de recibo)
 - 1.9.5. Becas, ayudas y subvenciones
 - 1.9.6. Solicitudes de intercambio de publicaciones
 - 1.9.7. Agendas, directorios y contactos Espacio P
 - 1.9.8. Correspondencia
 - 1.9.9. Solicitudes de patrocinio
 - 1.10. Documentación fundacional. Espacio P
 - 1.10.1. Asociación cultural Espacio P (1984)
 - 1.10.2. Asociación cultural de Videocreadores de Espacio P (1992)
 - 1.10.3. Asociación cultural Amigos de Espacio P
 - 1.11. Correspondencia
 - 1.12. Referencias bibliográficas sobre Espacio P
 - 1.13. Documentación gráfica
 - 1.13.1. Audiovisual
 - 1.13.2. Diapositivas y negativos
 - 1.13.3. Grabaciones sonoras
 - 1.13.4. Contactos y fotografías
 - 1.13.5. Planos
- 3.2. Sección 2: Fondo de Documentación Espacio P
-
- Desde su puesta en marcha, Espacio P quiso funcionar como centro de documentación e información sobre prácticas artísticas contemporáneas relacionadas con el video, la performance, la poesía visual, la música alternativa, el arte electrónico y el intermedia (Higgins. 1966 p.1). Se trataba de una de las escasas fuentes que, con anterioridad a la creación del actual MNCARS, tenían los artistas activos en España en aquella época. En un principio estaba compuesto por la biblioteca y la mediateca. Hoy debemos sumarle el archivo generado por más de quince años de actividad del propio Espacio P.
- 2.1. Publicaciones seriadas
 - 2.1.1. Boletines
 - 2.1.2. Revistas
 - 2.1.3. Fanzines

- 2.1.4. Prensa
- 2.2. Catálogos y folletos
- 2.3. Libros y monografías
- 2.4. Mediateca (grabaciones sonoras y recursos audiovisuales)
- 2.5. Postales y estampas (programas de mano, invitaciones...)
- 2.6. Dossiers temáticos
 - 2.6.1. Cine
 - 2.6.2. Danza
 - 2.6.3. Diseño
 - 2.6.4. Música
 - 2.6.5. Poesía
 - 2.6.6. Teatro
 - 2.6.7. Video
 - 2.6.8. Performance
- 2.7. Dossiers de artistas y entidades
 - 2.7.1. Artistas
 - 2.7.2. Galerías
 - 2.7.3. Asociaciones
 - 2.7.4. Entidades de gestión
 - 2.7.5. Museos
 - 2.7.6. Distribuidoras
 - 2.7.7. Festivales
 - 2.7.8. Espacios autogestionados
- 2.8. Dossiers de proveedores
- 2.9. Literatura gris (repertorios bibliográficos, instrucciones, manuales de usuario)

De manera paralela, el tratamiento archivístico de fondos análogos exigirá el desarrollo de un tesoro de prácticas artísticas contemporáneas que recoja los conceptos y herramientas surgidos del diálogo entre las artes plásticas, visuales, escénicas y sonoras, que ayudará a su mejor identificación.

La experiencia de clasificación y digitalización del archivo servirá además para avanzar en el desarrollo de los retos de la difusión de la cultura contemporánea en la red. Dicha difusión estará en sintonía con las políticas y recomendaciones de la UNESCO y de la Unión Europea para impulsar la cultura digital y, especialmente, la Recomendación de 27 de octubre de 2011, sobre la digitalización y la accesibilidad en línea del material cultural y la conservación digital,

que insta a los Estados miembro a redoblar sus esfuerzos para digitalizar contenidos culturales, implicando al sector privado. Asimismo, la Agenda Digital para España, reconoce la importancia estratégica de los contenidos digitales y especialmente en su punto 7. 2 que incluye el Plan de Impulso de la Economía Digital y los Contenidos Digitales.

4. Consideraciones y propuestas de futuro

Los contenidos digitales en Internet cuentan ya con su propio proyecto de archivo: Internet Archive (1996), organización sin ánimo de lucro destinada a la preservación de webs y recursos multimedia, pero se carece de una estrategia global para preservar la memoria del patrimonio digital anterior a la red Internet, caso del legado de Espacio P. De ahí que señalemos como prioridad la digitalización del Archivo de Espacio P y la creación de la web que le integrará en los protocolos de salvaguarda de Internet Archive.

Como propuestas de futuro se contempla la continuidad de esta iniciativa y desarrollo del cuadro de clasificación para las secciones: Fondo Personal Pedro Garhel; Fundación P/ Empaquetado (2000-2001)/ IEHC (1998-2005); Actividades en torno a P (2006-) y Materiales de Espacio P.

Una vez digitalizada la documentación y garantizada su conservación, se propiciará la consulta a la comunidad científica y a los ciudadanos, facilitando su recepción social y cultural.

La continuidad digital de este fondo documental, permitirá reflexionar y formular sugerencias metodológicas para ampliar el protocolo de digitalización de documentos y soportes específicos de la transición de la era analógica a la digital, en definitiva, contribuirá al desarrollo de la cultura digital y de la autogestión cultural en España..

Agradecimientos

Esta investigación se enmarca dentro del Proyecto de investigación "Archivo de Espacio P: Propuesta Metodológica para su Continuidad Digital" (2014-2017) (MINECO, Ref. HAR2013-44726-R). El presente proyecto da continuidad a las tareas ya emprendidas en 2008 para la recuperación historiográfica del artista Pedro Garhel desde el Centro Atlántico de Arte Moderno (CAAM) y el Gobierno de Canarias, con motivo de la muestra retrospectiva sobre el artista, en CAAM y La Regenta (Gran Canaria), DA2 (Salamanca) y TEA (Tenerife) (2010-2011).

Referencias

- Archivo Espacio P. Blog (2014). <https://archivoespaciop.wordpress.com/> (2015-07-16)
- Blesa, Pilar (2008). El Archivo del Museo Nacional de Arte de Catalunya (1997-2007): organización actual y plan-

- teamientos de futuro. // IV Encuentros de Centros de Documentación de Arte Contemporáneo. Vitoria/Gasteiz: Artium, 2008. 28-40
- Barros, Thiago Henrique Bragato; Moraes, João Batista Ernesto de (2012). Archival classification and knowledge organization: theoretical possibilities for the archival field. // *Categories, Contexts and Relations in Knowledge Organization: Proceedings of the Twelfth International ISKO Conference 6-9 August 2012 Mysore, India*, eds., A. Neelameghan, and K.S. Raghavan. Würzburg, Germany: Ergon, 2012. 272-276
- Castro Morales, Federico (2011). Sobre archivos y museos: a propósito de la retrospectiva de Pedro Garhel. // *Revista de Museología*. 51 (2011) 42-52.
- Cruz Mundet, José Ramón (1994). *Manual de Archivística*. Madrid: Pirámide, 1994. ISBN: 84-368-0860-6
- Diccionario de Terminología Archivística (1993). Madrid: Dirección General de Archivos Estatales, 1993.
- Felipe, Francisco y Ohlenschläger, Karin (2011). Conversación en torno al Espacio P y la videocreación de los años 80 en Madrid. // Nekane Aramburu y Carlos Trigueros (eds.). *Caras B de la Historia del vídeo arte en España* (2011) 80-88.
- García Melero, Luis Ángel (2005). La recopilación y conservación del patrimonio digital. // *Jornadas Sobre Bibliotecas Nacionales "las bibliotecas nacionales del siglo XXI"*, Valencia, 18, 19, 20 y 21 de mayo de 2005. <http://eprints.rclis.org/13830/> (2015-07-16)
- Higgins Dick (1966). *Intermedia*. // *Statment on Intermedia 1:1* (February 1966) <http://www.withoutbordersfest.org/2010/> (2015-07-16)
- Ohlenschläger, Karin (2011). *Pedro Garhel. Retrospectiva/retrospective*. Barcelona : CAAM-Gobierno de Canarias, 2011.
- Ribeiro, Fernanda (2014). The Use of Classification in Archives as a Means of Organization, Representation and Retrieval of Information // *Knowledge Organization*. 41:4 (2014) 319-326.
-
- Enviado: 2015-03-28. Segunda versión: 2015-07-16.
Aceptado: 2015-08-04.
-

Metateoria e Diplomática: a construção de um novo modelo de análise para a identificação de documentos de arquivo

Metateoría y diplomática: la construcción de un nuevo modelo de análisis para la identificación de los documentos de archivo

Metatheory and Diplomatics: building a new model of analysis for the identification of archival records

Natália Bolfarini TOGNOLI

Faculdade de Filosofia e Ciências – UNESP, Av. Hygino Muzzi Filho, 737 Marília – São Paulo – Brasil – CEP: 17525-900, nataliabtoynoli@yahoo.com.br

Resumen

Desde la década de los años 80, la Archivística se ha enfrentado a nuevos desafíos en lo que respecta a la organización del conocimiento archivístico. En este sentido, emerge el concepto de identificación, como parte del proceso de gestión documental. Este proceso cuenta con los estudios de la diplomática contemporánea, ya que busca la comprensión del contexto de producción por medio del análisis documental ofrecido por el análisis diplomático. En este sentido, el presente artículo presenta un nuevo modelo para la identificación de los documentos de archivo, con base en los métodos diplomáticos desarrollados en los siglos XVII, XIX y XX. Para eso, se contó con la metateoría como metodología para el conocimiento y análisis de los métodos diplomáticos. Los resultados demuestran un fuerte aporte metodológico de la metateoría a los estudios de la Organización del Conocimiento, así como un fuerte apoyo de la Diplomática a los estudios archivísticos.

Palabras clave: Diplomática. Metateoria. Identificación documental. Archivística. Documentos de archivo.

1. Introdução

As novas formas de produção documental, decorrente das novas tecnologias da informação e do aumento da complexidade burocrática nos sistemas administrativos, têm levado os profissionais da informação a encararem novos desafios no tocante à organização do conhecimento arquivístico.

Esse conhecimento, entendido como todo o conhecimento produzido por uma pessoa física ou jurídica e que está agrupado em um fundo documental, passa a ser produzido, portanto, em um contexto de constantes mudanças e rupturas paradigmáticas, o que leva os arquivistas a um repensar da teoria e dos métodos para organizar essa documentação.

Abstract

Since the 1980s Archival Science has faced new challenges regarding the organization of archival knowledge. In this way the concept of identification emerged as part of the record management processes. This process counts with the studies of contemporary diplomatics, once it aims to understand the records creation processes from the document analysis offered by the diplomatic method. This paper aims to present a new model for records identification, based on the diplomatic methods developed in 17th, 19th and XX centuries. Therefore, it counts with metatheory as a methodology for knowing and analyzing the diplomatic methods. The results show a strong methodological support of metatheory to the studies of Knowledge organization, as well as a strong support of Diplomatics to the archival studies.

Keywords: Diplomatics. Metatheory. Document identification. Archival science. Archival records.

Neste contexto nasce, na Espanha, no final dos anos 1980, a função da identificação, no âmbito da gestão documental, como uma resposta às necessidades da formulação de metodologias para enfrentar os desafios colocados aos arquivos contemporâneos.

Entendida como a primeira fase do tratamento arquivístico, a identificação encontra na Diplomática, mais precisamente em seu método, a ferramenta necessária para a compreensão dos conjuntos documentais, a partir da uma análise do documento, podendo ser considerada parte da chamada “análise documental”.

O método diplomático passa, portanto, a fundamentar a primeira etapa do tratamento arquivístico, entendida como a identificação do documento e de seu órgão produtor.

Dessa maneira, o presente artigo apresenta um novo modelo de identificação do documento de arquivo - capaz de lidar com qualquer tipo de documentação, em qualquer contexto de produção e organização de documentos - a partir da análise e fusão dos métodos diplomáticos desenvolvidos entre os séculos XVII e XXI.

Para tanto, conta-se com os estudos metateóricos, enquanto metodologia para a criação de um novo modelo, chamado, aqui, de "idealizado", seguindo o conceito definido por Weber (1904) que o entende enquanto uma acentuação unilateral de um ou mais pontos de vista e pela síntese de um fenômeno individual concreto, difuso, discreto, mais ou menos presente e ocasionalmente ausente, que são arranjados de acordo com aqueles pontos de vista unilateralmente enfatizados e com uma construção analítica unificada.

2. A identificação como função arquivística

Com a explosão documental gerada após a Segunda Guerra Mundial, fez-se necessário estudar os documentos em seu nascimento, visando à racionalização da produção. Neste contexto, emergiu o conceito de gestão documental, que visa a controlar o documento desde o momento de sua produção até sua destinação final, qual seja, eliminação ou guarda permanente.

Segundo Rodrigues (2015) a gestão documental se caracteriza como um conjunto de procedimentos aplicados para controlar os documentos arquivísticos durante todo o seu ciclo de vida, incidindo sobre o momento da produção e acumulação na primeira e segunda idade, ou seja, nos arquivos correntes e intermediários. Ainda segundo a autora, a introdução de programas de gestão de documentos na teoria Arquivística, levou à necessidade da criação de metodologias para resolver os problemas observados nos arquivos.

Assim, nasce na Espanha, em 1980, o conceito da identificação como uma ferramenta para a análise do documento e de seu órgão produtor, aplicada às massas documentais acumuladas, visando à elaboração de propostas para a avaliação e classificação dos documentos de arquivo.

São dois os objetivos de estudo da identificação: órgão produtor e a tipologia documental, como destacou Rodrigues (2015, p. 72):

Trata-se de uma tarefa de pesquisa sobre os elementos que caracterizam os dois objetos de estudos da identificação: órgão produtor, analisando o

elemento orgânico (estrutura administrativa) e o elemento funcional (competências, funções e atividades) e a tipologia documental, estudo que se realiza com base no reconhecimento dos elementos externos, que se referem à estrutura física, à forma de apresentação do documento (gênero, suporte, formato e forma) e internos.

É especificamente no tocante à tipologia documental, que a Diplomática Arquivística e seu método oferecem subsídios para sua análise. O método proposto pela Diplomática permite ao arquivista compreender os elementos internos e externos do documento de arquivo e as relações que eles guardam com o contexto.

3. A metateoria enquanto ferramenta metodológica para a idealização de um modelo de identificação

A metateoria tem sido amplamente utilizada no contexto da Ciência da Informação, principalmente no âmbito da Organização do Conhecimento, como demonstram os estudos de Vickery (1997), Svenonius (2004), Bates (2005) e Tennis (2008). Especificamente no contexto arquivístico, a Metateoria foi utilizada, primeiramente, por Tognoli (2014) como uma metodologia para melhor compreender a evolução da Diplomática desde o século XVII até o XXI. Entendida como a teoria sobre a teoria, a metateoria pode servir a quatro propósitos, segundo Ritzer (1991): (1) fornecer uma compreensão mais profunda do trabalho teórico existente (Mu); (2) fornecer uma perspectiva abrangente desse trabalho (Mo); (3) servir como um mecanismo para avaliação (Ma); e (4) servir como um prelúdio para um trabalho teórico futuro (Mp).

Ainda conforme Ritzer (1991), a partir da metateoria, há várias maneiras de se obter um conhecimento mais profundo e detalhado sobre a teoria da área de estudo, todas envolvendo uma dimensão interna/externa e intelectual/social. A interna se refere àquilo que existe dentro da área, que lhe é inerente, enquanto a externa se refere aos fenômenos que se encontram fora dela, mas que exercem algum impacto sobre a mesma. Por dimensão intelectual o autor entende tudo aquilo que se relaciona à estrutura cognitiva da área: teorias, ferramentas, ideias de outras disciplinas; por dimensão social entende-se a estrutura sociológica da área, conforme demonstrado no esquema a seguir (Figura 1).

Na primeira abordagem (Interna-Intelectual), é possível verificar os paradigmas pelos quais passam as ciências, escolas de pensamento ou correntes teóricas unidas a partir das teorias da área estudada, bem como as ferramentas e métodos utilizados para analisar as teorias e os

teóricos. A segunda (Interna-Social) foca os aspectos comuns entre as várias teorias da área, com ênfase em pequenos grupos de teóricos diretamente ligados e suas relações, bem como suas afiliações institucionais. A terceira (Externa-Intelectual) diz respeito aos conceitos, teorias e elementos emprestados de outras disciplinas, fazendo uso delas para analisar sua teoria. A quarta abordagem (Externa-Social) está diretamente ligada ao impacto da sociedade na construção da teoria.

Intelectual	
Paradigmas cognitivos	O uso de conceitos emprestados da:
Escolas de pensamento	Filosofia
Mudanças paradigmáticas	Economia
Escolas de pensamento	Linguística
Ferramentas metateóricas	Etc.
Teorias	
Interna	Externa
Paradigmas comuns	Impacto da sociedade
Colégios invisíveis	Impacto do social
Escolas	Instituições
Redes	Raízes históricas
<i>Background</i> Individuais	
Social	

Figura 1. Principais tipos de Mu

No entanto, é importante destacar que as abordagens não são excludentes, podendo-se encontrar os quatro tipos de Mu em um mesmo trabalho, como destacou Ritzer (1991):

Antes de proceder devo adicionar algumas notas cautelosas. Primeiro, as duas dimensões (interna-externa; intelectual-social) são uma continuação, sem linhas duras e estáticas entre seus polos. Segundo, os quatro tipos de Mu desenvolvidos ao transpassarem essas linhas não são rigidamente diferentes uns dos outros. Em outras palavras, em trabalhos metateóricos específicos, podem confluir dois, três, ou quatro tipos de Mu.

Com base nos estudos sobre o quadrante de Ritzer é possível concluir que o presente trabalho encontra-se fundamentado na primeira abordagem do quadrante (Interna-Intelectual), uma vez que aprofunda questões inerentes à

teoria e ao método da Diplomática para oferecer um estudo o mais detalhado possível, a partir das Escolas de pensamento, ou abordagens – aqui previamente identificadas como Diplomática Clássica, Diplomática Moderna e Diplomática Contemporânea –, e de seus respectivos métodos, a fim de constituir um método para a identificação de documentos de arquivo baseado em um tipo ideal. Metateoria, portanto.

4. A idealização de um modelo: o aporte da Diplomática à identificação dos documentos de arquivo

A Diplomática esteve sempre em contato com os documentos de arquivo. Desde o século XVI, os documentos analisados sob a luz da Diplomática eram essencialmente documentos arquivísticos. Esses documentos eram evidências que comprovavam propriedades de terras, direitos e deveres de cidadãos e fatos que ocorriam no passado. Já nos primeiros tratados de Diplomática, é possível identificar capítulos dedicados exclusivamente aos arquivos, notadamente por ser este o lugar onde se encontravam os documentos analisados.

Embora a Diplomática e a Arquivística tenham trabalhado sempre com o mesmo objeto de estudo, o documento de arquivo, ambas só foram colocadas efetivamente em contato na década de 1960, quando Bautier (1962) estabeleceu as *pièces d'archives* enquanto objeto de estudo da Diplomática, em resposta à chamada “crise diplomática” sugerida por alguns autores, na época, em decorrência de um exaurimento vivenciado pela disciplina no século XX.

Até 1960 o objeto de estudo da Diplomática estava ligado ao estudo da forma de documentos de natureza estritamente jurídica. A partir do discurso de Bautier, qualquer documento de arquivo, independente de sua natureza, época ou forma, passa a ser passível de análise pelo método diplomático.

No entanto, ao longo dos séculos, o método diplomático sofreu modificações, ao passo que os documentos se modificaram, assim como o sistema jurídico no qual são criados. Neste sentido, para a idealização de um modelo ideal de análise para a identificação dos documentos de arquivo, faz-se necessário abordar os métodos que sustentaram a disciplina nos últimos quatro séculos.

As primeiras formulações para a enunciação de um método surgiram na França, em 1681, com a publicação do tratado *De re diplomática*, de Jean Mabillon, que busca verificar a autenticidade de diplomas mantidos nos arquivos de

ordens religiosas. Confrontando e analisando diplomas inseridos em uma série cronológica ou em um conjunto específico de documentos de uma determinada época ou lugar, Mabillon cria um “método” histórico, analítico-comparativo que sustenta o estudo crítico dos documentos até os dias de hoje.

Seus elementos eram confrontados e analisados com base em um contexto pré-estabelecido. Para tanto, Mabillon (1681) definiu-os como caracteres intrínsecos e extrínsecos, descrevendo os primeiros como o conteúdo do documento que abrange, por sua vez, o estilo, o texto, os selos, a data e as assinaturas. Os caracteres extrínsecos compreendiam o material (matéria) sobre o qual eram escritos os documentos (pergaminho, papiro, tábuas de argila, chumbo, pedras) e o meio usado para inserir as escrituras ou símbolos nesse material (tinta).

Essas formulações podem ser consideradas o embrião do método diplomático que só foi efetivamente construído no século XVIII, na Áustria, por Theodor Von Sickel.

O autor mantém a divisão do documento em elementos internos e externos, aprofundando-os. Assim, os elementos externos são o material, os tipos de escritura, os selos e abreviaturas, e os comentários. Já os elementos internos apresentam uma evolução ao serem divididos em texto e protocolos, além da língua. Dessa forma, o conteúdo do documento é dividido em protocolo inicial, que engloba invocação e nome e título do autor; texto, englobando preâmbulo, nome e título do destinatário, exposição, dispositivo, corroboração e anúncio dos sinais de validação. O protocolo final, ou escatocolo é dividido em assinatura, datas e precação.

Os protocolos funcionam, portanto, como a moldura do documento. Eles são responsáveis por abrir e fechar o teor documental, e às vezes manterão a mesma fórmula quando se trata de um mesmo autor ou de um mesmo ato jurídico. Seu principal ofício dentro da estrutura do documento diplomático é fornecer os elementos de validade legal. Já o teor documental, ou o texto propriamente dito, traz o conteúdo que pode variar segundo a natureza do ato jurídico.

Vale dizer que ambos os métodos eram utilizados estritamente para verificar a autenticidade dos diplomas da Idade Média, de natureza estritamente jurídica, fosse para fins prático-jurídicos ou históricos.

A abertura do objeto de estudo da Diplomática a todos os documentos de arquivo, por Bautier, permitiu sua aproximação à Arquivística, dando origem, no final do século XX, a uma nova

abordagem que, baseando-se nos métodos aqui expostos, incorpora novos elementos para a análise dos documentos de arquivo.

Assim, o método proposto pela chamada Diplomática Arquivística, ou Diplomática Contemporânea, que tem como representante mais expressiva Luciana Duranti, recebe incorporações de novos elementos, criando um método à parte, com base naqueles propostos por Mabillon e Ficker, mas que pudesse ser aplicado aos documentos contemporâneos, principalmente àqueles gerados eletronicamente. Esse método mantém a divisão proposta por Mabillon, dos elementos internos e externos, e aquela proposta por Sickel, sobre os protocolos e texto. Assim, os elementos externos englobam meio, escrita, linguagem, sinais especiais e selos; enquanto os elementos internos englobam protocolo inicial, dividido em timbre, título do documento, datas, invocação, subscrição, inscrição, saudação, assunto, fórmula perpetuitatis e precação; o texto, dividido em preâmbulo, notificação, exposição, dispositivo e cláusulas finais e protocolo final dividido em corroboração, datas, precação, cláusula complementar, atestação, qualificação das assinaturas e notas secretariais.

Esse método sustenta, hoje, os estudos da Tipologia Documental e pode ser utilizado para a compreensão do contexto de criação dos documentos quando não é possível o conhecimento de outras fontes auxiliares para tanto. Ou seja, quando o arquivista não tem acesso ao organograma ou regimento interno da instituição, seu estudo pode partir diretamente do documento com o auxílio da análise diplomática. Dessa forma, a análise parte do documento e de suas partes para o fundo.

No entanto, alguns elementos devem ser incorporados ao método “contemporâneo”, para que ele possa ser efetivo na identificação dos documentos de arquivo.

Neste sentido, propõe-se, aqui, um tipo ideal de método diplomático, capaz de lidar com a documentação arquivística em seus mais diversos suportes, auxiliando o arquivista no momento da identificação dos documentos, bem como das funções e atividades que o geraram.

Especificamente no contexto deste artigo, entende-se por “tipo ideal” de método diplomático aquele formado a partir das formulações e dos métodos diplomáticos aqui apresentados, que funcione como uma síntese, arranjado com base nos elementos utilizáveis para analisar a forma de qualquer documento, em qualquer época, formando uma construção analítica unificada.

Nesse sentido, apresentam-se, a seguir, os elementos externos e internos da forma documental que devem ser considerados na crítica diplomática “ideal”:

Elementos externos	Elementos Internos
Material	Língua
Tipo de escrita	Estilo de linguagem
Qualidade de impressão	Protocolo Inicial (título/assunto/datas/ Invocação/titulação – nome e predicado do autor e destinatário/ saudação inicial)
Meio de registro	Texto (preâmbulo/exposição/ notificação/dispositivo/ sanção/corroboração/ anúncio dos sinais de validação)
Selos e Sinais	Protocolo Final (precação/saudação final/datas/subscrição)

Quadro 1. Elementos externos e internos da crítica diplomática ideal

Nos elementos externos, encontra-se o material, dividido em suporte (livro, pedra, tábuas de argila, disco rígido) e formato (A4, carta, impresso, digital). Os tipos de escrita compreendem aquelas hieroglíficas, cuneiformes, românicas, arábicas, góticas, etc. A qualidade de impressão visa a oferecer um melhor detalhamento das condições físicas do documento, analisando se ele possui rasuras ou deteriorações, por exemplo. Os meios de registro podem ser as tintas utilizadas quando se usa uma caneta, ou máquina de escrever, ou o software quando se trata de documentos eletrônicos. Os selos e sinais são considerados, aqui, como elementos externos por terem sido colocados após a feitura do documento, como os carimbos de autenticação, ou o registro de protocolo e número de classificação no arquivo, ou assinaturas eletrônicas, por exemplo.

Os elementos internos apresentam poucas variações com relação aos demais elementos das formulações analisadas. Aqui, a língua é o idioma no qual o documento é redigido, enquanto o estilo pode ser formal ou informal. A divisão entre texto e protocolos mantém-se, apresentando apenas algumas variações, como a inserção das datas e da saudação, tanto no protocolo inicial quanto no protocolo final. Inserem-se no protocolo inicial o assunto e o título do documento, e a titulação engloba tanto o autor e

seu predicado, quanto o destinatário do documento.

Entende-se, para a idealização do método, a participação de pelo menos três pessoas na formação do documento: autor (da ação/do documento), destinatário (da ação/do documento) e escritor (responsável pelo teor e pela articulação do texto). No entanto, vale ressaltar que todos devem estar respaldados nos conceitos de competência e responsabilidade para criar/redigir o ato.

Nesse contexto, após definir os elementos externos e internos para a análise do documento e as pessoas que contribuem para sua formação, apresenta-se, no quadro a seguir, a síntese do método diplomático idealizado:

Espécie documental

Tipo documental

Categoria Documental:

Dispositivo

Probatório

Informativo

Natureza do ato:

Público

Privado

Elementos externos:

Material

Tipo de escrita

Qualidade de impressão (visualização, integridade do documento)

Meio de registro

Selos e Sinais

Elementos internos:

Língua

Estilo de linguagem

Protocolo Inicial

(Título/Assunto/Datas/Invocação/Titulação
nome e predicado do autor e destinatário/
Saudação inicial)

Texto

(preâmbulo /exposição/notificação/dispositivo/sanção/
corroboração/ anúncio dos sinais de validação)

Protocolo Final

(precação/Saudação final/datas/subscrição)

Pessoas envolvidas na criação do documento:

Autor da ação

Autor do documento

Destinatário da ação

Destinatário do documento

Escritor

Testemunhas envolvidas

Estado de transmissão:

Pré-original

Original

Pós-original

Informações complementares:
 Fundo produtor/recebedor
 Grupo
 Série
 Notação

Quadro 2. Método diplomático idealizado

No método proposto, o primeiro elemento a ser identificado é a espécie documental, ou seja, o veículo redacional escolhido para o registro da ação. Segundo Camargo & Bellotto (1996), a espécie é a configuração que assume um documento de acordo com a disposição e natureza das informações nele contidas. Assim, a primeira coisa que o profissional da informação deve saber é qual é esse veículo, pois ele será determinante para o estudo da forma. Um exemplo de espécie que engloba tanto os documentos antigos quanto os contemporâneos é a “escritura”.

O segundo elemento, o tipo documental, é a espécie agregada da atividade que irá gerar o

documento, ou seja, seguindo o exemplo acima, quando se tem uma escritura de transação imobiliária, a “compra e venda de imóvel” corresponderá à atividade de registrar o ato de comprar e vender imóvel, e a escritura é o veículo redacional eleito e juridicamente válido para a inserção desse ato.

Em seguida é identificada a categoria documental, que diz respeito ao valor jurídico do conteúdo dos documentos. Desta feita, o documento dispositivo é aquele que nasce para fazer cumprir e dar validade à ação, enquanto o documento probatório simplesmente atesta um fato jurídico já completo, válido perfeitamente antes de sua documentação. Por último, o documento informativo esclarece as questões e/ou informações contidas em outros documentos.

A natureza do ato também é importante no momento da crítica diplomática, uma vez que determina a relação entre documento e autor. Para fins de esclarecimento, no método ideal, um documento é considerado público quando criado por uma pessoa pública, ou em seu nome, exercendo uma atividade jurídica no contexto público. Ao contrário, um documento é privado se criado por uma pessoa privada, ou em seu nome, no sistema jurídico no qual atua.

A identificação das pessoas envolvidas na criação do documento é fundamental uma vez elas são os sujeitos de deveres e direitos envolvidos no sistema jurídico no qual o documento é criado. Ao identificar os autores, destinatários, escritor e testemunhas, evidenciam-se justamente as relações entre elas e o sistema jurídico no qual atuam, fornecendo um estudo mais deta-

lhado do contexto de criação desses documentos.

O estado de transmissão determina o grau de autoridade do documento que está sendo analisado e, por essa razão, sua identificação é primordial na crítica diplomática. O documento original terá valor e peso muito maiores do que o rascunho ou cópia, por exemplo, por ser o primeiro e perfeito.

Por último, as informações complementares funcionam no contexto arquivístico, uma vez que é possível identificar os fundos/grupos/séries aos quais pertence o documento, assim como o código de classificação ou número de registro de protocolo, caso existam.

Logo, com base em um tipo ideal de método diplomático, seria possível analisar tanto os documentos antigos, quanto os documentos contemporâneos, uma vez que apresenta elementos comuns a todos os métodos analisados aqui, sem limitar cronologicamente o objeto de análise da Diplomática. Em adição, são incorporados alguns elementos da análise tipológica – como o estudo do fundo, espécie e tipo documental –, permitindo ao arquivista fazer uso do método sem maiores complicações e elementos desnecessários.

Considerando o método analisado enquanto estrutura ideal, seguindo o conceito weberiano de tipo ideal, propõe-se, aqui, ainda, um resumo desse método, que contemple tão somente os elementos necessários à análise dos documentos arquivísticos, excluindo, portanto, alguns elementos significativos somente à análise diplomática com o objetivo puramente de verificação da autenticidade dos documentos. Assim, apresenta-se, a seguir, o método para a identificação dos documentos de arquivo (Quadro 3).

Espécie documental:

Tipo documental

Categoria Documental:

Dispositivo

Probatório

Informativo

Natureza do ato:

Público

Privado

Elementos externos:

Material

Meio de registro

Elementos internos:

Protocolo Inicial (Título/Assunto/Datas/
Titulação - nome e predicado do autor e destinatário)
Texto (dispositivo)
Protocolo Final (datas/subscrição)

Pessoas envolvidas na criação do documento:

Autor da ação
Autor do documento
Destinatário da ação
Destinatário do documento
Escritor

Estado de transmissão:

Pré-original
Original
Pós-original

Informações complementares:

Fundo produtor/recebedor
Grupo
Série
Notação

*Quadro 3. Método para a identificação
de documentos de arquivo*

O método proposto oferece ao arquivista uma ferramenta para identificar as informações necessárias à reconstrução do contexto arquivístico, tendo como base o próprio conjunto documental. Tal método poderá funcionar como base metodológica para a identificação dos tipos documentais na constituição das séries no momento da identificação. Logo, o método seria eficiente na segunda fase da identificação, ou seja, no momento do estabelecimento das tipologias documentais.

5. Considerações

Buscou-se, neste artigo, apresentar uma nova ferramenta para a análise dos conjuntos documentais contemporâneos, no contexto da Arquivística. Neste sentido, a base metodológica oferecida pela Diplomática, permitiu uma aproximação de ambas as disciplinas a partir da década de 1960. Desde então, a Diplomática vem subsidiando o trabalho arquivístico no tocante à análise documental, quando o profissional da informação parte especificamente da identificação do documento, enquanto um dos objetos de estudo da identificação dos conjuntos documentais e do contexto de produção.

Para a construção dessa nova ferramenta metodológica, contou-se com os estudos metateóricos de Ritzer e com o conceito de tipo ideal formulado por Weber. Assim, os estudos de Ritzer nos permitiu enxergar o percurso histórico e conceitual trilhado pela Diplomática, bem como a evolução e constituição de seu método de análise documental. Da mesma maneira, o con-

ceito de tipo ideal de Weber, nos permitiu, após a análise metateórica, formular um método com base em todos os outros estudados, e considerá-lo ideal para analisar a documentação contemporânea.

É importante destacar, ainda, que, embora o método de análise proposto seja apenas idealizado, ele pode contribuir como uma ferramenta didática ao ensino da Diplomática Contemporânea nos cursos de Arquivologia, uma vez que fornece os elementos de análise para todos os documentos, de uma maneira mais clara e simplificada, além de imediata.

Dessa forma, abre-se espaço para os estudos posteriores que possam se dedicar à análise do “método ideal” ora proposto, às novas tipologias documentais que são constantemente criadas no universo documental, fortemente permeado pelos avanços tecnológicos.

Referências

- Bates, Marcia (2005). An introduction to metatheories, theories, and models. // Fischer Karen; Erdelez, Sandra; Mckechnie, Lynne (Eds). Theories of information behavior. Medford: Information Today, 2005, 1-24.
- Bautier, Robert-Henri (1961). Leçon d'ouverture du cours de diplomatique à L'Ecole des chartes. // Bibliothèque de l'Ecole des Chartes. 1961. 194-225.
- Camargo, Ana Maria de Almeida; Bellotto, Heloísa Liberalli (1996). Dicionário de Terminologia Arquivística. São Paulo: Secretaria da Cultura, 1996.
- Mabillon, Jean (1681) De re diplomatice libri sex. 1681. Paris.
- Ritzer, George (1991). Metatheorizing in Sociology. Lexington: Lexington Books, 1991.
- Rodrigues, Ana Célia (2015). Identificação como requisito metodológico para a gestão de documentos e acesso a informações na administração pública brasileira. // Ciência da Informação. 42:1(2015) 64-80.
- Sickel, Theodor Von (1867). Acta regum et imperatorum Karolinorum digesta et enarrata. Wien: Duck und Verlag Von Carl Gerold's Sohn, 1867.
- Svenonius, Elaine (2004). The epistemological foundations of knowledge representation. // Library Trends. 52:03 (Winter, 2004) 571-587.
- Tennis, Joseph T (2008). Epistemology, Theory, and Methodology in Knowledge Organization: toward a classification, metatheory, and research framework. // Knowledge Organization. 35:2 (2008)102-112.
- Tognoli, Natália Bolfarini (2013). A construção teórica da Diplomática: em busca de uma sistematização de seus marcos teóricos como subsídio aos estudos arquivísticos. Tese (doutorado). Programa de Pós-Graduação em Ciência da Informação – Faculdade de Filosofia e Ciências, Unesp, Marília, 2013.
- Weber, Max (1904). Objectivity in Social Science and Social Policy. // Shils, E. A.; Finch, H.A. (Eds). The Methodology of the Social Sciences. New York: Free Press, 1949. 77128.

Enviado: 2015-05-01. Segunda versão: 2015-09-01.
Aceptado: 2015-09-01.

Procedimiento de evaluación

Evaluation process

1. Introducción

El procedimiento de evaluación en Scire es el de revisión por pares mediante el sistema de doble ciego —los revisores no conocen el nombre de los autores, ni viceversa— semejante al de otras revistas científicas, aunque tiene algunas peculiaridades que permiten ajustarlo a las condiciones específicas de su campo científico y de sus lectores.

2. Objetivos

El objetivo del procedimiento de evaluación es que se puedan admitir con rapidez los trabajos buenos, mejorar los que pueden ser susceptibles de aceptación —esto es, que, siendo interesantes, contienen errores o lagunas que deben ser corregidas— y rechazar los que no son adecuados para la revista por su tema o calidad, redirigiéndolos hacia otras publicaciones adecuadas o haciendo sugerencias al autor para la posible reutilización y reorientación de su trabajo.

El procedimiento de revisión tiene que tener en cuenta a los autores de los artículos que revisa, con los que el revisor debe intentar ser justo, pero también debe responder ante los lectores y los demás autores que compiten por el espacio de la revista, así como, de forma más general, con el proyecto editorial, la propia comunidad científica y el conjunto de la sociedad.

3. Procedimiento

Cada artículo requiere al menos la opinión favorable de al menos dos revisores y cada nota breve de al menos uno. Si el director de la revista estima que el artículo entra dentro de los objetivos de Scire —que se encuentran públicamente disponibles en cada número y en la página web de la revista—, lo envía a dos revisores —o, si es necesario, a más— elegidos por el director. Los evaluadores pueden ser miembros del consejo científico, revisores propuestos por los miembros del consejo científico con anterioridad o con motivo del artículo en cuestión, o científicos de reconocido prestigio propuestos por el autor o autores del artículo. La elección se hace según su conocimiento del tema y el número de trabajos que han evaluado hasta el momento (para equilibrar la carga de trabajo entre ellos).

Si el director estima sin lugar a dudas que el tema del artículo no se adecua al de la revista, se lo comunica al autor con una explicación suficiente, y, si le es posible, le sugiere otra revista que pueda ser más adecuada. De esta forma, se ahorra tiempo del autor y de los revisores.

La agilidad del procedimiento de revisión es un aspecto fundamental, por lo que el director solicita a los revisores que envíen su valoración en dos semanas, y que, en caso contrario, renuncien expresamente al proceso de revisión. Sin embargo, si el revisor no realiza ninguna de estas acciones, el proceso se retrasa sin remedio. En este sentido, el autor debe tener en cuenta que el proceso de revisión tiene un carácter voluntario. La decisión de aceptar el trabajo depende, salvo en casos excepcionales, de la opinión de los revisores. En los casos en que no están de acuerdo, la decisión es tomada por el director de la revista, que, excepcionalmente, puede decidir enviar el artículo a un nuevo revisor.

4. Criterios de evaluación

Se solicita a cada revisor que indique su valoración sobre los siguientes aspectos en una escala del 1 al 5 (1 muy cuestionable, 2 marginal, 3 suficiente, 4 sólido, 5 excelente):

1. *Tipo de contribución*: Se indica si se trata de una contribución de tipo teórico o metodológico, un estado de la cuestión, una comunicación de resultados de investigación, una experiencia específica, una exposición con valor introductorio o docente sobre el tema o una noticia de interés. Si es una mezcla de varios tipos, es posible marcar más de uno.

2. *Adecuación del tema de la revista*: Aunque al enviar el artículo al revisor el director de la revista ya ha realizado un juicio sobre su adecuación al enfoque de la revista Scire, el revisor puede también valorar este aspecto, comparando el tema del artículo con la declaración de objetivos y alcance de la revista, disponible en cada número y en la página web.

3. *Interés e importancia del tema*: Grado en el que el artículo aborda una cuestión de gran interés para el área científica de Scire, ya sea tradicionalmente, en el momento actual o en el futuro previsible.

4. *Actualidad del tema*: Grado en el que el tema se sitúa en los frentes actuales de investigación. No es un aspecto imprescindible, pero sí se considera informativo.

5. *Novedad y originalidad teórica de la aportación*: Grado en el que el artículo aporta una novedad teórica a su campo de investigación; esto es, un nuevo problema, un nuevo enfoque, un nuevo modelo, una nueva relación o ley, una nueva hipótesis, un nuevo concepto o una interpretación o modificación original de otros anteriores. Debe tratarse lógicamente de una aportación relevante y fundamentada.

5. *Novedad y originalidad metodológica de la aportación*: Grado en el que el artículo aporta una novedad metodológica a su campo de investigación; esto es, un nuevo procedimiento de estudio, de análisis o de prueba, aunque sea a un problema conocido.

7. *Aplicabilidad práctica*: Grado en el que las ideas del artículo pueden servir para la mejora de procedimientos y actividades prácticas y, en definitiva, para el desarrollo y la innovación.

8. *Rigor metodológico y validez de los resultados*: Grado de coherencia, exactitud, precisión y cuidado con el que se aborda el tema de estudio; validez y actualidad de la metodología utilizada; rigor en la selección de las muestras y en la aplicación de la metodología a las mismas; grado de replicabilidad del estudio.

9. *Rigor en el estado de la cuestión y el reconocimiento del trabajo previo*: Hasta qué punto se ha realizado un estado de la cuestión suficiente y actualizado y se han manejado las referencias pertinentes para el problema abordado.

10. *Claridad expresiva, orden y facilidad de lectura*: El artículo debe estar redactado de forma clara y bien ordenada, sin redundancias ni ideas fuera de contexto; y debe ser fácilmente entendible por cualquier lector potencial de Scire. No debe faltar información necesaria ni deben existir lagunas sobre el problema abordado, el método, los resultados, la discusión, las conclusiones o las recomendaciones. Si el artículo contiene aparato matemático, debe ser comprensible por un lector no especialista.

11. *Adecuación de gráficos, ilustraciones, tablas y apéndices*: Los gráficos, ilustraciones, tablas y apéndices deben ser necesarios y relevantes, estar bien resueltos técnicamente y disponer de una leyenda suficientemente explicativa.

13. *Conformidad de las referencias a las normas de la revista*: Grado en el cuál las referencias del artículo son suficientes y se conforman a las normas de la revista Scire.

5. Familiaridad del revisor con el tema

Aunque el director de la revista envía los artículos para su revisión intentando que coincidan con el campo de especialidad del revisor, pueden producirse desajustes. El revisor puede rechazar la revisión sugiriendo o no un revisor alternativo, o puede efectuar la revisión indicando de 1 (mínimo) a 5 (máximo) su grado de familiaridad con el tema del artículo evaluado. Si no rellena esta información, se entenderá que su grado de familiaridad es alto (4) o muy alto (5).

Esta información sirve para que el director de la revista pueda estimar la necesidad de enviar el artículo a un nuevo revisor en caso de conflicto de pareceres entre los revisores.

6. Valoración final

El revisor puede realizar cinco tipos de propuestas finales:

1. Aceptar el artículo para su publicación sin restricciones ni cambios, porque es un trabajo de alta calidad.

2. Aceptar el artículo para su publicación una vez realizadas las correcciones indicadas, bien sea incondicionalmente, al ser un trabajo de calidad, o bien de forma condicionada a que haya espacio disponible en el número, por ser un trabajo de calidad marginal. Las correcciones solicitadas se indican en las instrucciones para el autor, son de carácter menor y su comprobación puede ser delegada en el comité editorial.

3. Someter a reevaluación el artículo una vez realizadas las correcciones indicadas en las instrucciones para el autor, porque son de importancia y amplitud, y deben ser comprobadas por el revisor original u otro revisor.

4. Rechazar el artículo por las razones indicadas en las instrucciones para el autor, donde el revisor especificará con claridad y amabilidad las razones por las que su trabajo debe ser rechazado y, si le resulta posible, realizará sugerencias sobre su mejora o lugares alternativos de publicación.

5. El revisor rechaza evaluar el artículo por su falta de familiaridad con el tema o por razones que especifica en los Comentarios confidenciales para el equipo directivo.

La labor de los revisores es fundamental para que una publicación científica realice su labor de manera adecuada. Scire agradece de forma muy especial el tiempo y esfuerzo que dedican, y reconoce su trabajo publicando periódicamente sus nombres.

Instrucciones para la presentación de trabajos

Instructions for authors

1. Cuestiones generales

Scire admite artículos con estados de la cuestión, discusiones de carácter teórico y metodológico, e informes de investigación sobre proyectos terminados, así como noticias sobre proyectos en curso, experiencias, notas y reseñas, que traten de la representación, normalización, tratamiento, recuperación y comunicación de la información y el conocimiento.

Las lenguas de publicación de Scire son el español, el inglés y el portugués.

Los artículos son sometidos a proceso de revisión por pares según el sistema de doble ciego, aunque los autores pueden proponer revisores libremente. El procedimiento es público, y puede ser leído en el apartado anterior a las instrucciones.

Los autores se comprometen a presentar y discutir públicamente y presencialmente sus trabajos en el congreso Scire correspondiente al año en curso.

Los trabajos deben ser originales y tendrán una extensión máxima de 20.000 caracteres en el caso de los artículos y de 12.000 caracteres en el caso de las noticias, experiencias, notas y reseñas. Excepcionalmente, estas cifras podrán ajustarse en virtud del interés del trabajo, aunque siempre con el acuerdo previo del editor.

Los autores deben presentar en una hoja aparte sus datos personales completos: nombre, cargo, dirección, teléfono, fax y correo electrónico.

Los artículos han de presentarse en tres copias impresas, debidamente maquetadas según la plantilla disponible en www.Scire.org con su gráficos incluidos. Irán precedidos de sus títulos, resúmenes informativos y palabras clave, todos ellos en inglés y español. Además, se deberá enviar por correo electrónico o en un disco compacto una copia en archivo de ordenador en formato Microsoft Word o RTF.

Todos los apartados, incluidos las notas, referencias y apéndices irán numerados secuencialmente. Los subapartados —cuyos títulos serán claros y sintéticos— deberán tener la forma “1.1. “, “1.2. “, “1.2.1. “ y así sucesivamente. No numere los resúmenes en español y en inglés, ni los apartados de notas, agradecimien-

tos y referencias. Evite un rango de encabezamientos superior a tres.

No se deben relatar datos analíticos repetitivos ni elencos de recursos en el texto principal. Por favor, preséntelos en forma de tablas o, si fuera imposible, en forma de apéndice.

El editor podrá realizar cambios de estilo para adecuar el trabajo a los requisitos de la revista.

Se recuerda al autor que la violación de derechos intelectuales y códigos éticos es un comportamiento inaceptable y puede constituir un grave delito.

2. Formatos de texto

No utilice subrayados, negritas ni versales.

Utilice mayúsculas sólo para acrónimos. No ponga puntos y espacios entre las letras que los forman.

Las cursivas se deben utilizar tan sólo para resaltar términos en otras lenguas, títulos de obras, palabras clave muy importantes que podrían servir a modo de títulos de un párrafo o conjunto de párrafos, y, con parquedad, para enfatizar los demás términos y frases a gusto del autor.

No utilice listas con salto de línea o de párrafo. Redacte párrafos completos según las reglas y los recursos gramaticales disponibles en su lengua de redacción. Puede indicar orden en las enumeraciones por medio de números o letras entre paréntesis, como, por ejemplo, (1), (2)...

3. Figuras y gráficos

Las figuras deben presentarse también en ficheros aparte en formato EPS o PDF con resolución de impresión de alta calidad (600 ppp) y, además, en el formato de la aplicación original con la que se creó. Tenga cuidado si genera imágenes a partir de la pantalla de su ordenador. Si recurre a imprimir pantalla en Windows obtendrá solo una resolución de 72 ppp. En ese caso, utilice una pantalla lo más grande posible, amplíe la ventana al máximo, copie con la tecla “impr pant”, pegue en el programa “Paint” y guarde el fichero; compruebe luego los resultados. Existen programas que permiten obtener imágenes de pantalla a una resolución acepta-

ble. En el cuerpo del artículo las imágenes portarán epígrafes indicativos de su contenido precedidos de la palabra “Figura” y un número correlativo (por ejemplo, “Figura 1. “).

4. Tablas

Deben realizarse con el mismo procesador de textos que el artículo. En texto portarán epígrafes indicativos de su contenido precedidos de la palabra “Tabla” y un numeral romano correlativo (por ejemplo, “Tabla I. “).

5. Notas

Las notas explicativas deben usarse excepcionalmente. No utilice el sistema automático de su procesador de textos. Refiéralas en texto tecleando su número secuencial entre paréntesis, e inclúyalas al final del trabajo en un apartado titulado “Notas” situado inmediatamente antes del dedicado a las “Referencias”, en orden secuencial y precedida cada una de ellas de su correspondiente número entre paréntesis seguido de tabulador. Las citas bibliográficas se realizarán como en el resto del texto; y la referencia completa se incluirá en el apartado de “Referencias”. Las páginas web se consideran referencias bibliográficas y deben tratarse como tales, evitando su cita como nota.

6. Citas bibliográficas en texto

Las referencias bibliográficas en texto se denotarán por el apellido del autor, el año del trabajo y, si es necesario, el número de página. Si conviene, el nombre y el apellido del autor podrán quedar fuera del paréntesis. Cuando se sucedan seguidamente varias referencias al mismo autor y obra, se podrá usar la abreviatura *ibidem*. Si dos autores de la lista de referencias poseen el mismo primer apellido, se referirán por sus dos apellidos y, si tienen ambos apellidos iguales, por el nombre también. Si dos o más trabajos del mismo autor son del mismo año, se distinguirán mediante letras del alfabeto secuenciales en minúsculas pospuestas a la fecha sin espacio. Si hay varias referencias dentro del paréntesis, se separarán por punto y coma si son de diferentes autores, y coma si son del mismo.

Así, por ejemplo, “Es una teoría sustentada por varios autores (Gallego, 1975; Fernández y Alonso, 1993) [...]”. “Según Jaime Pérez (1993, p. 24) se establecen [...]”. “El mismo autor (*ibidem*, p. 27) hace constar la conveniencia de [...]”. “Se han detectado diferencias en la replicación de la experiencia (Menéndez, 1994a, 1994b; Menéndez y Alonso, 1997)”.

7. Formato de las referencias bibliográficas

Las referencias bibliográficas se incluirán por orden alfabético al final del artículo bajo el título “Referencias”. Deberán seguir la norma UNE 50104:1996/ISO 690:1987, y, en cuanto a la puntuación, aproximarse lo más posible a las ISBD abreviadas, aunque sin espacio delante de punto, coma y punto y coma. Las referencias a páginas web deben incluirse también en este apartado, no en el texto ni en nota. Refiera adecuadamente en el texto del artículo toda la bibliografía utilizada; no incluya en el apartado referencias documentos que no haya incluido en el texto. Se presenta a continuación el formato para los principales tipos de materiales, con los elementos que pueden darse o no en cursivas, y seguidos de un ejemplo.

7.1. Monografías, congresos, informes, normas y tesis

Apellido, Nombre; ApellidoN, NombreN (función) (Año). Título: subtítulo. Edición. Lugar de publicación: editor; lugar de publicación: editor, fecha. ISBN número. Tesis doctoral.

Delclaux, Isidoro; Seoane, Julio (1982). *Psicología cognitiva y procesamiento de la información: teoría, investigación y aplicaciones*. Madrid: Ediciones Pirámide, 1982.

7.2. Artículos de publicaciones periódicas

Apellido, Nombre; ApellidoN, NombreN (función) (Año). Título: subtítulo. // Título de la publicación periódica. ISSN numero. Volumen: número (Fecha de publicación) primera página-última página.

Ellis, David (1992a). The physical and cognitive paradigms in Information Retrieval Research. // *Journal of Documentation*. 48:1 (March 1992) 45-46.

7.3. Capítulos de una monografía, de un informe o contribuciones a un congreso

Apellido, Nombre; ApellidoN, NombreN (función) (Año). Título: subtítulo. // Apellido, Nombre; ApellidoN, NombreN (función) (Año). Título: subtítulo. Edición. Lugar de publicación: editor; lugar de publicación: editor, fecha. ISBN número. Primera página-última página.

Markey, Karen (1990). Keyword searching in an online catalog enhanced with a library classification. // Bengtson, Betty G.; Hill, Janet Swan (eds.). *Classification of library materials: current and future potential for providing access*. New York: Neal-Shuman Publishers, 1990. 99-125.

Smith, Ph. J.; Beghtol, C.; Fidel, R.; Kwasnik, B. H. (eds.) (1993). *Proceedings of the 4th ASIS SIG/CR Classification Research Workshop*: Columbus, OH, Oct.24, 1993. Silver Spring, MD.: American Society for Information Science, 1993.

7.4. Páginas web y recursos en línea

Se tratarán como monografías, artículos o capítulos según corresponda y se añadirá al final detrás de punto el URL y la fecha de consulta entre paréntesis:

Referencia. URL (Fecha de consulta).

Sagredo Fernández, Félix; Espinosa Temiño, María Blanca (2000). *Del libro, al libro electrónico-digital. // Cuadernos de Documentación Multimedia*. 9 (2000). <http://www.ucm.es/info/multidoc/multidoc/revista/num9/cine/sagredo.htm> (2001-01-22).

8. Apéndice I. El resumen

Un resumen proporciona información sobre objetivos, alcance, metodología, resultados, conclusiones y recomendaciones. Por alcance se entiende la determinación de hasta qué punto se han tratado los objetivos. Los resultados, conclusiones y recomendaciones se distinguen unos de otros así: los resultados son las informaciones o datos específicos y concretos que se obtienen de la aplicación de la metodología; las conclusiones se derivan de los resultados bien al interpretarlos en el marco de conocimientos teóricos previos o bien al crear un marco de interpretación nuevo a partir de ellos a través de un proceso de abstracción; finalmente, las recomendaciones se refieren a los nuevos problemas y vías de estudio que plantean a la vista de los resultados y conclusiones.

El resumen se debe redactar de forma clara y concisa, sin redundancias ni perífrasis ni información difícil de interpretar, como neologismos o abreviaturas que no se desarrollan.

Es muy conveniente que su primera frase contenga en primer lugar la mención del tipo de documento. Por ejemplo, "Estado de la cuestión sobre la aplicación del modelo europeo de calidad a las bibliotecas universitarias", "Proyecto de aplicación del MARC21 a la catalogación de páginas web institucionales", etc.

9. Apéndice II. La redacción del artículo científico: algunos aspectos obvios frecuentemente descuidados

9.1. Sistemática

El artículo científico debe redactarse de forma muy organizada, con una estructura textual firme. En los informes de resultados de proyectos, se suele utilizar el esquema objetivos-metodología-resultados- conclusiones- recomendaciones. En los estados de la cuestión se aborda el asunto en orden cronológico o bien sistemático, según la ontología propia del área científica en cuestión; sigue la discusión, la síntesis, las conclusiones y las recomendaciones. Los artículos críticos suelen seguir el esquema tesis-síntesis-antítesis. Otros tipos documentales pueden exigir otras estructuras textuales, pero, en cualquier caso, dicha estructura debe ser explícita y aparente.

9.2. Claridad y concisión

El lenguaje del artículo científico debe ser claro y conciso. Excepcionalmente, en los trabajos de tipo ensayístico las personas con un alto dominio del lenguaje escrito pueden permitirse la utilización de recursos literarios.

9.3. Pertinencia y parsimonia

Debe evitarse proporcionar información ya conocida, salvo si constituye el punto central de la discusión. A ella cabe referirse por medio de las citas o, si estas no son necesarias, suponerla patrimonio de la comunidad profesional y científica. Sistematizar conocimiento es una función que se realiza solo en un tipo de artículo científico —la revisión o estado de la cuestión—; en los demás casos es contraproducente, aunque se deben citar las fuentes y las obras clave en la introducción del trabajo.

9.4. Civilidad

Es necesario citar a otros autores que han trabajado sobre el tema, demostrando que se conoce la literatura previa y que el trabajo se inserta civilizadamente en un determinado contexto científico.

Índice de autores

Author index

Alonso Berrocal, José Luis, 57
Bolfarini Tognoli, Natália, 75
Castro Morales, Federico, 69
Dal' Evedove, Paula Regina, 49
dos Santos, André Ynada, 13
Ferreira, Marcio, 27
Figuerola, Carlos G., 57
Fujita, Mariângela Spotti Lopes, 49
García López, Fátima, 69
Gil Urdiciain, Blanca, 35

Guimarães, José Augusto Chaves, 13
Magro, José L., 27
Martínez-Ávila, Daniel, 27
Matos, Daniela Fernanda de Oliveira, 13
Morales García, Ana María, 69
Ohlenschläger Ohlenschläger, Karin, 69
Ríos Hilario, Ana B, 35
Sales, Rodrigo de, 13
Zazo Rodríguez, Ángel F., 57

Índice de materias en español

Subject index in Spanish

América del Norte, 13
Análisis de contenido, 13
Análisis de dominio, 13
Archivística, 75
Archivos, 69
Arte electrónico, 69
Aspectos éticos, 27
Brasil, 13, 49
Catálogo de datos enlazados Mannheim, 35
Clasificaciones bibliográficas, 27
Concepto, 13
Conjunto de datos, vocabularios de valores y conjuntos de elementos de metadatos (VO-CABDATASET), 35
Continuidad digital, 69
Cuadro de clasificación, 69
Cultura digital, 69
Datos abiertos y vinculados (LOD), 35
Diagrama de la nube LOD, 35
Diplomática, 75
Documentos de archivo, 75

Edición colaborativa, 57
Espacio P, 69
España, 13
Estudios de usuarios, 57
Francia, 13
Galerías, bibliotecas, archivos y museos (GLAM), 35
Garhel, Pedro, 69
Identificación documental, 75
Indización, 49
ISKO, 13
Metateoría, 75
Metodología, 27
Organización del conocimiento, 13, 27, 49, 57
Performances, 69
Políticas de indización, 49
Representación del conocimiento, 49
Sociedad Internacional para la organización del conocimiento, 13
Teoría Crítica de Raza, 27
Wikipedia, 57

Índice de materias en inglés

Subject index in English

Archival records, 75
Archival science, 75
Archives, 69
Bibliographic classifications, 27
Brazil, 13, 49
Classification scheme, 69

Cloud diagram LOD, 35
Collaborative edition, 57
Concept, 13
Content analysis, 13
Critical Race Theory, 27

- Datasets, Value Vocabularies, and Metadata Element Set (VO-CABDATASET), 35
- Digital arts, 69
- Digital continuity, 69
- Digital culture, 69
- Diplomatics, 75
- Document identification, 75
- Domain analysis, 13
- Espacio P, 69
- Ethical aspects, 27
- France, 13
- Galleries, libraries, archives and museum (GLAM), 35
- Garhel, Pedro, 69
- Indexing, 49
- Indexing policy, 49
- International Society for Knowledge Organization, 13
- ISKO, 13
- Knowledge organization, 13, 27, 49, 57
- Knowledge representation, 49
- Linked Open Data (LOD), 35
- Mannheim Linked Data Catalog, 35
- Metatheory, 75
- Methodology, 27
- North America, 13
- Performances, 69
- Spain, 13
- User studies, 57
- Wikipedia, 57