

SCIRE

Representación y organización del conocimiento

SCIRE

Representación y organización del conocimiento

Vol. 26, n.º 1, enero-junio 2020

ISSN 1135-3716

Scire:

knowledge representation and organization

Vol. 26, n. 1, January-June 2020

Ibersid:

Red de Investigación
en Sistemas de Información
y Documentación

© 2020 Francisco Javier García Marco

Reservados todos los derechos. Prohibida la reproducción por cualquier medio de la totalidad o parte de la presente edición con fines comerciales 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.

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.

Se permite expresamente el autoarchivo de la versión electrónica por parte de los autores en sus repositorios personales e institucionales. Los autores pueden reproducir sus textos y figuras sin permiso expreso del editor indicando la fuente de publicación, y conservan los derechos de su original con la obligación de citar la presente edición.

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

Edita: Ibersid® con la colaboración de Prensas de la Universidad 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, ESCI, 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 762639. 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, ESCI, 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 762639. 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, UNESP, Brasil | Francisco Javier Martínez Mendez, Universidad de Murcia, España |
| José Augusto Chaves Guimarães, UNESP, Brasil | Francisco Javier García Marco, Universidad de Zaragoza, España | Álvaro Quijano Solís, Colegio de México, México |

Consejo científico / Scientific council

| | | |
|--|---|--|
| Dr. Isidro Aguillo Caño, IPP- CSIC, España | Vicente Pablo Guerrero Bote, Univ. de Extremadura, España | María del Carmen Negrete Gutiérrez, Universidad Autónoma de México |
| Tomás Baiget, EPI S. A., España | Víctor Herrero Solana, Univ. de Granada, España | José Luis Otal, Universidad Jaime I de Castellón, España |
| José Luis Bonal Zazo, Univ. de Extremadura, España | José María Izquierdo Arroyo, Universidad de Murcia, España | Manuel José Pedraza Gracia, Universidad de Zaragoza, España |
| Mercedes Caridad Sebastián, Universidad Carlos III de Madrid, España | María Pilar Lasala Calleja, Universidad de Zaragoza, España | María Pinto Molina, Universidad de Granada, España |
| Alberto Carreras Gargallo, Universidad de Zaragoza, España | Alfonso López Yepes, Universidad Complutense de Madrid, España | Gloria Ponjuán Dante, Universidad de La Habana, Cuba |
| Emilia Currás Puente, Universidad Autónoma de Madrid, España | José López Yepes, Universidad Complutense de Madrid, España | Blanca Rodríguez Bravo, Universidad de León, España |
| Constança Espelt Busquets, Universidad de Barcelona, España | Pedro Marijuán Fernández, Universidad de Zaragoza, España | José Vicente Rodríguez Muñoz, Universidad de Murcia, España |
| María Eulalia Fuentes Pujol, Universidad Autónoma de Barcelona, España | Bonifacio Martín del Brío, Universidad de Zaragoza, España | Adelaida Román Román, CINDOC (Madrid), España |
| Fernando Galindo Ayuda, Universidad de Zaragoza, España | Miguel Ángel Marzal García- Quismondo, Universidad Carlos III de Madrid, España | Francisco José Ruiz de Mendoza Ibáñez, Universidad de La Rioja, España |
| Blanca Gil Urdiciáin, Universidad Complutense de Madrid, España | José Antonio Moreira González, Universidad Carlos III de Madrid, España | Félix Sagredo Fernández, Universidad Complutense de Madrid, España |
| Isidoro Gil Leiva, Universidad de Murcia, España | Purificación Moscoso Castro, Universidad de Alcalá, España | Elías Sanz Casado, Universidad Carlos III de Madrid, España |
| Alan Gilchrist, Cura Consortium, Reino Unido | Félix Moya Anegón, Universidad de Granada, España | Carlos Serrano Cinca, Universidad de Zaragoza, España |

Revisores externos del número / External reviewers in this issue

Agradecemos la colaboración de José Luis Alonso Berrocal, Miguel Ángel Marzal García-Quismondo, Daniel Martínez-Ávila, María Jesús Martínez Pestaña, Orides Mezzaroba, Blanca Rodríguez Bravo, Aires Rover.

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

Tabla de contenidos en español.....9

Tabla de contenidos en inglés.....11

Artículos

*Evolución teórica del pensamiento
de Dagobert Soergel sobre los tesauros
(1972-2017)*

Mariângela Spotti Lopes FUJITA
Luciana Beatriz Piovezan RIO BRANCO
Roberta Caroline Vesú ALVES.....13

*Tendencias en procesos y sistemas
de organización del conocimiento de textos
narrativos de ficción*

Larissa de Mello LIMA
Patrícia de ALMEIDA
João Batista Ernesto de MORAES
Daniel MARTÍNEZ-ÁVILA
Mariana da Silva CAPRIOLI
Luana Calcete Vaz TENÓRIO.....27

*Herramientas y métricas aplicadas
al marketing editorial: del big data
a la inteligencia artificial*

Antonia SALVADOR BENÍTEZ
Juan Miguel SÁNCHEZ VIGIL.....35

*Herramientas, procesos y productos
de representación temática: un estudio
de las actas de los congresos de ISKO-Brasil
(2012-2019)*

Lais Pereira de OLIVEIRA
María Cláudia Cabrini GRÁCIO
Daniel MARTÍNEZ-ÁVILA.....47

*La Documentación al servicio de proyectos
I+D+i: la experiencia del centro
de documentación de Cibermov*

Rubén DOMÍNGUEZ-DELGADO
María-Ángeles LÓPEZ-HERNÁNDEZ.....57

*Avances en sistemas jurídicos de
recuperación de documentos*

Fernando GALINDO AYUDA.....63

Procedimiento de evaluación.....75

Instrucciones para la presentación de trabajos.....77

Índice de autores.....80

Índice de materias en español.....80

Índice de materias en inglés.....80

Tabla de contenidos en inglés

Table of contents in English

| | |
|---|----|
| <i>Table of contents in Spanish</i> | 9 |
| <i>Table of contents in English</i> | 11 |

Articles

*Theoretical evolution of thesauri:
a diachronic study of Dagobert Soergel
publications (1972-2017)*

Mariângela Spotti Lopes FUJITA
Luciana Beatriz Piovezan RIO BRANCO
Roberta Caroline Vesú ALVES..... 13

*Trends in Knowledge Organization Processes
and Systems for fictional narrative texts*

Larissa de Mello LIMA
Patrícia de ALMEIDA
João Batista Ernesto de MORAES
Daniel MARTÍNEZ-ÁVILA
Mariana da Silva CAPRIOLI
Luana Calcete Vaz TENÓRIO..... 27

*Tools and metrics in editorial marketing: from
big data to artificial intelligence*

Antonia SALVADOR BENÍTEZ
Juan Miguel SÁNCHEZ VIGIL..... 35

*Tools, processes, and products of subject
representation: a study of the ISKO-Brazil
conference proceedings (2012-2019)*

Lais Pereira de OLIVEIRA
Maria Cláudia Cabrini GRÁCIO
Daniel MARTÍNEZ-ÁVILA..... 47

*Library and Information Science at the service
of R&D projects: the experience of the
Cibermov documentation centre*

Rubén DOMÍNGUEZ-DELGADO
María-Ángeles LÓPEZ-HERNÁNDEZ..... 57

*Progress in juridical systems for document
retrieval*

Fernando GALINDO AYUDA..... 63

Evaluation process

75

Instructions for authors

77

Author index

80

Subject index in Spanish

80

Subject index in English

80

Evolução teórica sobre tesouro: estudo diacrônico das publicações de Dagobert Soergel (1972-2017)

Evolución teórica del pensamiento de Dagobert Soergel sobre los tesauros (1972-2017)

Theoretical evolution of thesauri: a diachronic study of Dagobert Soergel publications (1972-2017)

**Mariângela Spotti Lopes FUJITA, Luciana Beatriz Piovezan RIO BRANCO,
Roberta Caroline Vesú ALVES**

Departamento de Ciência da Informação da Faculdade de Filosofia e Ciências da UNESP, Campus de Marília, Universidade Estadual Paulista – UNESP, Brasil, mariangela.fujita@unesp.br, lbpiovezan@gmail.com, robertavesu@gmail.com

Resumen

Estudio de la evolución de la teoría acerca de los tesauros en el pensamiento del investigador Dagobert Soergel. Se analizó el discurso de Soergel en artículos que publicó sobre tesauros mediante la metodología del análisis de contenido de Bardin en muestra diacrónica. En particular se analizaron las categorías de construcción, funciones, mantenimiento, tecnologías y usos de los tesauros. Los artículos analizados muestran que Soergel se preocupa de los avances científicos y metodológicos de los tesauros en diferentes entornos tecnológicos. El investigador es innovador y sigue las tendencias sobre construcción, funciones, mantenimiento, tecnologías y usos de los tesauros, que apuntan a su desarrollo estructurado y su uso en nuevos ambientes tecnológicos y con interoperabilidad.

Palabras clave: Soergel, Dagobert. Tesauros. Evolución. Teoría.

1. Introdução

Os tesauros possuem uma evolução influenciada pelo contexto de uso, pelas tecnologias, como também pela diversidade cultural que se expressa pela linguagem.

As pesquisas sobre tesouro necessitam de investigações sobre seus fundamentos teóricos e metodológicos, para observar evolução técnica ou teórica e verificar aspectos negligenciados. Estudos de revisão de literatura temáticos são imprescindíveis pela análise evolutiva que se obtém e destaque de autores mais influentes. Por outro lado, a análise evolutiva de uma temática pode ser investigada pela trajetória de autores influentes na revisão de literatura e, sobretudo, na própria evolução do tema em questão.

Aspectos que fazem a diferença na evolução de tesauros podem ser explicados ou antecipados em publicações de Soergel, como por exemplo,

Abstract

Study on the evolution of thesaurus theory in Dagobert Soergel's discourse. The articles he published on the subject were analyzed using Bardin's content analysis with a diachronic sample. The following categories were considered: construction, functions, maintenance, technologies and uses of thesaurus in the researcher's discourse. The articles showed that Soergel is concerned with scientific and methodological improvement of thesauri in different technological environments. The researcher is innovative and follows the more innovative trends on the construction, functions, maintenance, technologies and uses of thesauri, which and points to their structured development and use in new interoperable technological environments.

Keywords: Soergel, Dagobert. Thesauri. Evolution. Theory.

relacionamento entre termos, software de construção e manutenção, o uso em ambientes tecnológicos e digitais, a interoperabilidade entre diferentes vocabulários, entre outros. Consideramos que as pesquisas e publicações do pesquisador Dagobert Soergel têm influência na elaboração de normas sobre tesauros, como por exemplo, na norma ANSI/NISO Z39.19-2005.

Dagobert Soergel nasceu em 7 de janeiro de 1940, em *Freiburg*, Alemanha, se formou em Física, em 1960, pela *University of Freiburg* e se tornou Mestre em Matemática e Física pela mesma universidade em 1964, depois se tornou Ph.D. em Ciências Políticas, também pela *University of Freiburg*. Atualmente é professor da área de Ciência da Informação no *Department of Library and Information Studies*, da *University of Buffalo*, em *New York, US*. Atua, também, como Professor Emérito em estudos da informação no *College of Information Studies*, na *University of*

Maryland. É membro da sociedade científica *International Society of Knowledge Organization (ISKO)* desde sua fundação.

Vale destacar que Soergel apresenta como interesse de pesquisa os temas: Armazenamento e recuperação de informação; Ontologias; Classificações; Taxonomias; Tesouros; e Design orientado para o usuário (Soergel, 2005).

A análise da evolução científica e metodológica de tesouros pela visão privilegiada de Dagobert Soergel é a proposta deste artigo com o objetivo de analisar o discurso do pesquisador em suas publicações. Para isso, foi realizada análise de conteúdo de Bardin em uma amostra diacrônica de suas publicações relacionadas à tesouros.

2. Tesouros

Os tesouros são instrumentos de controle de vocabulário com objetivo de representação de dupla função: explicitar o conteúdo de documentos na indexação e expressá-los em sistemas de informação na recuperação. Essa segunda função atua na expressão das necessidades de pesquisa de usuários para formulação de estratégias de busca (American National Standards Institute, 2010; International Organization for Standardization, 2011, 2013).

A norma ISO 25964-1 define o tesouro como um “vocabulário controlado e estruturado no qual conceitos são representados por termos, organizados para que as relações entre conceitos sejam explícitas e termos preferidos estejam acompanhados por entradas para sinônimos ou quase-sinônimos” (International Organization for Standardization, 2011, p. 12, tradução nossa).

Gollub (2011 citado por Mazzochi, 2018) considera que vocabulário controlado tem um significado mais restrito a qualquer conjunto de termos ou lista de termos controlados, enquanto que linguagens de indexação “são tipos específicos de vocabulários controlados que representam linguagens formalizadas projetadas e usadas para descrever o conteúdo do assunto dos documentos para fins de recuperação da informação”. O termo “vocabulário controlado” tem sido utilizado em normas internacionais para se referir a outros tipos de vocabulários que não o tesouro e que se relacionam entre si, seja para interoperabilidade semântica ou tecnológica.

Fujita et al. (2018) analisaram a importância do termo linguagem de indexação em Organização do Conhecimento e constataram que na Enciclopédia de Organização do Conhecimento da ISKO, o termo é considerado como parte da terminologia de organização do conhecimento (OC) porque é “Ainda empregado em um ambiente de

Biblioteconomia e Ciência da Informação/ Organização do Conhecimento” (Mazzochi, 2018).

Diferentes Sistemas de Organização do Conhecimento que “abrangem todos os tipos de esquemas que organizam e representam o conhecimento” (Carlan e Medeiros, 2011, p. 54), são utilizados em ambiente digital.

Segundo Berti Junior et al. (2017), exemplos desses Sistemas de Organização do Conhecimento tais como, vocabulários, classificações, taxonomias e tesouros, entre outros, são usados em serviços de informação de banco de dados e de bibliotecas digitais para garantir a correta interpretação semântica e interoperabilidade das terminologias e registros de informações.

Currás (2005) explica que em meados de 1940 as bibliotecas e centros de documentação especializados encontravam dificuldades de identificação e localização de informações, tornando-os obsoletos. A resposta para este problema foi a flexibilização da representação dos assuntos pelo surgimento em fins da década de 1950 do tesouro com sua estrutura sistemática. Neste período os tesouros experimentaram grande evolução teórica e metodológica e passaram a ser utilizados pela Ciência da Informação nos processos de indexação e de recuperação da informação, como o instrumento que reúne um conjunto controlado de termos de um determinado domínio. Dotado de relações sintáticas e semânticas entre termos, o tesouro permite que a linguagem utilizada nos documentos seja traduzida em uma linguagem mais estrita.

O controle de vocabulário exercido no tesouro colabora com a redução de ambiguidades, controle de sinônimos, relacionamentos entre termos, entre outros aspectos, para a recuperação precisa da informação. A organização dos termos é predominantemente por ordem alfabética e de relações hierárquicas, (Martins, 2013).

O tesouro, portanto, deve ter uma linguagem controlada, com base na linguagem dos documentos e na terminologia de uma área do conhecimento, como também se basear em necessidades de informação dos usuários, expressas em sua linguagem de busca no sistema de informação conforme seus perfis de interesse. Também, consiste no instrumento para organização do conhecimento que melhor se desenvolveu com relação à sua estrutura e funções e, por ser dotada de precisão e relevância, pode ser considerada a mais adequada para a representação e recuperação da informação.

A evolução dos tesouros acompanha as tendências teóricas para sua construção, também de acordo com tecnologias em face de cada tempo

e meio de recuperação da informação. Atualmente, a tendência é o uso dos tesouros para organização semântica e recuperação da informação com interoperabilidade entre tesouros, bem como com outros sistemas de organização do conhecimento.

Diante disso, serão destacados a seguir aspectos de uso, finalidade, construção e avaliação dos tesouros.

2.1. Usos e finalidades dos tesouros: aspectos teóricos e recursos tecnológicos

O tesouro tem a finalidade de representar de forma mais estrita e normalizada a linguagem dos documentos durante os processos de indexação e de pesquisa. Assim, se mantém a comunicação nos sistemas de unidades de informação para a recuperação de informação. Outra função é de tradução da linguagem natural dos textos para linguagem controlada, que exerce controle de vocabulário e evita dispersões semânticas e sintáticas. De acordo com Soler Monreal e Gil Leiva (2010, p. 310) o tesouro “[...] se fundamenta em terminologia consensual, normalizada e controlada”.

Além disso, orienta, tanto o indexador sobre os melhores termos para representar o assunto, quanto o pesquisador sobre qual termo escolher que corresponda à representação do assunto procurado. Tem como atribuição a organização de conceitos e termos de uma área do conhecimento para adequar a relação desses termos com a linguagem do usuário, o que contribui para comunicar ao indexador os critérios de busca do usuário (Dextre Clarke; Zeng, 2012; Soergel, 1975).

Para objetivos mais contemporâneos, os tesouros passam a interagir nos ambientes *web*, oferecendo maiores graus de precisão à organização do conhecimento em rede (Marcondes, 2011), e atuando de forma interoperável e intercambiável (Mendes; Reis; Maculan, 2015), o que exige mais qualidade na escolha e organização dos termos.

Em ambiente *web* de mecanismos de pesquisa, segundo García-Marco (2016a), o tesouro pode ser usado para melhorar a recuperação e se tornará cada vez mais relevante nesse meio global, devido à precisão que proporciona ao usuário. O tesouro garante a interoperabilidade e possibilita a alimentação de novos mecanismos semânticos da Internet, que está se tornando a Web Semântica. Também, a interconexão de sistemas de organização do conhecimento consiste em uma ótima abordagem para melhorar a recuperação na Internet (García-Marco, 2016d).

Conforme analisa García-Marco (2017), a interoperabilidade dos tesouros ocorre com outros sistemas de organização do conhecimento no contexto da Internet e Web Semântica. A interoperabilidade, segundo García-Marco (2017, p. 306, tradução nossa),

[...] surge na estrutura da rede semântica e no movimento de dados abertos, uma orientação que não é estranha, mas sim, junto ao trabalho íntegro realizado pelas equipes de desenvolvimento da ISO 25964 e do padrão W3C SKOS.

García-Marco (2017, p. 306, tradução nossa), se refere à os três modelos de mapeamento estrutural para interoperabilidade da norma ISO 25964-2,

- de unidade estrutural, na qual os conceitos e estruturas relacionais de ambos os sistemas são equivalentes, por exemplo, entre dois tesouros;
- *link* direto, no qual todos os vocabulários estão ligados;
- de *'hub'* ou eixo central, no qual os vocabulários são mapeados em ambas as direções em relação àquela que funciona como o núcleo do sistema de vocabulários estruturados.

Para a interoperabilidade de relacionamentos semânticos, segundo do mesmo autor, são importantes em três tipos de mapeamentos, os de equivalência, os hierárquicos e os associativos, entre outros mais específicos. Ainda, a interoperabilidade é recomendada pela norma ISO 25964-2 entre tesouros e os principais sistemas de organização do conhecimento, portanto, que pode ocorrer com: “classificações bibliográficas e de arquivo; taxonomias; sistemas de cabeçalho de assunto; ontologias; terminologias; listas de controle de autoridade; anéis de sinônimos” (García-Marco, 2017, p. 307, tradução nossa).

A norma ISO 25964-2 não normaliza os diferentes tipos de sistemas de organização do conhecimento, mas é um avanço na questão da interoperabilidade entre esses sistemas.

2.2. Construção de tesouros

Para a elaboração dos tesouros, entre outros aspectos, deve-se determinar o domínio ou campo do conhecimento. Conhecer o campo conceitual a ser representado, bem como estabelecer a terminologia e o sistema nocional que decorre da relação entre os termos e conceitos, pois são aspectos que resultam da metodologia de elaboração de tesouro (Cervantes, 2009).

O método indutivo para elaboração de tesouros permite obter a terminologia mediante a literatura corrente de uma área, elegendo os termos e conceitos de maior frequência ou ocorrência, que

também correspondam às perguntas dos usuários e após a coleta e organização dos termos é realizada a estruturação do domínio. O método dedutivo permite obter a terminologia por meio da consulta aos especialistas de um domínio, que fornecerão os termos para compor o tesouro, portanto, a escolha dos termos ocorre pela garantia dos produtores e utilizadores do sistema de informação, (Cervantes, 2009).

Em geral, as diretrizes normativas indicam a possibilidade de combinação empregando cada método em um estágio da construção (Cervantes, 2009). Assim, os termos podem ser revisados posteriormente para assegurar a consistência de sua escolha e do estabelecimento de relações.

As abordagens indutiva e dedutiva de construção de tesouros são denominadas por Garcia (2016c) como bottom-up ou ascendente e top-down ou descendente, respectivamente. Esta última oferece mais autonomia para a construção ao considerar necessidades de usuário e interoperabilidade em diferentes ambientes digitais (García-Marco, 2016c)

A elaboração de tesouros se relaciona com dois princípios para o levantamento de termos, a garantia literária e garantia de uso ou do usuário. A garantia literária assegura que os assuntos dos textos ou documentos de uma área sejam representados do modo mais preciso possível e a garantia de uso permite que a representação ocorra correspondendo ao que é importante para a recuperação da informação dos usuários (Barité et al., 2010).

Nesse contexto, o tesouro é formado, segundo Soergel (1975), por um conjunto de termos descritores e outros símbolos. Estes não fazem parte, mas organizam sua estrutura, relacionando o descritor apropriado a ser usado, por exemplo, para relacionar dois termos sinônimos. Ainda, apresenta como item obrigatório: “a lista de descritores, o vocabulário ou léxico do sistema, as relações entre os descritores (hierárquicas e associativas)” (Soergel, 1975, tradução nossa); e como itens opcionais, que pertencem à sintaxe do tesouro, a lista de indicadores como descritores relacionais, relacionamentos hierárquicos ou associativos entre indicadores de função ou entre descritores relacionais de função sintática (Soergel, 1975).

Além da seleção e normalização dos conceitos, são estabelecidas relações entre termos de tipo lógico-hierárquicas, gênero/espécie e todo/parte (entre conceitos) e relações semânticas (entre termos), de equivalência e associativas (entre conceitos) (Dextre Clarke, 2017).

As normas sobre tesouros, de modo geral, são derivadas das primeiras diretrizes de construção de tesouro, do *Committee On Scientific And Technical Information* (COSATI) em 1967 e da UNESCO em 1974 (Lancaster, 2002, p.51).

O processo de construção dos tesouros é guiado pelas normas internacionais, cuja evolução levou à atual ANSI/NISO Z39.19-2005 (American National Standards Institute, 2010) e, na sequência, as normas ISO 25964-1 “*Information and documentation. Thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval*” (International Organization for Standardization, 2011) e ISO 25964-2 “*Information and documentation. Thesauri and interoperability with other vocabularies - part 2: Interoperability with other vocabularies*” (International Organization for Standardization, 2013).

Fujita, et al. (2016), com base em Martins (2013), informa que a norma ANSI/NISO Z39.19 (2005) é a quarta edição da norma ANSI Z39.19-1974, cujo fator determinante de mudança foi a aplicação de tecnologias de informação para desenvolvimento de tesouros eletrônicos em função dos avanços da informação eletrônica e da Internet.

Esta norma se baseou em diferentes autores, incluindo Soergel, que contribuiu principalmente com critérios de avaliação e manutenção de vocabulários controlados, porque demonstram o desenvolvimento do tesouro no meio digital, entre outros aspectos.

Garcia Marco (2016b, p.36) analisa que a quarta edição da norma Z39.19-2005 foi precursora em realizar três principais inovações determinantes para a evolução dos tesouros atuais: abarcar todo tipo de recursos de informação, dos tradicionais aos eletrônicos; ampliar seu alcance para todos os vocabulários controlados, listas de termos controlados, anéis de sinônimos, taxonomias e tesouros; e, abordar o problema de interoperabilidade de vocabulários na internet.

As evoluções seguintes foram desenvolvidas pela norma britânica BS8723, publicada em cinco partes (BSI, 2005, 2007, 2008). Os principais avanços incluem: a interoperabilidade linguística para tesouros multilíngues, a distinção entre termos e conceitos, o desenvolvimento de um modelo de dados, a análise em facetas, o tratamento dos sistemas de classificação como ferramentas auxiliares do tesouro, uma estratégia de interoperabilidade entre linguagens documentárias e a abordagem de outras linguagens (listas de cabeçalhos de assuntos) e alguns outros sistemas de organização do conhecimento utilizados na internet tais como ontologias, taxonomias e anéis de sinônimos (García Marco, 2016b)

Portanto, as normas ISO 25964-1 e ISO 25964-2 apresentam aspectos evolutivos importantes advindos, da norma ANSI/NISO Z39.19-2005, “*Guidelines for the construction, format, and management of monolingual controlled vocabularies*” e, principalmente, da Norma BS8723, “*Structured vocabularies for information retrieval*”. A parte 1 reúne o tratamento dos tesouros monolíngues e multilíngues e a parte 2 é mais dedicada à interoperabilidade com outros vocabulários em contexto de web semântica e de dados abertos.

2.3. Avaliação e manutenção de tesouros

De acordo com a norma ANSI/NISO Z39.19-2005 “[...] os desenvolvedores de vocabulário controlado devem avaliar se os metadados são relevantes para sua comunidade e para o domínio do conhecimento, isso antes de desenvolver novos conjuntos ou esquemas com elementos de metadados.” (American National Standards Institute, 2010, p. 3, tradução nossa).

As razões para testar e avaliar, conforme a norma ANSI/NISO Z39.19-2005, são para determinar se o tesouro descreve o conteúdo e apresenta resultados de pesquisa adequados para os usuários (relevância e revocação) (American National Standards Institute, 2010).

Os seguintes métodos podem ser usados para medir a qualidade ou a eficácia do tesouro, segundo a mesma norma: a) Avaliação Heurística (especialistas avaliam em aspectos qualitativos e quantitativos); b) Modelagem de Afinidade (alguns usuários agrupam cartões com termos para que os resultados sejam analisados em relação à hierarquia dos termos existentes, entre outros aspectos); c) Teste de Usabilidade (conforme a literatura nas áreas de computação e arquitetura da informação sobre usabilidade e métodos de avaliação).

Os critérios de avaliação, expressos pelas perguntas abaixo conforme apontados na norma ANSI/NISO Z39.19-2005 e já mencionados por Soergel (2002), demonstram alguns dos cuidados para elaboração de tesouros (American National Standards Institute, 2010, p. 95, tradução nossa):

O vocabulário controlado contém termos sinônimos ou equivalentes sem indicar o relacionamento? A apresentação reflete a análise conceitual, por exemplo, a sequência de conceitos em nível hierárquico? As notas são fornecidas quando necessário? Todas as facetas necessárias estão incluídas? Os termos estão atualizados e são usáveis? (pontualidade dos termos). Os sinônimos de termos de entrada incluídos na lista alfabética são completos?

Outro fator de desenvolvimento do tesouro com qualidade consiste na manutenção, pois, a linguagem dinâmica do usuário e atualização constante dos documentos leva à necessidade de revisão e manutenção.

A manutenção do tesouro consiste em regulamentos para adicionar, modificar e excluir termos, de modo que cada termo tenha um responsável por sua elaboração e que um histórico fique registrado até a eliminação do termo. Vale destacar que, segundo a American National Standards Institute (2010, p. 96, tradução nossa) “[...] os vocabulários controlados refletem a linguagem e são, portanto, instrumentos dinâmicos [...]”. Segundo a norma ANSI/NISO Z39.19-2005, devem ocorrer revisões periódicas da terminologia por meio de políticas e procedimentos da instituição. Destaca-se que as versões eletrônicas têm a vantagem de serem acessadas e analisadas mais rapidamente.

3. Metodologia

A metodologia de análise de conteúdo de Bardin (2011) foi aplicada, conforme Amorin e Café (2016), em amostra de publicações científicas de Dagobert Soergel. A análise de conteúdo constituiu-se de três etapas metodológicas: a) pré-análise; b) exploração do material; c) tratamento de resultados com inferência e interpretação de aspectos sobre o tesouro apresentados pelo pesquisador. Assim, o pensamento de Soergel foi analisado, inferido e elucidado no que tange aos tesouros.

A etapa de pré-análise contribuiu para a seleção das obras do pesquisador, apresentando os seguintes passos e critérios:

a) *identificação da produção bibliográfica de artigos científicos do autor*: verificando o currículo do pesquisador Soergel disponibilizado em sua página web (<http://www.dsoergel.com/>), bem como suas publicações apresentadas no Google Acadêmico (<https://scholar.google.com/citations?user=4bJ6HcUAAA&hl=en>) para observação de suas características e estabelecimento da amostra e categorias;

b) *coleta de informações sobre os artigos publicados, como*: ano de publicação, título do artigo, título do periódico, indicação de coautoria e fonte para localização da publicação no período de 1972 a 2017;

c) *identificação e seleção de periódicos e eventos de maior relevância em que o autor publicou seus artigos*: Foi aplicado o critério de seleção de artigos com autoria única para evidenciar o pensamento de Soergel. O critério de seleção consistiu em considerar os periódicos da área de Ciência

da Informação, ativos até o presente com avaliação por pares às cegas, com destaque para *Knowledge Organization* (prev. *International Classification*) e *Journal of the American Society for Information Science and Technology*. Não foram considerados periódicos de áreas de assunto diversificadas ou amplas. Além disso, foram selecionados trabalhos publicados de conferências organizadas por sociedades de pesquisadores da área de Ciência da Informação e Organização do Conhecimento com destaque para *American Society for Information Science and Technology* (ASIS&T) e *International Society for Knowledge Organization* (ISKO). Sobre os artigos, quatro são artigos de periódicos, dois deles publicados no *Journal of the American Society for Information Science and Technology* (JASIS&T), dois no *Knowledge Organization* (KO) e três são trabalhos publicados em conferências, sendo estes, um do Workshop do *Special Interest Group – Classification Research* da *American Society for Information Science and Technology* (ASIS&T), outro proveniente da *ASIS&T Annual Meeting* e o terceiro do *International ISKO Conference*.

d) *identificação e seleção do tema a ser analisado nos artigos*: por meio de leitura rápida dos textos, verificando principalmente título, resumo e palavras-chaves para identificação do tema principal do artigo buscando entender termos tais como tesauro, linguagem de indexação, ontologia, esquemas de classificação, sistema de organização do conhecimento, terminologia, conceitos, para relacioná-los com linguagem de indexação e tesauro. Com base na fundamentação teórica desse estudo foram, principalmente, procurados os seguintes temas: elementos de construção, funções, manutenção, tecnologias e usos de tesouros e linguagens de indexação, utilizados, posteriormente, como categorias para análise do pensamento do autor nos artigos selecionados.

A etapa de exploração do material consistiu na identificação e análise, em sete trabalhos científicos de Soergel, de aspectos relacionados às cinco categorias destacadas na etapa anterior.

A etapa de tratamento dos resultados teve como enfoque a análise de inferência e interpretação de aspectos qualitativos segundo o pensamento de Soergel.

Foram definidas cinco categorias de análise (elementos de construção; funções; manutenção; tecnologias; e usos) após consulta às publicações com leitura flutuante, por serem aspectos recorrentes e abordados pelo autor nos trabalhos analisados. A etapa de exploração da amostra de 7 publicações possibilitou a identificação e análise

das categorias. Nesta etapa foram feitas análises, inferências e observações sobre os trabalhos publicados e coletadas citações, sempre com enfoque nas categorias de análise.

4. Discussão dos resultados

A pesquisa abrangeu todo o período de publicação do autor disponível para ser consultado em seu currículo online e em sua página de autor na ferramenta Google Acadêmico (2018), compreendendo um total 202 publicações entre os anos de 1972 a 2017. Porém, o período situado entre os anos de 1975 e 1990 e entre 2000 e 2014 não foram representados no estudo, dado que não foi identificado nenhum artigo de periódico ou conferência publicado pelo autor que tenha se adequadado aos critérios de seleção.

Os resultados da análise de conteúdo demonstram a inferência e interpretação sobre aspectos da linguagem de indexação e tesauro, relativos à construção, uso, avaliação e manutenção encontrados em sete artigos publicados por Dagobert Soergel. As análises são:

Primeiro Artigo – A Universal Source Thesaurus as a Classification Generator (1972): Estimulado pelo avanço das ações de cooperação no âmbito dos serviços de informação da segunda metade da década de 1960, apresenta a ideia de um *Universal Source Thesaurus* (UST) – tesauro de fonte universal, ou seja, um tesauro para ser uma fonte de informação universal, que seria utilizado como fonte para criação de linguagens especializadas – classificações, tesouros e demais linguagens de indexação e como linguagem de intercâmbio (*switching language*), ferramenta de integração entre diferentes linguagens para facilitar o trabalho cooperativo.

Para isso o tesauro deve apresentar as seguintes características segundo Soergel (1972, p.300, tradução nossa):

- (1) A estrutura de um tesauro universal.
- (2) O uso de um tesauro universal na construção de esquemas especiais.
- (3) O uso de um tesauro universal para concordância ou "troca de linguagem" entre diferentes esquemas de classificação, como base para catalogação cooperativa.
- (4) Organização e gestão de um dicionário de sinônimos de fonte universal.

Além disso, apresenta bases para o pensamento pró-interoperabilidade de sistemas no âmbito da interação de linguagens. Percebe-se a consonância com o pensamento de Dahlberg com relação à sua visão ontológica da organização do

conhecimento, porém o autor apresenta um pensamento mais transigente, ao adotar esta visão de uma linguagem de intercâmbio ou de reconciliação, que foi posteriormente retomada por Gnoli (2011) em sua “defesa” da herança ontológica de Dahlberg.

A noção do conceito como elemento central da linguagem de indexação também está presente no artigo, especialmente perceptível na afirmação de que a base para a construção de um UST e seu uso como linguagem de intercâmbio é a existência de compatibilidade conceitual. Assim “em outras palavras, a indexação cooperativa não requer uma classificação universal; a compatibilidade conceitual entre diferentes esquemas e a disponibilidade de uma concordância entre eles são suficientes” (Soergel, 1972, p. 7, tradução nossa).

Segundo Artigo – Automatic and semi-automatic methods as an aid in the construction of indexing languages and thesauri (1974): Apresenta um método automático baseado em estatística para determinação de termos e conceitos e o estabelecimento de relacionamentos terminológicos e classificatórios entre eles com base na frequência de ocorrência e coocorrência de termos e conceitos para construção e manutenção de tesouros. O autor apresenta este método como semi-automático, destacando que embora seja interessante, à época um método completamente automatizado para construção de tesouros não era viável.

Segundo o autor, a frequência coocorrência de termos pode ser utilizada para a determinação de estruturas terminológicas, como a definição dos relacionamentos de equivalência. Já a frequência coocorrência de conceitos, pode ser usada para a construção da estrutura classificatória. Uma das dificuldades apresentadas pelo autor seria a de que os softwares disponíveis não seriam capazes de identificar os conceitos complexos, formados por mais de uma palavra.

Para demonstrar as palavras do autor,

As seções anteriores estavam relacionadas (a) com a identificação de termos e suas inter-relações entre pares e (b) com a identificação de conceitos e suas inter-relações entre pares - as últimas poderiam ser chamadas de informações classificatórias “locais”. O próximo passo é a derivação automática de uma estrutura “global” (um esquema de classificação) para obter a imagem geral. Há duas tarefas inter-relacionadas:

- (a) Encontrar agrupamentos úteis de conceitos.
- (b) Encontrar um padrão de subdivisão do conjunto de documentos em classes não sobrepostas.

Existem dois métodos principais para executar estas tarefas:

- (1) métodos de agrupamento.
- (2) métodos teóricos de gráfico.

Até certo ponto, esses métodos se sobrepõem; pode-se até dizer que o método de *clustering* é um caso especial de métodos baseado na teoria dos gráficos. (Soergel, 1974, p. 38-39, tradução nossa).

Nesse sentido, Soergel explicou a derivação automática de esquemas de classificação para métodos automáticos de construção de tesouros.

Terceiro Artigo – Beyond facets: semantic roots and modifiers as elements of a conceptual morphology (1991): O autor expõe a proposta de uma premissa para a construção de linguagens de indexação baseada na análise facetada, em que por meio de uma morfologia conceitual, que parte da raiz semântica de cadeias de conceitos e de um sistema de modificadores para promoção de indexação mais flexível. Ou nas palavras do autor, sua sugestão é de “[...] uma classificação que consiste de um esqueleto bem-estruturado de conceitos centrais ou raízes semânticas e um sistema de modificadores cuidadosamente realizado. Muitos conceitos podem ser formados por meio da combinação destes elementos” (Soergel, 1991, p. 2, tradução nossa).

Segundo o autor, o sistema alcança a flexibilização da indexação ao permitir a pesquisa genérica por meio da raiz semântica, bem como a pesquisa mais específica por meio da combinação de raiz semântica acrescida de modificadores. Contudo, é um sistema ainda em estágio inicial.

Quarto Artigo – Software Support for Thesaurus Construction and Display (1994): Apresenta uma proposta de software de construção e visualização de tesouros TermMaster utilizado na construção do tesouro “Alcohol and Other Drugs”. Entre as funcionalidades que são atribuídas ao software estão a de criação de hierarquias de termos, o suporte para atribuição de notas aos termos, criação de variados tipos de relações semânticas, bem como, a atribuição de níveis de “importância” aos relacionamentos entre os termos, a possibilidade de customizar os relacionamentos para incluir tipos definidos pelo usuário e a combinação de diversos tesouros em um. Segundo Soergel (1994, p. 2, tradução nossa),

Um termo do tesouro está vinculado a outros termos por meio de diversos relacionamentos. TermMaster permite um grande número de desses relacionamentos, tornando possível, por exemplo, distinguir no banco de dados (não necessariamente na versão do usuário) entre Termo Sinônimo, e Termo Equivalente (quase-sinônimo) [...]. Notas de escopo são tratadas como relações com o texto; assim, um termo pode ter várias notas de escopo e pode haver

diferentes tipos de notas de escopo, como Notas de Histórico ou Notas Internas [...].

Alguns relacionamentos são mais importantes que outros. O TermMaster permite (mas não exige) as especificações de relacionamento com três níveis de importância que podem ser usados para orientar a inclusão de relacionamentos [...]. O programa pode ser facilmente customizado para incluir um conjunto de relacionamentos definido pelo usuário.

De acordo com o autor “o programa aproxima todas as relações que são conceitualmente iguais, mas expressas em diferentes termos enquanto estes termos forem mostrados como sinônimos nas fontes utilizadas” (Soergel, 1994, p. 3, tradução nossa).

A proposta de combinação de elementos de outros tesouros, especialmente a inclusão de relacionamentos entre conceitos de diferentes tesouros que estão em acordo com as propostas das diretrizes atuais para construção de tesouros (ISO, 25964-2011/2013), no tocante ao mapeamento de vocabulários controlados, base das atuais discussões sobre interoperabilidade semântica com interligação de vocabulários a partir da existência de relacionamentos inter-linguagens.

Quinto Artigo – *SemWeb*: proposal for an open, multifunctional, multilingual system for integrated access to knowledge about concepts and terminology (1996): O autor parte das recomendações feitas no Congresso Internacional da ISKO realizado na Polônia em 1995, que se dirigia a infraestrutura conceitual referente à organização da informação e do conhecimento na sociedade da informação do então eminente século XXI, para propor um sistema de acesso à base de conhecimento sobre conceitos e terminologia. De acordo com a proposta, o sistema faria a pesquisa em diversas bases, coletaria os dados e apresentaria ao usuário em uma interface comum integrada. Assim, o sistema seria projetado para ser usado por muitos níveis de usuários para melhorar o intercâmbio de informações. De acordo com Soergel (1996, p. 2, tradução nossa),

Este artigo trata da infraestrutura intelectual. Apresenta a visão e arquitetura básica de um sistema para acesso integrado a dados sobre conceitos e terminologia. O sistema reuniria os dados do usuário de uma variedade de fontes que até agora existem em grande parte em mundos separados, incluindo dicionários, tesouros e esquemas de classificação; ele utilizaria bases de conhecimento existentes que são acessíveis através da Internet ou em CD-ROM e em uma base de conhecimento distribuída integrada e comum que cresceria de forma incremental ao longo do tempo. As bases de conhecimento existentes seriam acessadas por meio de uma interface comum que procuraria várias bases

de conhecimento, agruparia os dados em um formato comum e os apresentaria ao usuário. A base de conhecimento distribuída integrada comum forneceria um ambiente no qual muitos colaboradores poderiam realizar projetos classificatórios e terminológicos de forma mais eficiente, com os resultados disponíveis em um formato comum. Com o tempo, os dados de outras bases de conhecimento poderiam ser incorporados na base de conhecimento comum, seja por transferência real (desde que os produtores da base de conhecimento estejam dispostos) ou por referência através de um link. De qualquer forma, tal incorporação requer trabalho intelectual, mas permite uma integração mais forte do que o acesso comum à interface com várias bases de conhecimento. Cada informação na base de conhecimento comum terá todas as suas fontes anexadas, fornecendo um mecanismo de reconhecimento que dá o devido crédito a todos os colaboradores. Todo o sistema seria projetado para ser usado por muitos níveis de usuários para melhorar a troca de informações.

As funções dessa base de conhecimentos, contendo conceitos e terminologia, segundo Soergel (1996, p. 4, tradução nossa), consistem em “[...] mapear um campo conceitual, relacionar os conceitos aos termos e fornecer definições, também para orientar e servir como ferramenta de referência”.

Além disso, as potencialidades do *SemWeb template*, está em organizar as informações dos conceitos e termos.

O modelo se concentra em informações sobre conceitos e termos individuais. [...] o sistema também deve mostrar estruturas conceituais e classificatórias de vários formatos (listas lineares, mapas bidimensionais, etc.), com navegação e navegação adequada ao usuário para que possa passar de uma visão geral para classificações detalhadas. (Soergel, 1996, p. 9, tradução nossa).

Sexto artigo – *The rise of ontologies or the reinvention of classification* (1999): O autor faz uma retomada dos objetivos e propósitos das classificações usadas em bibliotecas e cuja importância foi recentemente descoberta por outras áreas, tais como inteligência artificial, processamento de linguagens naturais, engenharia de softwares, levando à ascensão das ontologias. Aborda o processo de adoção por tais áreas de pequenas classificações de objetos que, ao passo que foram tornando-se mais complexas, denominaram-se ontologias. Soergel (1999) argumenta aqui em defesa da classificação, independente do nome que receba, porém, ressaltando que o desenvolvimento de instrumentos semelhantes em áreas diversas e sem comunicação, considerando especialmente o vasto conhecimento desenvolvido pela Ciência da Informação sobre Classificação,

poderia causar não apenas a reinvenção de ferramentas, mas também a duplicação de trabalho e a elaboração de ferramentas pobremente construídas. De acordo com Soergel (1999, p. 1, tradução nossa),

A classificação desempenha muitas funções, portanto, é reivindicada por muitos campos, mas a comunicação entre esses campos é fraca, levando a uma abordagem marcada por uma reinvenção fragmentada e dispendiosa.

Estruturas ontológicas e lexicais refletem a base do trabalho científico e acadêmico, da aprendizagem e da inteligência das máquinas. Elas servem para muitas funções críticas de pensamento, comunicação, organização e recuperação de informações por pessoas e máquinas. As funções das ferramentas que fornecem tais estruturas (dicionários, tesouros, ontologias / classificações) incluem o seguinte:

- Fornecer um roteiro semântico para campos individuais e os relacionamentos entre campos, fornecendo orientação e servindo como uma ferramenta de referência. Isso inclui as seguintes funções específicas: relacionar conceitos a termos e fornecer definições; esclarecer conceitos colocando-os no contexto de uma classificação / ontologia; relacionar conceitos e termos ou ícones em disciplinas, idiomas e culturas. [...]
- Recuperação de informações de suporte: fornecer suporte baseado em conhecimento de pesquisa do usuário final (árvores de menu, análise de faceta orientada de um tópico de pesquisa, navegação em uma hierarquia ou mapa conceitual para identificar conceitos de pesquisa, mapeamento de termos de consulta do usuário para descritores usados em um ou mais bases de dados ou para as múltiplas expressões de linguagem natural para pesquisa de texto livre); suporte a pesquisa expandida hierarquicamente; suporte a exibições bem estruturadas de resultados de pesquisa; fornecer uma ferramenta para indexação (controle de vocabulário, indexação centrada no usuário ou orientada a problemas).
- Fornecer a base conceitual para sistemas baseados em conhecimento.
- Fornecer a base conceitual para definição de elementos de dados e hierarquias de objetos em sistemas de software.
- Fazer tudo isso em disciplinas, idiomas e culturas.
- Servir como dicionário mono-, bi- ou multilíngue para uso humano e como base de dicionário / conhecimento para processamento de linguagem natural - tradução automática e compreensão de linguagem natural para extração de dados e abstração / indexação automática.

Além disso, anda segundo o autor,

A classificação tem sido usada há muito tempo em bibliotecas e sistemas de informação para fornecer orientação ao usuário, esclarecendo aspectos de sua necessidade de informação e estruturando re-

sultados de busca para navegação, funções amplamente ignoradas pela comunidade de recuperação da informação textual, mas agora recebe atenção crescente no contexto de buscar ajudar os usuários a lidarem com a vasta quantidade de informação na web. Recentemente, outros campos como IA [Inteligência Artificial], o processamento de linguagem natural e a engenharia de software descobriram a necessidade de utilizar a classificação, levando à ascensão do que esses campos chamam de ontologias. [...] uma vez que essas comunidades aumentaram sua consciência de que não existe apenas um problema de classificação, mas também de terminologia, as “ontologias” incluíam também vocabulários e se tornaram verdadeiros tesouros. (Soergel, 1999, p. 2, tradução nossa).

Para o autor, a ontologia é vista, de modo geral, como uma classificação de categorias básicas, entre outros aspectos. Além disso, afirma que existem muitos tipos de bases de conhecimento que utilizam conceitos e terminologia, que embora sejam desenvolvidos para diferentes propósitos, apresentam princípios e métodos muito semelhantes para sua construção, sendo que para essa reflexão o autor cita o próprio trabalho de 1996, que foi analisado nesta pesquisa no quinto artigo.

Sétimo artigo – Unleashing the power of data through organization: structure and connections for meaning, learning and discovery (2015): Neste artigo o autor endereça uma série de problemas em que a Organização do Conhecimento pode ser útil, desenhando uma possível ampliação do escopo da área, entre estas questões destacam-se, tais como, computação cognitiva, gestão de dados organizacionais, gestão de prontuários eletrônicos de pacientes, *linked data*, *big data*, interoperabilidade de dados, extração automática de informação, entre outras.

É destacado o fato de que os Sistemas de Organização do Conhecimento (SOC) permeiam as diversas problemáticas abordadas, emergindo a noção dos SOC como elementos centrais da Organização do Conhecimento. Além disso, segundo o pesquisador,

Resumindo (Soergel 2009, 3), “sistemas de organização do conhecimento (KOS) cobrem uma ampla gama de sistemas em camadas, atendendo a uma ampla gama de propósitos. Eles são conhecidos com nomes como ontologia, esquema de metadados, taxonomia, classificação, estrutura de diretórios da Web, plano de arquivamento, tesouros, dicionários, folksonomia e muito mais”. (Soergel, 2015, p. 402, tradução nossa).

Quanto aos SOC especificamente o autor discute a existência e a dispersão dessas ferramentas, que existem em diferentes tipos – como dicionários, sistemas de classificações, tesouros, ontologias, que possuem não apenas objetivos,

mas também características, padrões normativos e softwares (para criação e uso) diferentes, apresentando a unificação de SOC como requisito para a melhoria destas ferramentas e consequentemente para que a própria área de Organização do Conhecimento seja beneficiada.

Soergel (2015) coloca duas modalidades de unificação, a formal e a flexível, com o objetivo da integração dos dados de SOC com outros SOC, mas também com outros tipos de dados estruturados. Para o autor a unificação formal caracteriza-se por SOC usarem “[...] os mesmos princípios de organização, estrutura interna e apresentação para sistemas com diferentes tipos de entidades” e requer o estabelecimento de um único esquema padronizado e abrangente para representar todos os tipos de dados em qualquer SOC a partir de qualquer software (Soergel, 2015, p.419, tradução nossa). A unificação flexível, segundo o autor, de dois ou mais SOC,

[...] não requer fusão total e perda de identidade. Ela pode ser alcançada pela análise dos sistemas, descobrindo semelhanças e concordâncias onde existem, criando acordos onde for possível e explicando as diferenças que devem ser mantidas embora ainda estabelecendo relações de concordância parcial. Um único SOC pode e deve ser hospitaleiro para muitos pontos de vista (Soergel, 2015, p. 420, tradução nossa).

Para Soergel (2015) este processo de unificação, muito identificável com o que tem sido denominado como interoperabilidade, é o foco central do trabalho, sendo nas palavras do autor “a chave para liberar o poder dos dados pela integração de conjuntos de dados” (Soergel, 2015, p.419, tradução nossa).

A síntese das análises realizadas com os sete artigos, aponta que, em geral, as publicações analisadas de Soergel são marcadas pela proposição de novas metodologias e ferramentas de representação de conteúdo, destacando-se, por exemplo, o tesouro universal que servem como base para sistemas especialistas e de classificação (Soergel, 1972), métodos automático e semi-automático de construção e manutenção de tesouros (Soergel, 1974), elementos de construção de tesouros bem estruturados que contribuem para que sejam realizadas as funções de recuperação e uso (Soergel, 1991, 1994), o tesouro que passa a ser um elemento de construção de outros sistemas (Soergel, 1996, 1999, 2015). Esses fatores fortemente embasados nos desenvolvimentos computacionais e tecnológicos mais recentes, mas segundo cada época.

Vale destacar que o autor é muito cuidadoso ao apresentar esses aspectos em seus textos,

exemplificando-os quando parece ser mais pertinente, principalmente, quando explica os relacionamentos entre termos. Assim, apresenta uma escrita bem elucidativa sobre a construção dos tesouros e outros aspectos que abordou.

Especificamente, quanto às categorias analisadas destacam-se:

Os elementos de construção dos tesouros visam uma estrutura capaz de ser interoperável com outros sistemas de organização do conhecimento (Soergel, 1972, 2015). Além disso, apresenta diferentes metodologias de estruturação pelos relacionamentos entre termos e uso de dicionários (Soergel, 1991, 1994, 1996, 1999). Por exemplo, os tesouros em meio digital, segundo Soergel (1999), podem fornecer: relacionamentos mais ricos e definições, além de notas; e, apresentar links para dicionários linguísticos e de sinônimos, e outros documentos que embasam o conceito. Esses elementos, de modo geral, contribuem para eliminar ambiguidades e controle de sinonímia, segundo Souto (2003), além de contribuírem com a estrutura sintática e a semântica do tesouro (Soergel, 1975). Nesse contexto, as abordagens para desenvolvimento de tesouros segundo García-Marco (2016c) e outras, conforme Cervantes (2009), são imprescindíveis para demonstrar a representação do domínio de modo a expressar também as necessidades ou perguntas dos usuários, destinatários desse instrumento de organização elaborado para fins de recuperação da informação.

A manutenção do tesouro (Soergel, 1974), mostra que pode colaborar com a melhoria de relacionamentos terminológicos e classificatórios, entre outros aspectos. Nesse sentido, a manutenção regulamenta o desenvolvimento do tesouro diante de mudanças da linguagem do usuário (American National Standards Institute 2010; International Organization for Standardization, 2011). Além disso, a manutenção não pode ser mais vista à parte do processo de construção dos tesouros, pois é um instrumento de representação em constante desenvolvimento, principalmente quando seu uso apoia ambientes tecnológicos na Web, como nos mecanismos de pesquisa que, segundo García-Marco (2016a), melhora a precisão na recuperação.

O aspecto das funções demonstrou que tesouros contribuem para mapear campos de conceitos e servir de fonte de pesquisa, além de interoperar com outros sistemas de organização do conhecimento (Soergel, 1996, 1999, 2015). Nesse sentido, outras funções dos tesouros, propiciadas por novas tecnologias são, segundo Soergel (1999, p. 5, tradução nossa),

Transmitir significado, orientação e estrutura: auxiliando ao usuário a pensar em um problema e ajudar a formular a consulta, o que requer bons métodos para exibir uma boa estrutura de interface, mostrando a hierarquia, gráfico conceitual e facetas para consulta.

O pesquisador, sobretudo, enfatiza a interoperabilidade como aspecto importante na relação do tesouro com outros sistemas de organização do conhecimento que, de modo geral, segundo Mendes, Reis e Maculan (2015), demanda mais cuidado no relacionamento entre termos. Por isso, a preocupação do pesquisador com os relacionamentos adequados entre termos quando se utilizam diferentes métodos, padrões e tecnologias para garantir a qualidade e integridade do tesouro na interoperabilidade. Esse aspecto da interoperabilidade demanda muita interação entre tecnologias que propiciem a troca de informações de modo estruturado, com vocabulários interligados (García-Marco, 2017).

A *categoria de uso dos tesouros* apresentou a preocupação em oferecer elementos ricos para a recuperação da informação mediante as diferentes modalidades de pesquisas: genérica, específica e pesquisa expandida hierarquicamente, além da possibilidade de customização da busca para incluir relacionamentos definidos pelo usuário que possam ser utilizados na indexação e na construção de tesouros (Soergel, 1991, 1994, 1996, 1999). Esses aspectos contribuem para orientar a busca e recuperação da informação pelo usuário, como também para o indexador usar o tesouro na indexação conforme as necessidades dos usuários (Dextre Clarke; Zeng, 2012; Soergel, 1975). Ainda, os tesouros são contruídos para os usuários e segundo necessidades e perguntas dos usuários, nesse sentido, a construção do tesouro também é voltada para o seu uso (García-Marco, 2016c).

As *tecnologias utilizadas para desenvolvimento dos tesouros* evoluíram de sistemas automáticos e semiautomáticos, passando por *softwares* específicos, para sistemas que podem ser utilizados em conjunto na *web*, como as ontologias, além das tecnologias de inteligência artificial, *linked data*, entre outras visando a interoperabilidade (Mendes, Reis e Maculan, 2015; Soergel, 1974, 1994, 1996, 1999, 2015). Nesse contexto, a interoperabilidade entre instrumentos de representação como o tesouro ocorre em ambiente de Internet e Web Semântica, por oferecer condições de alimentação de novos mecanismos semânticos desses ambientes, por exemplo, com o padrão W3C SKOS (García-Marco, 2016d, 2017).

5. Considerações finais

A evolução dos tesouros no discurso de Dagobert Soergel mostra o aprimoramento científico e metodológico desse instrumento de representação, considerando os elementos de construção, funções, manutenção e tecnologias.

As publicações analisadas foram de grande importância para o entendimento desses aspectos porque demonstram as apropriações contemporâneas das contribuições de Soergel para o desenvolvimento dos tesouros, como por exemplo, o tesouro universal que serve de fonte para a construção de outros tesouros e antecipa a idéia de interoperabilidade, ou a ênfase sempre presente do emprego de métodos automáticos e semiautomáticos na construção de tesouros.

O aspecto das funções demonstrou que tesouros contribuem com sistemas especialistas e classificação, que permitem a pesquisa genérica por meio da raiz semântica, como também a pesquisa mais específica. As funções da base de conhecimentos sobre conceitos e terminologia consistem em mapear campo conceitual, fornecer definições, e servir de ferramenta de referência.

As funções dos sistemas de organização do conhecimento, conhecidos como ontologias, esquema de metadados, taxonomia, classificação, tesouros, entre outros, tem o objetivo da integração dos dados com outros sistemas de organização do conhecimento para ampliar e melhorar a organização do conhecimento e sua área científica, também deve ser hospitaleiro. Portanto, abrangem a contribuição para construção de outros sistemas de organização do conhecimento, como também para apresentar bases de conhecimentos, entre outros aspectos. Em geral, os aspectos da construção dos tesouros contribuem para a realização das funções de recuperação da informação, contribuição da elaboração de outros instrumentos de representação, entre outros aspectos.

A manutenção do tesouro é apontada em apenas um trabalho, mostrando que permite atualização do tesouro nos relacionamentos terminológicos e classificatórios, também na frequência de ocorrência e coocorrência de termos e conceitos, mostrando que assim o tesouro pode ser atualizado.

As tecnologias utilizadas para desenvolvimento dos tesouros e sistemas de organização do conhecimento evoluíram de sistemas automáticos e semiautomáticos, passando por *softwares* (sistemas ou programas) específicos, para sistemas que podem ser utilizados na *Web*, ainda aborda

as tecnologias apropriadas para inteligência artificial e ontologias, entre outras.

A categoria de uso dos tesouros demonstrou a preocupação com a recuperação da informação por meio de pesquisas (genérica e específica), a facilidade de ser customizado para incluir relacionamentos definidos pelo usuário, a intenção de ser projetado para muitos níveis de usuários, visando a melhora de intercâmbio de informações, além de mostrar que o uso pode ser a partir de recuperação por meio de pesquisa expandida hierarquicamente, ou com exibições bem estruturadas em resultados de pesquisa. Também, pode ser usado como ferramenta para indexação (controle de vocabulário, indexação centrada no usuário ou orientada a problemas).

Nesse sentido, o pesquisador Dagobert Soergel demonstra uma visão abrangente das linguagens de indexação como parte de um sistema com o qual troca insumos com outros sistemas de organização do conhecimento para o seu desenvolvimento.

Percebemos que as temáticas abordadas pelo autor vêm sendo de interesse da área com os recentes desenvolvimentos das ferramentas computacionais, o que também deu vida nova às discussões conceituais trazidas pelo autor, como a noção do conceito/termo como elemento central da linguagem de indexação, a construção de tesouros por meio de fusão de tesouros existentes e a importância dos dados semânticos não apenas para a recuperação da informação online, mas como para o desenvolvimento da própria web em suas funções.

É possível notar que o seu pensamento pode ser inovador em muitos aspectos, como por exemplo, quanto à interoperabilidade entre linguagens de indexação, que na visão do autor pode ser alcançada sem que seja necessário um apagamento das diferenças culturais de cada sistema envolvido. Também, na publicação mais recente, percebe-se que o autor passou a considerar a noção de sistemas de organização do conhecimento enquanto um termo abrangente que abarca os diversos instrumentos de representação do conhecimento, como o tesouro, aqui observado em especial, assim como os sistemas de classificação e as ontologias, entre outros.

Os elementos de construção dos tesouros são fundamentados pela necessidade de controle terminológico e de vocabulário adequados, em face às tecnologias de cada época, também visando a interoperabilidade.

As funções e uso dos tesouros demonstram, principalmente, a sua importância para a recuperação da informação e seu uso na construção de

outros sistemas de organização do conhecimento.

Em suma, os artigos analisados mostraram que Soergel apresenta inovações e acompanha tendências para elaboração, utilização e melhoria dos tesouros.

Referências

- American National Standards Institute; National Information Standards Organization. (2010). ANSI/NISO Z39.19-2005. Guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda: NISO Press, 2010. <https://www.niso.org/> (2017.10.16)
- Amorin, I. S.; Café, L. M. A. (2016). Os conceitos de comunidade discursiva, domínio e linguagem na análise de domínio Hjørlandiana // 17 Encontro Nacional de Pesquisa em Ciência da Informação. Salvador: UFBA, 2016. <http://www.ufpb.br/evento/lti/ocs/index.php/enancib2016/enancib2016/paper/view/3580> (2018.09.15).
- Bardin, L. (2011). Análise de conteúdo. São Paulo: Edições 70, 2011.
- Barité, M. et al. (2010). Garantia literária: elementos para uma revisão crítica após um século // *TransInformação*, 22:2, (maio/ago., 2010) 123-138 <http://www.scielo.br/pdf/tinf/v22n2/a03v22n2.pdf> (2018.07.26).
- Berti Junior, D. W. et al. (2017). Semiautomação de relações em tesouros: uma proposta para refinamento de relacionamentos semânticos a partir do tesouro Agrovoc. // *Informação & Informação*, 22:3 (set./out. 2017) 377-404. <http://www.uel.br/revistas/uel/index.php/informacao/article/view/28724> (2018.07.26).
- Carlan, E.; Medeiros, M. B. B. (2011). Sistemas de Organização do Conhecimento na visão da Ciência da Informação. // *Revista Ibero Americana de Ciência da Informação*, 4:2 (ago./dez.2011) 53-73. <http://periodicos.unb.br/index.php/RIC/article/view/6209> (2018.07.26).
- Cervantes, B. M. N. (2009). A construção de tesouros com a integração de procedimentos terminográficos. Universidade Estadual Paulista: Marília. Tese de Doutorado.
- Currás, E. (2005). Ontologías, taxonomía y tesouros: manual de construcción y uso. Gijón: Trea, 2005.
- Dextre Clarke, S. G. (2017). Thesaurus (for information retrieval) // Hjørland, B.; Gnoli, C. (Eds.). *Encyclopedia of knowledge organization*, 2017 <http://www.isko.org/cyclo/thesaurus#top> (2019.02.28).
- Dextre Clarke, S. G.; Zeng, M. L. (2012). From ISO 2788 to ISO 25964: the evolution of thesaurus standards towards interoperability and data modelling. // *Information Standards Quarterly*, 24:1 (2012). <http://eprints.rclis.org/16818/> (2017.10.16).
- Fujita, M. S. L. et al. (2016). A gestão de linguagens de indexação por softwares: avaliação para uso no controle de registros de autoridade em catálogos online de bibliotecas. // *Scire*, 23:2 9 (jul.-dic. 2016). 63-73. <https://www.ibersid.eu/ojs/index.php/scire/article/view/4430> (2018.07.30).
- Fujita, M. S. L. et al. (2018) Construction and evaluation of hierarchical structures of indexing languages for online catalogs of libraries: an experience of the São Paulo State University (UNESP). // *Knowledge Organization*, 45 (2018) 220-231.
- García-Marco, F. J. (2016a) Enhancing the Visibility and Relevance of Thesauri in the Web: Searching for a Hub in the Linked Data Environment. // *Knowledge Organization*, 43 (2016) 193-202.

- García-Marco, F. J. (2016b). Normas y estándares para la elaboración de tesauros de patrimonio cultural. // *El lenguaje sobre el patrimonio: estándares documentales para la descripción y gestión de colecciones*. Ministerio de Educación, Cultura y Deporte, Editors: Secretaría General Técnica. Centro de Publicaciones, Ministerio de Educación, Cultura y Deporte, 29-46.
- García-Marco, F. J. (2016c). Teaching Thesaurus Construction: A Top-Down Approach for LIS Undergraduate Programmes. // *AIKO*, 5 (2016) 546-554.
- García-Marco, F. J. (2016d). The evolution of thesauri and the history of knowledge organization: between the sword of mapping knowledge and the wall of keeping it simple. // *Brazilian Journal of Information Studies: Research Trends*. 10 (2016) 1-11.
- García-Marco, F. J. (2017). ISO 25964-2: interoperabilidad de los tesauros con otros sistemas de organización del conocimiento en contextos semánticos. // *Anuario ThinkEPI*, 11, (2017) 305-309.
- Gnoli, C. 2011. Ontological foundations in knowledge organization: the theory of integrative levels applied in citation order. // *Scire* 17:1: (2011) 29-34.
- Golub, K. 2011. Knowledge Organisation Systems. *Technical Foundations*, UK. 6. <http://technicalfoundations.ukoln.ac.uk/subject/knowledge-organisation-systems.html>
- Google Acadêmico (2018). Dagobert Soergel. <https://scholar.google.com.br/citations?hl=pt-BR&user=4bJ6HcUAAAAJ> (2018.10.03).
- International Organization for Standardization. (2011). ISO/FDIS 25964-1: Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval. Geneva: International Organization for Standardization, 2011.
- International Organization for Standardization. (2013). ISO/FDIS 25964-2: Information and documentation – thesauri and interoperability with other vocabularies - part 2: Interoperability with other vocabularies. Geneva: International Organization for Standardization, 2013.
- Lancaster, F. W. (2002). El control del vocabulario en la recuperación de información. Valencia: Universidad de Valencia, 2002.
- Marcondes, C. H. (2011) Um modelo semântico de publicações eletrônicas. *Liinc em Revista*, 7 (2011) 82-103. <http://revista.ibict.br/liinc/article/view/3290>
- Martins, M. C. B. (2013). Indexação e controlo da terminologia em bibliotecas do ensino superior politécnico em Portugal: o sistema no Instituto Politécnico de Portalegre. Universidad de Salamanca: Salamanca. Tese de Doutorado. <http://comum.rcaap.pt/handle/10400.26/7171> (2017.03.25).
- Mazzocchi, F. (2018). Knowledge organization systems (KOS). // *Knowledge Organization*, 45 (2018). 54-78.
- Mendes, P. R.; Reis, R. M. dos; Maculan, B. C. M. S. (2015). Tesauros no acesso à informação: uma retrospectiva. // *Revista ACB: Biblioteconomia em Santa Catarina*, 20:1 (2015) 49-66. <https://bit.ly/2Sxw0eb>. (2016.02.28).
- Soergel, D. (1972). A universal source thesaurus as a classification generator. // *Journal of the American Society for Information Science*, 23:5 (1972) 299-305. <http://www.dsoergel.com/cv/B11.pdf> (2018.07.19).
- Soergel, D. (1974). Automatic and semi-automatic methods as an aid in construction of indexing languages and thesauri. // *International Classification*, 1:1 (1974) 34-38.
- Soergel, D. (1975) Theoretical problems of thesaurus building with particular reference to concept formation. 1975. <http://www.dsoergel.com/cv/B12.pdf> (2018.06.26).
- Soergel, D. (1991). Beyond facets: Semantic roots and modifiers as elements of a conceptual morphology. // *Advances in Classification Research Online*, 2:1 (1991) 149-166. <http://www.dsoergel.com/cv/B36.pdf> (2018.07.19).
- Soergel, D. (1994). Software support for thesaurus construction and display. // *Advances in Classification Research Online*, 5:1 (1994) 157-184. <http://journals.lib.washington.edu/index.php/acro/article/viewFile/13784/11898> (2018.07.198).
- Soergel, D. (1996). SemWeb: Proposal for an open, multi-functional, multilingual system for integrated access to knowledge about concepts and terminology. // *Advances in Knowledge Organization*, 5 (1996) 165-173.
- Soergel, D. (1999). The rise of ontologies or the reinvention of classification. // *Journal of the American Society for Information Science*, 50:12 (1999) 1119-1120. <http://www.dsoergel.com/cv/B70.pdf> (2018.07.19).
- Soergel, D. (2002). Thesauri and ontologies in digital libraries: Tutorial. <http://www.dsoergel.com/cv/B63.pdf> (2018.06.26).
- Soergel, D. (2005). <http://www.dsoergel.com/> (2018.10.03).
- Soergel, D. (2015). Unleashing the power of data through organization: structure and connections for meaning, learning, and discovery. // *Knowledge Organization*, 42:6 (2015) 401-427.
- Soler Monreal, C.; Gil Leiva, I. (2010). Posibilidades y límites de los tesauros frente a otros sistemas de organización del conocimiento: folksonomías, taxonomías y ontologías. // *Revista Interamericana de Bibliotecología*, 33:2 (2010). <https://bit.ly/2NoVNrd> (2019.02.28).
- Souto, L. F. (2003). Recuperação de informações em bases de dados: usos de tesouro. // *Transinformação*, 15:1 (jan./abr.,2003) 73-81. <http://www.scielo.br/pdf/tinf/v15n1/06.pdf> (2019.02.28).

Enviado: 2019-04-01. Segunda versión: 2020-01-13.
Aceptado: 2020-06-04.

Tendências em processos e sistemas da organização do conhecimento de textos narrativos de ficção

Tendencias en procesos y sistemas de organización del conocimiento de textos narrativos de ficción

Trends in Knowledge Organization Processes and Systems for fictional narrative texts

Larissa de Mello LIMA (1), Patrícia de ALMEIDA (2), João Batista Ernesto de MORAES (3), Daniel MARTÍNEZ-ÁVILA (4), Mariana da Silva CAPRIOLI (5), Luana Calcete Vaz TENÓRIO (6)

(1) Universidade Estadual Paulista – Unesp – Marília – São Paulo, Brasil, larissalima.unesp@gmail.com (2) Universidade de Coimbra, Portugal, mebpatricia@gmail.com (3) Universidade Estadual Paulista – Unesp – Marília – São Paulo, Brasil, joao.batista@unesp.br (4) Universidad Carlos III de Madrid – Espanha, dmartine@bib.uc3m.es (5) Universidade Estadual Paulista – Unesp – Marília – São Paulo, Brasil/ Biblioteca Pública Municipal “João Martins Valença” – Marília – São Paulo – Brasil (6) Universidade Estadual Paulista – Unesp – Marília – São Paulo, Brasil, luanacalcete@gmail.com

Resumen

Se analizan los últimos trabajos publicados sobre organización del conocimiento de ficción, recuperados de la base de datos de ISKO Knowledge Organization Literature, con el objetivo de retratar la situación actual e identificar tendencias en el área. Se utilizan como categorías de análisis la distinción de Procesos de Organización del Conocimiento y Sistemas de Organización del Conocimiento de Hjørland. Para cada trabajo se extrajeron unidades textuales discursivas que apoyaron su categorización. Se concluye que a partir de 2010 hay una tendencia hacia los procesos de organización del conocimiento de ficción ya que incluso en los trabajos en que se trabajaron los sistemas de organización del conocimiento estos coexistieron con discusiones sobre los procesos.

Palabras clave: Textos narrativos de ficción. Procesos de Organización del Conocimiento. Sistemas de Organización del Conocimiento.

1. Introdução

De uma forma genérica, pode afirmar-se que a Organização do Conhecimento se preocupa com os modelos organizativos e representativos da realidade. Neste sentido, trata a descrição, representação, arquivamento e gestão; não só dos documentos tradicionalmente acumulados em unidades de informação, como também das suas representações, assuntos e conceitos (Hjørland, 2008; 2016). De acordo com vários autores, entre eles Hjørland e Broughton *et al.* (2005), a Organização do Conhecimento poderá ser dividida em duas correntes principais, a saber: Processos de Organização do Conhecimento (KOP), que incluem procedimentos os técnicos tais como catalogação, análise de assunto, indexação, classificação, etc.; e Sistemas de Organização do Co-

Abstract

We analyze the latest publications on fiction knowledge organization, retrieved from the ISKO database Knowledge Organization Literature, aiming to outline the current situation of the area as well as to identify trends. We used Hjørland's categories of Knowledge Organization Processes and Knowledge Organization Systems for the analysis. We extracted textual discursive units from each text to support their categorization. We conclude that since 2010 the trend is to work on knowledge organization processes of fiction as even those publications that dealt with fiction knowledge organization systems also discussed the processes.

Keywords: Fiction. Knowledge Organization Processes. Knowledge Organization Systems.

hecimento (KOS), relacionados com a elaboração de esquemas classificatórios, listas de cabeçalhos de assunto, tesouros, etc.

No que diz respeito à fundamentação de procedimentos técnicos do tratamento temático de conteúdo, o paradigma de análise amplamente difundido na tradição da Organização do Conhecimento é o relativo ao documento de cariz científico.

Contudo, García-Marco *et al.* (2010, p. 395) chamam a atenção para outros aspetos que começam a fazer-se presentes, nomeadamente os colocados pelos textos de cariz ficcional e narrativo:

The revolution of today is about multimedia information, and a great deal of such information does not suit into the frame of traditional scientific reports and articles. They are much nearer to fiction and narrative documents. In fact, narrative-based com-

munications are gaining considerable room in disciplines like education—with multimedia and video as increasing important tools in face-to-face, blended and virtual education—, psychology and even in fields like business and management in general, where the “story-telling” movement is gaining importance in marketing, corporative culture management and political communication.

Na verdade, nos últimos quarenta anos, a questão do tratamento documental de textos de ficção literária parece ganhar bastante relevância e o interesse de diversos autores, como provam os trabalhos de Pejtersen (1978, 1979, 1983, 1984, 1998), Beghtol (1986, 1992, 1994, 1995, 1997), Hayes (1992), Nielsen (1997), García-Marco (1997), Saarti (1999), Moraes (2006, 2008a, 2008b), entre outros.

Muito embora estes estudos compreendam um quadro temporal e uma tendência relativamente recente, a verdade é que a questão do tratamento documental das obras literárias já se coloca há mais de um século. Há registros de investigação nesta área datados de 1898, como bem lembra Eriksson (2005, p. 01):

Although it often may seem so, the classification and indexing of imaginative literature and fiction in particular is not an issue that has only been discussed in the last few decades. Presumably the first serious article on the subject appeared in 1898, when the noteworthy British librarian Ernest Baker wrote about the classification of fiction in the *Library World* [...] and in the beginning of the 20th century a prolonged discussion concerning the same issue took place in the American Library Association (ALA) after John Thomson had initiated a major fiction classification project at the Wagner Institute Branch of the Free Library of Philadelphia.

Os estudos mais recentes parecem concentrar o seu foco nas questões específicas relativas à indexação, classificação ou recuperação da informação. Deve-se acrescentar, ainda, que, na sua maioria, tais produtos documentais dão enfoque a aspetos externos e formais, em detrimento de aspetos internos temáticos de conteúdo. É muito frequente encontrar uma representação documental com recurso a categorias como género literário (romance, poesia, drama) ou nacionalidade (Romance Inglês, Romance Brasileiro). Por vezes, o tratamento documental é de tal forma genérico que apenas se encontra a referência “ficção”.

A propósito das possíveis categorias de análise do conteúdo de textos literários, García-Marco (2006, p. 35) chama a atenção para a possibilidade de existência de tópicos mais concretos e objetivos, bem como de outros mais abstratos e

interpretativos:

With the multimedia revolution, a growing percentage of content is leaving the relatively unambiguous path of scientific and journalistic genres and approaching the semantic jungles of creative literature. This kind of material requires the complex kind of content analysis that is characteristic of artistic disciplines, that can be summarized in the simplified model that was proposed by the Art historian Panofsky (1955) —also applicable analogically to the analysis of fiction—: description of common subjects; identification of the specific names and exemplars of these common subjects; and interpretation of the abstracts subjects that are being represented through the former.

Nas bibliotecas públicas e escolares, o documento objeto de análise não é apenas o de cariz científico. Ward e Saarti (2018) referem que o tratamento da ficção parece causar alguma consternação ou estranheza, ao contrário da metodologia amplamente definida para a não-ficção. Segundo os autores, este problema precisa ser ultrapassado, uma vez que as obras de ficção constituem a maior parte do acervo das bibliotecas públicas, com taxas de empréstimo bastante elevadas, exemplificando com a taxa de 44% relativa ao ano de 2013, no Reino Unido.

Nessas unidades de informação, o texto narrativo é objeto de especial atenção, seja no desenvolvimento de coleções, seja na organização das mesmas. Deverá, igualmente, ter-se em conta que as bibliotecas especializadas ou académicas, nomeadamente da área de Letras e Humanidades, abrigam um grande acervo de textos ficcionais. Neste âmbito, os documentos carecem de um tratamento temático adequado, de forma a servirem de suporte ao ensino e à pesquisa em cursos de Graduação e Pós-Graduação e em núcleos de investigação.

Apesar de necessária, quando a *matéria prima* são textos narrativos de ficção, a complexidade da análise documental aumenta consideravelmente, pois o modelo metodológico de identificação de conceitos disponível é aquele elaborado para textos científicos. Assim, a leitura documental das partes canónicas do texto, tais como título, subtítulo, resumo, etc. (Chaumier, 1988), poderá não ser suficiente para o devido tratamento dos textos narrativos ficcionais. Broughton (2004) refere que, em muitos casos, o título não ajuda na determinação do conteúdo e, pelo contrário, poderá ser mesmo enganador. A estrutura e as características dos documentos de cariz ficcional obrigam a que a Organização do Conhecimento considere outras formas de evidenciar o seu conteúdo.

Neste contexto, considera-se pertinente verificar quais as tendências para a análise documental

de textos narrativos de ficção nos últimos anos. Pretende-se estabelecer um panorama dos mais recentes estudos, a partir do ano 2010, bem como identificar possíveis tendências na Organização do Conhecimento. Para tal, apresenta-se, na sequência, a metodologia da pesquisa empírica, seguida de resultados e análise dos trabalhos recuperados.

2. Metodologia

Com o objetivo de recuperar os trabalhos que abordassem a questão no texto narrativo de ficção, foi utilizada a base de dados da ISKO *Knowledge Organization Literature* como universo de investigação, uma vez que se trata de uma base de dados especializada em Organização do Conhecimento. Para amostra, foram selecionados os resultados entre os anos 2010 a 2020, recuperados em três fases, com as seguintes expressões de busca: “Narrative texts”; “Fiction”; e “Narrative”.

Na análise, verificou-se se os trabalhos abordaram o tratamento dos documentos narrativos de ficção focando os Processos da Organização do Conhecimento (KOP), os Sistemas de Organização do Conhecimento (KOS), ou ambas categorias, seguindo a divisão de Hjørland e outros. Desta forma, os trabalhos foram analisados, sendo possível identificar as correntes da Organização do Conhecimento a que estes estejam vinculados.

Os resultados apresentam-se em quadro, de forma a facilitar a sua análise e sistematização. A fim de esmiuçar como se chegou à conclusão de que cada artigo refletia Processo de Organização do Conhecimento (KOP), Sistema de Organização do Conhecimento (KOS) ou ambos, recolheram-se algumas citações diretas. Estes excertos e respetivos comentários servem para clarificar a corrente com a qual se estabeleceu a relação.

3. Resultados

A análise dos resultados desenvolveu-se em três fases. Em cada uma das etapas, procede-se, de forma descritiva, à análise individual das expressões de busca.

O termo “Fiction” recuperou treze artigos, embora três deles tenham sido desconsiderados – dois deles por estarem escritos em língua coreana (Kim, 2012; Hosin, 2015) e um deles por tratar o termo ficção apenas no título do artigo de maneira conotativa (Trott e Tarulli, 2011). Assim, nesta primeira fase dos resultados, são analisados dez artigos com o termo “fiction”, sendo os dados expostos no Quadro 1:

It enriches contents by creating semantic connections between works. A data record is almost as if hung in a web made of ontological concepts and annotations where they define the web itself. (Hypén e Mäkelä, 2011, p. 280)

| <i>Artigo recuperado</i> | <i>KOP/S</i> |
|--|--------------|
| 1 – Hypén, K.; Mäkelä, E. (2011). An ideal model for an information system for fiction and its application: Kirjasampo and Semantic Web // <i>Library review</i> . 60:4 (2011) 279-292. | KOP e KOS |
| 2 - Bode, K.; Hetherington, C. (2015). Retrieving a World of Fiction: Building an Index—and an Archive—of Serialized Novels in Australian Newspapers, 1850–1914. // <i>The Indexer</i> . 33:2 (2015) 57-65. | KOP |
| 3 - Saarti, J., Hypén, K. (2010). From thesaurus to ontology: the development of the Kaunokki Finnish fiction thesaurus. // <i>The indexer</i> . 28:2 (2010) 50-59 | KOP e KOS |
| 4 - Moraes, J. B. E. de. (2012). Aboutness in fiction: methodological perspectives for knowledge organization // <i>Categories, contexts and relations in knowledge organization</i> . 13 (2012) 242-248. | KOP |
| 5- Gandier, Â. M.; Pinho, F. A. (2018). A importância da semântica discursiva para a análise documental: um estudo em texto ficcional // <i>Brazilian Journal of Information Studies: Research Trends</i> . 12:2 (2018) 13-23. | KOP |
| 6 - Ward, M.; Saarti, J. (2018). Reviewing, rebutting, and reimagining fiction classification // <i>Journal of Cataloging & Classification Quarterly</i> , 56:4 (2018) 317-329. | KOP |
| 7 - García-Marco, F.-J.; et al (2010). Knowledge organization of fiction and narrative documents: a challenge in the age of the multimedia revolution // <i>Paradigms and conceptual systems in knowledge organization</i> 12 (2010) 262-268. | KOP |
| 8 - Antonio, D. M.; Moraes, J. B. E. de. (2010). Analisis documental de obras de ficcion: aspectos metodológicos y de aplicabilidad [Documental analysis of fiction: methodological aspects of applicability] // <i>Scire: representacion y organizacion del conocimiento</i> . 16:2 (2010) 71-78. | KOP |
| 9 - Sundström, A. da S. S.; Moraes, J. B. E. de; Albuquerque, A. C. (2019). Filme de ficção para a Ciência da Informação: Um estudo sobre as abordagens de organização e representação temática // <i>Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação</i> . 24:54 (2019) 124-134. | KOP |
| 10 - Aagaard, H.; Viktorsson, E. (2014). Subject headings for fiction in Sweden: a cooperative development // <i>Cataloging and classification quarterly</i> . 52:1 (2014) 62-68. | KOS |

Quadro 1. Trabalhos recuperados com o termo “fiction”

O primeiro artigo recuperado com o termo “fiction” foca inicialmente os processos de Organização do Conhecimento, categorizando as conexões semânticas entre as palavras e voltando o olhar para o significado dos conceitos. Este processo é feito com o objetivo de fundamentar a criação de sistemas de organização de conhecimento de forma menos mecânica, levando em conta a dimensão ontológica dos conceitos.

Our model links the abstract concepts of work and author with specific documentary manifestations to represent fiction serialised in Australian newspapers in terms of processes rather than stable products. (Bode e Hetherington, 2015, p. 61)

O segundo artigo propõe um modelo que vincula conceitos abstratos a características específicas, no momento de representar textos de ficção em jornais australianos. Aqui, o interesse é delinear os processos de Organização do Conhecimento.

How can such traditional library tools as classification and indexing be combined with modern tools such as ontology-based annotation, tagging and user-evaluation? Defining the borderlines and the overlap between library work and the work done by the web communities will be an essential part of finding an answer to the challenges. (Saarti e Hypén, 2010, p. 54)

O terceiro artigo dá enfoque aos problemas nos Sistemas de Organização do Conhecimento e delinea a seguinte questão: Como é que as ferramentas tradicionais de indexação, voltadas para os sistemas de Organização do Conhecimento, podem interoperar a nível sintático com as ferramentas modernas de base ontológica, sem deixar de lado os processos de Organização do Conhecimento?

The subject analysis of narrative texts of fiction is complex, among other causes, since the methodological model of identification of concepts is elaborated for scientific texts. In such a way, it appears the proposal of theoretical and methodological use of the Generative Trajectory of Meaning postulate by Greimas with the intention of contribute for the identification of aboutness in narrative texts of fiction. (Moraes, 2012, p. 5)

O quarto artigo recuperado procura contribuir para a identificação dos aspetos temáticos dos textos narrativos de ficção, levando em consideração a sua complexidade. Para tal, os autores propõem a teoria da trajetória do significado de Greimas, focando, assim, nos processos de Organização do Conhecimento.

Este artigo se ocupa dos percursos temático e figurativo oriundos da Semântica Discursiva de Algirdas Julien Greimas, a serem aplicados em textos ficcio-

nais, de modo específico à crônica, que é considerado um dos gêneros literários da tradição (Gandier e Pinho, 2018, p. 19).

O quinto artigo trabalha numa metodologia teórico-aplicada, que leva em consideração a teoria da semântica discursiva, para servir de apoio à análise documental. Portanto, o foco aqui está nos processos de Organização do Conhecimento.

This article explores accepted definitions of fiction classification and puts forth three suggestions: a theoretical multi-warrant faceted classification, the use of controlled terminology in affective dimension searching (as well as the further integration of such searching in libraries), and the use of staff's tacit knowledge encapsulated as book reviews and blogging as a recommendation system that could be linked to the library catalog. (Ward e Saarti, 2018, p. 13)

O sexto artigo recuperado explora, a um nível teórico, as definições sobre classificação de obras de ficção, estando alinhado com os processos de Organização do Conhecimento.

The world of genres, subgenres and universals is very difficult to be represented in a one-dimensional classification that can be truly useful for narrative documents discovery. It would require a complex analysis and faceting, and, so, it results to be a very interesting field for experimenting with ontologies. (García-Marco, 2010, p. 25)

O sétimo artigo defende o uso de ontologias para dar conta da análise de textos narrativos de ficção, pelo que o foco recai nos processos de Organização do Conhecimento.

A análise documental de obras de ficção depende de procedimentos metodológicos que proporcionem os elementos necessários para a qualidade da indexação. (Antonio e Moraes, 2010, p. 72)

O oitavo artigo ressalta a importância dos procedimentos metodológicos adequados no processo de análise documental dos textos narrativos de ficção, estando o foco nos processos de Organização do Conhecimento.

O documento audiovisual em foco nesta pesquisa é o filme de ficção. Desse modo, questiona-se: como a Ciência da Informação está tratando a temática desses materiais? Assim, foi feito um levantamento bibliométrico em base de dados com o objetivo de analisar a literatura a respeito dos modos de tratamento temático de filmes de ficção no âmbito da Ciência da Informação e, também, entender como vem ocorrendo o tratamento temático desses materiais. (Sundström *et al.*, 2019, p. 52)

O nono artigo procura entender como se dá o tratamento temático do documento em formato audiovisual, portanto, é voltado para os processos de Organização do Conhecimento.

Subject headings make it easier to find and choose fiction. A consistente and standardized national system for subject headings can do a lot more. (Aagaard e Viktorsson, 2014, p. 67)

O décimo artigo narra a iniciativa do Comitê de Indexação de Ficção da Suécia e o desenvolvimento de um padrão nacional para a indexação de ficção. A procura de um sistema que abranja as singularidades dos textos narrativos de ficção coloca o foco nos sistemas de organização do conhecimento.

Na segunda fase da análise dos resultados, o termo “Narrative texts” recuperou apenas um artigo. Este encontra-se listado no Quadro 2 e analisado em seguida:

| <i>Artigo recuperado</i> | <i>KOP/S</i> |
|---|--------------|
| 1 Bell, H. K. (2012). Biographies as soft, narrative texts. <i>The Indexer: The International Journal of Indexing</i> , 30(3), 141-146. | KOP |

Quadro 2: Artigos recuperados com o termo “narrative texts”

Biographies constitute a genre for indexing, rather than a subject specialism. Indexers of human lives are out there alone, wanting the standard guidelines available to indexers in specialized subject fields, able to call only on sensitive reading of the text, subtle and flexible use of language, diligence and common sense to fulfill our task. (Bell, 2012, p. 145)

Este artigo destaca os documentos biográficos no momento de indexação, defendendo-se a necessidade de padrões e diretrizes, para que os indexadores possam efetuar o tratamento técnico deste gênero textual.

Na terceira fase da análise dos resultados, o termo “Narrative” recuperou quatro artigos, listados no Quadro 3.

A ideia de organizar e representar os topônimos para elaboração de linguagens documentárias vem ampliar os estudos na área da Ciência da Informação, dando um caráter interdisciplinar, uma vez que estimula uma relação com outras ramificações do conhecimento. (Oliveira *et al.*, 2012, p. 125).

O primeiro artigo procura organizar topônimos, para que seja possível elaborar linguagens documentárias consistentes no contexto da Ciência da Informação e Organização do Conhecimento. Entende-se, portanto, que o trabalho se insere na categoria processos de Organização do Conhecimento.

In particular, we recognize that the primary role of a taxonomy is to describe or narrate the natural relationships between concepts in a given document corpus. Therefore, a corpus-aware adaptation of a taxonomy should essentially distill the structure of

the existing taxonomy by appropriately segmenting and, if needed, summarizing this narrative relative to the content of the corpus. Based on this key observation, we propose A Narrative Interpretation of Taxonomies for their Adaptation (ANITA) for re-structuring existing taxonomies to varying application contexts and we evaluate. (Cataldi *et al.*, 2012, p. 125)

| <i>Artigo recuperado</i> | <i>KOP/S</i> |
|---|--------------|
| 1- Oliveira, M. O. E. de; Santos, B. P. dos; Oliveira, F. B. de. (2012). A representação dos identificadores geográficos na identidade cultural: um estudo em narrativas orais da área do marajó na amazônia paraense. // Guimarães, J. A. C.; Dodebeí, V. (org.). (2012). <i>Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade</i> . Marília, SP: SKO-Brasil: FUNDEPE. 137-142. | KOP |
| 2 Cataldi, M., Candan, K. S., Sapino, M. L. (2012). Narrative-based taxonomy distillation for effective indexing of text collections // <i>Data and knowledge engineering</i> . 72 (2012) 103-125. | KOP |
| 3 Candido, G. G.; Lima, L. M.; Moraes, J. B. E. de. (2013). <i>Aspectos teóricos del análisis de la estructura narrativa y su contribucion a la organizacion del conocimiento</i> // <i>Scire</i> , 19:2 (2013) 91-96. | KOP |
| 4- Moncada Landeta, R. (2016). Aplicacion de modelos semióticos al análisis de poesia narrativa popular ecuatoriana difundida en formatos audiovisuales. // <i>Cuadernos de Documentación Multimedia</i> . 27:1 (2016) 15-31. | KOP |

Quadro 3: Artigos recuperados com o termo “narrative”

O segundo artigo trabalha com taxonomias e com a sua capacidade de realizar a descrição das relações entre conceitos. Neste estudo, é proposta uma interpretação narrativa das taxonomias, pelo que se refere a processos de organização do conhecimento.

Os aspectos teóricos apresentados pelos autores Vladimir Propp, Ferdinand Saussure e A. J. Greimas proporcionam métodos que permitem identificar a estrutura narrativa. Ao serem utilizados em conjunto com a Análise Documental, que tem seu cunho na Ciência da Informação, possibilitam identificar a temática central do documento, auxiliando desta forma no processo de representação da informação. (Candido *et al.*, 2013, p. 96)

O terceiro artigo procura mapear os aspectos teóricos e metodológicos da estrutura narrativa dos textos, sistematizados por Vladimir Propp, Ferdinand Saussure e A. J. Greimas. Sendo assim, foca nos processos de organização do conhecimento.

Se expone una fundamentación teórica, lógica y metodológica de tres modelos semióticos: los elementos y funciones de la comunicación de Jakobson; el modelo actancial de Greimas; y los mundos: real, referencia y posible de Eco, aplicados al análisis de poesía narrativa popular difundida en distintos formatos audiovisuales. (Moncada Landeta, 2016, p.15)

O quarto artigo expõe, a um nível teórico, lógico e metodológico, três modelos semióticos que podem auxiliar no processo de análise, tendo como intuito a recuperação de poesia narrativa popular. Desta forma, o trabalho insere-se na categoria processos de organização do conhecimento.

4. Discussão dos resultados

Tal como apontado por Chaumier (1988), até à década de 2000, os estudos de Organização do Conhecimento registaram a influência da corrente norte-americana e dos trabalhos de Lancaster (2004). Desta forma, pode afirmar-se que foi notória uma corrente mais voltada para os sistemas de organização de conhecimento. Isto significa que o foco principal foi discutir o produto da indexação e o tempo hábil para a sua realização, de forma a otimizar a gestão da biblioteca como um todo. Sendo assim, a indexação foi encarada como uma técnica.

Tendo por base os dados recolhidos no presente estudo, foi possível perceber que os artigos recuperados na década de 2010 passaram preocupar-se mais com os Processos de Organização do Conhecimento. Verificou-se, igualmente, que as investigações e discussões voltadas para os sistemas de Organização do Conhecimento coexistem com as relativas aos processos de Organização do Conhecimento.

Em algum momento do seu desenvolvimento, todos os 15 trabalhos analisados abordaram a questão da importância dos processos na Organização do Conhecimento. Constatou-se também que apenas dois artigos registam uma discussão que associa os processos aos sistemas de Organização do Conhecimento, a saber:

- Hypén e Mäkelä (2011) propõem uma ideia de modelo para sistemas de informação voltados para a ficção com aplicabilidade do mesmo. As estruturas teóricas dos processos de Organização do Conhecimento são discutidas na construção do Sistema de Informação.
- Saarti e Hypén (2010) abordam a criação de um tesouro de ficção, enquanto sistema de Organização do Conhecimento. Para tal, os autores percorrem a teoria dos tesouros e ontologias, colocando o foco nos processos de Organização do Conhecimento.

Considera-se que este exercício colaborativo sistema/processo seja o desejável e ideal no contexto prático da generalidade das unidades de formação. Não será coincidência que ambos os trabalhos se refiram ao contexto finlandês, onde a indexação de ficção se encontra bem consolidada.

5. Considerações finais

Este estudo parece indicar que a década de 2010 trouxe uma nova tendência à Organização do Conhecimento de textos narrativos ficcionais. Neste contexto, foi possível deduzir que o foco da maioria dos trabalhos analisados é a forma como as etapas no processo de indexação podem ser detalhadas, procurando explicar-se o procedimento passo a passo. Pode-se afirmar, também, que as discussões sobre os sistemas de Organização do Conhecimento coexistem com as relativas aos processos de Organização do Conhecimento.

Dada a importância dos texto de ficção literária e tendo em conta o universo reduzido em análise neste estudo, considera-se necessária uma visão mais ampliada sobre este assunto. Neste sentido, sugerem-se estudos que abranjam outras bases de dados nomeadamente *Web of Science* ou *Google Scholar*, bem como outros períodos de tempo, em concreto um estudo diacrónico, tendo como ponto de partida o século XX.

Referências

- Aagaard, H.; Viktorsson, E. (2014). Subject headings for fiction in Sweden: a cooperative development // *Cataloging and classification quarterly*. 52:1 (2014) 62-68.
- Antonio, D. M.; Moraes, J. B. E. de (2010). Analisis documental de obras de ficcion: aspectos metodológicos y de aplicabilidad [Documental analysis of fiction: methodological aspects of applicability]. // *Scire: representacion y organizacion del conocimiento*. 16:2 (2010) 71-78.
- Bell, H. K. (2012). Biographies as soft, narrative texts. *The Indexer: The International Journal of Indexing*, 30:3 141-146.
- Beghtol, C. (1986). Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. // *Journal of Documentation*, London. 42:2 (June 1986) 84-113.
- Beghtol, C. (1992). Toward a theory of fiction analysis for information storage and retrieval. // Williamson, N.J.; Hudson, M. (1992). *Classification research for knowledge representation and organization*. Amsterdam: Elsevier.
- Beghtol, C. (1994). The classification of fiction: the development of a system based on theoretical principles. Metuchen, NJ: Scarecrow Press, 1994.
- Beghtol, C. (1995). Domain analysis, literary warrant, and consensus: the case of fiction studies. // *Journal of the American Society for Information Science*. 46:1 (1995) 30-44.

- Beghtol, C. (1997). Stories: applications of narrative discourse analysis to issues in information storage and retrieval. // *Knowledge Organization*. 24: 2 (1997) 64-71.
- Bode, K.; Hetherington, C. (2015). Retrieving a World of Fiction: Building an Index—and an Archive—of Serialized Novels in Australian Newspapers, 1850–1914. // *The Indexer*. 33:2 (2015) 57-65.
- Broughton, V. (2004). *Essential classification*. London: Facet Publishing.
- Broughton, V.; Hansson, J.; Hjørland, B.; López Huertas, M. J. (2005). *Knowledge Organization*. // Kajberg, L.; Lørring, L. (eds.), *European Curriculum Reflections on Library and Information Science Education*. Copenhagen: Royal School of Information Science (2005) 133–148.
- Candido, G. G.; Lima, L. M.; Moraes, J. B. E. de. (2013). Aspectos teóricos del análisis de la estructura narrativa y su contribución a la organización del conocimiento // *Scire*, 19:2 (2013) 91-96.
- Cataldi, M., Candan, K. S., Sapino, M. L. (2012). Narrative-based taxonomy distillation for effective indexing of text collections. // *Data and knowledge engineering*. 72 (2012) 103-125.
- Chaumier, J. Indexação: conceito, etapas e instrumentos. // *Revista Brasileira de Biblioteconomia e Documentação*. 21: 1/2 (jan./jun. 1988) 63-79.
- Eriksson, R. (2005). The classification and indexing of imaginative literature. 16th ASIS&T SIG/CR Classification Research Workshop, Charlotte: NC, October 29, 2005.
- Gandier, Â. M.; Pinho, F. A. (2018). A importância da semântica discursiva para a análise documental: um estudo em texto ficcional. // *Brazilian Journal of Information Studies: Research Trends*. 12:2 (2018) 13-23.
- García-Marco, F. J. (2006). Understanding the categories and dynamics of multimedia information: a model for analysing multimedia information. // *Proceedings of the Ninth International ISKO Conference*, Vienna. (July, 2006) 395-404.
- García-Marco, F.-J.; et al (2010). Knowledge organization of fiction and narrative documents: a challenge in the age of the multimedia revolution // *Paradigms and conceptual systems in knowledge organization* 12 (2010) 262-268.
- García-Marco, L. F.; García-Marco, F. J. (1997). El resumen de documentos literarios narrativos: algunas propuestas metodológicas. // *Organización del Conocimiento en Sistemas de Información y Documentación*. 2 (1997) 73-85.
- Hayes, S. (1992). Enhanced catalog access to fiction: a preliminary study. // *Library Resources & Technical Services*. 36:4 (1992) 441-59.
- Hjørland, B. (2008). What is Knowledge Organization (KO)? // *Knowledge Organization*. 35:2/3 (2008) 86-101.
- Hjørland, B. (2016). Knowledge organization. // *Knowledge Organization*. 43:6 (2016) 475-84.
- Hypén, K.; Mäkelä, E. (2011). An ideal model for an information system for fiction and its application: Kirjasampo and Semantic Web // *Library review*. 60:4 (2011) 279-292.
- Hosin L. (2015). Narrative Inquiry, as a Research Methodology for Library and Information Science. // *Journal of the Korean Society for Library and Information Science*. 49:1 (2015) 149-172.
- Kim, S. (2012). A research on close classification for Korean fiction in KDC: focusing on public library. // *Journal of the Korean society for library and information science*. 46:3 (2012) 57-78.
- Lancaster, F.W. (2004). *Indexação e resumos: teoria e prática*. 2. ed. Brasília: Briquet de Lemos, 2004.
- Moncada Landeta, R. (2016). Aplicación de modelos semióticos al análisis de poesía narrativa popular ecuatoriana difundida en formatos audiovisuales. // *Cuadernos de Documentación Multimedia*. 27:1 (2016) 15-31.
- Moraes, J. B. E. (2006) Guimarães, J. A. C. Análisis documental de contenido de textos literarios narrativos: en busca del diálogo entre las concepciones de aboutness/meaning y de recorrido temático/recorrido figurativo. // *Scire*, 12:1(2006) 71 - 84.
- Moraes, J. B. E. (2008a) Análise documental de crônicas: reflexões sobre uma trajetória de interlocução entre literatura, linguística e ciência da informação // Guimarães, J.A.C; Fujita, M.S.L. (org). (2008). *Ensino e pesquisa em Biblioteconomia no Brasil: a emergência de um novo olhar*. São Paulo: Editora Cultura. 129-144.
- Moraes, J. B. E.; Guimarães, J. A. C. (2008b) Análise documental de conteúdo de textos literários narrativos: em busca de um diálogo entre as concepções de aboutness/meaning e percurso temático/percurso figurativo // Gaspar, N.R.; Romão, L.M.S. (org). (2008) *Discurso e Texto: multiplicidade de sentidos na Ciência da Informação*. São Carlos, SP: EDUFSCar. 35-45.
- Moraes, J. B. E. de. (2012). Aboutness in fiction: methodological perspectives for knowledge organization // *Categories, contexts and relations in knowledge organization*. 13 (2012) 242-248.
- Nielsen, H. J. (1997). The nature of fiction and its significance for classification and indexing. // *Information Services & Use*. 17:2-3 (1997) 171-181.
- Oliveira, M. O. E. de; Santos, B. P. dos; Oliveira, F. B. de. (2012). A representação dos identificadores geográficos na identidade cultural: um estudo em narrativas orais da área do marajó na amazônia paraense. // Guimarães, J. A. C.; Dodebei, V. (org.). (2012). *Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade*. Marília, SP: ISKO-Brasil: FUNDEPE. 137-142.
- Panofsky, E. (1955). *Meaning in the visual arts: papers in and on Art History*. Garden City, N.Y.: Doubleday.
- Pejtersen, A. M. (1978). Fiction and library classification. // *Scandinavian Public Library Quarterly*. 11:1 (1978) 5-12.
- Pejtersen, A. M. (1979). The meaning of 'about' in fiction indexing and retrieval. // *Aslib Proceedings*. 31: 5 (May,1979) 251-257.
- Pejtersen, A. M.; Albrechtsen, H.; Sandelin, R.; Lundgren, L.; Valtonen, R. (1998). The Scandinavian book house: indexing methods and OPAC development for subject access to Scandinavian fiction literature. // *Advances in classification research*. 6 (1998) 99-113.
- Pejtersen, A. M.; Austin, J. (1983). Fiction retrieval: experimental design and evaluation of a search system based on users' value criteria: part 1. // *Journal of Documentation*. 39:4 (1983) 230-246.
- Pejtersen, A. M.; Austin, J. (1984). Fiction retrieval: experimental design and evaluation of a search system based on users' value criteria: part 2. // *Journal of Documentation*. 40:1 (1984) 25-35.
- Saarti, J. (1999). Fiction indexing and the development of fiction thesauri. // *Journal of Librarianship and Information Science*. 31: 2 (1999) 85-92.
- Saarti, J., Hypén, K. (2010). From thesaurus to ontology: the development of the Kaunokki Finnish fiction thesaurus. // *The Indexer*. 28:2 (2010) 50-59.
- Sundström, A. da S. S.; Moraes, J. B. E. de; Albuquerque, A. C. (2019). Filme de ficção para a Ciência da Informação: Um estudo sobre as abordagens de organização e representação temática // *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*. 24:54 (2019) 124-134.

Trott, B.; Tarulli, L. (2011). Readers' services and the library catalog: coming of age fiction? Or non-fiction? // Reference & user services quarterly. 51:2 (2011) 115-118.

Ward, M.; Saarti, J. (2018). Reviewing, rebutting, and reimagining fiction classification // Journal of Cataloging & Classification Quarterly. 56:4 (2018) 317-329.

Enviado: 2020-03-27. Segunda versión: 2020-05-22.

Aceptado: 2020-06-04.

Herramientas y métricas aplicadas al marketing editorial: del big data a la inteligencia artificial

Tools and metrics in editorial marketing: from big data to artificial intelligence

Antonia SALVADOR BENÍTEZ, Juan Miguel SÁNCHEZ VIGIL

Facultad de Ciencias de la Documentación (UCM), Santísima Trinidad, 37, 28010 Madrid,
asalvado@ucm.es, jmvigil@ucm.es

Resumen

El sector editorial es un paradigma en lo que se refiere al marketing, con un espacio específico en las secciones generales del organigrama de las empresas, al que se suman: edición, producción, comercial, comunicación y derechos. La función del marketing es transversal y envolvente al interrelacionar con el resto de elementos de la cadena de elaboración del libro (impreso o digital). Son objeto de estudio en este artículo las herramientas y métricas aplicadas al marketing editorial, tomando como referente tres modelos (Tekstum, Jellybooks y Bookmetrix), que desde diferentes planteamientos, y con el big data y la inteligencia artificial soportando sus actividades, tienen el denominador común de conocer los hábitos de lectura a través del análisis de la información, al objeto de conseguir una mayor y mejor representación y organización del conocimiento que atraiga y fidelice lectores.

Palabras clave: Edición. Marketing editorial. Organización del conocimiento. Big data. Inteligencia artificial. Tekstum. Jellybooks. Bookmetrix.

1. Introducción

Big data e inteligencia artificial son términos cuyas aplicaciones llevan tiempo impulsando y transformando no solo cómo accedemos, interpretamos y procesamos datos. En el sector del libro se dispone de aplicaciones y dispositivos de lectura digitales que proporcionan una gran cantidad de datos sobre el comportamiento del lector. Tabletas y lectores electrónicos recopilan datos a través de las aplicaciones, lo que permite saber y medir, por ejemplo, los tiempos promedio de lectura, el momento de la misma o cómo se interactúa con ciertos géneros. Estos datos facilitan información para conocer y comprender el comportamiento de los lectores, lo que derivará en los modelos y contenidos de publicación, y en la identificación y recomendaciones personalizadas; en definitiva, en una mayor y mejor organización del conocimiento.

Son objeto de estudio en este artículo las herramientas y métricas aplicadas al marketing editorial, tomando como referente tres modelos (Tekstum, Jellybooks y Bookmetrix) que tienen el de-

Abstract

The publishing sector is a paradigm in terms of marketing with a specific space in the organization chart of the companies, to which Edition, Production, Commercial, Communication and Rights are added. The function of marketing is transversal and interrelates with the rest of the elements of the book production chain (printed or digital). The tools and metrics applied to editorial marketing are studied in this article, taking as reference three models (Tekstum, Jellybooks and Bookmetrix), that from different approaches, and with big data and artificial intelligence supporting their activities, they have the common denominator of knowing reading habits through the analysis of information, in order to achieve a greater and better representation and organization of knowledge to attract and retain readers.

Keywords: Edition. Editorial marketing. Knowledge organization. Big data. Artificial Intelligence. Tekstum. Jellybooks. Bookmetrix.

nominador común de conocer los hábitos de lectura a través del análisis de la información. Para ello se examinan los parámetros e indicadores contemplados en cada caso a partir del testado y el análisis de los contenidos proporcionados en los website de cada una de las plataformas.

Sobre la planificación y el marketing en edición son varios los estudios de interés (Cole, 2003; Arqués Salvador, 2009; Gómez-Tarragona, 2010; Álvarez Ruiz, 2012; Escribano Hernández, 2012 y Fernández de la Puente-Campano, 2016), y en cuanto a métricas y big data se han tomado de referencia los clásicos (Valverde, 2016; González-Fernández, 2016, Zhao, Adib y Katabi, 2016; Serrano Cobos, 2014 y Barranco-Fragoso, 2012). Los datos de referencia relativos a la producción editorial, consumo y lectura proceden de la Panorámica de la edición de libros en España (Ministerio de Cultura) y de la Federación del Gremio de Editores de España.

Editores y autores disponen de perfiles en las redes sociales para gestionar y promover su presencia online e interactuar con los lectores.

También podemos ver esta tendencia en el crecimiento y el interés en colecciones digitales creadas cuidadosamente por los lectores en plataformas como Goodreads. Las acciones que derivan de estas revelaciones moldean e influyen en la creación de contenidos, así como en estrategias de marketing y comercialización con altas probabilidades de éxito. Los beneficios de estos datos son realmente positivos tanto para los lectores como para los editores: los primeros descubren más contenidos que responden mejor a sus necesidades, mientras que los segundos fomentan la creación de libros aún mejores, que a su vez generan beneficios al identificar aquello que interesa.

La tecnología permite convertir la información en inteligencia que tiende a redefinirse con la aparición de algoritmos. Estos datos pueden esconder y describir patrones de funcionamiento y relaciones, inapreciables a simple vista, pero que una vez tratados puede resultar una información muy útil como base para la toma de decisiones. El valor del big data surge cuando lo convertimos en información y, a través de la analítica, generamos información útil para la toma de decisiones.

Los datos masivos que aportan emails, encuestas, redes sociales, señales de móvil, conexiones, logins, entre otros muchos, ofrecen una información de altísimo valor, a veces incluso en tiempo real, para cualquier negocio, también para las editoriales. Chantal Restivo-Alessi, Directora Digital de Harper Collins, señala que la digitalización ha favorecido el acceso a los datos para conocer el comportamiento de los lectores, saber si se está llegando a la audiencia correcta, de la forma idónea y con los contenidos y formatos adecuados (Megapractical, 2016). A través de la venta directa, de las librerías online o mediante las plataformas sociales, el editor puede averiguar los gustos y hábitos de sus lectores, aspectos que le guiarán de manera efectiva en la gestión, desde el punto de vista de la estrategia de negocio, la promoción y el marketing.

Para ello es necesario definir qué queremos saber: hábitos, gustos, comportamientos, etc., es decir formular las preguntas estratégicas con cuyas respuestas, se puedan lograr los objetivos fijados. De este modo, las editoriales pueden testar y medir resultados, recabar otros nuevos y hacer los ajustes necesarios para evitar o reducir errores y alcanzar sus metas. Una de las principales aportaciones de los Big data es devolver a los clientes el valor de dicha información, ofreciéndoles servicios adaptados o productos totalmente nuevos que respondan mejor a sus intereses.

Prácticamente en cualquier área de una empresa o sector de actividad que nos imaginemos es posible encontrar aplicaciones para la analítica. En este artículo exploraremos de forma específica las relacionadas con el marketing en el sector editorial. Se presenta el estado actual del sector editorial, así como herramientas y tendencias de futuro que ponen en relación tres disciplinas que se van entrelazando a lo largo del texto: big data, o cómo y por qué se almacenan los datos y de dónde proceden; el papel de la Analítica de datos, o cómo extraer información útil de éstos mediante procesos matemáticos y estadísticos, y finalmente la inteligencia artificial, o cómo los datos pueden convertirse en inteligencia.

1.1. No todo es *big data*

Los big data figuran, junto con la impresión 3D, la inteligencia artificial, el desarrollo móvil o el Internet de las cosas, entre las principales tendencias tecnológicas dentro de la economía digital en los próximos cuatro años, según el informe anual de Gartner (2019). David Vivancos (2016) define big data como el “conjunto de tecnologías y procesos que están permitiendo capturar y almacenar cantidades ingentes de datos de distinta procedencia y tipologías, siendo la base tanto de la digitalización masiva del mundo analógico, como del almacenamiento de los propios datos generados en el mundo digital”. Estos contenidos comenzaron siendo textos en chats y foros, luego fotografías y vídeos en redes sociales y blogs, y en los últimos años, se han sumado los datos de nuestro entorno. Así es como surgió el denominado Internet de las cosas (Internet of Things, IoT), donde millones de dispositivos transmiten sus datos de funcionamiento en tiempo real. Los big data poseen diferentes dimensiones que se complementan y que se denominan las 5 V's: Volumen, Velocidad, Variedad, Veracidad y Valor.

El valor de una empresa está también en los datos que posee, especialmente en aquellas cuya base de operaciones es Internet. La información y los datos siempre han estado ligados al poder, pero ahora esta afirmación cobra una dimensión mucho mayor. Los datos que generan las empresas son el factor diferencial -si se usan correctamente- para entender mejor a los clientes, optimizar los procesos y mejorar la competitividad. Sin embargo, no se considerarían big data por su volumen (según el tamaño de la empresa pueden oscilar de decenas de Megabytes a cientos de Gigabytes). Salvo en grandes empresas de comercio electrónico, no se usa un big data para almacenar los datos, que suelen estar gestionados a través de un software de CRM (Customer Relationship Management) o gestor de relaciones

con los clientes, y por un software de ERP (Enterprise Resource Planning) o Planificador de Recursos. Tanto los CRM como los ERP suelen estar soportados por un gestor de bases de datos relacional que difiere de los usados por un big data, que trabaja con datos mucho más variados: textos, fotografías, vídeos, audio, series de datos temporales y muchos otros.

El marketing permite una visión global e integrada, que fusiona en tiempo real información de los CRM con los datos generados en los diversos canales de comunicación de la empresa, entre los que destacan los de las redes sociales. Así mismo pueden considerarse big data los contenidos audiovisuales, los datos de sensores, tanto del Internet de las cosas como de dispositivos wearables, de wikis y blogs de gran tamaño, gubernamentales, médicos, científicos, etc., en la mayoría de los casos con volúmenes de muchos Terabytes o Petabytes. Internet es el gran generador de contenidos y datos a mayor velocidad (Tabla I).

| Contenido | Volumen/minuto |
|--|----------------------------------|
| Emails enviados | 188 millones |
| Mensajes (WhatsApp y Facebook messenger) | 42 millones |
| Vídeos Youtube | 4.5 millones videos visualizados |
| Búsquedas en Google | 4 millones |
| Snapchat | 2.1 millones snaps creados |
| Netflix | 700 horas vistas |

Tabla I. Internet in a minute, 2019 (Elaboración propia a partir de www.visualcapitalist.com)

1.2. Analítica de datos e inteligencia artificial

La medición es la clave de cualquier proyecto, y en el contexto del marketing digital, permite predecir el comportamiento del consumidor. En internet medimos consumo, comparación de contenidos, tráfico web, leads, ventas, etc.; sin embargo, lo más difícil no es obtener el dato, sino ser capaz de interpretarlo adecuadamente y tomar las decisiones oportunas. La analítica web es la evaluación y explicación racional de los datos, provenientes o no del big data, con objeto de tomar decisiones estratégicas a nivel de negocio, producto o precio, proponer mejoras o analizar tendencias y comportamientos recurrentes en el tiempo. Consistiría en formular una serie de preguntas con objeto de obtener y analizar los datos de una empresa, que pueden ser de dos tipos (Tabla 2):

(a) Datos de situación/informes. Ofrecen una visión del estado actual de una organización y forman parte del llamado *business intelligence*, con objeto de unificar los procesos de almacenamiento y tratamiento de la información generada en la actividad de una empresa, y utilizarlos como base para la toma de decisiones (1).

(b) Datos tácticos y estratégicos. Ofrecen un conocimiento más profundo de la empresa para poder estar preparados y anticipar las decisiones de futuro más inmediato. Permiten entender qué cambios son necesarios para alcanzar el mejor resultado posible. En estos datos tácticos y estratégicos es donde el retorno de la inversión o ROI (*Return On Investment*) (2) de la analítica predictiva es mucho mayor, en ocasiones cerca del 300%.

En combinación con la analítica, el desarrollo de técnicas de visualización ha sido un aspecto fundamental para la interpretación de los datos, ya que precisamente, la mayor parte de nuestro cerebro está destinada al procesamiento de datos visuales.

| Tipos de datos | Objetivos | Preguntas clave |
|--|--|---|
| Datos de situación (Business intelligence) | Conocer el estado actual de la empresa | ¿Qué ha sucedido? |
| Datos tácticos y estratégicos | Comprender los datos de la empresa para estar preparados | ¿Por qué ha sucedido? |
| | Anticipar las decisiones de futuro | ¿Qué será lo próximo que ocurrirá? |
| | Qué cambiar para obtener el mejor resultado posible | ¿Qué es lo mejor y lo peor que puede suceder? |

Tabla II. Analítica de datos (Business intelligence)

El uso de gráficos 2D/3D, animaciones o incluso técnicas como la realidad virtual o aumentada están jugando un papel clave en el análisis de este tipo de datos, tan dinámico y multidimensional. De igual modo, ofrecen nuevas formas de presentar los resultados y métricas mediante animaciones y gráficos interactivos, que permiten al usuario modificar la visualización en función de los datos que le interesen.

Los algoritmos *deep learning* o de aprendizaje profundo de máquinas son los que han cosechado los mejores resultados en los últimos cinco años gracias principalmente a tres factores: la disponibilidad de cantidades ingentes de datos de diversas tipologías o un verdadero big data (contenidos de los que aprender); el crecimiento exponencial de la capacidad y velocidad de procesamiento de las máquinas (para aprender más rápido); y la aparición y desarrollo de una serie de algoritmos para simular el funcionamiento del cerebro humano dentro de una máquina, muchos de ellos ideados y descartados hace unas décadas, y que a día de hoy están demostrando ser excepcionalmente efectivos gracias a la existencia de los dos factores anteriores. En consecuencia, los puntos siempre se conectan a posteriori, y éste es el caso de las tres disciplinas que acabamos de introducir: el big data, la analítica y la inteligencia artificial.

2. El sector editorial en cifras

El sector editorial es una industria cultural que se organiza en torno a una cadena de valor cuyo producto final es el libro. Hasta ahora esta cadena era muy sencilla: autor, editor, productor, comercial y lector. Sin embargo, las nuevas tecnologías han alterado de forma radical la cadena, tanto en la producción como en el consumo al no haber un autor sino miles autores que escriben sus blogs o que autopublican gracias a la tecnología.

Internet ha reducido considerablemente el coste de creación de una editorial, y además es muy sencillo crear un espacio en la red, buscar autores dispuestos a escribir y vender sus creaciones. Lo que no ha cambiado es que sigue habiendo un lector, pero que ahora cuenta con infinidad de opciones para llegar a las lecturas y encontrar nuevos autores. El reto está, como antes, en la visibilidad.

El sector editorial ha tenido que hacer frente a un proceso de transformación como consecuencia de la irrupción digital. Este nuevo escenario ha multiplicado exponencialmente el número de fuentes a las que los lectores pueden acceder, surgiendo así un lector más participativo y más exigente. Ya no sólo se miden las ventas como

dato último, sino la capacidad de los libros para convencer y entretener, es decir, el grado de compromiso del lector con cada título. De ahí que la personalización se haya convertido en un aspecto clave en la evolución de muchas empresas del sector editorial para afrontar una mayor y más fuerte competencia.

El sector del libro en España constituye un importante motor económico (mueve anualmente más de tres mil millones de euros y emplea directa o indirectamente a más de 30.000 personas) posicionado entre las grandes potencias editoriales del mundo. Las cifras de la Panorámica de la edición de libros en España 2017 apuntan una tendencia positiva. El avance de Panorámica muestra descenso del 9,7% en la producción editorial, pasando de 89.962 en 2017 a 81.228 ISBN en 2018. Esta bajada se refleja en la edición tanto de libros en papel como digitales. La producción de libros impresos disminuye un 1,1% con respecto a 2017 y representan el 74,9% del total. La producción de libros impresos disminuye un 1,1%, con un total de 60.835 libros. La edición en otros soportes ha disminuido el 28,3%. La edición electrónica, incluida en la edición en otros soportes, desciende un 28,7%, representando el 25,0%. La mayor parte de los registros electrónicos (un 93,2%) son libros digitales. Estos muestran un descenso del 31,4%, representando el 23,3% del total de la producción. La mayor parte de los libros digitales se publican en formato Pdf, que suponen el 42,5%, seguidos por ePub (33%), Mobi (580 ISBN) y otros (4.054 ISBN). La edición de carácter público experimenta un descenso del 3,6% y representa el 9,2% del total; mientras que la edición privada, con el 90,8% del total, disminuye un 10,3%. El 24,8% de la producción privada fue editada por 90 empresas editoriales, que representan el 2,9% de las que tuvieron actividad en 2018.

Según el informe del Comercio Interior del Libro elaborado por el Gremio de Editores, la facturación en 2017 fue de 2.319,36 millones de euros, apenas un 0,1% más que en 2016 y sigue sostenido por los libros de papel siendo las librerías y cadenas de librerías los principales canales de venta de libros, representando más de la mitad de la facturación total (52,8%). La facturación del libro digital representa el 5,1% del total del sector, con un incremento del 1,6% respecto al año anterior. En cuanto a los géneros literarios, crecieron todos, tanto en adulto (1,0%) como en infantil (3,2%). Además crecen los libros científico-técnicos y universitarios (un 4,3% más), los de ciencias sociales y humanidades (2,2%), los libros prácticos (3,3%) y los cómics (4,2%). Los libros de divulgación general se mantuvieron, mientras que los diccionarios y enciclopedias descendieron. No obstante, la facturación se concentra en

empresas de Barcelona (50,8%) seguida de Madrid (42,3%).

La proliferación de sellos independientes es, junto a la concentración en grandes grupos como Planeta o Penguin Random House, la tendencia de las últimas décadas. Según el Observatorio de la Lectura y el Libro cada semana nacen de media en España seis nuevas editoriales.

3. Marketing editorial: herramientas y métricas

El análisis de datos en las empresas puede aportar valor a los usuarios (descuentos, experiencias formativas, entretenimiento, soluciones, apps, etc.). Empresas como TripAdvisor, Netflix o Amazon, gracias al big data y a la analítica, ofrecen promociones, paquetes descuento y precios dinámicos basados en datos, sistemas de recomendación que ayudan a entender mejor a los clientes, personalizar la oferta y mejorar la competitividad de las empresas.

La agregación de datos se revela como otro factor a considerar: navegación móvil, geolocalización, conexión e incluso el modo de teclear pueden ofrecer información valiosa sobre los consumidores. Sin embargo, la heterogeneidad y los diferentes perfiles de usuarios (hábitos, gustos, etc.) hacen que la información no sea uniforme. En este sentido, la agregación de datos de diferentes comportamientos y perfiles puede favorecer una mayor optimización para aumentar las ventas.

El conocimiento de los lectores y sus hábitos de consumo y lectura de libros constituyen importantes datos que las editoriales pueden gestionar para mejorar la oferta de servicios y contenidos para dirigirse a sus potenciales audiencias. La clave está en conocer qué datos sería necesario recopilar. Así mismo es importante identificar qué tipo de información dispersa puede ofrecer una nueva visión junto con los datos globales. Compartir dichos datos también puede ayudar a que los diferentes eslabones de la cadena mejoren sus márgenes de actuación de manera interna a la hora de pensar en las ventas a potenciales lectores.

Joe Wikert (2015), especialista en estrategia y desarrollo de negocios, ofrece un método para acotar la recogida de información sobre los hábitos de lectura y propone una batería de cuestiones clave: ¿Cuántas personas abren el e-book que han comprado? ¿Han leído el e-book de manera lineal o saltando páginas o capítulos? ¿En qué punto se abandona la lectura? ¿Cuál es la tasa de conversión de los capítulos gratuitos a la compra del e-book? ¿Qué horas del día y qué días de la semana son los más habituales para la

lectura? ¿Cuánto tiempo se dedica a la lectura de las citas/notas al pie? ¿Se llega a consultar el índice o tabla de contenidos? ¿Qué frases se subrayan, comentan y comparten con mayor frecuencia?

A estas preguntas se han incorporado otras también importantes, que completan los objetivos anteriores: tipo de formato elegido; origen de la plataforma de compra; dispositivo de lectura; app móvil preferida para leer y horarios de compra online.

Junto a los datos a los que pueden acceder las editoriales desde sus propios sistemas de venta y difusión, en los últimos años están surgiendo otras plataformas y herramientas anexas para ampliar el conocimiento del lector y conseguir una mejor definición de lo que quieren o buscan los potenciales lectores. Algunos de los datos que se pueden registrar son el tipo de dispositivo, horario y frecuencia de lectura, tasa de abandono, lugar del libro donde se abandona la lectura, además de todas las valoraciones y comentarios que comparten los lectores sobre un libro. De cómo los editores utilicen estas herramientas para interactuar con el lector, y de cómo interpreten y aprovechen todo este big data dependerá su éxito.

3.1. Tekstum: La experiencia emocional de la lectura como base de la recomendación

Ideada por los economistas Marc Santandreu y Marc Martínez y apoyada por el Ministerio de Industria, Energía y Turismo, Tekstum es una potente plataforma de recomendación de libros basada en los big data que analiza la experiencia emocional que resulta de la lectura de un libro utilizando la inteligencia artificial y las opiniones, valoraciones y comentarios de los lectores en blogs, redes sociales y plataformas literarias. Con toda esta información, el equipo de Tekstum ha creado un diccionario de más de 20.000 palabras específicas del mundo editorial para identificar los sentimientos que provoca una obra. El análisis de estos datos permite descubrir la experiencia real de los lectores, entender por qué les gusta un libro e identificar oportunidades de venta.

Tekstum está dirigida tanto a lectores, escritores, editores y empresas relacionados con la industria editorial, así como agentes literarios, bibliotecarios u otros profesionales que trabajen en o con este sector. Permite optimizar tanto la venta de los libros como la venta de experiencias, centrándose tanto en las necesidades de los lectores como en las necesidades de los editores. El sistema proporciona un análisis científico en tiempo real de las opiniones, comentarios y reseñas que

los lectores realizan en plataformas literarias, blogs y redes sociales, que van más allá de simples listas de los más vendidos, metadatos editoriales o puntuaciones con estrellas que no contemplan los matices que pudieran existir entre ellas. Opera en las redes y aprovecha los datos utilizando la información contextual de comentarios y tweets para ofrecer información cualitativa del mercado del libro a los distintos agentes del sector, mejorar los resultados de las recomendaciones y fortalecer la proporción de éxitos de editores, autores, ejecutivos de marketing y ventas. Con este objetivo proporciona información relacionada con los libros y su impacto en el mercado que va desde un análisis de la popularidad y el sentimiento que despierta un libro (3), un perfil del autor y la medición de su reputación digital, basado en el número de seguidores en Twitter o Facebook; las consultas realizadas en Wikipedia sobre un título específico, las personas más influyentes en las redes sociales (*Top Influencers*) así como el origen, sexo y edad de los lectores o audiencias potenciales, entre otros aspectos.

El informe incluye algunos datos sobre el análisis emocional y el impacto del libro, información sobre el autor y su popularidad, su número de seguidores en Twitter o “Me gusta” en Facebook y la media de páginas vistas al mes en la Wikipedia.

El acceso mediante registro ofrece la misma información que la versión gratuita pero abierta a la selección de libros que desee el usuario, que puede solicitar el análisis de un libro introduciendo algunos datos (por ejemplo: autor, título o ISBN), de modo que el informe se genera de forma automática bajo demanda.

Puede realizarse el análisis de ocho libros al mes, sin perjuicio de que los interesados puedan acordar el análisis de un mayor número de libros. El modelo de informe incluye la siguiente información: Puntuación media del análisis emocional, reputación del autor, tasa de retención o de impacto social (Top of mind), popularidad, noticias o conceptos relacionados con el libro analizado en la red (Sensación de nube), información sobre el autor, número de seguidores en Twitter, número de “Me gusta” en Facebook, visitas a la página de Wikipedia, demografía de los lectores, género, edad, ubicación, y “Top Influencers” en Twitter, es decir, personas de referencia relacionadas con el libro en las redes sociales. También se puede conocer el número de visualizaciones en Youtube y menciones en otras plataformas de medios de comunicación (prensa, radio, televisión) así como una comparativa entre libros similares o relacionados para contrastar datos.

Las cifras que ha acumulado esta plataforma ponen de relieve el impacto que ha tenido en el sector (134.021.533 de libros publicados hasta 2016; 50.000.000 reseñas en plataformas literarias, 6.000 tweets/segundo y 4.000.000 millones de likes/minuto en Facebook en 2016) con clientes recurrentes que generan (Grupo Planeta, Strands, Sinerrata y The Spanish Bookstage que generan tracción (4).

El Big Data emocional de Internet nos descubre que libros están triunfando. Se miden las emociones positivas, negativas o neutras que está despertando cualquier publicación. Y su evolución a lo largo de un año. A partir de aquí se crea una nube de palabras que se convierte en el ADN emocional del libro. Una vez estos datos han sido tratados, los editores los pueden usar para decidir los próximos lanzamientos. Además, ha diseñado un nuevo método de recomendación de libros. Actualmente el lector puntúa y se guía a través de los libros más vendidos, pero el que propone Tekstum es un recomendador 100% emocional. Marc Santandreu, cofundador de Tekstum comenta (Trinidad, 2015):

¿Qué es la Literatura? subjetividad, emociones, sentimientos y eso las ventas no lo dicen. Por eso nosotros buscamos esta información subjetiva y la convertimos en datos objetivos para que los editores puedan tomar decisiones, como si reeditan un libro o no, o cómo enfocar la campaña de marketing.

Según un estudio de la compañía norteamericana Deloitte, las consultas hechas desde dispositivos móviles influyen un 64% en las ventas de libros en los establecimientos de los Estados Unidos. El objetivo de Tekstum es que su aplicación se utilice para consultar los sentimientos que despierta un libro justo antes de que los lectores realicen una compra (5).

3.2. Jellybooks y el compromiso de los lectores

Jellybooks (<https://www.jellybooks.com>), empresa londinense pionera en analizar el modo en que leemos y la interacción del lector con el libro, ha desarrollado proyectos para la recuperación de información editorial, por ejemplo: si se ha leído el libro, los dispositivos elegidos para leer, horarios, frecuencia de lectura, tiempo en terminarlo, en qué punto se ha abandonado la lectura o si el lector lo recomienda, entre otros aspectos. El objetivo es entender mejor la conexión entre el lector y el libro, y detectar de un modo más directo qué títulos son los más adecuados para cada tipo de lector.

Para la recopilación de estos datos de lectura Jellybooks ha desarrollado un software inteligente (candy.js) incrustado en libros electrónicos, especialmente en formato ePub. El procedimiento

para la recogida de datos es través de campañas en las que se invita a participar a una muestra de lectores (entre 200 y 500 personas) a los que se informa que van recibir un libro electrónico con un software de seguimiento integrado para registrar sus datos de lectura. Hay dos tipos de campañas: dirigidas al análisis de un solo título, o campañas multitítulo en la que se ofrece a los lectores una selección de la que pueden elegir dos títulos.

En la mayoría de los casos, se trata de libros que aún no han sido publicados. En otros se someten a prueba libros en tapa dura para edición de bolsillo; o bien, cuando los editores quieren decidir sobre la edición de una secuela de un libro o similar. Conviene precisar que todos los libros distribuidos por Jellybooks (copias pre-pub, fragmentos o ejemplares gratuitos) están sujetos a derechos de autor y no está permitido copiar, imprimir o compartir. Cada documento tiene un identificador único y el software de seguimiento añade marcas de agua digitales que permiten seguir el rastro del libro y evitar acciones de piratería y descarga ilegal.

La lectura puede realizarse con aplicaciones compatibles (iBooks, Ebook Reader, Adobe Digital Editions, Bookshelf, Azardi y Cloudshelf Reader) (6), y la duración de una campaña oscila entre cuatro y seis semanas, periodo que se registra si el lector "abre" el libro, si lo termina, si lo lee en una sola sesión o en varias, el horario y frecuencia de lectura, el dispositivo utilizado y la forma en que avanza o se desplaza por los contenidos.

La tasa de libros abandonados -y el punto en que se abandona su lectura- es un dato importante para la editorial, ya que puede saber qué es lo que hace que la historia, llegada a un punto, deje de funcionar. También los hábitos de lectura constituyen un indicador.

El compromiso con un libro es uno de los requisitos previos que motiva a los lectores a recomendar un título. De acuerdo con una Encuesta de Nielsen Global Trust In Advertising, el 92% de los usuarios tiene en cuenta las recomendaciones de amigos o familiares, lo que las convierte en fuentes de confianza con una gran influencia para los consumidores (Nielsen, 2016). También recurren a expertos y líderes para obtener orientación para facilitar el proceso de toma de decisiones. Según el estudio Anual de Redes Sociales (IAB España, 2018), el 72% de los usuarios de redes sociales entre 16 y 55 años en España siguen perfiles de personas que marcan tendencias y lideran opiniones, es decir, los denominados *influencers*.

Las redes sociales se han convertido en el nuevo boca a boca, el lugar donde los usuarios comparan sus opiniones y experiencias con un producto

determinado, sin embargo The Keller Fay Group -agencia especializada en interacción, y social media para grandes marcas- ha demostrado que solo el 10% del boca a boca se produce online. Para lo demás se siguen utilizando los contactos directos, donde herramientas privadas como WhatsApp han capitalizado la conversación social relevante (Fernández de la Puente-Campano, 2016). En consecuencia, la medición de la tasa de terminación y el compromiso del lector se convierten en posibles indicadores para la previsión de qué título podría construir un fuerte boca a boca Rhomberg (2015). Lo interesante de esta herramienta de monitorización es que, a diferencia de Google Analytics, puede trabajar offline, de este modo los datos son capturados y almacenados localmente hasta que el usuario se conecta a Internet y envía los datos recopilados con un solo click.

La medición y analítica web es fundamental para Jellybooks, ya que proporciona un análisis basado en datos para medir el compromiso del lector con la mayor rapidez y, por tanto, la comercialización y la campaña de publicidad pueden ser ejecutadas con el máximo efecto. En este sentido es ante todo una herramienta de marketing para autores y editores.

La analítica dirigida a los autores mide cómo los lectores se involucran con los libros (compromiso), su satisfacción y su probabilidad de recomendarlo a los demás. Pone a prueba un libro antes y después de su publicación, a través de grupos de discusión virtuales, copias de lectura avanzada o libros gratuitos. También mediante el seguimiento de lectura de fragmentos, para detectar si el lector se engancha y compra una copia completa.

En el caso de los editores la analítica proporciona conocimientos aplicables a las necesidades de la audiencia para tratar de evitar errores: cómo gestionar los escasos recursos para publicidad y promoción; cómo optimizar el título, la descripción y presentación del libro; y dónde posicionarlo en términos de audiencia para lograr un fuerte boca a boca. Editores expertos como Random House, Elsevier, MacMillan y Bonnier ya están haciendo uso de estas herramientas de medición para aumentar sus ventas y la cuota de mercado de sus títulos. En la mayoría de los casos estos editores utilizan la analítica antes de que el libro sea publicado para decidir sobre los gastos de marketing y dónde enfocar los recursos de promoción y publicidad.

Jellybooks es la primera y hasta ahora única empresa que permite a los editores recoger los datos de aplicaciones de terceros para entender el comportamiento de los lectores, a diferencia de Amazon, Apple y Google, que no comparten los datos

que recogen. Trabaja con tres indicadores clave de rendimiento (KPI) aplicados al sector editorial: Tasa de finalización de la lectura, Factor de recomendación y Factor de ajuste de la cubierta.

(a) *Tasa de finalización de la lectura*: Porcentaje de lectores que inician y finalizan un libro. Es uno de los KPI básicos para medir la satisfacción del lector con el contenido y se calcula dividiendo el número de lectores que terminaron el libro entre el número total de lectores que envió la lectura del libro en cualquier punto o capítulo del libro. Generalmente, si los lectores no terminan una novela, es que su grado de satisfacción no era muy grande. Las tasas de finalización no parece que sean dependientes del género. Un resultado por debajo del 25% indica un compromiso débil; entre 25-50% es un buen resultado y se logra en la mayoría de los libros; y un resultado superior al 50% es muy bueno, especialmente para los de no ficción, e indica que un libro destaca por encima del resto. Una tasa por encima del 75% es excepcional, cifra que sólo han alcanzado un 5% de los libros analizados por Jellybooks, siendo algunos de ellos los mayores éxitos de ventas de los últimos años.

(b) *Factor recomendación*: Este indicador es una adaptación del concepto *Net Promoter Score* (NPS) y consiste en preguntar a los lectores si recomendarían el libro a sus amigos. Las cuatro respuestas posibles son: "Sin duda lo recomendaría a un amigo" (estos serían los promotores); "Posiblemente lo recomendaría a un amigo" (neutrales); "Probablemente no se lo recomendaría a un amigo" (detractores débiles) y "Sin duda no lo recomendaría a un amigo" (detractores fuertes). Para calcular este KPI sólo se considera la respuesta de los lectores que han terminado el libro, y se obtiene restando el porcentaje de detractores débiles y fuertes, siendo el resultado el

número de promotores. Un libro con factor de recomendación positiva tiene más promotores que detractores. Un porcentaje entre el 30% y el 60% sigue siendo un resultado muy bueno, mientras que un título con más del 60% de promotores tiene el potencial de un hit viral. Como venimos señalando, el factor de recomendación es un indicador clave del potencial del boca a boca de un libro. Libros con altas tasas de finalización y alto factor de recomendación son los principales candidatos para una fuerte campaña de marketing y relaciones públicas.

(c) *Factor de ajuste de la cubierta*: La cubierta es una de las razones más comunes por las que un contenido tiene que vender. Al igual que ocurre con el diseño de los carteles de cine -en los que convergen los intereses dispares del director, el productor y el distribuidor- en el diseño de la cubierta de un libro también se produce esta tensión entre las propuestas del autor y el editor. Para el autor y el lector, la portada evoca el contenido y lo sintetiza. En las condiciones de saturación que definen el actual mercado editorial, el principal mérito de la cubierta es la diferenciación, tanto en la mesa de novedades como en la plataforma de venta online. Este indicador permite conocer si, en opinión de los lectores, el contenido del libro coincide con la expectativa planteada en la cubierta.

Todos estos datos son procesados, agregados y analizados y se presentan en gráficos interactivos con diferentes códigos de color para identificar si los resultados son pobres (azul), buenos (amarillo), muy buenos (verde) o excelentes (púrpura). Este método de visualización facilita la interpretación a cualquier usuario, ya sea un editor, un experto en marketing o un gestor de medios sociales. Se trata de datos inteligentes no de grandes datos, que alcanzan su potencial máximo cuando se comparten con el equipo y el autor.

| Tipo de datos | Fuente | Métrica Bookmetrix |
|---------------|-----------|---|
| Citas | Springer | Número de citas/año Listado de trabajos que citan el libro/capítulo Actualización semanal |
| Menciones | Altmetric | Número total de menciones online/propias del libro/capítulo Actualización en tiempo real |
| Lectores | Mendeley | Datos demográficos y ubicaciones geográficas Actualización semanal |
| Reseñas | Springer | Número de reseñas del libro seleccionado Actualización semanal |
| Descargas | Springer | Número total de descargas de capítulos Actualización mensual |

Tabla III. Datos, fuentes y métricas de la plataforma Bookmetrix

3.3. Bookmetrics y el impacto de las publicaciones científicas

Bookmetrix (www.bookmetrix.com) es un proyecto de la editorial Springer Science + Business Media Altmetric que reúne un conjunto de métricas de rendimiento para buscar, conocer y comparar el impacto de las publicaciones científicas.

Ofrece una visión global del alcance, uso y número de lectores de libros o capítulos sueltos proporcionando métricas a diferentes niveles: citas, menciones en línea, estadísticas de lectores, reseñas y descargas. Bookmetrix recupera los metadatos de los libros y capítulos de Springer y CrossRef y agrega las métricas asociadas a dichos libros y capítulos a partir de diversas fuentes (Tabla III).

Los informes son públicos y fácilmente accesibles e incluyen la información del libro, un listado de todos los capítulos y la visualización de las métricas asociadas al libro y/o capítulos.

4. Neuromarketing y tendencias editoriales

Junto a los grandes volúmenes de datos cargados de racionalidad, lo cierto es que en la mayoría de las decisiones de compra los argumentos basados en la lógica suelen quedar en segundo plano ante los mecanismos emocionales. En los últimos tiempos se están generalizando aproximaciones a las técnicas de persuasión como el neuromarketing, que trata de explicar que el comportamiento del consumidor no responde a normas ni pautas razonables, sino que adopta expresiones muy complejas basadas en la emocionalidad.

Una de las grandes aliadas en este contexto ha sido la neurociencia, puesto que un mejor conocimiento del funcionamiento del cerebro humano puede ayudar a entender mejor las verdaderas necesidades y gustos del consumidor, y a comunicarnos mejor con él. De ahí su desarrollo y aplicación en el marketing (neuromarketing) y la comunicación (neurocomunicación).

El Laboratorio de Inteligencia Artificial del Instituto de Tecnología de Massachusetts (MIT), elaboró EQ-Radio (<http://eqradiocsaail.mit.edu>), dispositivo que puede revolucionar estas prácticas. Se trata de una herramienta que puede utilizar las señales *wireless* para detectar nuestro estado emocional con la misma precisión que un cardiograma, midiendo la variabilidad del pulso cardíaco y la respiración. Este dispositivo determina con un 87% de probabilidad si una persona está feliz, triste, enfadada, etc. Los datos son ejecutados mediante una serie de algoritmos y procesados por una máquina de auto-aprendizaje que

busca parámetros que coincidan con el comportamiento de una persona que siente una emoción específica.

Con ella se pretende humanizar la interacción de los usuarios con la tecnología con la mayor precisión posible. Se espera que facilite la labor a todo tipo de sectores -también al editorial- a la hora de determinar qué quieren los consumidores, no sólo por sus gustos e intereses sino también, en función de su estado de ánimo. Por el momento EQ-Radio necesita la participación voluntaria de los usuarios, pero su desarrollo es una evidencia más de las posibilidades que ofrece la tecnología para profundizar en el conocimiento de las personas, clientes potenciales de cualquier servicio o producto.

En consecuencia, una webcam podría bastar para que mientras leemos un libro, asistimos a un concierto online o miramos una obra de arte, la tecnología sepa mediante nuestros gestos si nos gusta o no, y qué sensación nos produce. Esta tecnología se utiliza a través de cámaras de vigilancia para saber la satisfacción de sus clientes, aplicable a las librerías para conocer en tiempo real las emociones predominantes de los visitantes: reacciones a cubiertas de libros, precios, experiencias, servicios, sonido e iluminación para favorecer el bienestar, mejorar la experiencia de compra y aumentar las ventas. Todo para lograr cierta estimulación y también para la anticipación.

Art19 (art19.com), es una compañía de podcasts que proporciona servicios a editores y anunciantes para dirigirse a las audiencias creando mensajes de audio personalizados y ayudando a los editores a rentabilizar estas grabaciones gracias al análisis del comportamiento del usuario. Otra experiencia piloto es a2o (www.a2o.santillana.com), que aprovecha la potencia de los datos para ofrecer experiencias de aprendizaje personalizadas en el ámbito docente.

En un mercado competitivo donde la personalización es la clave - y no es posible personalizar sin datos personales- el empleo de las tecnologías (redes sociales, aplicaciones móviles, e-commerce, etc.) y la interacción con los clientes resulta indispensable para la estrategia empresarial (7). Sin embargo, el mismo usuario que demanda productos y servicios personalizados, con frecuencia se escandaliza si las empresas le trackean (8) para ofrecerle la personalización que él mismo solicita, y lo que resulta aún más grave, reconoce que no lee las políticas de privacidad (Salvador Benítez, 2019).

En España, la nueva Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos y Garantía

de los Derechos Digitales (LOPD GDD) -conocida como LOPD 2018- da cobertura al Reglamento General de Protección de Datos (RGPD), de obligado cumplimiento en todos los países de la Unión Europea. En el caso de España -donde la protección de datos es un derecho fundamental protegido por el artículo 18.4 de la Constitución- el RGPD incorpora novedades, tanto en el régimen de consentimiento y tratamiento de los datos como en la introducción de nuevas figuras y procedimientos (Salvador Benítez, 2019).

5. Conclusiones

De la investigación y análisis sobre el tema propuesto, se han obtenido las conclusiones que se indican:

Los avances tecnológicos y el abaratamiento del almacenamiento junto con el desarrollo de las técnicas estadísticas están permitiendo extraer conocimiento derivado del análisis de datos masivos procedentes de distintas fuentes, con un coste mucho menor y con mayor riqueza y exactitud para la toma de decisiones. En su mayoría se trata de datos personales, asociados a las operaciones que los consumidores llevan a cabo en cualquier sector, incluido el editorial. Nos encontramos inmersos en la economía del dato y términos como big data o inteligencia artificial se han vuelto cotidianos en los planes estratégicos de cualquier organización, independientemente de la industria de la que se trate.

El big data aplicado al sector editorial es una "fuente" para generar información que, analizada y gestionada, permite conocer la penetración en los diferentes mercados, la fijación de precios, las posibilidades de éxito de autores noveles, etc. Gracias a los datos masivos y la analítica, las empresas ofrecen precios dinámicos basados en datos, y sistemas de recomendación que cada vez conocen más al consumidor y pueden personalizar la oferta.

En edición, el control de recursos hace necesaria una inversión en análisis de datos. Teniendo en cuenta que una de las principales características de los big data es la velocidad, la posibilidad de rastrear el comportamiento e información de los potenciales lectores de forma rápida y eficiente permite reducir los riesgos. Más allá de disponer de una u otra solución tecnológica, para llevar a cabo una buena estrategia de analítica, es importante que las organizaciones se rodeen de un capital humano apto para esta tarea. Entre los perfiles clave destacan el Científico de datos (*Data Scientist*) en la parte más especializada, y el CDO (*Chief Data Officer*) o Jefe de Datos, en la parte más estratégica.

A nadie escapa que la retención masiva de datos de carácter socioeconómico, demográfico, financiero y de comportamiento con propósitos analíticos puede tener graves consecuencias para la privacidad y libertad individual de los consumidores. Desde una perspectiva de seguridad y privacidad, la normativa reguladora, la tecnología que sustenta las plataformas y las empresas, se enfrentan al reto de garantizar que los usuarios tengan el suficiente control sobre sus datos para prevenir un uso indiscriminado a la vez que se mantiene su utilidad para extraer conocimiento, patrones y, en suma, valor como activo empresarial.

La cadena lineal de valor del libro en la que intervienen autor, agente, editor, impresor, distribuidor, librero, crítico y lector se ha sustituido por una red de valor, con relaciones más fluidas entre los eslabones. Las nuevas formas relacionales en el sector parecen tender a generar redes en relación al contenido más que a la posición en la cadena de valor, que se sustituye por una red de valor. En un entorno en el que todos podemos publicar contenidos autoeditados (denominados *non traditional publishing*), las editoriales dejan de tener el control sobre la producción y se convierten fundamentalmente en mediadores culturales multimedia que explotarán los contenidos en todos los formatos a su alcance. En consecuencia, las herramientas como las que aquí se analizan pasan a ser decisivas en la toma de decisiones.

La materialidad del libro físico ha dado paso al contacto personal y la interacción e intercambio de opinión con otros lectores. El poder de la prescripción es cada vez mayor y la recomendación es un extraordinario motor de venta de libros, quizá el principal. Para que podamos recomendar una lectura, el contenido debe provocar emociones que nos empujen a hablar del libro con otros y lo conviertan en un objeto social. Nos dirigimos hacia una web cada vez más personal, que se adapta a lo que busca el usuario, que aprende de las páginas que visita, que sirven para alimentar la inteligencia colectiva. Las preferencias que el usuario haya reconocido como propias son la base para recomendar qué libros tiene que leer o qué lugares tiene que visitar alguien que está conectado con él.

Notas

- (1) El *business intelligence* no se debe confundir con el Big Data, ya que podría no serlo, dependiendo del volumen y de otros criterios.
- (2) El ROI es uno de los principales medidores por los que se rigen las actividades de una empresa. Aunque en ocasiones es difícil de medir, trata de relacionar la inversión

realizada para poner en marcha una actividad determinada y los resultados económicos que ha generado.

- (3) El sistema utiliza un algoritmo propio basado en tecnología big data e inteligencia artificial que descubre las emociones y sentimientos que un libro transmite a sus lectores. Para ello se enseña una máquina a identificar e interpretar diferentes sentimientos dentro de un texto utilizando técnicas de Procesamiento del Lenguaje Natural (PLN). Este método se denomina "Análisis de Sentimiento" y permite identificar y clasificar automáticamente la opinión de un texto como positivo, negativo o neutro (es decir, que no contiene ningún sentimiento). Es capaz de clasificar automáticamente el sentimiento porque busca patrones ocultos en grandes volúmenes de datos cualitativos y luego usa esta información para aprender características comunes en el uso del lenguaje que indican un sentimiento específico. Por lo tanto, cada patrón es indicativo de un sentimiento particular: consisten en patrones gramaticales, tendencias lexicales, tonalidad, uso de vocabulario particular, Coloquialismos, frases, etc. Al aprender de estos patrones textuales, se puede predecir la opinión de un texto desconocido debido a sus similitudes con otros contextos indicativos de un sentimiento específico.
- (4) Si hay un factor importante para una *startup* es la tracción. Se trata de una métrica del progreso que demuestra la validación real de un modelo de negocio, cómo está siendo aceptado por los clientes y si hay una demanda real en el mercado. En definitiva, la tracción es un elemento clave para tomar nuevas decisiones sobre el futuro de la compañía y para buscar inversores que hagan crecer el proyecto. La base es tener un buen producto que cubra problemas reales de los clientes.
- (5) Las aplicaciones no compatibles con Jellybooks son: Kindle, Nook, Kobo, Overdrive Mediaconsole, Tolino, Aldiko, Google Play Books y Moon+Reader.
- (6) En la actualidad la plataforma Tekstum no se encuentra activa por falta de financiación. El análisis y testado se realizó durante el periodo que permaneció operativa.
- (7) La gestión de la información de los usuarios constituye quizá el aspecto más desconocido, controvertido e incierto de las redes y plataformas sociales. La investigación sobre la materia ha generado varios trabajos entre los que señalamos el estudio realizado por Inteco y la Agencia Española de Protección de Datos (2009); Salvador Benítez y Gutiérrez David (2010); Díaz Buck (2013); Roura y Ferré (2014), Salvador, Olivera y Sánchez (2014), Salvador Benítez (2019).
- (8) El término se refiere básicamente a seguir puntualmente los "movimientos" de los usuarios. Las analíticas de Google agregan este elemento denominado Heatmaps.

Referencias

- Álvarez Ruiz, Antón (2012). *La magia del planner. Cómo la planificación estratégica puede potenciar la comunicación persuasiva*. Madrid: ESIC.
- Arqués Salvador, Neus (2009). *Marketing para escritores. Cómo publicar, promocionar y vender tu libro*. Barcelona: Alba editorial.
- Barranco-Fragoso, Ricardo (2012). *¿Qué es big data?* IBM. <http://www.ibm.com/developerworks/ssa/local/im/que-es-big-data> (2020-01-20).
- Cole, David (2003). *Marketing editorial. La guía*. México: Fondo Nacional para la Cultura y las Artes.
- Díaz Buck, Anid Vanessa (2013). *La autorregulación en redes sociales como forma de garantizar los derechos de intimidad, privacidad y protección de datos personales*. // Derecom. ISSN: 1988-2629. 13 (marzo-mayo 2013), 125-143.
- Escribano Hernández, Asunción (2012). *La retórica publicitaria editorial. El arte de vender un libro*. Madrid: Arco Libros
- Fernández de la Puente-Campano, Jaime (2016). *Plan de marketing digital. Claves para triunfar*. Madrid: The Valley Digital Business School.
- Gartner, Inc. (2019). *Top 10 Strategic Technology Trends for 2020. A Gartner Special Report*. <https://www.gartner.com/en/doc/432920-top-10-strategic-technology-trends-for-2020>. (2020-01-20)
- Gómez-Tarragona, Daniel (2010). *Marketing editorial. Cómo satisfacer las necesidades de los lectores de libros*. Madrid: Ediciones Pirámide
- González-Fernández Villacencio, Nieves (2016). *Métricas de la web social para bibliotecas*. Barcelona: UOC.
- González, Manuel (2017). *Tekstum, el mejor analista de libros*. <https://www.elreferente.es/innovadores/tekstum-10innovadores-30616> (2020-01-20).
- Instituto Nacional de Tecnologías de la Comunicación; Agencia Española de Protección de Datos. (2009). *Estudio sobre la privacidad de los datos personales y la seguridad de la información en las redes sociales online*. https://www.csirtcv.gva.es/sites/all/files/downloads/estudio_intecoaeprd_privacidad_redes_sociales.pdf (2020-05-29).
- España (2018). *Ley Orgánica 3/2018, de 5 de diciembre, de Protección de datos personales y garantía de los derechos digitales* // BOE. 294 (6-12-2018), 119788-119857. <https://www.boe.es/eli/es/lo/2018/12/05/3> (2020-05-29)
- Megapactical (2016). *Cómo crear valor en el sector editorial a través de los Big Data*. <https://www.megapactical.com/noticias/como-crear-valor-en-el-sector-editorial-a-traves-de-los-big-data> (2020-01-20).
- Nielsen Global Trust In Advertising (2016). *Winning Strategies for an evolving Media Landscape*. <https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/global-trust-in-advertising-report-sept-2015-1.pdf> (2020-01-20)
- Panorámica de la edición de libros en España 2017 (2018). Madrid: Secretaría General Técnica. Centro de Publicaciones. Ministerio de Educación, Cultura y Deporte.
- Rhomberg, Andrew (2015). *Jellybooks: Tracking Reader Engagement for Better Marketing*. // *Publishing perspectives*. (August 26, 2015). <https://publishingperspectives.com/2015/08/jellybooks-tracking-reader-engagement-for-better-marketing/>
- Roura, Elisabeth; Ferré Pavia, Carme. (2014). *Aspectos legales de las redes sociales: protección de datos y uso consciente de la tecnología*. // Ferré Pavia, C. (Ed.) (2014). *El uso de las redes sociales: ciudadanía, política y comunicación. La investigación en España y Brasil*. Universitat Autònoma de Barcelona: Institut de la Comunicació, 92-115.
- Salvador Benítez, Antonia (2019). *Entre la privacidad y el negocio. Condiciones de uso y política de datos en redes sociales* // Derecom. ISSN: 1988-2629. 27 (Septiembre-Marzo 2019), 139-161.
- Salvador Benítez, Antonia; Gutiérrez David, Estrella. (2010). *Redes sociales y medios de comunicación. Desafíos legales*. // *El Profesional de la Información*, ISSN: 1699-2407. 19:6 (Noviembre-Diciembre 2010) 667-674.
- Salvador Benítez, Antonia; Olivera Zaldúa, Maria; Sánchez Vigil, Juan Miguel (2014). *Redes sociales y medios de comunicación. Políticas de uso de contenidos audiovisuales*. // Sierra Sánchez, Javier; García García, Francisco (2014). *Tecnología y narrativa audiovisual*. Madrid: Fra-gua, 717-735.

Serrano Cobos, Jorge (2014). Big data y analítica web. Estudiar las corrientes y pescar en un océano de datos. // El profesional de la información. ISSN: 1699-2407. 23:6 (Noviembre-Diciembre 2014), 561-565.

Trinidad, Meritxell. (2015). Un sistema para medir la emoción que provoca un libro. *elmundo.es* (04-06-2015). <https://www.elmundo.es/economia/2015/06/02/556c9bab22601d913e8b4591.html> (2020-01-20).

Valverde, Jaime (2016). *Redes sociales. Cómo diseñar una estrategia de éxito*. Madrid: The Valley Digital Business School.

Vivancos, David (2016). *Big Data: Hacia la inteligencia artificial*. Madrid: The Valley Digital Business School.

Wikert, Joe (2015). [Blog]. Here's how reader analytics can help publishers. https://jwikert.typepad.com/the_average_joe/2015/06/heres-how-reader-analytics-can-help-publishers.html (2020-01-20).

Zhao, Mingmin.; Adib, Fadel.; Katabi, Dina. (2016). Emotion recognition using wireless signals. // Chen, Yingying; Gruteser, Marco (2016). *MobiCom'16. The 22nd Annual International Conference on Mobile Computing and Networking*. New York: Association for Computing Machinery, 95–108. <http://equadro.csail.mit.edu/files/equadro-paper.pdf> (2020-01-20).

Enviado: 2020-03-29. Segunda versión: 2020-06-04.
Aceptado: 2020-06-05.

Instrumentos, processos e produtos do Tratamento Temático da Informação: um estudo nos anais da ISKO-Brasil (2012-2019)

Herramientas, procesos y productos de representación temática: un estudio de las actas de los congresos de ISKO-Brasil (2012-2019)

Tools, processes, and products of subject representation: a study of the ISKO-Brazil conference proceedings (2012-2019)

Lais Pereira de OLIVEIRA (1), Maria Cláudia Cabrini GRÁCIO (1), Daniel MARTÍNEZ-ÁVILA (2)

(1) Universidade Estadual Paulista Júlio de Mesquita Filho, Avenida Hygino Muzzi Filho, 737, Marília, São Paulo, lais-pereira.oliveira@unesp.br, cabrini.gracio@unesp.br (2) Departamento de Biblioteconomía y Documentación, Universidad Carlos III de Madrid (UC3M), España, dmartine@bib.uc3m.es

Resumen

Se analizan los trabajos sobre los instrumentos, procesos y productos del tratamiento temático de la información recogidos en las actas de los congresos de ISKO-Brasil, el congreso brasileño de organización y representación del conocimiento. El estudio es de carácter descriptivo, y combina metodologías cualitativa (análisis de contenido), cuantitativa (bibliométrica) y bibliográfica. Los resultados que los trabajos sobre tesauros y políticas de indización predominan en el apartado de instrumentos; la indización y clasificación, en procesos; y la notación, en productos. Destacan sobre todo los trabajos sobre construcción y tipología de tesauros en el apartado de instrumentos; y los de tratamiento documental de archivos e imágenes en procesos.

Palabras clave: Organización de la información. Tratamiento temático de la información. Descripción temática. Producción científica. Brasil.

Abstract

The approaches to the tools, processes, and products of subject representation in the ISKO-Brazil conference proceedings are analyzed. It is a descriptive, qualitative, and bibliographic research that works with the ISKO Brazil conference proceedings as the object of study and employs bibliometric and content analysis methods to obtain results on the scientific production. The results reveal the predominance of thesauri and indexing policies, in relation to tools; indexing and classification, in relation to processes; and notation, in relation to products; and more specifically, construction and typology of thesauri, in relation to tools; and document processing in archives and images, in relation to processes.

Keywords: Information organization. Subject representation. Thematic description. Scientific production. Brazil.

1. Introdução

O tratamento temático da informação (TTI) pode ser considerado um macroprocesso que apresenta os atributos de assunto dos documentos e que tem um papel mediador fundamental junto à busca e recuperação. Objetiva, portanto, representar e recuperar documentos a partir do conteúdo que trazem consigo (Medeiros *et al.*, 2016).

Dessa forma, exige a extração de conceitos que são representativos do documento (Braz e Carvalho, 2017), com vistas a viabilizar sua recuperação pelo usuário da informação. Ação esta que torna o TTI extremamente complexo ou, nas palavras de Sousa (2013), uma área desafiadora.

Além disso, descerram-se sobre o trato temático, diversas designações. Entre estas, análise documental, análise temática, análise de assunto, descrição de conteúdo e tratamento temático da informação (Café e Sales, 2010).

O TTI acaba por assumir, ainda, uma constituição teórica subjetiva, percebida em sua incumbência e na própria literatura do campo (Oliveira e Martínez-Ávila, 2019; Medeiros *et al.*, 2016; Dias e Naves, 2013). Correntes distintas também se estabelecem em seu entorno, sendo estas: a catalogação de assunto, a indexação e a análise documental (Guimarães, 2008; Guimarães, 2009; Guimarães *et al.*, 2011).

Ademais, o TTI tem recebido um tratamento distinto na literatura, em detrimento do tratamento descritivo. Aspecto esse que fez com que, até mesmo a catalogação de assunto, tenha sido menos estudada, comparada com a catalogação descritiva (Martinho e Fujita, 2010). Nesse cenário, considera-se importante a realização de investigações que tenham como objeto de estudo o Tratamento Temático da Informação, no intuito de contribuir para a compreensão das suas particularidades (Oliveira e Martínez-Ávila, 2019).

Além do que essa é uma forma de se buscar sua consolidação em termos teóricos e aplicados.

Nesse contexto, esta investigação objetiva analisar as abordagens acerca dos instrumentos, processos e produtos do tratamento temático da informação nos anais da ISKO-Brasil. Considerando seu entorno revestido de complexidade, subjetividade, diversidade de designações, correntes distintas e diferenças de tratamento na literatura em detrimento do trato descritivo, busca, assim, contribuir para o debate relativo à problemática teorizante acerca do TTI.

Deve-se considerar, nessa delimitação, a necessidade de se visualizar como o TTI tem construído seu referencial teórico-metodológico no contexto da própria International Society for Knowledge Organization (ISKO) (Guimarães, 2008) e também do capítulo brasileiro, ISKO-Brasil.

A primeira edição da ISKO-Brasil, como evento independente, desassociado do GT-2 do ENANCIB, foi realizada no ano de 2011, tendo os anais sido divulgados em 2012. Por congregar, desde 2011, investigadores brasileiros da área de organização do conhecimento, considera-se que a ISKO-Brasil constitui um fórum significativo para a caracterização das investigações sobre as instâncias particulares do TTI na comunidade brasileira. Nessa medida, parte-se do pressuposto que tal fórum descerra também, questões elementares acerca dos instrumentos, processos e produtos do TTI que permitem, em alguma medida, configurar as peculiaridades desse entorno.

O estudo pode contribuir, então, com o desenvolver teórico em tratamento temático da informação, especialmente sobre seus instrumentos, processos e produtos base. Também pode agregar em âmbito prático, seja ressaltando aspectos centrais na realização das operações do TTI, seja esclarecendo acerca de seus componentes característicos.

2. Tratamento Temático da Informação

O Tratamento Temático da Informação, base substancial da organização da informação (OI) (Café e Sales, 2010) e parte das operações do trato informacional (Fujita, 2013), é uma atividade intelectual, de caráter subjetivo (Santos *et al.*, 2019), que advém do desenvolvimento da organização do conhecimento (OC) (Felipe *et al.*, 2019).

Para concretização do tratamento temático, o bibliotecário recorre à análise de assunto, voltando-se, basicamente, à compreensão e à extração do mesmo (Vieira *et al.*, 2017). Dessa forma, analisa, descreve e representa conteúdo

a fim de armazenar e recuperar a informação (Guimarães *et al.*, 2011).

No âmbito da organização do conhecimento, o tratamento temático encontra terreno fértil à sua concretização. Afinal, marcos históricos no desenvolvimento da OC, ligados às fases classificatória, tesáurica e classificatória revisitada, refletem questões específicas do TTI (Guimarães *et al.*, 2011). Além do que a organização do conhecimento integra o universo epistemológico do tratamento temático (Guimarães, 2008).

No âmbito da organização da informação, o tratamento temático corresponde à dimensão dedicada ao conteúdo. A ela se opõe o tratamento descritivo, voltado ao aspecto físico dos itens. Na prática, porém, a distinção entre ambas as vertentes da OI não acontece (Bräscher e Guimarães, 2018), sendo comum sua condução conjunta sem que se perceba a individualidade de cada uma.

Entretanto, tratamento descritivo e temático incorrem em diferentes atividades documentárias. Além do que se circundam de distintas ferramentas na geração de objetos caracterizadores da informação, peculiares de cada dimensão representada.

Sendo assim, pode-se distinguir o TTI como o conjunto de instrumentos, processos e produtos de trato do conteúdo (Oliveira e Martínez-Ávila, 2019). Reúne, pois, operações que contemplam o tratamento do conteúdo dos itens informacionais (Santos *et al.*, 2019), assim como atributos caracterizadores e aparatos de apoio à consecução destes, com foco na acessibilidade temática (Dal'Evedove e Fujita, 2013). Dito de outra forma, o tratamento temático é (Oliveira e Martínez-Ávila, 2019, p. 88):

[...] cotidianamente colocado em prática em ações biblioteconômicas de representação do assunto contido nos documentos, com emprego de instrumentos específicos e a consequente geração de produtos característicos de cada processo técnico desse universo.

De modo que, no seu desenvolvimento habitual, aborda-se o assunto dos documentos (Sousa e Fujita, 2013), com a constante preocupação de possibilitar seu acesso (Sousa, 2013). Para tanto, o TTI (Oliveira e Martínez-Ávila, 2019, p. 90):

[...] desenvolverá os processos de classificação, indexação e elaboração de resumos. Em decorrência destes gerará notações, descritores e resumos, que viabilizarão cada qual a seu jeito, a representação temática da informação contida no documento.

Ademais, com vistas a gerar os produtos documentários do tratamento temático, na forma de

notações, termos e resumos documentários, empregam-se diversos instrumentos. Dentre os quais, sistemas de classificação, tesouros, política de indexação e norma de resumo. Cada um desses serve aos propósitos de um dos processos técnicos do TTI, respectivamente, a classificação, a indexação e a elaboração de resumos, conforme ilustrado na Figura 1.

Tal constituição, congregando no entorno temático de trato informacional um rol de processos, produtos e instrumentos, reflete uma concepção diligente, estruturada e procedimental da ação organizativa por assunto. Representa, portanto, uma perspectiva do TTI, concernente a suas operações, ferramentas e atributos resultantes, visto aqui em seu enlace pragmático com a organização da informação e do conhecimento.

Em suma, pode-se afirmar que processos, produtos e instrumentos, com teor particular, se desenvolvem no âmbito do TTI, que historicamente resultou da necessidade pragmática de trato documental (Guimarães, 2008). Contudo, para além do aspecto prático, este também pode ser considerado uma área de estudos (Guimarães *et al.*, 2011), posto que precisa se ater a um referencial específico para própria concretização da sua dimensão aplicada. Desse modo, essa “ação prática fundamental à busca e recuperação por assunto” (Oliveira e Martínez-Ávila, 2019, p. 90) que é o TTI, viabiliza-se a partir de instrumentos, processos e produtos característicos do universo temático. Mas, ao mesmo tempo, decorre de incursões teóricas, como a que hora se estabelece, para lhe dar sustentação.

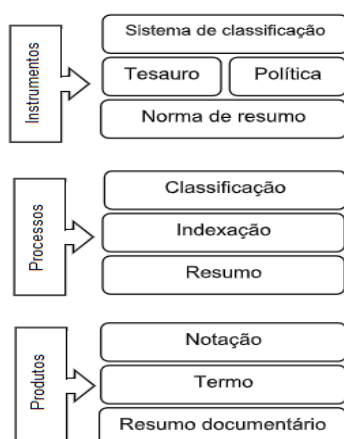


Figura 1. Mapa do Tratamento Temático da Informação

3. Metodologia

O estudo caracteriza-se como descritivo, de natureza quali-quantitativa. Também é de ordem bi-

bliográfica, com emprego de bibliometria e análise de conteúdo, descerrando teor analítico-interpretativo sobre uma frente específica da literatura da Ciência da Informação, que é a organização do conhecimento. Nesta, pauta-se pela abordagem em tratamento temático da informação.

A coleta de dados foi realizada nos anais do Congresso Brasileiro em Organização e Representação do Conhecimento (ISKO-Brasil). A escolha dessa literatura como objeto da investigação se justifica pela intenção de perscrutar o desenvolvimento do tema TTI no principal fórum científico nacional brasileiro com foco em OC, realizado também em diversos outros países.

O evento da ISKO constitui-se, pois, como um fórum científico internacional de teor representativo (Guimarães, 2008) e, objeto de interesse na investigação sobre organização do conhecimento (Guimarães *et al.*, 2015). Configura-se mediante capítulos específicos mundo afora.

Ainda que a ISKO-Brasil tenha sido, inicialmente, abrigada junto a duas edições do Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB) – em 2007 e 2009 – o capítulo brasileiro propriamente dito teve início, de forma independente, no ano de 2011. Os anais dessa primeira edição foram lançados em 2012.

Dessa forma, a ISKO-Brasil conta, desde então, com cinco edições do evento, a saber: 2011, 2013, 2015, 2017 e 2019. O Brasil foi também sede de uma edição internacional do congresso. Esta, porém, não compõe o corpus de análise deste estudo, considerando que tem por objetivo o mapeamento específico do universo brasileiro e por se entender que internacionalmente o tratamento temático se afigura a partir de outras concepções e perspectivas. A análise de conteúdo contemplou trabalhos sobre tratamento temático recuperados junto aos anais eletrônicos dos anos de 2012, 2013, 2015, 2017 e 2019, acessados na página institucional da ISKO-Brasil. Foram analisados, assim, todos os cinco anais desse evento.

O Quadro 1 apresenta as informações descritivas dos anais da ISKO-Brasil. Observa-se que além dos diferentes temas norteadores do evento, cada um foi realizado em uma localidade do país, a cada biênio, compondo cinco edições. Por serem os anais o objeto deste estudo, opta-se por indicar a edição 1 com o ano de 2012, quando foram lançados, não 2011, quando foi realizado o evento.

Outra característica da ISKO-Brasil é a divisão dos trabalhos em três dimensões principais, conforme pode ser observado no Quadro 2. São elas: a epistemológica, a aplicada e a política.

| Ano | Organizadores | Título | Local |
|------|--|--|---------------------|
| 2012 | Guimarães, José Augusto Chaves; Dodebei, Vera | Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade | Brasília (DF) |
| 2013 | Dodebei, Vera; Guimarães, José Augusto Chaves | Complexidade e organização do conhecimento: desafios de nosso século | Rio de Janeiro (RJ) |
| 2015 | Guimarães, José Augusto Chaves; Dodebei, Vera | Organização do conhecimento e diversidade cultural | Marília (SP) |
| 2017 | Pinho, Fabio Assis; Guimarães, José Augusto Chaves | Memória, tecnologia e cultura na organização do conhecimento | Recife (PE) |
| 2019 | Barros, Thiago Henrique; Bragato; Tognoli, Natalia Bolfarini | Organização do conhecimento responsável: promovendo sociedades democráticas e inclusivas | Belém (PA) |

Quadro 1. Anais da ISKO-Brasil

Nota-se que as três primeiras edições contemplavam, na nomenclatura, a designação das dimensões enquanto: epistemológica, aplicada e social, cultural e política da organização e representação do conhecimento. Nas duas últimas, por sua vez, sintetizam-se como epistemológica, aplicada e política e social da organização do conhecimento. Houve, ainda, nas edições de 2012 e 2015, outras dimensões e seções, além das principais supramencionadas.

Vale ressaltar que as dimensões epistemológica, aplicada e política e social, enquanto subtemas do capítulo brasileiro da ISKO, adotados desde a primeira edição do congresso no Brasil, contemplam, cada qual, um dado aspecto da organização do conhecimento. Assim, a dimensão epistemológica volta-se às bases conceituais, históricas e metodológicas da OC, assim como aos seus diálogos interdisciplinares e à sua produção científica. A aplicada pauta-se em modelos, formatos, instrumentos, produtos e estruturas em OC. A política e social, por sua vez, relaciona-se à formação e atuação profissional, ética, contextos, cultura e identidade em OC (Guimarães e Dodebei, 2012).

Corporificada nessas três principais dimensões, a ISKO-Brasil assume um caráter próprio, contemplando aspectos reflexivos de ordem teórica e prática em organização do conhecimento, a partir de sua configuração nacional. As abordagens epistemológica, aplicada, política e social

constituem-se, portanto, como uma influência à comunidade de OC no Brasil. Apesar disso, enquanto representante brasileira da ISKO (Sales, 2017), a ISKO-Brasil acaba por aproximar, também, o debate internacional em OC da realidade brasileira (Vieira e Karpinski, 2019).

| Ano | Dim. principais | Outras dimensões |
|------|--|---|
| 2012 | Epistemológica; Aplicada; e Social, Cultural e Política da ORC | Pesquisa em ORC no Brasil: uma análise dos cursos de pós-graduação; Pesquisa em ORC no Brasil: uma análise dos grupos de pesquisa; Visibilidade científica e interlocução internacional em ORC |
| 2013 | Epistemológica; Aplicada; e Social, Cultural e Política da ORC | — |
| 2015 | Epistemológica; Aplicada; e Social, Cultural e Política da ORC | Perspectivas da Pesquisa em ORC no Brasil; Pesquisa em Análise Documentária no Brasil: a influência de Jean Claude Gardin; Resumos dos trabalhos ganhadores do prêmio ISKO Brasil 2015, a serem publicados, na íntegra, na revista Knowledge Organization; Resumos dos trabalhos ganhadores do prêmio ISKO Brasil 2015, a serem publicados, na íntegra, na revista Brazilian Journal of Information Science |
| 2017 | Epistemológica; Aplicada; e Política e Social da OC | — |
| 2019 | Epistemológica; Aplicada; e Política e Social da OC | — |

Quadro 2. Dimensões da ISKO-Brasil

Quanto aos procedimentos de investigação, em um primeiro momento, a partir do acesso individual aos anais, realizou-se a prospecção sobre os sumários de cada um deles, em busca de produções versadas em TTI junto às três dimensões principais. Em seguida, selecionaram-se manualmente entre estes, durante a leitura dos títulos, os trabalhos concernentes aos instrumentos, processos ou produtos do tratamento temático da informação.

Por fim, procedeu-se à confirmação da aderência temática dos trabalhos selecionados, mediante leitura do resumo ou da introdução dos mesmos, com a concomitante base analítico-interpretativa sobre o conjunto amostral. Esta desenvolvida prioritariamente por meio da bibliometria, pela análise dos indicadores de produção. Em suma, aplicou-se análise de conteúdo para: avizinhar as

produções do tema TTI; eleger aquelas versadas em seus instrumentos, processos ou produtos; estabelecer as concernentes, de fato, ao assunto; e agrupá-las, por fim, nas categorias temáticas, considerando seu enfoque.

Teve-se, como critério central de seleção dos trabalhos, a presença – nos títulos – das palavras e expressões designativas dos instrumentos, processos e produtos do TTI, conforme apresentado na Figura 1 (seção 2), considerando-se também, termos assemelhados, observados em algumas produções, a saber: notação ou número de chamada, sistema ou esquema de classificação, tesouro ou vocabulário controlado. Nesse ponto, evidencia-se uma limitação do estudo bibliográfico, que considerou primordialmente a designação presente no título dos trabalhos. Em alguma medida, portanto, enfatizaram-se produções que delinearão de forma estrita o teor de suas abordagens em tratamento temático da informação, excluindo-se outras que porventura tenham perpassado de forma breve sobre os elementos do mesmo, mas, sem explicitá-los no título.

4. Resultados e discussões

A Tabela I apresenta a distribuição dos estudos publicados nos anais da ISKO-Brasil, segundo as três principais dimensões, ao longo das cinco edições. Nela, se observa uma tendência ascendente de participação dos investigadores brasileiros nos congressos deste capítulo da ISKO, em que, no cômputo geral, aproximadamente metade das investigações apresentadas estiveram focadas na dimensão aplicada da OC, reiterando a visão do tratamento temático enquanto ação prática (Oliveira e Martínez-Ávila, 2019).

| | <i>Epistemológica</i> | <i>Aplicada</i> | <i>Política</i> | <i>Total</i> |
|-------|-----------------------|-----------------|-----------------|--------------|
| 2012 | 7 (28%) | 11 (41%) | 7 (28%) | 25 (100%) |
| 2013 | 14 (31%) | 14 (31%) | 17 (38%) | 45 (100%) |
| 2015 | 20 (32%) | 35 (56%) | 7 (11%) | 62 (100%) |
| 2017 | 13 (28%) | 21 (46%) | 12 (26%) | 46 (100%) |
| 2019 | 12 (21%) | 35 (63%) | 9 (16%) | 56 (100%) |
| Total | 66 (28%) | 116 (50%) | 52 (22%) | 234 (100%) |

Tabela I. Total de trabalhos dos anais da ISKO-Brasil, por ano e dimensão

Outrossim, a prevalência dos estudos na dimensão aplicada está regularmente distribuída ao longo de todas as edições do evento, com exceção do ano de 2013, em que a dimensão política concentrou um maior número de trabalhos apresentados que as outras duas dimensões (epistemológica e aplicada). Corroborá-se,

então, a visão da ISKO-Brasil enquanto expressivo fórum científico em organização do conhecimento, especialmente na investigação de questões prático-laborais desse âmbito.

Do total de 234 trabalhos publicados nos anais da ISKO-Brasil (Tabela I), 43 (18,4%) atenderam o critério estabelecido neste estudo, a saber: a presença das palavras e expressões designativas dos instrumentos, processos e produtos do TTI nos títulos dos trabalhos, assim como daquelas correlatas. Esse corpus de análise está mais concentrado na dimensão aplicada, uma vez que os trabalhos que versaram sobre o tema estão distribuídos da seguinte forma: 16,7% (11 de 66) da dimensão epistemológica; 24,1% (28 de 116) da dimensão aplicada; e 7,7% (4 de 52) da dimensão política.

É, portanto, na dimensão aplicada da ISKO-Brasil que está a maior concentração de trabalhos sobre instrumentos, processos e produtos do TTI. Em seguida, destaca-se a epistemológica e, com menor número, a política. Esse cenário demonstra que a comunidade brasileira da área de organização do conhecimento tem investigado sobretudo questões práticas acerca dos elementos do tratamento temático.

Ademais, a presença quantitativamente significativa (20%) de trabalhos com títulos expressando explicitamente seu foco em aspectos relacionados aos instrumentos, processos e produtos do TTI, no principal evento científico da área de OC, pode ser considerada um indicativo da integração dessa temática com a organização do conhecimento, da qual, aliás, advém (Felipe *et al.*, 2019).

Assim sendo, o fato dos investigadores brasileiros se sentirem à vontade para explicitar suas investigações acerca dos elementos do tratamento temático na ISKO-Brasil, reforça a integração OC e TTI no mesmo universo epistemológico (Guimarães, 2008). O que também sinaliza para uma tendência investigativa do próprio evento, na medida em que se configura como lócus científico sobre operações, atributos e aparatos do trato temático.

A Tabela II apresenta a visualização diacrônica desse conjunto analítico nas três dimensões principais nas edições da ISKO-Brasil. A partir da Tabela 2, é possível observar que a porcentagem de investigações focadas nos aspectos aplicados do TTI mantém-se estável ao longo das edições, ao variar entre 20 e 29% dos trabalhos apresentados nesta dimensão. Por outro lado, com exceção da 2ª edição (2013), as investigações que focam aspectos epistemológicos do TTI vêm apresentando regularmente uma tendência ascendente. Vale destacar, ainda, que os trabalhos sobre instrumentos, processos e produtos do TTI, na dimensão política, embora surjam apenas

a partir das últimas edições (2017 e 2019) da ISKO-Brasil, apresentam-se também de forma percentualmente crescente.

| Ano | Epistem. | | Aplic. | | Política | | Total | |
|-------|----------|-----|--------|-----|----------|-----|-------|-----|
| | nº. | % | nº. | % | nº. | % | nº. | % |
| 2012 | 1 | 14% | 3 | 27% | 0 | 0% | 4 | 16% |
| 2013 | 1 | 7% | 3 | 21% | 0 | 0% | 4 | 9% |
| 2015 | 3 | 15% | 7 | 20% | 0 | 0% | 10 | 16% |
| 2017 | 2 | 15% | 6 | 29% | 1 | 8% | 9 | 20% |
| 2019 | 4 | 33% | 9 | 26% | 3 | 33% | 16 | 29% |
| Total | 11 | 17% | 28 | 24% | 4 | 8% | 43 | 18% |

Tabela II. Total de trabalhos tratando do TTI nos anais da ISKO-Brasil, por ano e dimensão

Desse modo, ainda que junto às cinco edições da ISKO-Brasil observe-se uma presença constante e significativa de investigações de natureza aplicada em TTI, os estudos de âmbito epistemológico e político têm mostrado paulatinamente um crescimento no cenário brasileiro em OC. Questão relevante para diversificação das discussões em tratamento temático e viabilização de um lastro científico em organização do conhecimento com abordagens variadas.

Em suma, por se tratar de uma área desafiadora (Sousa, 2013) com diversas designações (Café e Sales, 2010), o tratamento temático acaba demandando incursões de ordem teórica e reflexões que ampliem sua compreensão. Ao mesmo tempo, cada vez mais avizinham-se questões de ordem sociocultural no âmbito da análise e da representação de assunto. Tal conjuntura ajuda a entender o motivo do crescimento dos estudos epistemológicos e políticas sobre instrumentos, processos e produtos do TTI na ISKO-Brasil.

Há que se considerar, ainda, a subjetividade do TTI em termos teóricos e práticos (Oliveira e Martínez-Ávila, 2019; Medeiros *et al.*, 2016; Dias e Naves, 2013), aspecto que incita a busca por seu amadurecimento. E este vem sendo buscado de alguma forma, na medida em que sobrelevam-se investigações sobretudo aplicadas, seguidas pelas epistemológicas e políticas, junto ao capítulo brasileiro da ISKO.

Contraopondo os dados da Tabela II com o Quadro 1, pode-se associar o aumento das investigações nas dimensões epistemológica e política com o desenho sobre os temas norteadores de cada edição da ISKO-Brasil. Assim, é a partir de 2015 que o foco do evento passa a ser mais voltado a questões de diversidade, cultura, memória e inclusão em organização do conhecimento.

Justamente nessas ocasiões, os trabalhos tratando do TTI nas dimensões epistemológica e política, se tornam crescentes.

A Tabela III apresenta o tipo de autoria registrada nos trabalhos apresentados na ISKO-Brasil que tratam de questões associadas aos instrumentos, processos e produtos em TTI, a fim de avaliar se o conhecimento que vem sendo construído nessa temática em âmbito brasileiro tende a ser decorrente de estudos colaborativos.

| | Epist. | | Aplicada | | Política | | Total | |
|------------|--------|-----|----------|-----|----------|-----|-------|-----|
| | nº. | % | nº. | % | nº. | % | nº. | % |
| Individual | - | - | 3 | 11 | 1 | 25 | 4 | 9 |
| Dupla | 7 | 64 | 15 | 54 | 2 | 50 | 24 | 56 |
| Tripla | 2 | 18 | 8 | 29 | - | - | 10 | 23 |
| Quádrup. | 2 | 18 | 2 | 7 | 1 | 25 | 5 | 12 |
| Total | 11 | 100 | 28 | 100 | 4 | 100 | 43 | 100 |

Tabela III. Tipo de autoria nos trabalhos tratando do TTI nos anais da ISKO-Brasil, por dimensão

Há uma tendência de os trabalhos serem desenvolvidos em coautoria, dado que somente 4 (9%) são decorrentes de autoria individual. Ainda, a coautoria dupla configura a maioria (56%) dos trabalhos colaborativos. Esses resultados são semelhantes aos encontrados por Oliveira, Castanha e Grácio (2020), ao observarem que os artigos de autoria simples compreendem uma pequena parte dos resultados científicos socializados em periódicos brasileiros da área de Ciência da Informação, uma vez que 80% dos artigos publicados foram decorrentes de colaboração científica, em especial, em coautoria dupla resultante da relação de orientação acadêmica.

Considerando que o universo aqui analisado consiste de um congresso, em que investigadores, usualmente, comunicam seus estudos a fim de avaliar o interesse da comunidade científica, críticas e sugestões para suas melhorias (Vanz e Silva Filho, 2019), pode-se tomar como uma hipótese plausível que, também em relação ao escopo do presente estudo, as autorias duplas podem configurar reflexo de investigações de mestrado e doutorado desenvolvidas nos programas de pós-graduação brasileiros. Esse indicativo sugere que questões associadas aos instrumentos, processos e produtos em TTI têm despertado interesse de jovens investigadores.

O Gráfico 1 apresenta a distribuição dos trabalhos analisados de acordo com o elemento do tratamento temático mais abordado na ISKO-Brasil.

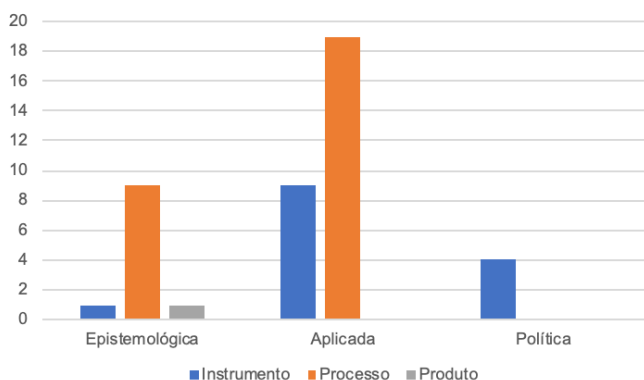


Gráfico 1. Elementos do TTI

Nota-se, pois, o predomínio de investigações sobre processos do tratamento temático (28), seguidas pelas que se voltam aos seus instrumentos (14) e produtos (1). No caso dos instrumentos e processos do TTI, estes marcam presença nas dimensões epistemológica e aplicada, enquanto que os produtos, somente na primeira. A dimensão política é a única a contar com trabalhos sobre instrumentos do tratamento temático.

Sendo objetivo do tratamento temático representar para recuperar documentos a partir do conteúdo (Medeiros *et al.*, 2016), pode-se afirmar que essa perspectiva prático-laboral acaba encontrando eco também nas investigações teóricas no tema, apresentadas na ISKO-Brasil, grandemente voltadas ao universo de seus processos e instrumentos, como expõe o Gráfico 1. Ou seja, são estudos dedicados, sobretudo, à ação aplicada via processos técnicos, amparados por instrumentos documentários específicos.

Acaba sendo ressaltada, ainda, a visão do TTI como área de estudos (Guimarães *et al.*, 2011). Isso por conta das investigações sobre os instrumentos do tratamento temático, presentes na ISKO-Brasil. Investigações dessa ordem demonstram a preocupação de teorizar e conhecer os aparatos estruturantes que apoiam a ação aplicada de tratar informação a partir do seu teor.

Por fim, sendo o TTI um conjunto de instrumentos, processos e produtos de trato do conteúdo (Oliveira e Martínez-Ávila, 2019), é possível afirmar que os estudos apresentados na ISKO-Brasil têm explorado quase que sua completude de aspectos. À exceção dos produtos, pouco representados no capítulo brasileiro. A Tabela IV apresenta a distribuição dos trabalhos sobre instrumentos, processos e produtos do TTI, por ano de publicação. Observa-se que ao longo de 2015 e 2019, houve um maior número de trabalhos sobre processos do tratamento temático. Já os instrumentos e produtos, foram mais trabalhados

nas edições de 2017 e 2019. Marcadamente, porém, foram os processos do TTI que tiveram mais presença desde o início do evento.

| Ano | Instrumento | Processo | Produto |
|-------|-------------|----------|---------|
| 2012 | 1 | 3 | - |
| 2013 | 2 | 2 | - |
| 2015 | 1 | 9 | - |
| 2017 | 4 | 4 | 1 |
| 2019 | 6 | 10 | - |
| Total | 14 | 28 | 1 |

Tabela IV. Elementos do TTI por ano

Dessa forma, diante da diversidade de correntes teóricas (Guimarães, 2008; Guimarães, 2009; Guimarães *et al.*, 2011) que o TTI apresenta, quais sejam a catalogação de assunto, a indexação e a análise documental, os dados da Tabela 4 apontam para uma conjuntura investigativa do mesmo modo diversa na ISKO-Brasil, contemplando instrumentos, processos e produtos, de forma variada. Ainda assim, sobressai a abordagem de ordem mais aplicada no tema, essencialmente processual e instrumental. Esta pode, aliás, ajudar a estabelecer apontamentos sobre as três correntes sobre as quais se costuma caracterizar o tratamento temático.

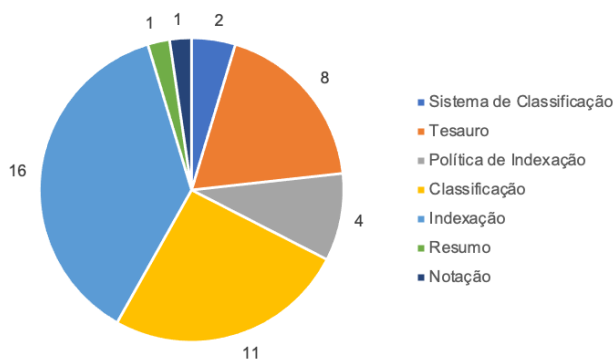


Gráfico 2. Tipos de instrumentos, processos e produtos do TTI

O Gráfico 2 apresenta os tipos de instrumentos, processos e produtos do TTI que são objeto da investigação na ISKO-Brasil. Destacam-se, no contexto dos instrumentos, o tesouro e a política de indexação. Entre os processos, a indexação e a classificação. E em termos de produtos, a notação. As categorias norma de resumo, termo e resumo documentário, não figuram no Gráfico 2, por não aparecerem nas investigações da ISKO-

Brasil. Contudo, diante da necessidade de compreender as particularidades do tratamento temático (Oliveira e Martínez-Ávila, 2019), é primordial que esse conjunto passe a ser, também, explorado. Entende-se, pois, que mesmo essa constatação da ausência de estudos sobre tais instrumentos e produtos do TTI, no capítulo brasileiro da ISKO, representa um aspecto em potencial ao sinalizar para uma lacuna que precisa ser perscrutada pelos investigadores dedicados à vertente do trato temático.

| | <i>Epistemológica</i> | <i>Aplicada</i> | <i>Política</i> |
|--------------------------|-----------------------|-----------------|-----------------|
| Sistema de classificação | 1 | - | 1 |
| Tesouro | - | 6 | 2 |
| Política de indexação | - | 3 | 1 |
| Norma de resumo | - | - | - |
| Classificação | 4 | 7 | - |
| Indexação | 5 | 11 | - |
| Resumo | - | 1 | - |
| Notação | 1 | - | - |
| Termo | - | - | - |
| Resumo documentário | - | - | - |

Tabela V. Tipos de instrumentos, processos e produtos do TTI por dimensão

Apesar disso, percebe-se junto ao Gráfico 2, o destaque de categorias como a indexação, a classificação, o tesouro e a política de indexação. Mais uma vez, reitera-se aqui a constatação do Gráfico 1, que evidencia processos e instrumentos como os elementos mais abordados nos trabalhos sobre TTI, apresentados no capítulo brasileiro da ISKO.

A Tabela V apresenta a distribuição dos trabalhos sobre cada instrumento, processo e produto do TTI, por dimensão. Conforme a Tabela V, o tesouro e a política de indexação são abordados nas dimensões aplicada e política da ISKO-Brasil; e os sistemas de classificação, na epistemológica e na política. Já a indexação e a classificação, insurgem nas dimensões epistemológica e aplicada; e o resumo, na aplicada. Por fim, a notação aparece na dimensão epistemológica.

Tem-se, no geral, mais trabalhos explorando a classificação, a indexação, o tesouro, a política de indexação e, em menor número, o sistema de classificação. Por ser base substancial da organização da informação (Café e Sales, 2010), é compreensível que o TTI se afigure nos trabalhos

da ISKO-Brasil prioritariamente a partir de dois de seus processos e, ainda, dos instrumentos que lhes dão sustentação. Afinal, representa a base aplicada a partir da qual é enxergado, assim como reforça sua derivação histórica da ação pragmática de trato documental (Guimarães, 2008).

| | Item | Abordagem | Nº. | |
|---|-----------------------------------|--|-------------------------|---|
| Instrumentos | Sistema de classificação | Classificação Decimal de Dewey e Classificação Decimal Universal | 1 | |
| | | Influências filosóficas | 1 | |
| | Tesouro | Diversidade cultural | 1 | |
| | | Estruturação no modelo SKOS | 2 | |
| | | Modelo de conversão | 1 | |
| | | Tesouro de direitos humanos | 1 | |
| | | Tesouro funcional | 1 | |
| | | Tesouro jurídico | 2 | |
| | Política de indexação | Base de dados | 1 | |
| | | Biblioteca universitária | 1 | |
| | | Diretrizes multiculturais | 1 | |
| | | Estruturação de vocabulário controlado | 1 | |
| | Classificação | Classificação de arquivos | 7 | |
| | | Classificação de informações estatísticas | 1 | |
| Classificação de interesse do leitor | | 1 | | |
| Classificação Decimal de Dewey | | 1 | | |
| Classificação profissional e não-profissional | | 1 | | |
| Indexação | | Análise documental | 1 | |
| | | Base de dados | 1 | |
| | | Contribuição teórica e prática | 1 | |
| | | Estudos semânticos | 1 | |
| | | Indexação de arquivos | 1 | |
| | Indexação de fotografias | 1 | | |
| | Indexação de imagens | 3 | | |
| | Indexação de publicações seriadas | 1 | | |
| | Indexação de teses e dissertações | 1 | | |
| | Indexação multimodal | 1 | | |
| Processos | Indexação social | 1 | | |
| | ISKO-Brasil | 1 | | |
| | Nanopublicações | 1 | | |
| | Pós-verdade | 1 | | |
| | Resumo | Elaboração de resumos | 1 | |
| | Produtos | Notação | Ordenação de documentos | 1 |

Tabela VI. Abordagens nos instrumentos, processos e produtos do TTI

A Tabela VI mostra as abordagens em cada instrumento, processo e produto do TTI. Observam-se 33 abordagens distintas, a maioria (20) delas

centrada nos processos, uma pequena parte (12) voltada aos instrumentos e, somente uma relativa aos produtos do TTI. Tais dados são, antes de mais nada, representativos da instância prática do tratamento temático, cujo desenvolvimento cotidiano se faz a partir de ações de representação do assunto e com instrumental específico (Oliveira e Martínez-Ávila, 2019).

Como pode ser observado, entre os instrumentos, destacam-se as abordagens relativas à estruturação de tesouro no modelo SKOS e ao tesouro jurídico. Classificação de arquivos e indexação de imagens são as abordagens destacadas entre os processos e a ordenação de documentos, entre os produtos. Nessa medida, configura-se, junto à ISKO-Brasil, um entorno de investigação em tratamento temático sobre construção e tipologia de tesouro e, trato documental na forma de arquivos e imagens.

Há que se ressaltar que o TTI é parte das operações do trato informacional (Fujita, 2013), estando focado na análise, na descrição e na representação do conteúdo (Guimarães *et al.*, 2011). Incidem, nos anais da ISKO-Brasil, estudos com abordagens que contemplam justamente esse entorno caracterizador, ou seja, para fundamentar operações de trato informacional por assunto mediante constituição de vocabulários controlados, permitindo sua classificação, indexação e ordenação.

5. Conclusão

A investigação buscou analisar as abordagens acerca dos instrumentos, processos e produtos do tratamento temático da informação nos anais da ISKO-Brasil.

Constatou-se o predomínio de trabalhos sobre o tesouro e a política de indexação, enquanto instrumentos do TTI. Estes, junto às dimensões aplicada e política. No âmbito dos processos, sobressaem trabalhos sobre a indexação e a classificação, em sua maioria junto à dimensão aplicada. No que diz respeito aos produtos, o destaque é para um trabalho sobre notação, junto à dimensão epistemológica.

Quanto às abordagens, destacam-se: junto aos instrumentos, o tesouro, em termos de sua construção e tipologia; e junto aos processos, a classificação de arquivos e a indexação de imagens. Tem-se, em suma, um cenário de investigação científica sobre constituição de vocabulário controlado e, classificação, indexação e ordenação de informação.

Nessa medida, o pressuposto central da investigação, de que a ISKO-Brasil descerra questões elementares acerca dos instrumentos, processos

e produtos do TTI que permitem, em alguma medida, configurar aspectos pontuais nesse entorno, foi confirmado. Afinal, é justamente sobre indexação, classificação, tesouro e política de indexação, o maior enfoque das investigações na ISKO-Brasil, com a preocupação de explorar a realização dos processos e a construção dos instrumentos.

Dessa forma, em alguma medida, a trajetória de investigação em tratamento temático da informação, no capítulo brasileiro da ISKO, deixa seu contributo para a perspectiva internacional em organização do conhecimento. Afinal, ao longo de cinco edições da ISKO-Brasil, evidencia-se um lastro investigativo no tema TTI, essencialmente pautado em seus instrumentos e processos, junto ao contexto da OC.

Além disso, a própria comunidade brasileira tem sido bastante ativa na ISKO internacional. Aspecto esse que vem sendo reconhecido por esse grupo de investigadores, sobretudo a partir de Smiraglia (2011, 2013a, 2013b, 2017 e 2018) e que acaba por representar, também, um relevante aporte e ponto de contribuição.

Estudos futuros podem se dedicar a explorar as abordagens em TTI junto a outros fóruns científicos em Biblioteconomia e Ciência da Informação. Outra perspectiva de investigação diz respeito ao aprofundamento analítico sobre instrumentos, processos e produtos do tratamento temático em âmbito formativo e aplicado.

Referências

- Bräscher, Marisa; Guimarães, José Augusto Chaves (2018). Tratamento temático da informação (TTI): influência dos paradigmas físico, cognitivo e social em artigos de revisão de literatura no período de 1966-1995. // *Liinc em Revista*. 14:2 (nov. 2018) 241-258.
- Braz, Márcia Ivo; Carvalho, Evanise Souza de (2017). Práticas em tratamento temático da informação: interfaces de ensino e aprendizagem. // *Revista Brasileira de Biblioteconomia e Documentação*. 13:n. esp. (2017). 2496-2509.
- Café, Lígia Maria Arruda; Sales, Rodrigo de (2010). Organização da informação: conceitos básicos e breve fundamentação teórica. // *Robredo, Jaime; Bräscher, Marisa (orgs.)*. Passeios pelo bosque da informação: estudos sobre representação e organização da informação e do conhecimento. Brasília: IBICT, 2010. 115-129.
- Dal'Evedove, Paula Regina; Fujita, Mariângela Spotti Lopes (2013). Estudo sociocultural da comunidade discursiva do tratamento temático da informação em bibliotecas universitárias. // *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*. 18:36 (jan./abr. 2013) 23-50.
- Dias, Eduardo Wense; Naves, Madalena Martins Lopes (2013). Análise de assunto: teoria e prática. 2. ed. rev. Brasília: Briquet de Lemos, 2013.
- Felipe, Carla Beatriz Marques; Pinho, Fabio Assis; Medeiros, Wagner Oliveira de (2019). Mapeamento da produção Norte e Nordeste sobre indexação de imagens: um estudo a partir dos repositórios institucionais. // *Barros,*

- Thiago Henrique Bragato; Tognoli, Natalia Bolfarini (orgs.). *Organização do conhecimento responsável: promovendo sociedades democráticas e inclusivas*. Belém: Ed. da UFPA, 2019. 342-348.
- Fujita, Mariângela Spotti Lopes (2013). A importância teórica e prática da indexação na fundamentação científica da organização e representação do conhecimento. // Dodebei, Vera; Guimarães, José Augusto Chaves (orgs.). *Complexidade e organização do conhecimento: desafios do nosso século*. Rio de Janeiro: ISKO-Brasil; Marília: FUNDEPE, 2013. 147-159.
- Guimarães, José Augusto Chaves (2008). A dimensão teórica do tratamento temático da informação e suas interações com o universo científico da International Society for Knowledge Organization (ISKO). // *Revista Ibero-Americana de Ciência da Informação*. 1:1 (jan./jun. 2008) 77-99.
- Guimarães, José Augusto Chaves (2009). Abordagens teóricas de tratamento temático da informação: catalogação de assunto, indexação e análise documental. // *Ibersid*. 3 (2009) 105-117.
- Guimarães, José Augusto Chaves; Dodebei, Vera Lúcia Doyle Louzada de Mattos (2012). Introdução. // Guimarães, José Augusto Chaves; Dodebei, Vera Lúcia Doyle Louzada de Mattos (orgs.). *Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade*. Marília: ISKO-Brasil; FUNDEPE, 2012. 12-20.
- Guimarães, José Augusto Chaves; Ferreira, Gustavo M.; Freitas, Maria Fernanda 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. // Pérez Pais, Carmen; González Bonome, María de los Ángeles (eds.). *20 años del Capítulo Español de ISKO: actas del X Congreso ISKO-Espanha*. Ferrol: Universidade da Coruña, 2011. 181-194.
- Guimarães, José Augusto Chaves; Matos, Daniela Fernanda de Oliveira; Santos, André Ynada dos; Sales, Rodrigo de. 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 // *Scire*. 21:2 (jul./dic. 2015) 13-26.
- Martinho, Noemi Oliveira; Fujita, Mariângela Spotti Lopes (2010). La catalogación de materias: apuntes históricos sobre su normalización. // *Scire*. 16:2 (2010) 61-70.
- Medeiros, Graziela Martins de; Vital, Luciane Paula; Bräscher, Marisa (2016). Tratamento temático da informação em documentos arquivísticos: estudo dos anais da ISKO e do GT2 do ENANCIB. // *Tendências da Pesquisa Brasileira em Ciência da Informação*. 9:1 (jan./ago. 2016).
- Oliveira, Caroline G.; Castanha, Rafael; Grácio, Maria Cláudia C. Autoria dupla na produção científica da área da Ciência da Informação: um estudo nos periódicos brasileiros Qualis A1 e A2 (2013-2017). // *Encontro Brasileiro de Bibliometria e Cientometria (EBBC)*, 2020.
- Oliveira, Lais Pereira de; Martínez-Ávila, Daniel (2019). Tratamento temático da informação: análise sobre abordagens e enfoques em artigos científicos da área de Ciência da Informação. // *Informação em Pauta*. 4:2 (jul./dez. 2019) 83-100.
- Sales, Rodrigo de. Diferentes perspectivas nos contextos do GT2 da ANCIB e da ISKO-Brasil. // *Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB)*, 2017.
- Santos, Raimunda Fernanda dos; Silva, Jefferson Higinio da; Albuquerque, Maria Elizabeth Baltar Carneiro de; Oliveira, Henry Poncio Cruz de (2019). Implicações da Pós-verdade na indexação de recursos informacionais. // Barros, Thiago Henrique Bragato; Tognoli, Natalia Bolfarini (orgs.). *Organização do conhecimento responsável: promovendo sociedades democráticas e inclusivas*. Belém: Ed. da UFPA, 2019. 85-94.
- Smiraglia, Richard P. (2011). ISKO 11's Diverse Bookshelf: An Editorial. // *Knowledge Organization*. 38:3 179-186.
- Smiraglia, Richard P. (2013a). ISKO 12's Bookshelf-Evolving Intension: An Editorial. // *Knowledge Organization*. 40:1 3-10.
- Smiraglia, Richard P. (2013b). The epistemological dimension of knowledge organization. // Guimarães, José Augusto Chaves; Dodebei, Vera Lúcia Doyle Louzada de Mattos (orgs.). *Complexidade e organização do conhecimento: desafios de nosso século*. Rio de Janeiro: ISKO-Brasil; Marília: FUNDEPE, 2013. 17-25.
- Smiraglia, Richard P. (2017). ISKO 14's Bookshelf: Discourse and Nomenclature-An Editorial. *Knowledge Organization*. // 44:1 3-12.
- Smiraglia, Richard P. (2018). ISKO 15's Bookshelf: Dispersion in a Digital Age-An Editorial. // *Knowledge Organization*. 45:5 343-357.
- Sousa, Brisa Pozzi (2013). Representação temática da informação documentária e sua contextualização em biblioteca. // *Revista Brasileira de Biblioteconomia e Documentação*. 9:2 (jul./dez. 2013) 132-146.
- Sousa, Brisa Pozzi; Fujita, Mariângela Spotti Lopes (2013). A classificação bibliográfica no contexto do tratamento temático da informação: um estudo com protocolo verbal individual em bibliotecas do Instituto Federal de Educação, Ciência e Tecnologia (IF's). // *Revista ACB: Biblioteconomia em Santa Catarina*. 18:1 (jan./jun. 2013) 796-813.
- Vanz, Samile Andréa de Souza; Silva Filho, Rubens da Costa (2019). O protagonismo das revistas na comunicação científica: histórico e evolução. In: Carneiro, F.F.B.; Ferreira Neto, A.; SANTOS, W. (orgs.). *Comunicação científica em periódicos*. 1.ed. Curitiba: Appris, 2019. 19-44.
- Vieira, Ana Paula da Fonseca; Oliveira, Lais Pereira de; Cunha, Tatielle Marques (2017). Incursões sobre o tratamento temático da informação: estudo da política de indexação em bibliotecas universitárias goianas. // *Informação em Pauta*. 2:1 (jan./jun. 2017) 28-49.
- Vieira, Keitty Rodrigues; Karpinski, Cezar (2019). Indexação: uma análise da produção científica no capítulo brasileiro da International Society for Knowledge Organization (ISKO-Brasil). // Barros, Thiago Henrique Bragato; Tognoli, Natalia Bolfarini (orgs.). *Organização do conhecimento responsável: promovendo sociedades democráticas e inclusivas*. Belém: Ed. da UFPA, 2019. 145-155.

Enviado: 2020-03-27. Segunda versão: 2020-05-22.
 Aceptado: 2020-06-04.

La Documentación al servicio de proyectos I+D+i: la experiencia del centro de documentación de Cibermov

*Library and Information Science at the service of R&D projects:
the experience of the Cibermov documentation centre*

Rubén DOMÍNGUEZ-DELGADO, María-Ángeles LÓPEZ-HERNÁNDEZ

Facultad de Comunicación de la Universidad de Sevilla, Avenida Américo Vespucio, s/n, C.P. 41092, Sevilla,
rdd@us.es, alhernan@us.es

Resumen

Se presenta la experiencia de creación y gestión de un centro de documentación al servicio de los integrantes del proyecto I+D+i: *Ciberactivismo, ciudadanía digital y nuevos movimientos urbanos* (Cibermov), como ejemplo práctico de la importante labor que puede desempeñar la documentación y sus especialistas en este tipo de proyectos. Explicamos las diversas funciones desempeñadas por este servicio, centrándonos especialmente en el desarrollo y gestión de una base de datos de fuentes bibliográficas y audiovisuales seleccionadas en torno al ámbito temático del proyecto.

Palabras clave: Proyectos I+D+i. Documentación. Centros de documentación. Bases de datos. Ciberactivismo. Ciudadanía digital. Movimientos sociales.

Abstract

The experience of creating and managing a documentation center at the service of researchers in the R&D project *Cyberactivism, digital citizenship and new urban movements* (Cibermov), is presented as a practical example of the important work that can be carried out by library and information science specialists in this type of projects. We explain the various functions performed by this, focusing especially on the development and management of a database of selected bibliographic and audiovisual sources around the thematic scope of the project.

Keywords: R&D projects. Library and information science. Documentation centers. Databases. Cyberactivism. Digital citizenship. Social movements.

1. Introducción y objetivo

Un proyecto I+D+i puede concebirse, desde un punto de vista documental, como un sistema de información, de vida temporal, integrado por un grupo de investigadores organizados en torno a una temática concreta y en el que (con respaldo de unos recursos económicos y materiales, además de los humanos) se requiere, en primer lugar, de unas entradas de documentos con información pertinente y relevante (exploración documental), a partir de las cuales, tras la correspondiente experimentación, se generarán, en último lugar, nuevos documentos de salida con información acerca de la temática específica (resultados de investigación), en forma de artículos científicos, libros, capítulos, tesis doctorales o patentes, entre otros formatos, que suponen un aporte de nuevo conocimiento que contribuirá al progreso.

En este tipo de sistemas, la exigencia de obtener tales resultados en un tiempo muy limitado (normalmente en torno a tres o cuatro años) convierte a la gestión de la información y a la disciplina de la documentación en claves para la eficacia, la eficiencia, y, en definitiva, el éxito, del proyecto. Y es que, como apunta Kisielnicki (2014, p. 57),

la gestión de proyectos I+D+i requiere, entre otras competencias, de grandes habilidades y conocimientos sobre información moderna, tecnologías de comunicación, sistemas de gestión de información y herramientas avanzadas que contribuyan al desarrollo de la organización.

Así, los documentalistas, como en cualquier otro ámbito informativo, pueden también en este marco, como afirman Fuentes Pujol y Arguimbau Vivó (2008, p. 55), “organizar la dinámica de los flujos de entrada, procesamiento y salida de los datos de manera eficiente, rápida y precisa (...) Analizar de forma rigurosa los diferentes tipos de documentos, superando los inconvenientes detectados y aportando metodologías robustas”. También consideran estos autores que los documentalistas serían necesarios para “difundir las iniciativas científicas y tecnológicas y los resultados obtenidos”.

Según Repiso (2012), se trataría de un perfil similar al de “bibliotecario integrado”, aunque, no obstante, este tendría “la complicación de que debe integrarse en una mezcla de grupos complejos, con diferentes dinámicas de trabajo e información de diferentes tipologías”.

Por todo lo expuesto, y dado que no existe una amplia bibliografía relativa a la importante labor que puede desarrollar la Documentación en proyectos I+D+i – sobre todo a nivel de casos de aplicación práctica, a excepción del informe de *Research Information Network* (2011) sobre proyectos de investigación colaborativos, en general –, nos planteamos con este trabajo dar a conocer nuestra experiencia positiva de creación y gestión de un centro de documentación propio dentro del proyecto I+D+i *Ciberactivismo, ciudadanía digital y nuevos movimientos urbanos* (referencia CSO2016-78386-P), financiado por el Ministerio de Economía y Competitividad del Gobierno de España y de ejecución en el periodo 2016-2020.

2. El centro de documentación del proyecto I+D+i Cibermov: funciones

Al inicio de este proyecto, integrado por 17 miembros, se decidió que dos de ellos, especialistas en Documentación, creasen y gestionasen un centro de documentación vinculado al mismo y al que se asignarían las siguientes funciones principales e inmediatas:

- El diseño y desarrollo de una base de datos relacional, bibliográfica y audiovisual, que diese acceso inmediato a los investigadores a fuentes pertinentes y relevantes sobre las temáticas del proyecto.
- La búsqueda, selección y registro de dichas fuentes en la base de datos, que sería, cada tres meses, actualizada y sometida a controles para evitar duplicaciones y obsolescencia de registros.

Asimismo, el centro de documentación asumiría las siguientes funciones específicas:

- Recopilar y registrar en la base de datos los resultados de la investigación que se fuesen publicando, en los distintos formatos y en el transcurso del proyecto, por parte de sus miembros.
- Elaborar estudios bibliométricos en torno al ciberactivismo, la ciudadanía digital y los nuevos movimientos urbanos, inexistentes hasta la fecha y que consideramos de gran importancia, siguiendo la estela de otros muchos investigadores que han hecho uso de este método cuantitativo para mapear el conocimiento en sus respectivas áreas (Tijssen y Van Wijk, 1998; Gu, 2004; Van Leeuwen, 2006; Junquera y Mitre, 2007; Sun y Grimes, 2016; Lino de Araujo, Ávila, Battisti y Cristine, 2017; Liu, Tian, Kong, Lee y Xia, 2019).
- Realizar envíos bisemanales a los integrantes del proyecto, por correo electrónico, de una

selección de diez fuentes, de publicación reciente, recomendadas por el centro de documentación, de entre todas las que este va registrando en la base de datos.

- Dar respuesta y asesoramiento a los usuarios en sus dudas y necesidades informativas.

3. Desarrollo, características y particularidades de la base de datos Cibermov

Las bases de datos se utilizan principalmente para recuperar referencias específicas relativas a la información sobre un tema determinado, publicada en diversos tipos de documentos primarios, por lo que consideramos que disponer de una propia, Cibermov (AA.VV., 2020), con información seleccionada en base a los objetivos del proyecto I+D+i y a las necesidades de sus usuarios, supondría un importante ahorro de tiempo y esfuerzo, sobre todo en la fase de exploración documental, para sus integrantes, quienes podrían centrarse así más en el resto de etapas posteriores de la investigación científica.

También consideramos que la base de datos serviría como una importante herramienta de difusión del conjunto de investigaciones, en los distintos formatos, desarrolladas en el marco del proyecto. Y no solo para su equipo, sino que se decidió, en consonancia con las propuestas de Siso-Calvo y Arquero-Avilés (2018), facilitar acceso abierto a este recurso documental.

Además, esta herramienta sería de gran utilidad, como aseguran Galbán y Vázquez (1988, p. 369), Urdín y Morillo (2000, p. 380), De Souza (2013, p. 108) o Efrain-García y García-Zorita (2016, e122, p.2), para la realización, en el ámbito de estudio, aún joven, del proyecto, de los estudios bibliométricos asignados como objetivo a su centro de documentación.

En cuanto al diseño de la base de datos, esta se estructuró del siguiente modo: por un lado, una vista de usuario (Figura 1), con un cuadro de búsqueda de fuentes de información mediante etiquetas (palabras claves temáticas), títulos, autor/es o año de publicación/difusión; por otro lado, una vista de administrador (Figura 2) desde la que registrar nuevas fuentes, consignando en los siguientes campos la información que correspondía, en cada caso (según la categoría documental): título, resumen, publicación, número de páginas, DOI, enlace al documento original, congreso, ISSN/ISBN, año, lugar de edición, idioma, editor/revista, edición/volumen, fecha de publicación, páginas, autor/es, categoría (artículo científico, libro, capítulo, actas, documento audiovisual...) y etiquetas temáticas.

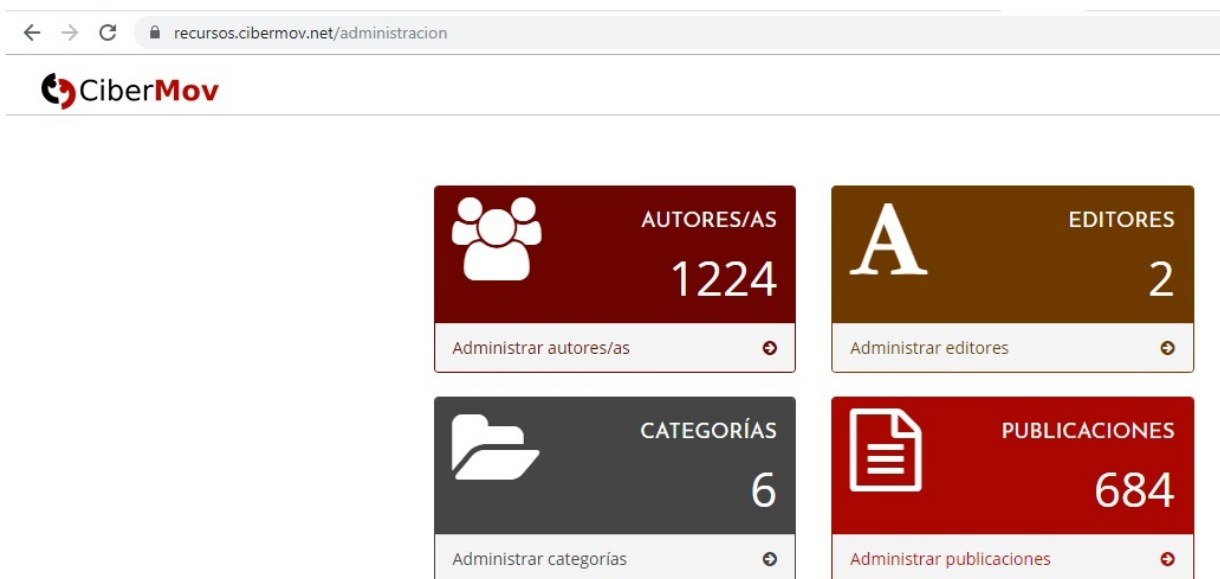


Figura 1. Vista principal de administrador (en apéndice)

> ADMINISTRAR PUBLICACIONES

Datos publicación | Selección autores/as | Selección categorías | Selección etiquetas | Subir imagen

| | |
|--|----------------------|
| Título * | ISBN/ISSN |
| Introduzca el título de la publicación | |
| Resumen | Año |
| | |
| Publicación | Lugar de edición |
| | |
| Núm. Páginas | Idioma |
| | |
| DOI | Editor/Revista |
| | |
| Enlace al texto | Edición/Volumen |
| | |
| Congreso | Fecha de publicación |
| | |
| | Páginas |
| | |

LIMPIAR CANCELAR GUARDAR

Figura 2. Campos del registro (en apéndice)

Por otro lado, con respecto a la selección de las fuentes que serían registradas en la base de datos, se tuvieron en cuenta, además de las necesidades de los usuarios, los objetivos del proyecto I+D+i y los asuntos legales, otros elementos importantes, para Molina Salinas (2019, pp.

42-43), cuando se recopila documentación para una base de datos especializada, que son: disponibilidad documental (a través de acceso abierto o bajo permisos, adquiribles por todos los miembros del proyecto), idiomas que se pretenden cubrir (inglés, francés, italiano, portugués y español

en nuestro caso) y criterios de selección por parte de especialistas en la temática (en nuestro caso, el investigador principal y los coordinadores del proyecto).

Así, se determinó, en primer lugar, y en lo que respecta a las fuentes bibliográficas, hacer una búsqueda de todas las revistas científicas, nacionales e internacionales, que tiene suscritas la Universidad de Sevilla, a la que está vinculada el proyecto, sobre Comunicación, Ciencias Políticas y Sociología. Recopiladas todas las revistas con difusión online (más de 550), y buscados y añadidos los datos de indexación de las revistas en Journal Citation Report y Scimago Journal Ranking/Scopus – por si esto pudiese ser un dato a tener en cuenta en la selección –, los resultados le fueron remitidos a los responsables, quienes decidieron que el centro de documentación centrara su atención especialmente en 45 de esas revistas.

Posteriormente, el centro de documentación decidió, junto con el investigador principal y los coordinadores del proyecto, expertos en la temática del mismo, las palabras claves (o, en su caso, descriptores) que se emplearían para buscar y recuperar las fuentes, tanto en las revistas nacionales como en las internacionales. En total, se sugirieron los 24 términos siguientes: por un lado, y en español, ciberactivismo, cibercultura, ciberdemocracia, ciberpolítica, ciudadanía digital, democracia 4.0, movimientos sociales, tecnopolítica y videoactivismo; por otro lado, y en inglés, *alternative networks, clicktivism, critical communication, cyberactivism, cyberculture, cyberdemocracy, democracy 4.0, digital citizenship, internet activism, netactivism, social media action, social movements, social network, technopolitics y videoactivism*.

En cuanto al periodo de búsqueda de información, se decidió, siguiendo las indicaciones de los responsables del proyecto, que esta se hiciera retrospectivamente desde 2007, justo un año antes del estallido mundial de la crisis económica, hasta 2020. Teniendo en cuenta los anteriores parámetros, esto es, 45 revistas y 24 palabras claves, el centro de documentación realiza 1080 operaciones de búsqueda, trimestralmente, en los buscadores o, en su defecto, en los índices de las revistas seleccionadas, para que la base de datos sea operativa.

Por último, y en lo que respecta al material audiovisual activista seleccionado y registrado en la base de datos, el centro de documentación decidió realizar las búsquedas en las siguientes plataformas: *Youtube, Vimeo, Nodo50, Filmin, Metromuster y 3boxmedia*, habiéndose recopilado y

registrado hasta ahora más de cuarenta documentos españoles online en abierto (en su mayoría con licencia *Creative Commons*).

En la actualidad, los miembros de este centro seguimos trabajando en la búsqueda, selección y registro documental en nuestra base de datos, con casi 700 documentos hasta el momento, a la vez que damos cumplimiento al resto de funciones encomendadas a este servicio

Notas

- (1) Este trabajo forma parte del proyecto de investigación de I+D+i *Ciberactivismo, ciudadanía digital y nuevos movimientos urbanos* (referencia CSO2016-78386-P), financiado por el Ministerio de Economía y Competitividad del Gobierno de España.

Referencias

- AA.VV. (2020) *Cibermov* [base de datos]. <https://www.recur-sos.cibermov.net/> (2020-03-03).
- De Souza, Claudia Daniele (2013). A organização do conhecimento: Estudo bibliométrico na base de dados ISI Web of Knowledge. // *Biblios*. eISSN 1562-4730. 53, 20-32. <http://biblios.pitt.edu/ojs/index.php/biblios/article/view/108> (2020-05-12).
- Efrain-García, Preiddy y García-Zorita, Carlos (2016). Desktop Scientometrics: una metodología para el uso de datos procedentes de WoS mediante el programa estadístico R. // *Revista Española de Documentación Científica*. eISSN 1988-4621. 39:1 (Enero-Marzo 2016) e122. <http://redc.revistas.csic.es/index.php/redc/article/view/327> (2000-05-12).
- Fuentes Pujol, Eulàlia; Arguimbau Vivó, Llorenç (2008). I+D+i: Una perspectiva documental. // *Anales de Documentación*. ISSN 1697-7904. Nº 11 (2008) 43-56. <https://revistas.um.es/analesdoc/article/view/24801> (2020-02-19).
- Galbán, Carmen y Vázquez Valero, Manuela (1988). Las bases de datos como fuentes de información para estudios bibliométricos. // *Boletín de la ANABAD*. ISSN 0210-4164. 38:1-2 (1988) 369-381. <https://dialnet.unirioja.es/servlet/articulo?codigo=2094772> (2020-02-26).
- Gu, Yinian (2004). Global knowledge management research: a bibliometric analysis. // *Scientometrics*. ISSN 1588-2861. 61:2 (Octubre 2004) 171-190. <https://doi.org/10.1023/B:SCIE.0000041647.01086.f4> (2020-02-26).
- Junquera, Beatriz; Mitre, María (2007). Value of bibliometric analysis for research policy: A case study of Spanish research into innovation and technology management. // *Scientometrics*. ISSN 1588-2861. 71:3 (Junio 2007) 443-454. <https://doi.org/10.1007/s11192-007-1689-9> (2020-02-25).
- Kisielnicki, Jerzy (2014). Project management in research and development // *Foundations of Management*. ISSN 2080-7279. 6:3 (2014) 57-70. <https://doi.org/10.1515/fman-2015-0018> (2020-02-21).
- Lino De Araújo, Alcione; Ávila Rodrigues, Bethânia; Battisti Telles, Leomara; Cristine S. Vaz, Mônica (2017). A bibliometric analysis of the Scielo database: a Brazilian chart of solidarity economy. // *Scientometrics*. ISSN 1588-2861. 112:1 (Julio 2017) 1-20. <https://doi.org/10.1007/s11192-017-2382-2> (2020-05-13).
- Liu, Jiaying, Tian, Jiahao; Kong, Xiangjie; Lee, Ivan; y Xia, Feng (2019). Two decades of information systems: a

- bibliometric review. // *Scientometrics*. ISSN 1588-2861. 118:2 (Febrero 2019) 617-643. <https://doi.org/10.1007/s11192-018-2974-5> (2020-05-12).
- Molina Salinas, Claudio. Una base de datos de saber terminológico para las artes espaciales. // *Scire*. ISSN 1135-3716. 25:2 (Julio-Diciembre 2019) 35-44. (2020-02-22).
- Repiso, Rafael (2012). La gestión de la información en los proyectos de colaboración universitarios. // *Blok de BiD* (2012). <http://www.ub.edu/blokdebid/es/content/la-gesti%C3%B3n-de-la-informaci%C3%B3n-en-los-proyectos-de-colaboraci%C3%B3n-universitarios> (2020-02-24).
- Research Information Network (2011). Information handling in collaborative research: An exploration of five case studies. A Research Information Network Report. British Library, 2011. http://www.rin.ac.uk/system/files/attachments/Collaborative_research_report.pdf (2020-02-25).
- Siso-Calvo, Brenda; Arquero-Avilés, Rosario (2018). Acceso abierto, promoción y difusión digital de proyectos I+D+i: el caso de la Universidad Complutense de Madrid. // *Ibersid*. ISSN 1888-0967. 12:2 (2018) 25-34. <https://www.ibersid.eu/ojs/index.php/ibersid/article/view/4565> (2020-02-22).
- Sun, Yutao; Seamus, Grimes (2016). The emerging dynamic structure of national innovation studies: a bibliometric analysis. // *Scientometrics*. ISSN 1588-2861. 106:1 (Enero 2016) 17-40. <https://doi.org/10.1007/s11192-015-1778-0> (2020-05-13).
- Tijssen, Robert y Van Wijk, Erik (1998). The global scientific base of information and communication technologies: bibliometric analysis of ICT research work. // *Scientometrics*. ISSN 1588-2861. 42:1 (Mayo 1998) 41-60. <https://doi.org/10.1007/BF02465011> (2020-02-25).
- Urdín, Carmen y Morillo, Fernanda (2000). Producción científica de Andalucía en las bases de datos Science Citation Index e índice español de ciencia y tecnología. // *Revista Española de Documentación Científica*. eISSN 1988-4621. 23:4, 379-394. <http://redc.revistas.csic.es/index.php/redc/article/view/327> (2000-05-12).
- Van Leeuwen, Thed (2006). The application of bibliometric analysis in the evaluation of social science research. Who benefits and why it is still feasible. // *Scientometrics*. ISSN 1588-2861. 66:1 (Enero 2006) 133-154. <https://doi.org/10.1007/s11192-006-0010-7> (2020-02-25).

Enviado: 2020-03-31. Segunda versión: 2020-05-14.
Aceptado: 2020-06-04.

Avances en sistemas jurídicos de recuperación de documentos

Progress in juridical systems for document retrieval

Fernando GALINDO AYUDA

Facultad de Derecho, Universidad de Zaragoza, Pedro Cerbuna s/n, cfa@unizar.es

Resumen

Los sistemas de recuperación de documentos están en continua mejoría: mediante la incorporación de utilidades dependientes de la expansión de su uso, el avance de las aplicaciones tecnológicas y el claro delimitado de sus funciones. El trabajo propone mediante la consideración de varios ejemplos, desarrollados en el ámbito jurídico, tomarlos como referencia para promover el perfeccionamiento de tales sistemas en otros ámbitos.

Palabras clave: Sistemas de información documental. Derecho de la información. Documentación jurídica.

Abstract

Document retrieval systems are in continuous improvement: through the incorporation of utilities dependent on the expansion of their use, the advancement of technological applications and the clear separation of their functions. The article proposes by considering several examples, developed in the legal and the judicial field, to take them as a reference to promote the improvement of such systems in other areas.

Keywords: Document management systems. Information law. Law LIS.

1. Introducción

Es conocido el pujante desarrollo de los mecanismos de recuperación de documentación. En la actualidad, incluso, utilizando términos procedentes del desarrollo tecnológico, se habla de ellos como “sistemas de organización del conocimiento” (1).

La puesta en realidad de estos sistemas trae numerosas consecuencias prácticas si tenemos en consideración, especialmente, el hecho de que, a diferencia de la organización de “sistemas tradicionales” como los constituidos por las bibliotecas pensadas para proporcionar un acceso general a la información contenida en libros, hoy se utilizan, incluso, los recursos de la denominada Inteligencia Artificial, que posibilitan un acceso particularizado a todo tipo de información proporcionada por los sistemas de organización del conocimiento.

De poner de relieve algunas de las soluciones y dificultades dadas a dichas realidades versa este artículo, teniendo como referencia lo que sucede en un ámbito de acción como es el jurídico, en el que tienen lugar experiencias desde hace algunos años. La opción se toma porque si hay un terreno de claro éxito en relación a la creación, desarrollo e innovación de sistemas de recuperación de documentación, este es el jurídico. De hecho, los sistemas de recuperación de documentación jurídica surgieron a partir de los años ochenta del siglo XX, prácticamente desde el comienzo del desarrollo de aplicaciones informáticas en forma de bases de datos (Bing, 1978). El éxito es

tal que, en la actualidad, en la mayor parte parte de los países democráticos, la información jurídica: textos legales y jurisprudenciales (sentencias judiciales), es accesible únicamente en formato digital. Este es el caso en España (2). En lo relativo a la tramitación de expedientes y la elaboración de sentencias por la Administración de Justicia en Juzgados y Tribunales, otra especie de sistemas de recuperación de documentación jurídica, ello ocurre, también en España, desde los años ochenta del siglo pasado (3).

Es por lo dicho que aquí nos centramos en lo siguiente. En el próximo apartado (2) reseñamos, brevemente, las características y exigencias generales del conocimiento tal y como propone, desde hace tiempo, la “biología de la cognición”. En el siguiente apartado, (3), nos ocupamos de explicar las necesidades concretas de las actividades jurídicas que precisan acceder y usar documentación por la obligación existente en el Estado de Derecho de “motivar” o “justificar” todo tipo de decisión jurídica. En el siguiente apartado, (4), describimos algunas de las características fundamentales, y sus limitaciones, de los sistemas de recuperación de documentación jurídica en funcionamiento y desarrollo. A continuación, apartado (5), expondremos las lecciones aprendidas de la experiencia jurídica, y su posible aplicación al desarrollo y uso de los sistemas de organización del conocimiento en ámbitos generales. Al final (6) se establece una conclusión.

2. Biología de la cognición

Si hablamos de “sistemas de organización del conocimiento” es importante tener siempre presente, en cualquier reflexión hecha al respecto, las características del conocimiento humano que han sido puestas de relieve desde los años ochenta del siglo XX por el estudio denominado “biología de la cognición”.

Según dice este estudio el conocimiento debe considerarse, atendiendo a las características del funcionamiento del cerebro, en su particularidad / dependencia / interrelación con el contexto de vida de cada ser humano. Afirmación que está en contra de las hipótesis modernas, establecidas desde el Renacimiento y la Ilustración, que igualaban las características racionales del conocimiento en todos los seres humanos, proponiéndoles a éstos como método general de estudio y ordenación de la realidad el uso de las categorías establecidas por la ciencia.

Así es, sin ir más atrás, Maturana y Varela hablaban, ya en 1984, de que el conocimiento de los hombres y de los animales es “autopoietico”, es decir, se va haciendo propio de cada hombre en relación a su historia y contexto (Maturana, 2003):

Los seres vivos no son únicos ni en su determinación, ni en su acoplamiento estructural, sin embargo. Lo que les es propio es que en ellos la determinación y el acoplamiento estructural se realizan en el marco de la continua conservación de la autopoiesis que los define, sea ésta de primero o segundo orden, y el que todo en ellos quede subordinado a esta conservación. Así, incluso la autopoiesis de las células que componen un metacelular se subordinan a la autopoiesis de éste como sistema autopoietico de segundo orden. Por lo tanto, todo cambio estructural ocurre en un ser vivo necesariamente acotado por la conservación de su autopoiesis, y serán perturbaciones aquellas interacciones que gatillen en él cambios estructurales compatibles con dicha conservación, e interacciones destructivas las que no. El continuo cambio estructural de los seres vivos con conservación de su autopoiesis está ocurriendo a cada instante, continuamente, de muchas maneras simultáneamente. Es el palpitir de toda la vida.

Aportaciones que, puede comprobarse, están respaldadas y aumentadas por obras más recientes (Auletta, 2011).

Todo lo cual implica aceptar que, en todos los ámbitos, en relación al conocimiento humano, los sistemas que proporcionen documentación o información han de estar contruidos de tal forma que sean capaces de satisfacer la pluralidad y particularidad del conocimiento concreto requerido por aquellos que demanden información a los sistemas. Actitud que, ha de destacarse, es muy diferente a la que promueven los sistemas bibliotecarios, destinados a proporcionar acceso

a documentación general, en especial libros, a los ciudadanos (4).

3. La motivación

Si hay un ámbito de trabajo (conocimiento y acción) en el que hace ya mucho tiempo es preciso atender a particularidades es el jurídico. Ello es así porque en este ámbito siempre es preciso resolver problemas concretos, que suceden entre personas, atendiendo a textos aportados por los particulares y por el sistema jurídico (normas e instituciones). Lo anterior supone que se ha de contar para solventar cada caso o problema, indispensablemente, con conocimiento establecido, desde diferentes perspectivas, sobre documentación de carácter jurídico (leyes, sentencias judiciales, estándares, códigos de práctica...). Ello es así por las exigencias democráticas propias del Estado de Derecho, que se demandan de cualquier actividad profesional de carácter jurídico, tal y como expresamos en este apartado.

En ello coinciden: la consideración de cuál es la función propia de las soluciones jurídicas (3.1) y el mismo concepto de Derecho (3.2), reflexiones que ponen en realidad la exigencia constitucional de la motivación de cualquier propuesta jurídica (3.3). El apartado lo cierra una conclusión (3.4).

3.1. La función de las soluciones jurídicas

En la actualidad se viene a coincidir en que, tal y como establecen las normas y principios jurídicos democráticos, la resolución de los casos, cometido propio del ordenamiento e instituciones jurídicas, tan sólo puede estar regida por el logro de la justicia / equidad, o, lo que es lo mismo, por el alcance del acuerdo consensuado sobre las resoluciones jurídicas, basado en el texto de las leyes, los contenidos de los principios jurídicos y la aceptación sobre las circunstancias del caso de todos aquellos que participan en la resolución del conflicto: juez o jueces, abogados, testigos, peritos y ciudadanos que sean responsables y protagonistas de la actividad desarrollada (Vesting, 2015; Robles, 2015).

3.2. El concepto de Derecho

Lo anterior es la razón en la que se basa una definición del Derecho, que resume otras ampliamente aceptadas, propuesta a comienzos de los años noventa del siglo pasado y que sigue en vigor. Es una definición que expresa la complejidad que tiene el fenómeno jurídico en los sistemas democráticos. La definición dice así (Alexy, 1992):

El Derecho es un sistema normativo, (1) llamado a la justicia, (2) integrado por la totalidad de las normas que pertenecen a una Constitución, reconocida

socialmente, y no son absolutamente injustas, así como por la totalidad de las normas que han sido dictadas en virtud de esa Constitución, muestran un mínimo de realidad social o posibilidad de realidad social y no son absolutamente injustas, y por (3) los principios y los argumentos normativos especiales en los que se fundamenta y/o debe fundamentarse el procedimiento de aplicación jurídica para satisfacer las exigencias de justicia.

3.3. La motivación

En las aproximaciones y concepto expresados en los sub-apartados 3.1 y 3.2 se da respuesta concreta a la exigencia básica del Estado de Derecho de la necesidad de la motivación o, lo que es lo mismo, de la fundamentación basada en textos de la puesta en acción de las actividades jurídicas profesionales (5). O lo que es lo mismo de la necesidad de acceder a textos jurídicos como actividad común y previa a la de cualquier otra actividad jurídica. Dichos textos son: códigos, constituciones, reglamentos, sentencias judiciales..., y el resto de documentos que son de uso preciso para cualquier actividad jurídica. Sea esta actividad cual sea: la aplicación del Derecho, la interpretación del Derecho, la creación de nuevas normas o la creación de doctrinas jurídicas por la "ciencia" del Derecho.

La motivación, entendida como obligación de acceder a textos jurídicos, es una de las características fundamentales que diferencia al régimen jurídico liberal del propio del Antiguo Régimen. La verdad es que en este último no era fácil el hallazgo de la Ley vigente, e incluso, en muchas ocasiones, no existía: la solución procedía de la voluntad arbitraria del Monarca, sus gobernantes o los jueces (6). De ahí la necesidad de cambio en forma de generalización de la motivación. Cambio jurídico liberal que se concreta en el establecimiento por los representantes de los ciudadanos de las Constituciones y Códigos que contengan claramente los textos jurídicos y los mecanismos, procedimientos e instituciones, precisas para crear, aplicar y resolver en forma justificada, democráticamente, los conflictos surgidos con su puesta en acción.

El problema se hizo complejo cuando resulta, como en la actualidad sucede, que en la práctica existen numerosas fuentes del Derecho y textos jurídicos de referencia obligatoria como los originados, entre otras razones, por: el establecimiento de Estados federales que cuentan en la Federación y en los Estados miembros con diferentes organismos capaces de emitir normas, aplicarlas y juzgar casos en su ámbito de competencia; la necesidad de juzgar asuntos que han estado regidos por leyes procedentes del Antiguo

Régimen; la aparición de normas que son promulgadas en el ámbito de acción propios del desarrollo industrial o tecnológico; la existencia de comportamientos regidos por reglas aprobadas por la práctica, usos, costumbres o creencias sociales del correspondiente grupo... (Galindo, 1993). De ahí que el acceso a textos jurídicos sea una necesidad apremiante.

3.4. Conclusión

Por todo lo hasta aquí dicho no es de extrañar que, cuando hablamos de sistemas de conocimiento en el ámbito jurídico, sea conveniente exponer los rasgos más característicos, mejor las funcionalidades, de los sistemas de acceso a documentación jurídica que auxilian a la realización de actividades jurídicas utilizando herramientas informáticas.

4. Los sistemas de recuperación de documentación jurídica

4.1. Introducción

En virtud de lo expresado no es de extrañar que desde los años ochenta del siglo pasado, en cuanto las Tecnologías de la Información y la Comunicación (TICs) pudieron proporcionar acceso a una numerosa y gran heterogeneidad de datos (numéricos y alfabéticos), se desarrollaran bases de datos, sistemas de recuperación de documentación jurídica, sistemas de tratamiento de textos, aplicaciones... que facilitaban el acceso a textos jurídicos con el fin de auxiliar a resolver casos o conflictos sociales concretos (7).

Una vez que estos sistemas tienen como objetivo proporcionar con suficientes garantías de fiabilidad, con respecto a las fuentes utilizadas, los textos precisos para poder actuar profesionalmente en relación a cualquier materia, la posibilidad de acceder a los mismos por los diferentes agentes jurídicos, e incluso por los ciudadanos, dichos sistemas pueden ser considerados instrumentos indispensables para acceder a textos jurídicos y por tanto para realizar cualquier actividad profesional jurídica. O lo que es lo mismo: cabe predicar de ellos la función de satisfacer los principios del ordenamiento a los que hacíamos referencia en el anterior apartado 3. En definitiva: pueden proporcionar varias hipotéticas fundamentaciones a las actividades profesionales de las juristas realizadas en su labor de resolver conflictos (8).

En el presente apartado vamos a referirnos a la evolución de las aplicaciones informáticas diferenciando dos ámbitos: el acceso a textos jurídicos (4.2), y la elaboración de sentencias (4.3). Finalmente establecemos una conclusión (4.4).

4.2. Acceso a textos jurídicos

Sin duda en la actualidad existen aplicaciones o sistemas informáticos en uso y desarrollo que mejoran el acceso a textos jurídicos utilizando recursos que permiten realizar un extenso aprovechamiento y estudio de las fuentes jurídicas gracias a las posibilidades que proporcionan la innovación y potencia de los programas y las TICs. Una de estas herramientas es la denominada “Jurimetría”, comercializada en España por la empresa Wolters Kluwer (<https://jurimetria.laleynext.es/>, 2020-04-21).

Esta herramienta, proporciona acceso, como respuesta al esquema de un caso presentado al programa en forma de “palabras clave” o textos significativos que lo caractericen, a textos normativos, jurisprudencia o comentarios doctrinales, que puedan ser utilizados como referencia para la motivación o fundamentación de la actividad jurídica concreta por la que se haga la consulta. El acceso a estos textos se produce por la comparación que el sistema, mediante el uso de diferentes recursos de programación, establece entre los términos de la cuestión y los que se contienen o caracterizan a los textos que proporciona como respuesta. Las materias contenido de los textos jurídicos son: Civil, Social y Contencioso administrativa, próximamente lo será también la de contenido Penal. Se hace un estudio de estas funcionalidades comparando las que ofrecen otros sistemas en Galindo (2019a).

La novedad de la respuesta facilitada por el programa “Jurimetría” reside en que, además de lo expresado, puede facilitar como referencia sentencias sobre la materia consultada en las que hayan participado el juez o magistrado que, previsiblemente, ha de juzgar el caso por el que se realiza la consulta, o los abogados que también lo hayan hecho. También puede indicar empresas o Administraciones públicas concretas que hayan podido ser parte en casos similares.

Estas respuestas son suministradas porque el sistema cuenta no sólo con las palabras que integran la información textual contenida en los argumentos de las sentencias, incluyendo el sentido de las resoluciones (si son estimatorias o no de las peticiones de las partes, la posición adoptada en las mismas sobre una u otra interpretación de las leyes...), sino también con los nombres de los jueces o magistrados, los abogados, las empresas y las instituciones públicas que hayan participado en los procesos cuyo resultado, decisión o sentencia, es contenido en la base de datos (9).

De la fiabilidad del contenido de la documentación tratada da cuenta el hecho de que la em-

presa está obligada a obtenerlo del Consejo General del Poder Judicial, órgano de gobierno del poder judicial español, a través del cauce establecido al respecto: su Centro de Documentación Jurídica (CENDOJ).

El programa también puede facilitar datos sobre la previsible duración media de los procesos e incluso el grado de congestión o de admisibilidad de los recursos en el Juzgado o Tribunal que pueda encargarse de la resolución del caso por el que se realiza la consulta. Igualmente cabe ofrecer una historia judicial de los asuntos de carácter similar, organizados por materias, en cuya resolución haya participado un juez o magistrado concreto a lo largo de su carrera profesional. El contenido de esta información es el recogido en las Estadísticas Judiciales que publica el Consejo General del Poder Judicial. El programa realiza una combinación de los datos de las Estadísticas y la información y todos los datos de las sentencias, información que es proporcionada en todo caso por el Consejo.

Con la descripción de estas funcionalidades observamos que el autor de la consulta puede recopilar no sólo información y documentación, sino también contar con un mayor número de perspectivas o estrategias que las que proporcionan los sistemas de recuperación documental tradicionales, limitados a contrastar los argumentos característicos de un caso con los de otros casos, más o menos similares al primero, ya resueltos.

Con lo cual cabe decir que el jurista usuario se encuentra con un instrumento que, al contrario de lo que sucede con otras aplicaciones o programas, potencia su libertad de trabajo como profesional, y la posibilidad de contrastar informaciones a las que con anterioridad tenía un difícil acceso, mejorando con ello la calidad jurídica de su labor.

El programa, a la vez, cuenta con seguridad con respecto al contenido de la información ofrecida: como hemos dicho el sistema, que es responsabilidad de una empresa, da acceso a documentación oficial, sistematizada y recopilada por el Consejo General del Poder Judicial, y a datos estadísticos generados desde 1995 por los propios sistemas de información utilizados en la Administración de Justicia, que están supervisados por el mismo Consejo General del Poder Judicial y, en lo que le corresponde en lo relativo a programas e infraestructura, el Ministerio de Justicia (9).

4.3. Elaboración de sentencias

Objetivo distinto al de los sistemas mencionados en el anterior sub-apartado (4.2) es el de los sistemas de recuperación de documentación que

proporcionan acceso a modelos de tramitación procesal, y, por lo mismo, auxilian a la elaboración de expedientes judiciales satisfaciendo las correspondientes normas. Todo lo cual no impide recordar que su objetivo final sigue siendo la resolución de conflictos concretos mediante sentencias.

Muy distinto a lo expresado es lo que sucede cuando el uso de las tecnologías quiere modificar las mismas características de la actividad jurídica, al ser ésta sustituida por un programa o un sistema y sus fines y características propias. Este es el caso cuando se habla de aplicación automática del Derecho, o, en general, de decisiones jurídicas automáticas / automatizadas / codificadas, incluyendo las decisiones realizadas en la Administración de Justicia con el fin de resolver un caso o problema.

A continuación, nos ocupamos de estas materias distinguiendo entre la elaboración judicial de sentencias (4.3.1) y la elaboración automática de sentencias (4.3.2).

4.3.1. *Elaboración judicial de sentencias*

Aquí vamos a fijarnos en dos experiencias que muestran el alcance que tiene el uso de sistemas de recuperación de documentos en la Administración de justicia: las que dan alguna muestra de lo que ocurre en la Unión Europea a través del portal “e-justice” (4.3.1.1), y la reforma española de la oficina judicial que, a la vez, mantiene las exigencias democráticas y acomoda su organización y distribución del trabajo a las exigencias de una Administración de Justicia que utiliza las TICs en los procesos de elaboración de sentencias (4.3.1.2).

4.3.1.1. *E-justice y la transmisión digital de documentos judiciales*

La Unión Europea hace una presentación de los diferentes sistemas judiciales existentes en los países que la integran utilizando un sistema informático: la página Web o portal “e-justice”. La página actual es la versión preliminar de la que será la definitiva: <https://beta.e-justice.europa.eu/home> (2020-04-07). En este trabajo se va a considerar el contenido de esta última página.

El sistema es una recopilación documental que contiene rica y significativa información sobre cómo está constituida y funciona la Administración de Justicia en los diferentes países que integran la Unión así como en la propia Unión Europea considerada como organización política. Información que es suministrada y actualizada por la Comisión Europea, los respectivos Ministerios de Justicia de los países miembros y varias organizaciones jurídicas.

El portal no contiene únicamente información sobre “materias jurídicas” de carácter dogmático tradicional (Civil, Penal, Social...), sino que contiene información sobre materias más concretas: “Derecho de familia y sucesorio”, “Créditos pecuniarios”, “Procedimientos judiciales”, “Emprender acciones judiciales”, “Sus derechos”, “Legislación y jurisprudencia”, “Registros mercantiles de insolvencia y de la propiedad”, “Encontrar un profesional del Derecho” y “Formación, agencias y redes judiciales”. Cada materia está subdividida en las correspondientes submaterias.

A nuestros efectos: presentar algunos datos característicos de los sistemas de organización del conocimiento en el ámbito judicial, interesa destacar que, con relación a los procesos de auxilio a la elaboración de las decisiones judiciales, en la materia “Procedimientos judiciales”, sub-materia Asuntos civiles, se encuentra información sobre: “Cooperación judicial en materia civil entre los países de la UE con el fin de mejorar la interoperabilidad entre sus sistemas judiciales.”

Esta información se subdivide en:

1. Notificación de documentos: transmisión oficial de documentos jurídicos. La página informa a quien la accede sobre lo siguiente: “Si participa usted en un proceso judicial y necesita enviar o recibir documentos judiciales o extrajudiciales, en esta página puede encontrar información nacional sobre la manera de proceder”.
2. Tratamiento electrónico de asuntos y comunicación electrónica con los tribunales. La página proporciona a quien la accede: “Información sobre la situación actual de las notificaciones por vía electrónica a los órganos jurisdiccionales nacionales y las efectuadas por estos”.

Como puede observarse las expresiones y aclaraciones indican que las respectivas secciones de la página Web se limitan a documentar el estado de la realización de comunicaciones electrónicas con tribunales y juzgados, y las de estos entre sí y con respecto a juristas y particulares.

Obviamente en los lugares (páginas Web) correspondientes se pone de relieve que existen numerosas diferencias entre el estado de la cuestión existente en unos y otros países, y su aceptación o no depende de numerosos factores. De todas formas, cabe observar que ya el enunciado de los asuntos tratados da cuenta de la tendencia existente hacia la informatización de procedimientos judiciales, e incluso la articulación del denominado “expediente judicial electrónico”. Ello se manifiesta inequívocamente en el hecho de que el portal asuma como categorías del mismo que es

posible comunicarse con los sistemas de tratamiento de documentación en formato electrónico tanto en lo referido a enviar documentos a los mismos como en lo relativo a recibirlos.

4.3.1.2 La Oficial judicial: nueva organización respetando el ordenamiento y las exigencias de las TICs

El portal también da cuenta de otra tendencia existente contraria a la expresión “decisiones judiciales automatizadas”. Ello se hace en lo que respecta a la organización de la Administración de Justicia en España, una vez que la correspondiente página Web del portal “e-justice” da cuenta de que en este país se ha producido una reforma de la organización de la Administración de Justicia realizada con una doble finalidad: 1) la acomodación de la misma a las exigencias propias de las TICs y la implantación, con ello, por ejemplo, de un “expediente judicial electrónico”, y 2) la satisfacción de los principios de acción propios de las decisiones judiciales en un estado democrático, es decir que ellas sean realizadas por jueces en un contexto concreto con el fin de poner en realidad la justicia / equidad.

La reforma consiste en que, fijándonos en la organización actual de la Oficina judicial en España, esta ya no es, como era lo propio de Juzgados y Tribunales tradicionalmente: una oficina formada por un juez, un secretario y un determinado número de funcionarios que trabajan de modo dependiente en torno a la figura del Juez, sino una nueva organización que establece distintos sistemas de trabajo racionales y homogéneos, con el fin de procurar que la actividad judicial se desempeñe con la máxima eficacia a la vez que con un adecuado reparto de responsabilidades. En concreto la nueva organización de Oficina Judicial está conformada por tres tipos de unidades procesales diferentes, atendiendo a su funcionalidad:

1. Las Unidades Procesales de Apoyo Directo, que asisten a jueces y magistrados en el ejercicio de las funciones que les son propias, realizando las actuaciones necesarias para el exacto y eficaz cumplimiento de cuantas resoluciones de justicia / equidad dicten jueces y magistrados.
2. Los Servicios Comunes Procesales que, bajo la dirección de un letrado de la Administración de Justicia, asumen labores centralizadas de gestión y apoyo en actuaciones derivadas de la aplicación de las leyes procesales: el “expediente judicial electrónico”.
3. Las Unidades Administrativas que, sin estar integradas en la Oficina Judicial, dirigen, ordenan y gestionan los recursos humanos, los medios informáticos y los medios materiales.

Todo ello está regulado en la Ley Orgánica 19/2003, de 23 de diciembre, de modificación de la Ley Orgánica 6/1985, de 1 de julio, del Poder Judicial, que define a la Oficina Judicial como “la organización de carácter instrumental que sirve de soporte y apoyo a la actividad jurisdiccional de jueces y tribunales”. Con lo cual el nuevo modelo organizativo que introduce la Oficina Judicial diferencia claramente, lo que no hacía el modelo antiguo, los tres tipos de actividad que se realizan en el ámbito de la Administración de Justicia: la jurisdiccional, la elaboración de la sentencia, que recae en jueces y magistrados; la procedimental, que corresponde a los letrados de la Administración de Justicia (antiguos Secretarios de Justicia) y a los servicios de apoyo y procesales; y la administrativa, incluida la relacionada con el uso de las TICs, que recae en el Ministerio de Justicia o en las Comunidades Autónomas con competencias asumidas.

De ahí quepa decir que esta regulación moderniza la Administración de Justicia atendiendo a la división del trabajo y a la satisfacción de la regla constitucional del principio de separación de poderes, al establecer una clara separación entre la facultad judicial de juzgar y la administrativa, propia del poder ejecutivo que está centrada en dotar de infraestructura suficiente al ejercicio de la función jurisdiccional. Esta última función se mantiene independiente como prescriben las normas, constitución y principios jurídicos que están incluidos en las anteriores. Por ello en modo alguno la regulación plantea la decisión judicial automática aun cuando si establece un plan de acción para el encaje de las innovaciones tecnológicas que se implanten en Juzgados y Tribunales en diferentes funciones administrativas.

4.3.2. Elaboración automática de sentencias

Cabe hacer referencia aquí a un reciente desarrollo: la “elaboración automática de sentencias judiciales”, que se menciona como supuesto de investigación y como acción efectivamente realizada o en fase de realización en el futuro, gracias al auxilio de las técnicas de inteligencia artificial.

Ello sucede especialmente en Estados Unidos, país de derecho de “common law” y en el que tiene una gran relevancia la aplicación del desarrollo tecnológico a muy distintas actividades. Es de interés a estos efectos recoger aquí, en forma resumida, el estado de la cuestión, que responde a varias experiencias e iniciativas que tienen lugar en dicho país, y que queda recogido en el trabajo de Re&Solow-Niederman (2019).

En este sub-apartado nos referimos a lo siguiente: a la caracterización que se viene a hacer

de la justicia de inteligencia artificial, también llamada codificada (4.3.2.1), a la relevancia de la regla del precedente en los ensayos de justicia codificada (4.3.2.2), y a los límites que presenta esta justicia de inteligencia artificial (4.3.2.3).

4.3.2.1. *La justicia de inteligencia artificial o codificada*

En el trabajo mencionado (Ibidem, 253-254), se define a la justicia de inteligencia artificial o codificada, confrontándola a la “justicia de equidad”, de la siguiente forma:

La justicia codificada se refiere a la aplicación rutinizada de procedimientos normalizados a un conjunto de hechos...En general, la justicia codificada... aspira a establecer por adelantado el conjunto total de variables jurídicamente pertinentes, descontando otros hechos y circunstancias que se pueden detectar en procedimientos individualizados. El objetivo básico de esa normalización es reducir el espacio para la discreción humana en la adjudicación, disminuyendo así las oportunidades de arbitrariedad, parcialidad y despilfarro, aumentando al mismo tiempo la eficiencia, la coherencia y la transparencia. En resumen, la justicia codificada ve los vicios de la discreción, mientras que la justicia equitativa ve sus virtudes. En consecuencia, la justicia codificada tiende a eliminar la necesidad de cualquier explicación, restricción o legitimación, aparte de la adhesión a los propios procedimientos normalizados. En otras palabras, el poder y la autoridad del juez son emitidos como no discrecionales y derivados de cualquier entidad una vez creado el algoritmo legal relevante.

Esta definición da muestra de las limitaciones de este tipo de justicia de inteligencia artificial, justamente al indicarse en la comparación recogida en el texto que es insuficiente para la satisfacción de la justicia / equidad que es lo que caracteriza en el Estado de Derecho a la elaboración de sentencias judiciales. Recuérdese que, como expresábamos más arriba (3.1), en la actualidad se viene a coincidir en que, tal y como establecen las normas y principios jurídicos democráticos, la resolución de los casos, cometido propio del ordenamiento e instituciones jurídicas, tan sólo puede estar regida por el logro de la justicia o, lo que es lo mismo, por el alcance del acuerdo consensuado sobre las resoluciones jurídicas, basado en el texto de las leyes, los códigos, los contenidos de los principios jurídicos y la aceptación sobre las circunstancias del caso de todos aquellos que participan en la resolución del conflicto: juez o jueces, abogados, testigos, peritos y ciudadanos que sean responsables e incluso parte de la actividad desarrollada.

Estas funciones son aceptadas, genéricamente, tanto en los países del ámbito del derecho continental como en los de “common law”, pese a lo

cual está surgiendo la justicia de inteligencia artificial que, acabamos de ver, está caracterizada especialmente, en cambio, por la puesta en práctica de algo más limitado: la promoción de la codificación. Parecería, por ello, que la justicia de la inteligencia artificial tuviera como objetivo hacer realidad la idea de codificación presente en los países de common law como señalamos a continuación.

Recordemos que la codificación del derecho es una referencia que se mantiene como regla practicada en los países de derecho continental desde las revoluciones liberales del siglo XVIII: son las leyes en forma de códigos y constitución las que han de ser puestas en práctica como referencia o motivación en las sentencias judiciales de dichos países. Esto es distinto a lo que sucede en los países de common law en los que nunca se ha contado con un derecho codificado y si, en cambio, desde la Edad Media, con la puesta en práctica de la “regla del precedente” cuyo contenido “codificador”, al que se refería la definición sobre inteligencia artificial recogida en este apartado, resumimos a continuación.

4.3.2.2. *La relevancia de la regla del precedente*

La relevancia queda referida a que en los países de common law existe la regla denominada “regla del precedente”, que indica que un precedente o autoridad es una regla o un principio establecido en un anterior caso legal que es vinculante o persuasivo para una Corte u otro órgano judicial para decidir en casos posteriores con similares cuestiones o hechos (Cross&Harris, 2012). Regla que está presente directamente en la definición de justicia codificada o de inteligencia artificial que se mencionaba con anterioridad.

Como hemos dicho esta regla del precedente no es propia del derecho continental, una vez que en estos países los casos: 1) han de estar motivados por la ley únicamente, y, además, 2) ser resueltos por la justicia / equidad / el consenso. Ello es así porque la puesta en acción de la codificación desde el siglo XVIII ha puesto de relieve sus limitaciones y la necesidad de acudir a otras referencias complementarias. En todo caso en este tipo de derecho continental las resoluciones de casos anteriores son, siempre, sólo referencias no obligatorias, la referencia obligatoria es el texto de la ley.

4.3.2.3. *Límites de la justicia de inteligencia artificial*

A lo hasta aquí expresado ha de añadirse que existen fuertes limitaciones en las experiencias hechas sobre la justicia de inteligencia artificial justamente por el contenido de su definición que

resumíamos con anterioridad (4.3.2.1). Los límites o críticas básicas surgidas a partir de las experiencias y, además, hechas desde la consideración de los objetivos propios de las sentencias judiciales en el Estado de Derecho que referíamos más arriba (3.1) están reducidos a las siguientes consecuencias de las sentencias de inteligencia artificial. En concreto su generación de: “incomprensión”, “datificación”, “desilusión” y “alienación”. A continuación exponemos algunos de dichos límites (ver: Galindo, 2019b, en posiciones elaboradas a partir de las experiencias expuestas en Re&Solow-Niederman, 2019).

Incomprensión, porque las sentencias pueden “funcionar de maneras que son difíciles o imposibles de comprender para los seres humanos”, dado el método de inteligencia artificial utilizado predominantemente: el aprendizaje de máquina. Este método implica el uso de la técnica del aprendizaje profundo que carece de razonamiento lógico explícito o inferencias causales, que son las que marcan las explicaciones humanas convencionales.

Datificación, que consiste en la necesidad de que las sentencias precedentes tengan la capacidad de producir datos que puedan ser leídos por un ordenador y, por tanto, sean expresadas en forma numérica. Esto significa que los procesos judiciales automáticos están / estarán regidos por las reglas de la codificación numérica y no por el contenido y expresión de los textos legales. Es decir, por las reglas matemáticas y no por las normas aprobadas por los ciudadanos y sus representantes, en forma muy diferente a lo que prescriben los ordenamientos democráticos que predicán de algo que es justo “si es legalmente válido”, o como decíamos más arriba: si está motivado.

Desilusión. Estas propuestas muestran que las sentencias de IA pueden generar desilusión con respecto al sistema judicial tradicional al mostrar, aparentemente, el sistema de IA sus contradicciones y sesgos. Ello podría ser positivo, sólo que la verdad hasta ahora es que algo ocurre en sentido contrario con respecto a experiencias habidas con la justicia de IA. Esto es así por el hecho de que lo que hasta ahora se ha mostrado es, justamente, desilusión con respecto al funcionamiento de los sistemas de IA en el ámbito de la justicia. Lo muestran estudios que prueban que el “sesgo” está presente en sistemas de conocimiento que toman como absolutas, ciertas y positivas resoluciones sobre casos, precedentes, que estaban basados en visiones propias de sociedades en las que los valores dominantes estaban impregnados de prejuicios con respecto a actividades de personas (personas de raza negra, gitanos...) caracterizadas con notas distintas

a las de la sociedad de “blancos”, castas o grupos dirigentes (Belloso, 2019).

Alienación. Por el alejamiento de la deliberación y la discusión participada por los ciudadanos y los agentes jurídicos que acarrea la justicia de inteligencia artificial, que es contraria a la puesta en acción de los principios de participación propios del Estado de Derecho, se mire por donde se mire.

4.4. Conclusión

Los ejemplos han servido para mostrar que es posible estudiar y poner en realidad actividades jurídicas como la del acceso a textos jurídicos o la elaboración de sentencias a partir de sus particularidades propias, y que ello lo es aun cuando elementos que forman parte de la misma actividad, como son los programas de ordenador destinados a la recuperación de textos jurídicos, son contruados atendiendo a las características técnicas que el programa de referencia menciona.

O lo que es lo mismo: la consideración y uso de productos técnicos no afectan a las funciones jurídicas propias de la actividad concreta para con determinado caso que, como prescribe el ordenamiento, ha de ser particularizada por la puesta en acción de la “justicia” y el resto de los valores propios de los ordenamientos democráticos. Otra cosa es el uso de productos técnicos que disminuyen o eliminan la satisfacción de los principios propios del Estado de Derecho, estos productos o no pueden utilizarse o han de ser acomodados a esta satisfacción.

5. Los sistemas de organización del conocimiento

5.1. Introducción

Ha sido visto hasta aquí que la aceptación de la función y evolución de los sistemas de organización del conocimiento en el ámbito jurídico (expuestas a través de varios ejemplos supra en el apartado 4), depende directamente de su capacidad de potenciar la satisfacción de las necesidades de sus usuarios, en concreto los juristas, en el momento de realizar las actividades que les son precisas para resolver casos o conflictos sociales reales, teniendo como horizonte el logro de la idea de justicia propia de sociedades democráticas (explicada supra en el apartado 3). Todo ello, a la vez que son herramientas adecuadas para cualquiera que las utilice, sea el que sea el respectivo horizonte / interés de conocimiento del jurista usuario, y la forma o el procedimiento en que haya sido adquirido (en conformidad con lo establecido supra en el apartado 2). Es decir,

siempre que los sistemas faciliten optar libremente por los materiales justificativos (la documentación) de una u otra composición de la posición profesional que el jurista usuario precise por que quiera adoptarlos en relación a la propuesta de solución de un concreto caso.

En este momento ya cabe preguntarse y discurrir sobre lo siguiente: el establecimiento y práctica de los sistemas de organización del conocimiento en el ámbito jurídico reseñados... ¿pueden extenderse a otros ámbitos? En este apartado respondemos a la cuestión tomando en consideración lo que se expresa en los siguientes subapartados: 1) la situación de los sistemas de organización del conocimiento utilizados en la puesta en acción del “gobierno electrónico” o la “administración electrónica” (5.2); 2) algunas de las características generales de algunos sistemas de organización del conocimiento desarrollados en otros ámbitos (5.3); 3) las exigencias básicas para el establecimiento de otros sistemas de organización del conocimiento atendiendo a lo desarrollado, sustancialmente, en el ámbito jurídico (5.4), y 4) la conclusión (5.5).

5.2. Sistemas de conocimiento de uso en el gobierno electrónico y la administración electrónica

Lo expuesto en apartados anteriores con respecto a los sistemas de organización de la información de carácter jurídico, no está referido únicamente a la elaboración de sentencias judiciales o a la actuación genérica de los juristas como profesionales del derecho. Si aquí nos hemos limitado al estudio de los sistemas de organización del conocimiento en relación a dichos ámbitos, ello es porque el procedimiento es un adecuado objeto común de referencia para caracterizar las actividades y el funcionamiento de todas las instituciones jurídicas a las que incumbe el concepto de Derecho citado con anterioridad (supra 3.1 y 3.2). Instituciones implicadas por el concepto que tienen distinto ámbito de actuación: el del poder legislativo, el del ejecutivo, el de cualquiera de los que forman parte del ámbito judicial e, incluso, el de las soluciones extrajudiciales (Kyritsis, 2015).

Aquí no hemos entrado en el estudio de estas concreciones porque ello requeriría ampliar innecesariamente los límites del trabajo, lo que no impide reconocer que su funcionamiento en los ámbitos jurídicos reseñados requiere satisfacer las exigencias cumplidas por los sistemas documentales antes mencionados. Por tanto, aquí nos limitamos a decir que a los sistemas de organización del conocimiento que se utilizan en cualquier institución jurídica, a los que también se denomina genéricamente “Gobierno electrónico” o,

mejor, “Administración electrónica”, les son de aplicación las consideraciones establecidas en los anteriores apartados y sub-apartados de este artículo. Ello incluye a los sistemas de organización del conocimiento en relación a archivos, registros, bases de datos personales, expedientes administrativos y protocolos, por ejemplo.

Es por ello que aquí nos vamos a referir a otros sistemas de organización del conocimiento que no tienen por objeto directo o indirecto resolver casos conflictivos, sino proporcionar información sobre determinada actividad, cuestión u ocurrencia. Estamos hablando, en concreto, de información detallada sobre una materia que una persona espera encontrar en una biblioteca u otro sistema de información general: un museo, un libro, una página Web... De ello nos ocupamos en el siguiente sub-apartado.

5.3. Otros ámbitos

Como es obvio en otros ámbitos de conocimiento no cabe tomar como referencia el contenido que han de satisfacer los sistemas de carácter jurídico. En esos otros ámbitos ha de atenderse, por lo general, a solicitudes que se realicen demandando información sobre materias concretas que no estén orientadas por el encuentro de soluciones a conflictos determinados. Observando este tipo de sistemas comprobamos que, en ocasiones, han tenido una evolución similar a la indicada en los de carácter jurídico, en otras no, como vamos a mostrar a continuación. Tomamos como referencia, a modo de ejemplo, sistemas que proporcionan acceso al contenido existente en: 1) museos, 2) bibliotecas, o, 3), información almacenada en Internet para cuyo hallazgo se utilice un sistema de organización del conocimiento como el que proporciona Google mediante su buscador.

5.3.1. Museos

El contenido y características de un sistema de organización del conocimiento que acceda a un museo tiene particular interés por el hecho de que la información que contiene, obras de arte, en el ejemplo concreto al que vamos a recurrir: cuadros o lienzos, tiene una gran relevancia el conocimiento del usuario. Es decir, cada interesado que acceda al contenido de la información tiene un conocimiento y sensibilidad diferente al de otros usuarios: son los que ha ido adquiriendo a lo largo de su particular experiencia de vida (10), como expresábamos supra en el apartado 2.

Esto es lo que ocurre en la organización del contenido del Museo del Prado que está presente en el sistema concreto que integra su página Web: <https://www.museodelprado.es/> (2020-04-20).

Esta página presenta el contenido de las colecciones que almacena en una forma clásica. Esto es: compilándolas atendiendo a las clasificaciones de las mismas que hacen manuales y expertos en arte (<https://www.museodelprado.es/coleccion>, 2020-04-20). También, y esto es lo que aquí nos interesa destacar, la página permite que cada visitante que lo desee realice una sistematización de los lienzos atendiendo a sus propios gustos, sistematizaciones que luego pueden ser seguidas por cualquier otro usuario que las quiera utilizar para realizar su acceso al contenido del museo (<https://www.museodelprado.es/mi-prado>, 2020-04-20). Es la página denominada “Mi Prado”.

Con ello constatamos que este sistema ofrece, como ocurría en el sistema jurídico “Jurimetría”, al que nos referíamos supra en el sub-apartado 4.2, la posibilidad de acceder a la información desde: 1) los procedimientos tradicionales: conceptos, palabras claves y textos, y 2) diferentes perspectivas adoptadas por diferentes tipos de usuarios con respecto al acceso a la información almacenada por el sistema.

5.3.2. Bibliotecas

Existen bibliotecas, sistemas de organización del conocimiento, que no acceden únicamente a fichas bibliográficas sino también a contenidos digitales: textos, imágenes, mapas, páginas Web... Se destacan dos: Europea (<https://www.europeana.eu/es>, 2020-04-20), creada por la Comisión Europea, y la Biblioteca digital mundial (World digital library: <https://www.wdl.org/> (2020-04-20), creada por UNESCO. En estas bibliotecas se puede acceder a la información por conceptos, palabras clave y frases. También por colecciones de los documentos almacenados. En Europea existe un mayor número de colecciones temáticas, la biblioteca incluye, también, colecciones orientadas a que sus usuarios sean profesores o estudiantes de diferentes niveles educativos.

La evolución de estos sistemas reside en su virtualidad de dar acceso a textos, imágenes y páginas Web. Las búsquedas pueden acotarse por subtemáticas o filtros en los que se encuadran la totalidad de los documentos, que los sistemas dan como respuesta al primer contenido de la cuestión o búsqueda de información hecha al sistema. El usuario puede elegir entre estas subtemáticas a efectos de encontrar una información más próxima a aquella por la que realice la consulta.

5.3.3. Información contenida en Internet

Un sistema de organización del conocimiento de carácter general es el constituido por el “motor”

de búsqueda de los diferentes sistemas de acceso a contenidos de Internet existentes en páginas Web. El más significativo por la extensión de su uso es el que se utiliza por la empresa Google. Aquí vamos a explicar las características básicas de su función.

Su particularidad consiste en proporcionar información contenida en Internet que sea respuesta a preguntas hechas por el usuario atendiendo tanto a la coincidencia del contenido de lo expresado en la respuesta con lo que este explicita en la cuestión concreta, palabras o frases, como, por ejemplo, a la situación geográfica en la que está situado el ordenador o teléfono móvil desde el que se hace la pregunta, a las respuestas que se han dado al usuario con relación a una misma o parecida cuestión hecha al buscador por el mismo usuario con anterioridad,... y otros factores contemplados por el conjunto de algoritmos de búsqueda de Google (<https://www.google.com/search/howsearchworks/> (18.4.2020).

Este sistema de búsqueda no es de uso obligatorio una vez que cada usuario puede disponer que el funcionamiento del mismo se realice en el mismo sentido que ha hecho en otras ocasiones o no, seleccionando la correspondiente configuración del programa que controla el sistema de búsqueda Google que utilice.

5.4. Exigencias básicas

Con lo aquí reseñado puede comprobarse que en ámbitos diferentes al jurídico los sistemas de organización del conocimiento también han experimentado diferentes evoluciones, acomodadas a lograr la satisfacción de lo que se persigue con dichos sistemas y la evolución alcanzada por las técnicas establecidas por quienes los diseñan y utilizan. La cuestión es: ¿podrían establecerse sistemas de organización del conocimiento de sistemas documentales no jurídicos que pudieran dar una satisfacción semejante a las cuestiones de sus usuarios como las que les dan los sistemas de organización del conocimiento de los sistemas de documentación de carácter jurídico a sus usuarios?

Se intenta dar respuesta a esta pregunta en el proyecto de investigación “Posibilidades y requisitos de los sistemas de organización del conocimiento para la interoperabilidad entre las instituciones de la memoria y el sector del turismo cultural”. El proyecto incluye entre sus objetivos dar respuesta a la pregunta mencionada (García-Marco; Galindo; Lasala; López del Ramo, 2020).

En concreto este proyecto, todavía en curso, quiere averiguar cómo organizar sistemas de acceso a documentación en relación a preguntas

hechas por usuarios determinados a materias que les interesen. Para ello se ha seleccionado como ejemplo concreto de estudio el acceso a sistemas de información que se ocupan de proporcionar información / ayuda a personas que, por uno u otro motivo, quieren realizar el Camino de Santiago desde un lugar concreto hasta la ciudad de Santiago de Compostela en Galicia.

Como puede observarse este no es un ejemplo jurídico: no se trata de resolver o dar solución a casos conflictivos. Se trata de proporcionar información variada, accesible por Internet, a libros, escritos, documentos, imágenes, mapas, sistemas de información geográfica, previsiones del tiempo, rutas, alojamientos, restaurantes, medios de locomoción, estados de las carreteras... a personas que, por una u otra razón, quieran realizar dicho Camino.

El ejemplo es similar al jurídico únicamente en el siguiente sentido: así como los sistemas de organización de la documentación jurídica están diseñados atendiendo al cometido profesional de los juristas, que está tasado en los países democráticos por las reglas establecidas para la solución de casos conflictivos, en el ejemplo se trataría de diseñar sistemas que proporcionaran información de interés, suficientemente adecuada y precisa, a aquellas personas que se propongan realizar o estén realizando dicho Camino de Santiago.

Son varias las exigencias a satisfacer: por un lado, en lo que respecta a la información a proporcionar por el sistema, información existente que es numerosa y accesible por Internet; por otro, los tipos de peregrinos que podrían estar interesados en la adquisición de tal información. Centrándonos en lo último: se hace preciso un estudio sociológico sobre el tipo de peregrinos e interesados en realizar el camino y los temas, intereses, valores o creencias, por los que cada uno de dichos tipos lo realiza y le concierne. Con estos datos podrán construirse sistemas más precisos que los que tienen un propósito general. Se propone que, a partir de modelos iniciales, se pueda construir con las experiencias de los usuarios sistemas modélicos de organización del conocimiento destinados a la recuperación de información.

5.5. Conclusión

En este momento ya cabe dar una respuesta a la pregunta con la que comenzábamos el presente apartado: el establecimiento y práctica de los sistemas de organización del conocimiento en el ámbito jurídico reseñados... ¿pueden extenderse a otros ámbitos? Existen señales de que la respuesta es positiva. Así lo indican los ejemplos estudiados que dan cuenta de:

1. Cómo un museo como el del Prado cuenta con sistemas de acceso que facilitan tanto la utilización de procedimientos de acceso clásicos como nuevos procedimientos (Mi Prado) que permiten elaborar y utilizar sistemas de acceso personales, basados en el propio conocimiento, y cuyo uso puede ser generalizado.
2. Las bibliotecas consideradas acceden a textos, imágenes, mapas, páginas Web... utilizando procedimientos clásicos (conceptos o palabras), que pueden introducir modificaciones que permiten concretar la búsqueda a cuestiones particulares.
3. Un motor de búsqueda como el de Google recupera información almacenada en Internet considerando los usos y gustos de quien realiza la búsqueda, a la vez que permite, alternativamente, la búsqueda al margen de prejuicios anteriores.
4. Existen proyectos de I+D+i que utilizan procedimientos estándar para estudiar aplicaciones concretas preparadas para grupos de usuarios típicos, como ocurre en los sistemas de documentación jurídica...

Por todo ello los ejemplos jurídicos pueden ser tomados como ejemplo para permitir el avance de los sistemas de organización del conocimiento desarrollados en otros ámbitos, en forma respetuosa con los objetivos y fines propios de estos últimos sistemas.

6. Conclusión final

El artículo ha revisado aspectos significativos de la función y su evolución que cumplen los sistemas de recuperación de documentación en el ámbito jurídico. Sistemas que cabe considerar están destinados a la organización del conocimiento en este ámbito. Ámbito que tiene un objeto limitado: dar soluciones justas, propias de países democráticos, a problemas concretos. Soluciones que han de ser plurales, una vez que la información que suministran para resolver los casos ha de ser utilizada por las diferentes posiciones que han de ocupar quienes participan en la resolución de los casos.

El grado de expansión e innovación de estos sistemas da cuenta de que pueden tomarse como ejemplo para la modernización de otros sistemas de organización del conocimiento, que, aun cuando tengan que satisfacer fines distintos a los jurídicos, también han de proporcionar información abierta a su disposición a todos los integrantes de la sociedad, sea la que sea su forma de aproximación al conocimiento.

Notas

- (1) Un estado de la cuestión sobre estos sistemas de organización del conocimiento se encuentra en: Suárez, 2017.
- (2) Esto sucede desde 2009 en relación al Boletín Oficial del Estado, que se publica solamente en formato digital. Las sentencias judiciales, especialmente las de órganos colegiados, son publicadas desde antes: 1997 en formato digital por el Consejo General del Poder Judicial, por su Centro de Documentación Judicial (CENDOJ). El CENDOJ es la única instancia que provee acceso a esta información a todo interesado: tanto por Internet al particular, que tenga interés por acceder a una u otra sentencia, como a la empresa que quiera tratar el contenido de la jurisprudencia y proporcionarla a sus clientes.
- (3) Existen referencias de los pasos dados desde 1983 (Generalitat de Catalunya, 1983). Con relación a lo que ocurre al respecto en la Unión Europea y los países que la integran informa el portal "e-justice", <https://beta.e-justice.europa.eu/home?plang=es&action=home> (2020-03-31).
- (4) "La misión de la Biblioteca Nacional de España, desde hace más de 300 años, es la de preservar y transmitir el patrimonio bibliográfico y documental español. Su voluntad es revertir en la sociedad el inmenso legado cultural que atesora, potenciando al máximo el acceso, uso y reutilización de la información que conserva y genera, en beneficio de la sociedad", <http://www.bne.es/es/LaBNE/Transparencia/index.html> (2020-04-01).
- (5) El art. 120.3 de la Constitución Española 1978 establece, sucintamente, que: "Las sentencias serán siempre motivadas, lo que supone que las sentencias tienen que dar o explicar las razones o motivos que se han tenido en cuenta para adoptarse en los términos que se han hecho."
- (6) Ver una descripción de la situación en la introducción de Beccaria de su libro sobre los delitos y las penas, publicado en el siglo XVIII (Beccaria, 1774).
- (7) Existían precedentes. Ya en 1929 se constituyó la empresa característica de España, Aranzadi, constituida con el fin de dar acceso en forma sistemática a la legislación y la jurisprudencia publicados en formato papel. El acceso se producía mediante el uso de índices temáticos y cronológicos de la documentación.
- (8) Obviamente cuestión distinta es la justicia, o incluso la idoneidad o no, de la actividad de optar por el acceso a unos u otros textos jurídicos. Es evidente que la atribución de justa a esta actividad estará en relación a su virtualidad, es decir a si es realizada o no en conformidad con las características propias del tipo de actividad jurídica para la que el acceso a textos jurídicos sea llevado a cabo.
- (9) En cumplimiento de la legislación sobre protección de datos no se proporcionan los nombres de los particulares que hayan formado parte del proceso.
- (10) El filósofo Gadamer se fijó en las particularidades que tiene la observación de cada cuadro por el visitante a un museo que las tomó como punto de referencia (Gadamer, 1999, 12-13, 127), junto a las características de la interpretación del derecho señalada por Betti (Ibidem, 397-399) (Betti, 1955), para fundar su filosofía hermenéutica que tan gran repercusión en la comprensión de la heterogeneidad y pluralidad del conocimiento humano a lo largo del siglo XX y el actual XXI ha tenido y tiene.

Agradecimientos

El artículo está apoyado por el proyecto "Posibilidades y requisitos de los sistemas de organización del conocimiento

para la interoperabilidad entre las instituciones de la memoria y el sector del turismo cultural", financiado por el Programa estatal de investigación desarrollo e innovación orientada a los retos de la sociedad (CSO2015-65448-R MINECO/FEDER), Convocatoria 2015.

Referencias

- Alexy, Robert (1992). *Begriff und Geltung des Rechts*. Freiburg: Alber, 1992. 201.
- Auletta, Gennaro (2011). *Cognitive Biology: Dealing with Information from Bacteria to Minds*. Oxford: Oxford University Press, 2011.
- Beccaria, Cesare (1774). *Tratado de los delitos y de las penas*. Madrid: Joachin Ibarra, 1774. 1-5.
- Belloso, Nuria (2019). Algoritmos predictivos al servicio de la Justicia: ¿una nueva forma de minimizar el riesgo y la incertidumbre?. // *Revista da Faculdade Mineira de Direito*. 22:43, 1-31.
- Betti, Emilio (1955). *Teoria generale della interpretazione*. Milan: Giuffrè, 1955.
- Bing, Jon (1978). *Legal Information Retrieval Systems: The Need For and the Design of Extremely Simple Retrieval Strategies*. // *The John Marshall Journal of Information Technology & Privacy Law*. 1:1, 379-399.
- Cross, Rupert & Harris, James W. (2012). *El precedente en el Derecho inglés*. Madrid: Marcial Pons, 2012.
- Gadamer, Hans George (1999). *Verdad y método*. Octava edición. Salamanca: Sígueme, 1999. Volumen 1.
- Galindo, Fernando (2019a), ¿Inteligencia Artificial y Derecho? Sí, pero ¿cómo?. // *Revista Democracia Digital e Governo Eletrónico*. 1:18, 39-43.
- Galindo, Fernando (2019b). *Inteligencia Artificial y acceso a documentación jurídica: sobre el uso de las TICs en la práctica jurídica*. // *Revista Democracia Digital e Governo Eletrónico* 1:18, 156-165.
- Galindo, Fernando (1993). *El acceso a textos jurídicos*. Zaragoza: Mira, 1993. 19-41.
- García-Marco, F.J.; Galindo, F.; Lasala, P.; López del Ramo, J. (2020). *Advancing the interoperability of the GLAM+ and cultural tourism sectors through KOS: perspectives and challenges*. // Artículo aceptado para su presentación en la 16th International ISKO conference, 2020.
- Generalitat de Catalunya (1983). *Curso de Gestión Automatizada en el Ámbito de la Justicia*. Barcelona: Department de Justicia, Generalitat de Catalunya, 1983.
- Kyrtisis, Dimitrios (2015). *Shared Authority. Court and legislatures in legal theory*. Oxford: Hart Publishing, 2015. 160-164.
- Maturana, Humberto (2003). *El árbol del conocimiento: las bases biológicas del entendimiento humano*/Humberto Maturana y Francisco Varela. V ed. Buenos Aires: Lumen, 2003. 89.
- Re, Richard M, & Solow-Niederman, Alicia (2019). *Developing Artificially Intelligent Justice*. // *Stanford Technology Law Review*. 22:1, 242-289.
- Robles, Gregorio (2015), *Teoría del Derecho*. Pamplona: Aranzadi, 2015. 49-51.
- Suárez-Sánchez, Adriana (2017), *Sistemas para la organización del conocimiento: definición y evolución histórica*. // *e-Ciencias de la Información*. 7:2, 1-17.
- Vesting, Thomas (2015). *Rechtstheorie*. München: Beck, 2015. 18-19.

Enviado: 2020-05-05. Aceptado: 2020-05-11.

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; 12.000 caracteres para los estudios de caso; y 6.000 caracteres para las noticias, proyectos en curso, 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 un documento Word o RTF debidamente maquetados según la plantilla disponible en www.iberid.org con sus gráficos incluidos. Irán precedidos de sus títulos, resúmenes informativos y palabras clave, todos ellos en inglés y español. Se incluirán en ficheros aparte los gráficos e imágenes en su archivo fuente para su eventual tratamiento, y, en cualquier caso, con resolución de impresión.

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, agradecimientos 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 aceptable. 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. 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 (Pérez, 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. ISBN número. Tesis doctoral.

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

7.2. Artículos de publicaciones periódicas

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

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

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> (2001-01-26).

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, antítesis y síntesis. 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. Civildad

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

Almeida, Patrícia de, 27
Alves, Roberta Caroline Vesú, 13
Caprioli, Mariana da Silva, 27
Domínguez-Delgado, Rubén, 57
Fujita, Mariângela Spotti Lopes, 13
Galindo Ayuda, Fernando, 63
Grácio, Maria Cláudia Cabrini, 47
Lima, Larissa de Mello, 27

López-Hernández, María-Ángeles, 57
Martínez-Ávila, Daniel, 27, 47
Moraes, João Batista Ernesto de, 27
Oliveira, Lais Pereira de, 47
Rio Branco, Luciana Beatriz Piovezan, 13
Salvador Benítez, Antonia, 35
Sánchez Vigil, Juan Miguel, 35
Tenório, Luana Calcete Vaz, 27

Índice de materias en español

Subject index in Spanish

Bases de datos, 57
Big data, 35
Bookmetrix, 35
Brasil, 47
Centros de documentación, 57
Ciberactivismo, 57
Ciudadanía digital, 57
Derecho de la información, 63
Descripción temática, 47
Documentación, 57
Documentación jurídica, 63
Edición, 35
Evolución, 13
Inteligencia artificial, 35
Jellybooks, 35

Marketing editorial, 35
Movimientos sociales, 57
Organización de la información, 47
Organización del conocimiento, 35
Procesos de Organización del Conocimiento, 27
Producción científica, 47
Proyectos I+D+i, 57
Sistemas de información documental, 63
Sistemas de Organización del Conocimiento, 27
Soergel, Dagobert, 13
Tekstum, 35
Teoría, 13
Tesauros, 13
Textos narrativos de ficción, 27
Tratamiento temático de la información, 47

Índice de materias en inglés

Subject index in English

Artificial Intelligence, 35
Big data, 35
Bookmetrix, 35
Brazil, 47
Cyber-activism, 57
Databases, 57
Digital citizenship, 57
Document management systems, 63
Documentation centers, 57
Edition, 35
Editorial marketing, 35
Evolution, 13
Fiction, 27
Information law, 63
Information organization, 47

Jellybooks, 35
Knowledge organization, 35
Knowledge Organization Processes, 27
Knowledge Organization Systems, 27
Law LIS, 63
Library and information science, 57
R&D projects, 57
Scientific production, 47
Social movements, 57
Soergel, Dagobert, 13
Subject representation, 47
Tekstum, 35
Thematic description, 47
Theory, 13
Thesauri, 13