





# ***SCIRE***

Representación y organización del conocimiento



# SCIRE

Representación y organización del conocimiento

Vol. 23, n.º 1, enero-junio 2017

ISSN 1135-3716

Scire:

knowledge representation and organization

Vol. 23, n. 1, January-June 2017

Ibersid:

Red de Investigación  
en Sistemas de Información  
y Documentación

© 2017 Francisco Javier García Marco y los autores

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

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, 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.

## Agradecimientos

---

Este número ha contado con una Ayuda para la edición de revistas científicas del Vicerrectorado de Investigación de la Universidad de Zaragoza, del año 2016.

## Redacción, distribución y canje

---

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

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

## Suscripciones y números sueltos

---

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

---

# Scire

knowledge organization  
and representation

---

## Scope and aims

---

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

## Dissemination

---

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

## Acknowledgments

---

This issue has received a grant for the edition of scientific journals of the Vicerectorate of Research of the University of Saragossa, Spain, for the year 2016.

## Contact address

---

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

Tel.: int + 34 976 762239. Fax: 34 976761506.  
E-mail: [scire@unizar.es](mailto:scire@unizar.es)

## Subscriptions

---

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

## Editor

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

### Consejo de redacción / Editorial council

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

### Consejo científico / Scientific council

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

### Revisores externos del número / External reviewers in this issue

Agradecemos enormemente la colaboración altruista y desinteresada de los doctores José Luis Alonso Berrocal, José Javier Callau Puente, Juan Antonio Comeche, Cristina Faba-Pérez, José Luis Herrera Morilla, Cristina Herrero Pascual, Daniel Martínez Ávila, Michela Montesi, Luis Miguel Moreno Fernández, Blanca Rodríguez Bravo, Carmen Villar y Ángel Zazo Rodríguez.

#### Candidaturas al consejo científico

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

#### Scientific council membership policy

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



---

# Tabla de contenidos en español

Table of contents in Spanish

---

<i>Tabla de contenidos en español</i> .....	9	<i>Procedimiento de evaluación</i> .....	73
<i>Tabla de contenidos en inglés</i> .....	11	<i>Instrucciones para la presentación de trabajos</i> .....	75
<b>Artículos</b>		<i>Índice de autores</i> .....	78
<i>Propuesta de directrices para la elaboración de la política de indización en las bibliotecas universitarias: un estudio sociocultural con protocolo verbal</i>		<i>Índice de materias en español</i> .....	78
<i>Paula Regina DAL'EVEDOVE</i> <i>Mariângela Spotti Lopes FUJITA</i> .....	13	<i>Índice de materias en inglés</i> .....	78
<i>Representación del conocimiento en historia clínica electrónica interoperable: el caso de la Historia Clínica Digital del Sistema Nacional de Salud de España</i>			
<i>Gonzalo MARCO CUENCA</i> <i>José Antonio SALVADOR OLIVÁN</i> .....	25		
<i>Método para la extracción masiva de canales de sindicación</i>			
<i>Manuel BLÁZQUEZ OCHANDO</i> .....	39		
<i>Producción científica e inserción internacional de la revista Scire en el periodo 2006-2014</i>			
<i>Ely Francina Tannuri de OLIVEIRA</i> <i>Bruno Henrique ALVES</i> <i>Marcos Rodrigues do PRADO</i> <i>Maria Aparecida PAVANELLI</i> .....	47		
<i>La catalogación de materias y la indización en bibliotecas universitarias: un estudio comparativo con protocolo verbal</i>			
<i>Roberta Cristina Dal'Evedove TARTAROTTI</i> <i>Mariângela Spotti Lopes FUJITA</i> .....	57		
<b>Estudios de caso, investigaciones en curso, artículos breves y comunicaciones</b>			
<i>Pedro Garhel y su implicación en "Espacio P": propuesta de un cuadro de clasificación para el archivo personal de un artista polivalente</i>			
<i>Federico CASTRO MORALES</i> <i>Fátima GARCÍA LÓPEZ</i> <i>Ana María MORALES GARCÍA</i> <i>Karin OHLENSCHLÄGER</i> .....	67		



---

# Tabla de contenidos en inglés

Table of contents in English

---

<i>Table of contents in Spanish</i> .....	9	<i>Evaluation process</i> .....	73
<i>Table of contents in English</i> .....	11	<i>Instructions for authors</i> .....	75
<b>Articles</b>		<i>Author index</i> .....	78
<i>A proposal of guidelines for the development of an indexing policy in university libraries: a verbal-protocol-based socio-cultural study</i>		<i>Subject index in Spanish</i> .....	78
<i>Paula Regina DAL'EVEDOVE</i> <i>Mariângela Spotti Lopes FUJITA</i> .....	13	<i>Subject index in English</i> .....	78
<i>Knowledge representation in electronic interoperable health records: the case of the Digital Health Record of the Spanish National Health System</i>			
<i>Gonzalo MARCO CUENCA</i> <i>José Antonio SALVADOR OLIVÁN</i> .....	25		
<i>A method for the massive extraction of syndication channels</i>			
<i>Manuel BLÁZQUEZ OCHANDO</i> .....	39		
<i>Scientific production and international insertion of the journal Scire in the 2006-2014 period</i>			
<i>Ely Francina Tannuri de OLIVEIRA</i> <i>Bruno Henrique ALVES</i> <i>Marcos Rodrigues do PRADO</i> <i>Maria Aparecida PAVANELLI</i> .....	47		
<i>Subject cataloguing and indexing in university libraries: a verbal-protocol comparative study</i>			
<i>Roberta Cristina Dal'Evedove TARTAROTTI</i> <i>Mariângela Spotti Lopes FUJITA</i> .....	57		
<b>Research in advance, brief articles and notes</b>			
<i>Pedro Garhel and its involvement in "P Space": proposing a classification scheme for a versatile artist</i>			
<i>Federico CASTRO MORALES</i> <i>Fátima GARCÍA LÓPEZ</i> <i>Ana María MORALES GARCÍA</i> <i>Karin OHLENSCHLÄGER</i> .....	67		



---

# Proposta de diretrizes para elaboração de política de indexação em bibliotecas universitárias: um estudo sociocultural com protocolo verbal

*Propuesta de directrices para la elaboración de la política de indización en las bibliotecas universitarias: un estudio sociocultural con protocolo verbal*

*A proposal of guidelines for the development of an indexing policy in university libraries: a verbal-protocol-based socio-cultural study*

---

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

(1) Universidade Federal de São Carlos, Departamento de Ciência de Informação, Rodovia Washington Luís, Km 235, CEP. 13.565-905, São Carlos, SP, Brasil. Correio eletrônico: dalevedove@ufscar.br. (2) Universidade Estadual Paulista, Av. Hygino Muzzi Filho, 737, Campus Universitário, CEP. 17525900, Marília, SP, Brasil. Correio eletrônico: fujita@marilia.unesp.br

## Resumen

Se analiza la política de indización desde la perspectiva humana del tratamiento temático de la información, teniendo en cuenta las indicaciones y conceptos teóricos; detalles de la práctica de la profesión relacionada con la actividad de indización; y la realidad de las bibliotecas universitarias. Para ello, se aplicó la metodología del protocolo verbal en grupo a indizadores profesionales. Como resultado, se presentan ocho directrices para la elaboración de políticas de indización en las bibliotecas universitarias brasileñas.

**Palabras clave:** Tratamiento temático de la información. Políticas de indización. Aproximación sociocultural. Protocolo verbal. Bibliotecas universitarias. Brasil.

## Abstract

A set of eight guiding guidelines for the preparation of indexing policies in Brazilian university libraries is presented, as a result of a verbal-protocol study among a group of indexers from Brazilian university libraries.

**Keywords:** Subject cataloguing. Indexing policies. Verbal protocol. Sociocultural approach. University libraries. Brazil.

## 1. Apontamentos iniciais

Nos últimos anos houve um aumento significativo de pesquisas brasileiras sobre política de indexação, com destaque para os trabalhos apresentados por Carneiro (1985); Cervantes, Fujita e Rubi (2008); Dal'Evedove (2014); Dal'Evedove e Fujita (2015); Fujita e Gil Leiva (2009); Fujita e Santos (2016); Gil Leiva e Fujita (2012); Guimarães (2004); Kochani, Boccato e Rubi (2012); Lousada et al. (2011); Nunes (2004); Redigolo et al. (2012); Rubi (2008); Rubi, Fujita e Boccato (2012a e 2012b), entre outros. O interesse da comunidade científica da área de Organização e Representação do Conhecimento pela temática decorre, principalmente, da complexidade da atividade de indexação com fins de representação e recuperação da informação, bem como pela relevância e influência deste recurso para a formação e atuação profissional de bibliotecários.

Por delimitar os procedimentos de indexação, contemplando os princípios, métodos e orientações para a condução da atividade com efeitos

na recuperação, a política de indexação interessa ao sistema de informação e seus respectivos sujeitos: profissionais que realizam o tratamento temático de documentos, gestores de sistemas de recuperação da informação e usuários (Dal'Evedove e Fujita, 2012, p. 13). Na prática, as pesquisas na temática centram esforços para o aprofundamento de aspectos teóricos, metodológicos e pedagógicos considerando não apenas o bibliotecário indexador, uma vez que em bibliotecas universitárias, por exemplo, outros atores sociais estão atrelados direta ou indiretamente aos resultados provenientes da elaboração e aplicação deste recurso (Dal'Evedove, 2010).

Dessa forma, as relações humanas atuantes na representação de assunto precisam ser consideradas para que os produtos e serviços informacionais gerados a partir do processo de tratamento temático da informação não sejam abandonados à espontaneidade da subjetividade ou "bom senso" profissional; alheios às necessidades informacionais dos usuários; ou mesmo apartados da literatura especializada. Essa enorme tarefa

não poderá ser realizada a menos que, sob uma vigência coletiva, os fenômenos, manifestações e entendimentos que cercam este fazer profissional sejam observados a partir da percepção dos atores sociais que compõem esse universo.

Em pesquisa realizada por Dal'Evedove (2014) intitulada *Tratamento temático da informação em abordagem sociocognitiva: diretrizes para elaboração de política de indexação em bibliotecas universitárias* (1), a autora centrou esforços para os desafios da política de indexação frente à diversidade cultural. Tomou-se como base a concepção de que a política de indexação precisa refletir o conjunto de ideias, valorações, atitudes e conceitos peculiares dos atores sociais que compõem o Tratamento Temático da Informação no campo dos acontecimentos discursivos, para que suas diretrizes sejam coerentes com os princípios teóricos e metodológicos da área de Organização e Representação do Conhecimento, com a prática cotidiana da profissão e, sobretudo, com a realidade específica do contexto de bibliotecas universitárias que disponibilizam seus produtos e serviços informacionais por meio de catálogos coletivos *online*. Nessa direção, a proposta foi contribuir com diretrizes de políticas de indexação elaboradas mediante investigação da vertente humana do processo em contextos de bibliotecas universitárias.

Esta ideia ampara-se na necessidade de primar pela garantia cultural às informações construídas como prática social em domínios informacionais específicos (Beghtol, 2002). Com efeito, passa-se a considerar o contexto sociocultural que envolve o processo de tratamento temático da informação em razão do processo de conhecimento individual ser socialmente condicionado (Fujita, 2007).

Na referida pesquisa o Tratamento Temático da Informação foi delineado no campo dos acontecimentos discursivos, quais sejam: a) fundamentos teóricos e metodológicos do corpo de docentes da área de Organização e Representação do Conhecimento; b) contexto sociocognitivo do bibliotecário indexador; e c) características e princípios culturais dos usuários. Buscou-se, para tanto, investigar as diferentes vertentes que compõem o processo em abordagem sociocultural, delimitadas em: *vertente científica* composta pelos docentes que ministram disciplinas relacionadas à área de Organização e Representação do Conhecimento no âmbito do Tratamento Temático da Informação; *vertente profissional* constituída pelos profissionais bibliotecários indexadores que atuam em contexto de bibliotecas universitárias; e *vertente de uso* formada pela comunidade usuária.

Considerando-se como fatores o cenário que envolve a pesquisa nas Ciências Sociais e, especificamente, a busca por resultados proveitosos, a linha de raciocínio adotada foi a do método fenomenológico. Enquanto um modelo compreensivo em detrimento da mera explicação de fenômenos, o método em destaque torna “[...] explícita a estrutura e o significado implícito da experiência humana” (Sanders, 1982, p. 353). O ponto chave da Fenomenologia é compreender a essência dos fenômenos em análise a partir dos significados que as experiências vividas têm para os sujeitos; considerando sua individualidade e suas relações sociais estabelecidas e experienciadas em contextos diversos.

A primeira parte do estudo consistiu na elaboração e aplicação de três questionários eletrônicos, cada qual direcionado para um grupo pesquisado, a fim de identificar: a) o conhecimento e concepção dos docentes sobre os princípios teóricos e metodológicos da política de indexação; b) a prática cotidiana da comunidade profissional atuante no processo de tratamento temático da informação e a questão da política de indexação no contexto das bibliotecas universitárias; e c) a percepção dos usuários quanto da precisão dos produtos informacionais resultantes da recuperação da informação realizada nos catálogos coletivos *online*. O teor das questões em cada um dos questionários pode ser consultado em Dal'Evedove (2014). Do total dos questionários respondidos, obteve-se um grupo social composto por doze docentes, quatro profissionais e cinquenta usuários. Isto quer dizer que 29,3% do total dos sujeitos identificados na vertente científica participaram da pesquisa, sendo 13,8% na vertente profissional e 67,6% na vertente de uso. O número de atores sociais identificados foi 144, sendo o total de sujeitos participantes 66. O baixo retorno dos questionários enviados aos docentes foi justificado por apenas dezesseis sujeitos participantes que alegaram falta de tempo e excesso de trabalho. Os bibliotecários indexadores que não responderam ao questionário não se manifestaram.

A análise conjunta e comparativa dos questionários viabilizou a obtenção de resultados que refletem a realidade do processo na perspectiva das diferentes vertentes investigadas. Os resultados obtidos contribuíram diretamente para: a) conhecer o universo discursivo-reflexivo da política de indexação na perspectiva dos atores sociais testados e, de modo concreto, distinguir os pontos favoráveis e adequados para o contexto de bibliotecas universitárias; e b) incorporar elementos de cada uma das vertentes investigadas e de seus respectivos atores sociais na composição de diretrizes para a definição de políticas de indexação para o contexto de bibliotecas universitárias.

Com base nestes resultados, foram traçados novos elementos de política de indexação para bibliotecas universitárias. Como forma de verificar a pertinência e exequibilidade de tais indicações pela perspectiva da prática profissional, utilizou-se no segundo momento da pesquisa o Protocolo Verbal em Grupo. A aplicação desta técnica introspectiva de coleta de dados serviu para identificar as opiniões, impressões e conhecimentos dos sujeitos participantes sobre os elementos de política de indexação estabelecidos por meio da análise dos questionários.

Diante do exposto, a proposta neste artigo é apresentar os resultados advindos com a aplicação do Protocolo Verbal em Grupo com profissionais bibliotecários indexadores, a fim de expor as diretrizes para definição de política de indexação obtidas ao término da pesquisa, de modo a divulgá-las para a comunidade científica e profissional.

## 2. Procedimentos metodológicos

A utilização do Protocolo Verbal em Grupo como técnica introspectiva para coleta de dados serviu para verificar junto à sociedade profissional se os elementos estabelecidos com base nos resultados dos questionários eram compatíveis com a realidade dos contextos de bibliotecas universitárias e com a prática cotidiana do bibliotecário indexador, com vistas a sua adequação.

Reconhece-se, portanto, a importância do sujeito no processo de construção do conhecimento, visto que a realidade é construída socialmente e entendida como o percebido, interpretado e/ou comunicado. A importância do levantamento destas manifestações reside no fato de que os bibliotecários indexadores precisam estar confortáveis diante destas diretrizes para que venham a utilizá-las no momento da elaboração de uma política de indexação específica para a biblioteca universitária em que atuam. Sendo assim, a escolha do Protocolo Verbal em Grupo justifica-se em decorrência de que (Rubi, 2008, p. 18)

[...] quem conhece a política de indexação é o profissional que a faz e somente ele poderá iniciar e dar continuidade ao processo de construção de novos conhecimentos sobre a indexação e, conseqüentemente, sobre sua política, dentro do sistema de informação.

De modo geral, nesta modalidade de Protocolo Verbal os sujeitos participantes verbalizam o conhecimento processual individual para desvendar determinada atividade de leitura, alternando-se o conhecimento prévio com a leitura e discussão do texto apresentado de modo aleatório (Nardi, 1999).

Para que as medidas tomadas para a aplicação do Protocolo Verbal em Grupo não fossem avulsas, foram adotados os procedimentos metodológicos elaborados por Nardi (1999), sendo exposta uma contextualização geral sobre as etapas realizadas em cada tópico dos procedimentos anteriores, durante e posteriores à aplicação, conforme segue:

I. Procedimentos anteriores às sessões de aplicação da coleta de dados:

- *Definição do universo da pesquisa:* A aplicação da técnica ocorreu com três bibliotecários indexadores atuantes na Biblioteca do Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas – UNICAMP (área de exatas).

Proposta de diretrizes de política de indexação para bibliotecas universitárias

**Fase 1: Preparação** As características e objetivos institucionais deverão ser claramente identificados, servindo de orientação para a definição e cumprimento dos demais pontos da política de indexação.

Cada biblioteca universitária deverá estabelecer orientações de conduta ética por meio das quais os profissionais bibliotecários atuantes no tratamento temático da informação deverão se pautar.

Realizar um estudo sobre a cultura organizacional da biblioteca universitária que seja coerente com a demanda da comunidade usuária.

Prever um estudo de usuário que contemple questões relacionadas ao funcionamento do processo de busca nos sistemas de recuperação da informação, a fim de se conhecer as formas de uso e significação da linguagem pela comunidade usuária.

**Fase 2: Desenvolvimento** Delimitar com precisão o nível de abrangência que a indexação será realizada. Recomenda-se que os níveis de análise de assunto sejam encaminhados a partir da concepção orientada ao usuário, na qual o bibliotecário leva em consideração o conhecimento especializado da comunidade usuária sobre os assuntos para realizar a representação conceitual do documento.

Realizar a representação do conteúdo informativo do documento a partir de regras de compartilhamento que preservem a garantia cultural, epistêmica e ética da informação.

Adequar os assuntos do documento no momento da representação conceitual de acordo com as características da comunidade usuária.

Prever a criação de diferentes representações do mesmo documento como forma de amparar usuários inseridos em domínios de conhecimento específicos. Recomenda-se que o catálogo da biblioteca ofereça produtos com valor agregado como forma de demonstrar valorização e reconhecimento da importância de seus usuários no provimento de informações.

*Tabela 1. Texto utilizado na aplicação do Protocolo Verbal em Grupo (Dal'Evedove, 2014)*

- *Seleção do texto-base*: O material discutido pelos sujeitos de pesquisa foi elaborado mediante a análise comparativa dos resultados provenientes dos questionários aplicados com docentes, profissionais e usuários que permitiram a delimitação de alguns elementos de política de indexação para o contexto de bibliotecas universitárias (Tabela I).
- *Definição da tarefa*: Ocorreu a recepção dos sujeitos participantes pela pesquisadora para uma apresentação formal entre as partes; e conversa informal entre a pesquisadora com os sujeitos participantes para expor os objetivos e relevância da pesquisa, bem como para realizar a familiarização dos sujeitos com a técnica.

## II. Procedimentos durante a coleta de dados:

- *Gravação do “Pensar Alto” do grupo de sujeitos participantes durante a discussão do Texto-Base*: A gravação do “Pensar Alto” dos sujeitos durante a discussão do material ocorreu por meio de um gravador digital.
- *Entrevista retrospectiva (optativa)*: O intuito desta manobra foi esclarecer dúvidas sobrepostas durante a execução da atividade em análise e incitar discussões de tópicos pouco explorados ou não abordados pelo grupo durante a discussão principal.

## III. Procedimentos após o término das sessões de coleta de dados:

- *Transcrição literal das gravações das falas dos sujeitos participantes*: Realizou-se a transcrição dos dados, na íntegra, com a identificação das fontes individuais de cada sujeito participante. Para uma melhor exposição, todas as falas individuais foram numeradas em turnos, o que facilitou a análise dos dados e manteve a natureza contínua da interação entre os sujeitos de pesquisa. Na parte prática das transcrições, utilizou-se o *Express Scribe Transcription Software* que, por ser um *software* destinado para auxiliar os digitadores na transcrição de textos falados para escritos, contribuiu para melhorar a confiança e validade das transcrições, não havendo perda de dados.
- *Leitura detalhada dos dados em busca de fenômenos significativos e recorrentes para a construção das categorias de análise*: Ocorreu uma leitura detalhada das transcrições do protocolo verbal por parte da pesquisadora na busca por aspectos relevantes e fenômenos significativos que permitissem a construção das categorias para análise dos dados coletados. Esta etapa foi baseada na leitura para observação.

- *Construção das categorias de análise*: A construção das categorias de análise pautou-se nos referenciais teóricos abordados e discutidos em Dal'Evedove (2014) e nos depoimentos dos sujeitos participantes.
- *Retorno aos dados para retirar trechos da discussão que exemplifiquem cada categoria de análise*: Houve, neste momento, uma releitura das transcrições das coletas de dados para a retirada de trechos da discussão que melhor exemplificassem cada fenômeno (objeto de estudo) abordado nas categorias de análise e seus tópicos referenciais.
- *Análise dos dados*: Concluídas as etapas anteriores, a análise dos dados foi baseada nas declarações realizadas pelos sujeitos de pesquisa em relação aos elementos contidos nas diretrizes apresentadas como produto da discussão social. A análise qualitativa dos dados obtidos ficou a cargo da Análise de Conteúdo (Bardin, 2009) como forma de obter a fruição da realidade dos dados coletados mediante a concepção dos atores sociais testados.

## 3. Apresentação e discussão dos resultados

Os elementos de política de indexação discutidos foram elaborados em Dal'Evedove (2014) considerando-se, para tanto, as indicações teórico-metodológicas presentes na literatura especializada da área de Organização e Representação do Conhecimento e dos resultados obtidos por meio dos questionários. Os elementos foram apresentados adotando-se as duas primeiras fases da elaboração de política de indexação: preparação e desenvolvimento; uma vez que a avaliação enquanto terceira e última etapa é o momento em que o modelo de política de indexação será discutido pela equipe da biblioteca universitária para indicação de ajustes e melhorias e, também, para completar a proposta de diretrizes (Rubi, 2008).

Dentre todas as falas transcritas do protocolo verbal, são utilizadas aquelas que melhor explicitam e/ou exemplificam os temários abordados nas respectivas categorias de análise. Os trechos significativos de cada categoria estão sublinhados, sendo dispostos de maneira complementar com uma síntese analítica acerca dos exemplos apresentados, com destaque aos aspectos referentes à política de indexação. A ênfase na análise dos dados coletados foi direcionada na tentativa de identificar as opiniões, impressões, concepções e comportamentos pessoais e coletivos dos profissionais bibliotecários investigados acerca dos elementos explorados durante a discussão grupal.



### 3.1. Elementos de política de indexação do ponto de vista profissional

A primeira etapa é o momento em que, depois de verificadas todas as condições de preparação para o desenvolvimento da política de indexação, toda e qualquer decisão e/ou diretriz a ser estabelecida deve corresponder aos objetivos e à filosofia da biblioteca, da organização na qual está inserida e dos usuários atendidos (RUBI, 2008). Em sintonia com tal indicação, o primeiro item apresentado para uma avaliação pontual dos bibliotecários indexadores comporta questões de gestão, cultura organizacional e comportamento profissional, a saber:

#### 3.1.1. *As características e os objetivos institucionais deverão ser claramente identificados, servindo de orientação para a definição e cumprimento dos demais pontos da política de indexação*

Os sujeitos dedicaram grande parte da discussão sobre a importância desta fase da política de indexação, o que demonstra uma preocupação pelo estabelecimento de um recurso que considere o contexto de informação, suas características e particularidades. Na ótica dos sujeitos, torna-se de grande relevância que as características e os objetivos da biblioteca universitária sejam considerados na política de indexação, sendo preciso, para tanto, conhecer o usuário para que o bibliotecário indexador tenha condições de elaborar um recurso que seja coerente e passível de aplicabilidade no âmbito da biblioteca universitária. Seus discursos indicam que a definição precisa das características e objetivos institucionais servirão de orientação para o estabelecimento e cumprimento dos demais pontos da política de indexação. O posicionamento dos sujeitos participantes reafirma o pensamento de que a política de indexação deve ser compreendida como uma decisão administrativa por parte das bibliotecas universitárias, sendo representada por meio de uma filosofia que reflita os objetivos da instituição.

[19 Sujeito 2] Agora, eu acho que a pessoa mais importante para desenvolver uma política de indexação são os usuários e a instituição como um todo. Eu acho que tem que levar em consideração qual é a função dela, qual são os objetivos, o público que vai atender o nível de especificidade que tem que ter.

Além disso, os sujeitos indicam que a política de indexação deve estar regulamentada e publicada. A literatura especializada da área recomenda esta prática e defende a necessidade do registro formal da política de indexação em manuais de indexação (Rubi, 2008), como forma de salvaguardar a sua continuidade, avaliação e

aperfeiçoamento ao longo do tempo e, sobretudo, para que todos os profissionais envolvidos com a atividade tenham conhecimento claro acerca “[...] das decisões tomadas, suas razões e seu contexto” (Guimarães, 2000, p. 56).

Esta confluência de ideias atesta que o momento que antecede a política de indexação é crucial para cada contexto de informação identificar com precisão todos os aspectos a serem considerados nestas diretrizes, ação que viabiliza o cumprimento de tais indicações no cotidiano da atividade. Nunes (2004, p. 57) esclarece que sendo a política de indexação um documento específico de cada biblioteca universitária, “[...] há que se admitir diferentes gradações dessa política. Isso é absolutamente esperado. O que não é esperado é a ausência completa de políticas formalmente enunciadas [...]”. Na ótica dos participantes a publicação da política de indexação possibilita que outras instituições sejam beneficiadas, como forma de divulgar o trabalho realizado pela instituição, contribuindo com diálogos e trocas de conhecimentos entre os bibliotecários indexadores no cumprimento de suas ações profissionais.

[18 Sujeito 2] É importante ter um documento que comprove e também você divulgar a forma que você faz a catalogação, porque você vai unir esforços para ver o que outra instituição está fazendo, assim trazer melhorias para a política e divulgar o que você faz, trocar ideias entre pares, com bibliotecas da mesma área.

Na prática, consideram que o ato de regulamentar e publicar a política de indexação contribui para que esforços sejam minimizados e a política de indexação seja aperfeiçoada.

#### 3.1.2. *Cada biblioteca universitária deverá estabelecer orientações de conduta ética por meio das quais os profissionais bibliotecários atuantes no tratamento temático da informação deverão se pautar*

O enunciado foi bem aceito pelos sujeitos participantes ao demonstrarem a relevância da inclusão desta questão na política de indexação, como forma de contribuir para uma possível padronização e conduta profissional no tratamento temático da informação. Sobre isto, cabe reafirmar que o bibliotecário indexador realiza o processo em um contexto cultural e ideológico e, por isso, suas crenças pessoais interferem na representação do conhecimento (Pinho, 2010). Dessa forma, prever na política de indexação questões de orientação e conduta ética é contribuir para que a atuação profissional seja coerente com os objetivos institucionais. Para os sujeitos participantes, tal indicação permite que os bibliotecários indexadores tenham uma visão mais adequada da realidade profissional e reflitam sobre

suas ações na indexação. No geral, consideram que diretrizes que tenham essa conotação sejam necessárias na política de indexação por contribuírem para uma mudança de ações e postura profissional frente aos desafios que lhes são colocados todos os dias na prática cotidiana da profissão.

[35 Sujeito 1] Agora nesse ponto: cada biblioteca universitária deverá estabelecer orientações de conduta ética por meio das quais os profissionais bibliotecários atuantes no tratamento temático da informação deverão se pautar, isso aqui também é importantíssimo, porque se ninguém seguir o mesmo padrão e falar a mesma linguagem vai ficar sem padrão.

### 3.1.3. Realizar um estudo sobre a cultura organizacional da biblioteca universitária que seja coerente com a demanda da comunidade usuária

Esta indicação foi considerada de extrema relevância pelos sujeitos participantes como forma de identificar o contexto em que a biblioteca universitária está inserida e, conseqüentemente, definir com maior precisão a função da biblioteca e os objetivos a serem assumidos por ela na atribuição de suas atividades. É notória a postura reflexiva dos sujeitos sobre a importância de a cultura organizacional ser trabalhada em uma política de indexação direcionada para bibliotecas universitárias, sendo estas “imbuídas da responsabilidade de mapear e atender às necessidades e intenções de seus usuários” (Carvalho, 2004, p. 169) com rigor e excelência. Ademais, é oportuno ressaltar que a política de indexação requer avaliações contínuas, especificamente no que tange à linguagem documental quanto do seu “[...] desempenho durante a busca por assunto na recuperação da informação que possa orientar seu aprimoramento e seu uso no processo de representação para indexação documental” (Boccatto, 2009, p. 236).

Para os sujeitos participantes, o estudo da cultura organizacional viabiliza a identificação dos elementos essenciais que fundamentam a política de indexação na biblioteca universitária, diante de suas particularidades e realidade próprias. Afirmam ainda que, assim como a universidade preserva a sua cultura, a biblioteca universitária também precisa preservar a sua cultura organizacional e estender essa preocupação no momento do estabelecimento de sua política de indexação, resgatando o histórico da instituição para se delimitar com maior propriedade as iniciativas que deram certo e as que sofreram mudanças ao longo dos anos.

[31 Sujeito 1] Mas assim, é importante na política também resgatar o histórico, o passado porque tem

gente nova que não conhece as tentativas que deram erradas e as que deram certo, então é importantíssimo documentar o histórico da atividade, da instituição para resgatar a cultura da organização porque é ela que vai definir o caminho que a instituição tem que direcionar para elaborar a política.

[35 Sujeito 1] Agora, realizar um estudo sobre a cultura organizacional da biblioteca universitária que seja coerente com a demanda da comunidade usuária também é algo muito importante.

[36 Sujeito 2] A cultura é quem fornece os elementos essenciais que darão um norte, que fundamenta qualquer política para uma biblioteca no ambiente da universidade, qual o contexto em que a biblioteca está inserida.

### 3.1.4. Prever um estudo de usuário que contemple questões relacionadas ao funcionamento do processo de busca nos sistemas de recuperação da informação, a fim de se conhecer as formas de uso e significação da linguagem pela comunidade usuária

Para os bibliotecários indexadores é preciso conhecer as opiniões e interesses dos usuários para definir os critérios a serem empregados no estabelecimento da política de indexação. Essa postura compromissada por parte dos sujeitos indica que a presença do usuário parece ser vista como essencial quando se considera a elaboração de política de indexação. Outra questão é de que o bibliotecário precisa conhecer a área de especialização da biblioteca universitária para poder compreender a linguagem adotada pelo usuário.

No geral, verifica-se que os sujeitos participantes são favoráveis de que os sistemas de informação sejam planejados de modo a respeitar as particularidades de cada biblioteca universitária, de modo a viabilizar a troca de experiências entre usuários e bibliotecários, pois

[...] deve haver um esforço do bibliotecário no que diz respeito à representação adequada do conteúdo do documento. No entanto, isso às vezes não ocorre, cabendo ao usuário, geralmente especialista no assunto, auxiliar o bibliotecário nessa tarefa (Rubi, 2008, p. 156).

Mesmo não sendo especialista em todos os assuntos de que trata, o bibliotecário indexador interage com o documento a partir de linguagens documentais especializadas, considerando a intenção do sistema de informação e a estrutura textual do documento. Neste sentido, os sujeitos participantes entendem que conhecer o processo de busca e as formas de uso e significação da linguagem pelos usuários contribui para que a linguagem documental adotada pelo sistema seja

atualizada e a representação conceitual corresponde ao assunto designado pelo usuário no momento da busca.

[21 Sujeito 2] Agora, conhecer a área do usuário eu acho que é muito importante, eu acho que todo o bibliotecário deveria conhecer a área que ele trabalha pelo menos o básico. O bibliotecário precisa ter o conhecimento para poder conversar, ter a mesma linguagem do usuário.

[36 Sujeito 2] Prever um estudo de usuário que contemple questões relacionadas ao funcionamento do processo de busca da informação pela comunidade usuária, a fim de se conhecer as formas de uso e significação da linguagem pelos usuários.

Depois de verificadas todas as condições de preparação para a elaboração da política de indexação, a fase de desenvolvimento é a etapa composta de decisões e diretrizes que deverá corresponder aos objetivos e à filosofia da biblioteca, da organização na qual está inserida e dos usuários atendidos. Neste ponto da discussão, a primeira indicação a ser considerada na elaboração da política de indexação compreende:

### 3.1.5. *Delimitar com precisão o nível de abrangência que a indexação será realizada*

A preocupação dos sujeitos participantes foi demonstrar situações vivenciadas na prática profissional que atestam a relevância desta delimitação no momento do tratamento temático da informação, considerando, dentre outros fatores, a recuperação no sistema de informação. Cunha (2000, p. 81) esclarece que “a política de indexação seguida pela biblioteca é que irá delinear quais níveis de representação da informação serão adotados em um determinado acervo”. No geral, a tendência entre as bibliotecas universitárias é representar o assunto dos documentos no nível mais geral, o que acarreta a recuperação de um grande número de documentos (Rubi, 2008).

Nas diretrizes de política de indexação contidas na literatura especializada, o momento da especificidade é considerado crucial, pois se a precisão não for contemplada, o usuário despenderá de um esforço maior para selecionar o item documental que satisfaça as suas intenções de busca (Rubi, 2008). Diante disso, a política de indexação é o local apropriado para se definir com precisão o nível de especificidade do documento a ser representado na biblioteca universitária.

[38 Sujeito 2] Já na questão do desenvolvimento, quando o texto fala que é preciso delimitar com precisão o nível de abrangência que a indexação será realizada é bem aquilo que eu falei antes, era isso que eu estava querendo dizer para vocês, eu acho que a gente não tem isso aqui, embora nós tenhamos muita coisa que a gente adota e usa na prática,

ainda não está definido isso, porque tem gente que mantém termos gerais.

### 3.1.6. *Realizar a representação do conteúdo informativo do documento a partir de regras de compartilhamento que preservem a garantia cultural, epistêmica e ética da informação*

A postura assumida pelos sujeitos participantes foi favorável com o enunciado. Consideram, portanto, que a garantia cultural, epistêmica e ética da informação são questões que merecem atenção na política de indexação. Apesar de ao longo da discussão social a cultura no âmbito do contexto de bibliotecas universitárias ter sido uma questão bastante abordada, outros pontos do enunciado em destaque não foram amplamente discutidos pelos sujeitos; fato que prejudicou uma análise mais verticalizada sobre a questão da garantia epistêmica e ética da informação na concepção dos atores sociais testados.

[40 Sujeito 2] Nossa, olha aqui, na questão de preservar a garantia cultural é algo muito importante, olha só de novo a questão da cultura, como isso é importante na política de indexação, em qualquer política [...].

### 3.1.7. *Adequar os assuntos do documento no momento da representação conceitual de acordo com as características da comunidade usuária*

Na discussão social deste item, os sujeitos participantes reconhecem a importância de serem consideradas as características e preferências da comunidade usuária, especialmente no momento da busca no catálogo da biblioteca. Para os sujeitos, esta indicação deveria ser considerada no momento da indexação, pois o contato com o usuário permite conhecer os termos utilizados e o modo como a busca é realizada no sistema de recuperação da informação. Como resultado, tem-se a possibilidade de adequar os assuntos do documento no momento da representação conceitual de acordo com as características da comunidade usuária. Sobre isto, Boccato (2009, p. 231) alerta que:

O uso adequado da linguagem documentária no processo de representação para indexação está vinculado ao processo de avaliação constante da busca por assunto, proporcionando elementos norteadores para o aperfeiçoamento e a adequação de um sistema de organização do conhecimento que reflita o contexto informacional, cultural e social em que está inserido.

Considerar as características da comunidade usuária na atribuição de termos descritores é entender as possibilidades de recuperação de informações significativas no catálogo da biblioteca universitária, o que amplia o grau de satisfação

dos usuários no uso. Os sujeitos reconhecem a valia deste ser um item considerado na política de indexação para bibliotecas universitárias.

[21 Sujeito 1] Deveria ser um item a ser considerado no momento da indexação, mas ela não está formalizada [...]. Às vezes o usuário indica que um ou outro termo não é mais usado, então a gente aproveita esse momento para atualizar.

[20 Sujeito 2] Por isso que é importante o contato com o usuário porque a gente acaba conhecendo como ele pensa, como ele busca, quais os termos que eles utilizam. Assim, você acaba conhecendo a linguagem deles.

[22 Sujeito 2] O catalogador tem que ter essa preocupação de estar voltado para a recuperação da informação quando eu vou preencher os cabeçalhos, eu vou dar a ênfase no assunto principal, tudo para que o aluno não fique entrando em títulos que ele abre o material e não tem nada a ver com a busca dele.

### 3.1.8. Prever a criação de diferentes representações do mesmo documento como forma de amparar usuários inseridos em domínios do conhecimento específicos

Neste item um dos sujeitos participantes foi contrário com a indicação de criar diferentes representações do mesmo documento como forma de amparar usuários que possuem necessidades informacionais específicas. No entanto, constatou-se nesta pesquisa que a abordagem sociocultural defende que a atividade humana é dotada de um sistema de significação em um dado contexto cultural e social, sendo mediada por instrumentos carregados de significação cultural.

No campo da informação, esses princípios coadunam quando se considera que sua preocupação enquanto ciência é tornar a informação significativa para os que dela necessitam, o que implica uma efetiva preocupação sobre as condições, padrões e regras de uso enquanto fatores que tornam a informação significativa, ou seja, considerar o sujeito e suas relações na coletividade. Segundo Carvalho (2004, p. 38), “na sociedade contemporânea, oferecer serviços e produtos com valor agregado representa uma demonstração de valorização e reconhecimento da importância de seus usuários”. Este pensamento foi compartilhado por um dos sujeitos participantes que se mostrou favorável com o enunciado ao informar ser esta uma preocupação prioritária no contexto da biblioteca universitária.

Apesar de não ser objetivo da pesquisa se debruçar sobre as implicações que esta prática pode gerar nas atividades e serviços da biblioteca universitária, é oportuno salientar que esta indicação foi considerada válida pelos demais profissionais investigados. Diante disso, entende-se que

tal indicação tem condições de ser examinada no momento da definição de uma política de indexação, pois toda informação inserida no sistema de recuperação de informação deve representar o grupo social para quem ela é destinada, sem desvios ou dissociações.

[56 Sujeito 1] Isso é uma opção ótima, porque se for representar o mesmo documento de diferentes maneiras eu concordo, porque eu acho que é uma prioridade pensar nessa opção.

No geral, os resultados obtidos com a discussão do material elaborado sobre os elementos a serem observados na definição de política de indexação em bibliotecas universitárias foram consistentes e permitem constatar a viabilidade da adoção dessas indicações. Os bibliotecários indexadores refletiram diante dos pontos apresentados e, com exceção do último item da segunda categoria de análise, todos foram favoráveis quanto à adoção dessas indicações para a elaboração e/ou manutenção da política de indexação em e para bibliotecas universitárias. Não houve manifestações sobre a necessidade de reformular alguma das indicações expostas. Apesar do incentivo da pesquisadora, os sujeitos participantes não apontaram nenhuma alteração proeminente a ser realizada nas diretrizes apresentadas na discussão social.

Por um lado, considera-se que a aceitação das diretrizes apresentadas à comunidade profissional imprima uma considerável credibilidade ao material proposto. Contudo, os enunciados foram respondidos por um pequeno grupo de bibliotecários indexadores; resultado que abre precedentes para que alguns dos pontos aqui trabalhados tenham sido observados superficialmente pelos sujeitos participantes. Isto não invalida que as diretrizes sejam consistentes e oportunas de serem assumidas no momento da elaboração e/ou adequação da política de indexação, mas indica que o material elaborado e apresentado nesta pesquisa seja discutido por um número maior de bibliotecários indexadores, como contributo à questão e, também, como forma de identificar os aspectos que podem ser considerados inapropriados em outros contextos.

## 4. Diretrizes para elaboração de política de indexação em bibliotecas universitárias

Ao passo que as diretrizes foram consideradas proveitosas e passíveis de uso por parte da comunidade profissional investigada, tem-se um total de oito diretrizes de orientação que permeiam as fases de preparação e desenvolvimento a serem consideradas na definição de uma política de

indexação para o contexto de bibliotecas universitárias, quais sejam:

- As características e os objetivos institucionais deverão ser claramente identificados, servindo de orientação para a definição e cumprimento dos demais pontos da política de indexação;
- Cada biblioteca universitária deverá estabelecer orientações de conduta ética por meio das quais os profissionais bibliotecários atuantes no tratamento temático da informação deverão se pautar;
- Realizar um estudo sobre a cultura organizacional da biblioteca universitária que seja coerente com a demanda da comunidade usuária;
- Prever um estudo de usuário que contemple questões relacionadas ao funcionamento do processo de busca da informação, a fim de se conhecer as formas de uso e significação da linguagem pela comunidade usuária;
- Delimitar com precisão o nível de abrangência que a indexação será realizada;
- Realizar a representação do conteúdo informativo do documento a partir de regras de compartilhamento que preservem a garantia cultural, epistêmica e ética da informação;
- Adequar os assuntos do documento no momento da representação conceitual de acordo com as características da comunidade usuária; e
- Prever a criação de diferentes representações do mesmo documento como forma de apoiar usuários inseridos em domínios do conhecimento específicos.

A tarefa de propor novas diretrizes de política de indexação decorre em razão das bibliotecas universitárias apresentarem diferentes realidades; isto é, terem objetivos e características próprias. Muitas vezes, o que é adequado para uma biblioteca universitária pode não ser pertinente para outra. Sendo assim, optou-se por seguir a linha de propor diretrizes pensando-as como algo flexível, em que cada biblioteca universitária deve adequar estas indicações de acordo com os objetivos a serem alcançados no cumprimento de suas atividades. Com efeito, ressalta-se que todas as diretrizes aqui apresentadas visam tornar a política de indexação mais coerente quanto à aplicação dos conhecimentos teóricos do processo diante das especificidades da prática cotidiana da profissão realizada em contextos diversificados por domínios de conhecimento especializados.

Diante da carência de estudos que realizem menções ou busquem uma confluência direta entre os atores sociais ativos no estabelecimento, execução e resultado do processo de tratamento temático da informação realizado em contexto de bibliotecas universitárias, esta pesquisa abre caminhos alternativos para se trabalhar com a política de indexação, além de novas diretrizes de orientação para a elaboração e/ou avaliação deste recurso em bibliotecas universitárias brasileiras.

## 5. Considerações finais

A pesquisa se desenvolveu valendo-se da importância da inclusão na política de indexação de elementos que sustentem e amparem a informação construída como prática social, isto é, que, além de promoverem sintonia entre as concepções teóricas e a prática profissional, preservem os valores, crenças e predisposições específicas da comunidade usuária.

O caminho percorrido foi propício por trazer à tona novas diretrizes de política de indexação que, além de contribuir para que o tratamento temático da informação em contexto de bibliotecas universitárias seja realizado de modo coerente, atrelando os objetivos institucionais aos de sua comunidade usuária, abrem novas possibilidades investigativas de se trabalhar com questões relativas à política de indexação.

Adotou-se o posicionamento de que a sistematização de diretrizes de política de indexação em e para bibliotecas universitárias deve ser conduzida com base nas indicações que apontam para a necessidade de se considerar o usuário como sujeito inserido em um domínio do conhecimento. Então, significa dizer que para ser adequadamente assimilada é preciso admitir que a informação seja situacional e, portanto, os produtos e serviços informacionais gerados nas atividades de organização e representação da informação precisam contemplar esta característica. Logo, a cultura dos usuários precisa ser refletida nos produtos e serviços informacionais gerados, considerando-os como sujeitos coletivos inseridos em um contexto cultural e social de atividade.

Diante desta nova configuração, buscou-se ao longo da pesquisa apontar que a política de indexação acentua de forma regulamentada que, em se tratando da representação do conteúdo do documento, o interesse da comunidade da biblioteca universitária figura acima de outros. Mais do que isso, quando bem delimitada, a política de indexação contribui para que a representação do documento torne a informação significativa para o usuário ao estabelecer passos de orientação para que a informação seja vinculada ao contexto

cultural e social em que ela é tratada, organizada e socializada.

Dentre outros aspectos, o conjunto de diretrizes de orientação para a definição de política de indexação apresentado aqui propicia uma relação mais harmoniosa entre o 'saber' e o 'fazer' no que tange à prática do processo, como forma de salvaguardar o objetivo da biblioteca universitária enquanto instituição direcionada à socialização do conhecimento – atendimento efetivo das necessidades informacionais de seus usuários. Portanto, defende-se que tais diretrizes devam ser assumidas no momento da elaboração e/ou reformulação da política de indexação para o contexto de bibliotecas universitárias que possuam sistemas de recuperação de informação automatizados.

O encaminhamento de pesquisas que busquem novas reflexões sobre as questões que envolvem a política de indexação em abordagem sociocultural é um direcionamento eficaz para se estabelecer, com um mínimo de suporte necessário, diretrizes que preservem as características e particularidades dos sistemas de informação. Os resultados apresentados devem servir como suporte e direcionamento para novas pesquisas avancem nas reflexões aqui encabeçadas e propiciem formas inovadoras de se pensar a política de indexação para biblioteca universitária enquanto contexto social de natureza complexa e dinâmica.

Reconhecendo a responsabilidade que cerca toda e qualquer discussão sobre política de indexação, não há nesta pesquisa a pretensão de esgotar a questão, nem seria tal empreendimento possível, pois esta prática contempla o acompanhamento de fenômenos que se revestem de aspectos sociais, cognitivos e culturais – o fazer profissional, o sistema de informação e o usuário. Todavia, admite-se a necessidade de novas alternativas de produção do conhecimento sobre política de indexação que sejam construídas coletivamente, para que este seja um recurso funcional. Por fim, espera-se que os resultados expostos sirvam aos propósitos práticos da comunidade profissional atuante em bibliotecas universitárias.

## Notas

- (1) Pesquisa de doutorado realizada no Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista (UNESP/Campus de Marília, Brasil) defendida em 2014.

## Referências

Bardin, L. (2009). *Análise de Conteúdo*. Lisboa, Portugal; Edições 70, LDA, 2009.

- Beghtol, C. (2002). A proposed ethical warrant for global knowledge representation and organization systems. // *Journal of Documentation* (London). 58:5 (2002) 507-532.
- Boccatto, V. R. C. (2009). Avaliação do uso de linguagem documental em catálogos coletivos de bibliotecas universitárias: um estudo sociocognitivo com protocolo verbal. 299 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2009.
- Carneiro, M. V. (1985). Diretrizes para uma política de indexação. // *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte. 14:2 (set. 1985) 221-241.
- Carvalho, I. C. L. (2004). A socialização do conhecimento no espaço das bibliotecas universitárias. Rio de Janeiro: Interciência, 2004.
- Cervantes, B. M. N.; Fujita, M. S. L.; Rubi, M. P. (2008). Terminologias em política de indexação. // *Ibersid*. (2008) 211-221.
- Cunha, M. B. da. (2000). Construindo o futuro: a biblioteca universitária brasileira em 2010. // *Ciência da Informação* (Brasília). 29:1 (jan./abr. 2000) 71-89.
- Dal'Evedove, P. R. (2010). A perspectiva sociocognitiva no tratamento temático da informação em Bibliotecas universitárias: aspectos inerentes da percepção profissional. 301 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.
- Dal'Evedove, P. R. (2014). O tratamento temático da informação em abordagem sociocultural: diretrizes de política de indexação em bibliotecas universitárias. 266 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília. 2014.
- Dal'Evedove, P. R.; Fujita, M. S. L. (2015). A pesquisa sobre indexação no Brasil: avanços e desafios. // *Scire*. 17:2 (jul.-dic. 2015) 49-56.
- Fujita, M. S. L.; Gil Leiva, I. (2009). Política de indexación en latinoamerica. // *Ibersid*. (2009) 155-162.
- Fujita, M. S. L.; Santos, L. P. (2016). Política de indexação em bibliotecas universitárias: estudo diagnóstico e analítico com pesquisa participante. // *Transinformação*. 28:1 (jan./abr 2016) 59-76.
- Gil Leiva, I.; Fujita, M. S. L. (orgs.) (2012). *Política de indexação*. São Paulo: Cultura Acadêmica; Oficina Universitária. 2012.
- Guimarães, J. A. C. (2004). As políticas de indexação como elementos para a gestão do conhecimento nas organizações. // Vidotti, S. A. B. G. (Org.). *Tecnologia e conteúdos informacionais*. São Paulo: Polis, 2004. 43-52.
- Guimarães, J. A. C. (2000). Políticas de análisis y representación de contenido para la gestión del conocimiento en las organizaciones. // *Scire* (Zaragoza). 6:2 (jul./dic. 2000) 48-58.
- Kochani, A. P.; Boccatto, V. R. C.; Rubi, M. P. (2012). Política de indexação para clippings: otimização do tratamento temático em coordenadorias de comunicação social brasileiras. *Revista Interamericana de Bibliotecología*. 35:3 (set.-dic 2012) 257-273.
- Lousada, M. et. al. (2011). Políticas de indexação no âmbito da gestão do conhecimento organizacional. // *Informação & sociedade*. 21:1 (2011) 191-202.
- Nardi, M. I. (1999). A metáfora e a prática de leitura como evento social: instrumentos do pensar a Biblioteconomia do futuro. 272 f. Tese (Doutorado em Linguística Aplicada e Estudos da Linguagem) – Pontifícia Universidade Católica, São Paulo, 1999.
- Nunes, C. O. (2004). Algumas considerações acerca da ausência de políticas de indexação em bibliotecas brasileiras. // *Biblos* (Rio Grande do Sul). 16 (2004) 55-61.

- Pinho, F. A. (2010). Aspectos éticos em representação do conhecimento em temáticas relativas à homossexualidade masculina: uma análise da precisão em linguagens de indexação brasileiras. 149 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.
- Redigolo, F. M. et. al. (2012). Elementos de política de indexación para bibliotecas universitarias del área médica. // *Scire* (Zaragoza). 18 (2012) 75-86.
- Rubi, M. P.; Fujita, M. S. L.; Boccato, V. R. C. (2012a). Del conocimiento tácito al explícito: elaboración de un manual de política de indexación en bibliotecas universitarias. // *Scire* (Zaragoza). 18:2 (2012a), 67-74.
- Rubi, M. P.; Fujita, M. S. L.; Boccato, V. R. C. (2012b). Elaboração do manual de política de indexação na formação continuada do catalogador. // Gil Leiva, Isidoro; Fujita, Mariângela Spotti Lopes (Org.). *Política de indexação*. São Paulo: Cultura Acadêmica, 2012b. 217-227.
- Rubi, M. P. (2008). Política de indexação para construção de catálogos coletivos em bibliotecas universitárias. 169 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília. 2008.
- Sanders, P. (1982). Phenomenology: a new way of viewing organizational research. // *The Academy of Management Review*. 7:3 (1982) 353-360.
- 
- Enviado: 2016-04-04. Segunda versão: 2016-06-29.  
Aceptado: 2016-06-29.
-





---

# Representación del conocimiento en historia clínica electrónica interoperable: el caso de la Historia Clínica Digital del Sistema Nacional de Salud de España

*Knowledge representation in electronic interoperable health records: the case of the Digital Health Record of the Spanish National Health System*

---

Gonzalo MARCO CUENCA (1), José Antonio SALVADOR OLIVÁN (2)

(1) Departamento de Biblioteconomía y Documentación. Universidad Complutense de Madrid, España, gmarco@ucm.es  
(2) Departamento de Ciencias de la Documentación e Historia de la Ciencia. Universidad de Zaragoza, España, jaso@unizar.es

## Resumen

La Historia Clínica Digital del Sistema Nacional de Salud (HCDSNS) es uno de los pilares fundamentales para conseguir la interoperabilidad efectiva de la historia clínica electrónica dentro del Sistema Nacional de Salud. El presente artículo analiza el proceso de normalización y los estándares semánticos de referencia que están permitiendo a este sistema, con diez años de vida, aproximarse a una representación del conocimiento de la historia clínica electrónica. En primer lugar, se expone la base teórica y legal en la que se ha apoyado su desarrollo; a continuación, se delimitan y analizan los estándares de referencia propuestos para alcanzar la interoperabilidad semántica; y, finalmente, se revisan los recursos más significativos que han sido desarrollados para tal fin.

**Palabras clave:** Historias clínicas electrónicas. Interoperabilidad semántica. SNOMED CT. Modelos clínicos. Terminologías clínicas. Ontologías. Normalización. Historia Clínica Digital. HCDSNS. Sistema Nacional de Salud (España). España.

## 1. Introducción

En el contexto de la información sanitaria, los sistemas de historia clínica electrónica son un pilar básico para el conocimiento de la actividad clínico-asistencial. La historia clínica se compone de un conjunto de documentos que recopilan datos, valoraciones y otras observaciones sobre la situación clínica de un paciente y su evolución a lo largo del proceso asistencial.

Estos registros documentales, de forma independiente, atienden a contextos específicos asistenciales y se originan a partir de la relación de los propios pacientes con los profesionales sanitarios que los atienden, constituyéndose como la principal evidencia de la actividad asistencial y como un recurso indispensable para asegurar una prestación eficiente de los servicios sanitarios de una forma continuada. Por este motivo, la

## Abstract

The Electronic Health Records of the National Health System (HCDSNS) of Spain is one of the fundamental pillars for achieving the effective interoperability of digital medical records within the National Health System. This article analyses the process of standardization and the semantic standards of reference that are allowing this ten-year old system to represent knowledge in electronic medical records. Firstly, the theoretical and legal basis that support the system development is exposed. Below are identified and analysed the reference standards proposed to achieve semantic interoperability, and finally the most significant resources that have been developed for this purpose are reviewed.

**Keywords:** Electronic Health Records. Semantic interoperability. SNOMED CT. Clinical models. Clinical terminologies. Ontologies. Standardization. HCDSNS. National Health System (Spain). Spain.

historia clínica se convierte en el eje de la información clínica del paciente y en un instrumento fundamental para la comunicación entre los diferentes profesionales.

Sin embargo, la historia clínica electrónica puede encontrarse, en el conjunto de nuestro Sistema Nacional de Salud (SNS), fragmentada en diversos sistemas de información autónomos y aislados; dicha situación viene a dificultar la tarea de los profesionales sanitarios, pudiendo provocar un acceso incompleto a cierta información que podría ser relevante para la evaluación y toma de decisiones clínicas y, por consiguiente, conducir a una inadecuada atención y un riesgo potencial para la salud y seguridad de los pacientes.

El hecho es que cada proveedor de salud ha introducido, a partir de sus propias necesidades,

mejoras considerables en sus sistemas de información con el objetivo de optimizar la gestión de la información clínica y administrativa de los pacientes, siendo desarrolladas atendiendo a necesidades y entornos locales que han hecho que la información de un mismo paciente pueda encontrarse distribuida en diferentes lugares, de forma confinada, y sin posibilidad de compartirse.

Hay que tener en cuenta que las administraciones sanitarias, tanto a nivel nacional, como autonómico, tienen entre sus responsabilidades hacer efectivo el derecho de los ciudadanos a la protección de la salud. Tanto la Constitución Española, en su artículo 43, como la Ley 14/1986, de 25 de abril, General de Sanidad, en su título preliminar, recogen este derecho a la protección de la salud. Concretamente la Ley 14/1986, en su artículo 1.2, establece que son titulares de este derecho todos los españoles y los ciudadanos extranjeros que tengan su residencia en el territorio nacional (Ministerio de Sanidad y Consumo, 2004). Para ello, han de facilitar los recursos necesarios para que los profesionales sanitarios puedan realizar su trabajo con las máximas garantías. Existen preceptos legales que establecen los mecanismos necesarios para llevar a efecto el intercambio electrónico de la información sanitaria, concretamente la *Ley 16/2003 de cohesión y calidad del SNS* en su artículo 56 y la *Ley 41/2002 básica reguladora de la autonomía del paciente* en su disposición adicional tercera vienen a señalar al Ministerio de Sanidad como el ente coordinador de las acciones que deben hacer posible el intercambio de la información clínica dentro del SNS, proveyendo el acceso a ciudadanos y profesionales, en los términos estrictamente necesarios que garanticen la calidad de la asistencia y la confidencialidad e integridad de la información.

Cabe también destacar la *Directiva 2011/24/UE del Parlamento y del Consejo Europeo* relativa a la aplicación de los derechos de los pacientes a una asistencia sanitaria transfronteriza que tiene el objetivo de garantizar la movilidad de los pacientes, establecer unas reglas para facilitar su acceso a una asistencia sanitaria segura y de alta calidad en la Unión Europea y promover la cooperación en materia de asistencia sanitaria entre los Estados miembros, respetando plenamente las responsabilidades de estos en la organización y prestación de dicha asistencia.

La lógica asistencial, pues los pacientes se desplazan y pueden requerir asistencia fuera de su Comunidad Autónoma de origen, los derechos de los pacientes y usuarios, la necesidad y responsabilidad de los profesionales y las obligaciones legales resultan ser razones ineludibles para

afrontar el reto de la interoperabilidad de la historia clínica electrónica en nuestro país (Etreros y otros, 2009a). Para solucionar este problema de intercambio y comunicación de la historia clínica electrónica entre los diferentes agentes del SNS, el Ministerio de Sanidad comenzó en el año 2006 la puesta en marcha del proyecto de Historia Clínica Digital del Sistema Nacional de Salud (HCDSNS) (Ministerio de Sanidad y Política Social, 2010).

La HCDSNS pretende responder a esta necesidad garantizando a los ciudadanos y a los profesionales sanitarios el acceso a aquella información clínica relevante para la atención sanitaria de un paciente desde cualquier lugar del SNS, asegurando a los ciudadanos que el acceso a sus datos queda restringido a quien está autorizado para ello. Este proyecto fue enmarcado en sus inicios en el conjunto de proyectos que contribuían a desarrollar la estrategia de *Sanidad en Línea* enunciada en el Plan de Calidad para el SNS y que pretendían contribuir a la generalización del uso de las nuevas tecnologías en el sistema sanitario para mejorar la atención a pacientes y ciudadanos (Etreros y otros, 2009b). Posteriormente, también ha sido incluido entre los proyectos clave para la estrategia de salud electrónica para el sistema sanitario, entre los que consideran que la utilización de las Tecnologías de la Información y de la comunicación son imprescindibles para afrontar los retos sanitarios (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2014).

El proyecto mantiene desde su origen un marcado propósito asistencial, aunque como se ha resaltado se circunscribe entre los proyectos vigentes del SNS que mantienen como intención principal alcanzar la interoperabilidad —como *la Tarjeta Sanitaria Individual y la Receta Electrónica del SNS*—, entendiendo la misma no solamente como una cuestión tecnológica, sino como una cualidad necesaria y exigible al propio sistema de información para poder ofrecer los servicios que son requeridos. Tendríamos también que decir que la HCDSNS es en sí misma un ambicioso proyecto en el que se fijan las bases comunes para la interoperabilidad de la historia clínica a nivel nacional, así como las infraestructuras necesarias para acceder a ella desde cualquier sistema de salud autonómico. Para ello se exige la adopción de elementos de interoperabilidad entre todos los servicios de salud y la puesta en común de criterios de normalización para la información, conjuntamente con el uso de estándares y políticas como la principal estrategia para conseguirla.

Diez años después, la HCDSNS se ha convertido en un sistema real, alcanzando una cobertura

significativa para los perfiles de servicio previstos, emisores y/o receptores de la información, en la mayor parte de los Servicios de Salud del SNS. Hay que señalar como un importante logro que la cobertura de población con referencias HCDSNS en cifras absolutas es de 35.726.423 ciudadanos, y en relación con la población con Tarjeta Sanitaria Individual es del 77,48% (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2017).

Actualmente, el sistema permite intercambiar documentos en diferentes formatos, esencialmente en formatos de documento portátil o PDF, en un total de 16 Comunidades Autónomas y el IN-GESA —el Instituto Nacional de Gestión Sanitaria encargado de la gestión de las prestaciones sanitarias en el ámbito territorial de las Ciudades de Ceuta y Melilla—, lo que implica que el servicio se encuentra activado al menos con el perfil de emisor en 17 Servicios de Salud sobre un total de 18 (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2017). Sin embargo, aún no se ha superado la barrera de un intercambio eficaz y completo a nivel de datos estructurados que facilite la interoperabilidad semántica, donde se consiga alcanzar una transferencia del conocimiento clínico de una forma consistente, multilingüe y normalizada y preparada para que los sistemas puedan ser capaces de producir inferencia.

## 2. Interoperabilidad y normalización de la historia clínica electrónica

En los proyectos de interoperabilidad sanitaria se deben atender diferentes requisitos o factores que los aproximen a su viabilidad. Estos elementos mantienen una relación directa con las diferentes dimensiones de la interoperabilidad: *técnica, sintáctica, semántica y organizativa* (Muñoz Carrero, 2013). Hay que considerar también la importancia que adquiere, dentro de cualquier proyecto de intercambio de información, el establecimiento de un marco común para la interoperabilidad, entendiendo el mismo como un enfoque acordado por las organizaciones que desean trabajar unidas hacia la prestación conjunta de servicios públicos. Dentro de su ámbito de aplicabilidad, este marco debe definir el conjunto de elementos comunes de uso consensuado, tales como: *principios, políticas, directrices, normas, vocabularios, especificaciones y prácticas* (Deloitte, 2013).

La información clínica se construye a partir del registro de los hechos, lo que da origen a datos que se van ordenando y agregando en los sistemas. En la medida que esta información se elabora y contextualiza se transforma en conocimiento que la organización puede aprovechar

(Indarte y Vero, 2014). Para que esto sea posible, la normalización de los registros clínicos resulta ser un paso previo necesario y también una exigencia para alcanzar la interoperabilidad semántica.

En el contexto de la HCDSNS, la normalización de estos registros clínicos se ha tenido que realizar alcanzando un consenso profesional e institucional que ha venido a determinar el alcance de los datos a intercambiar, su estructura y el origen de su contenido (Ministerio de Sanidad y Política Social, 2010). Esta labor ha quedado reflejada con efecto legal en el *Real Decreto 1093/2010, por el que se aprueba el conjunto mínimo de datos de los informes clínicos a intercambiar en el SNS* y que viene a definir los datos imprescindibles que deben contener los informes clínicos de uso frecuente en el SNS, cualquiera que sea su soporte, electrónico o papel. Para poder ver la trascendencia de esta norma hay que echar la vista atrás y observar que desde la *Orden Ministerial de 6 de septiembre de 1984, por la que se regulaba la obligatoriedad del informe de alta*, no se había generado ninguna otra norma de estas características a nivel nacional.

La definición de este conjunto mínimo de datos comenzó siendo un compromiso de colaboración y de trabajo en el que participaron expertos de 32 organizaciones, entre las que se encontraban sociedades científicas, asociaciones de ciudadanos, miembros designados por las Comunidades Autónomas y por el Ministerio de Sanidad. Los profesionales debían tener un perfil predominantemente sanitario, con amplia experiencia asistencial y encontrarse en activo. Además, debían ser usuarios de alguno de los sistemas de información en producción dentro del SNS. También se incluyó entre estos grupos de trabajo a gestores de sistemas de salud y a responsables de unidades de admisión y documentación clínica. De esta forma, se nombraron 60 representantes para constituir un total de 7 grupos de trabajo (Ministerio de Sanidad y Política Social, 2010).

Participantes	Resultados
32 organizaciones	8 informes clínicos
60 representantes	432 variables definidas

Tabla I. Participantes y resultados en el desarrollo del conjunto mínimo de datos para los informes clínicos del SNS.

El plan de trabajo trazaba como requerimiento principal la definición de un conjunto mínimo de datos que pudiera garantizar la interoperabilidad de la historia clínica electrónica dentro del

alcance determinado para la HCDSNS. El reto, por tanto, consistía en conseguir por consenso una normalización de los modelos de datos a compartir entre todos los agentes del SNS que hiciese posible la interoperabilidad y, fundamentalmente, consiguiese un marco común para la interoperabilidad semántica de la información.

En tal sentido, el propio Real Decreto 1093/2010 viene a decir *“uno de los pasos imprescindibles para alcanzarla es definir, por acuerdo, el conjunto de datos que, por su relevancia, deben estar contenidos en los diferentes informes clínicos que describen los procesos de atención sanitaria realizados a ciudadanos concretos en cualquier centro o servicio del Sistema Nacional de Salud. Esta homogeneidad es uno de los elementos de normalización que facilitan el intercambio entre sistemas diferentes al servicio de los ciudadanos”*.

Con esta contribución el SNS se ha provisto de un conjunto de datos relevantes para la representación de la información clínica de una forma sistematizada y homogénea garantizando unos niveles óptimos para que se pueda intercambiar entre los diversos centros y dispositivos asistenciales que integran el sistema. Además, se han definido para las distintas variables que integran cada modelo clínico las propiedades y los tipos de valores admitidos o recomendados.

Hay que subrayar que esta disposición legal se ha convertido en un excelente ejemplo del resultado de transformación de un conocimiento tácito, basado en la experiencia de los profesionales, en un conocimiento explícito, formalizado y documentado, lo que ha permitido reconocer el contexto, los participantes, las acciones y relaciones de un conjunto relevante de actos clínicos para que puedan ser procesables y transmisibles por medio de los sistemas de información. No obstante, siendo un documento imprescindible para el desarrollo de la HCDSNS hay que matizar que se trata de una definición funcional a muy alto nivel, que no concreta estándares para la comunicación de la historia clínica electrónica y tampoco para la representación y organización del conocimiento clínico con fines asistenciales.

Por este motivo y para fijar las bases de una estrategia compartida en torno al desarrollo de la interoperabilidad semántica, el Ministerio de Sanidad publicó, en el año 2014, el documento *“Políticas de Normalización de Terminologías. Principios y estrategias de interoperabilidad semántica para los sistemas de información clínicos del SNS”* (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2014b). Estas políticas formulan una serie de principios y estrategias comunes para el

SNS con objeto de alcanzar una interoperabilidad semántica efectiva, tratando de definir aquellos principios básicos, estrategias factibles y directrices claras de normalización terminológica que faciliten la consecución de la interoperabilidad de los sistemas de información clínicos, así como la descripción de los mecanismos de coordinación a aplicar para obtener el máximo rendimiento del esfuerzo compartido de desarrollo y despliegue de los recursos de información.

El documento expone una serie de principios comunes entre los que se incluyen la conveniencia de no duplicar esfuerzos innecesariamente, que los elementos utilizados en la normalización deben partir de modelos de referencia definidos y la necesidad de compartir recursos de terminología que sean reutilizables.

Principio	Fundamento
1	Se necesita que la información clínica sea interoperable
2	No se deben duplicar esfuerzos innecesariamente
3	Los elementos utilizados en la normalización deben partir de modelos de referencia definidos
4	Las políticas de normalización necesitan una infraestructura de servicios compartida
5	Se necesitan recursos compartidos y reutilizables de terminología
6	Contribuir a la implementación de las políticas de normalización en el SNS

Tabla II. *Principios de interoperabilidad semántica para los sistemas de información clínicos del SNS (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2014b)*

En el caso de las estrategias de interoperabilidad, las mismas se concretan en el desarrollo de modelos, terminologías, enlaces terminológicos, extractos de documentación clínica y sistemas de mensajería, así como en la relación entre todos estos componentes. Además, el documento especifica que para poder llevar a cabo las estrategias formuladas, se pondrán a disposición de los agentes del SNS diferentes recursos y servicios, los cuales irán acompañados de forma general de canales para su difusión, planes de despliegue y adopción, planes de formación y documentación.

Estrategia	Fundamento
1	Creación y difusión de modelos de información clínica de referencia
2	Creación de una red colaborativa de servidores terminológicos
3	Formación sobre los recursos de normalización terminológica
4	Regulaciones
5	Cooperación coordinada
6	Certificación

Tabla III. Estrategias de interoperabilidad semántica para los sistemas de información clínicos del SNS (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2014b)

Las dos primeras estrategias se encuentran actualmente en un estado muy avanzado y toman como base dos estándares de referencia, como son: la norma UNE-EN 13606 para la comunicación de la historia clínica electrónica y la terminología clínica Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT).

### 2.1. Normalización de los modelos clínicos

La norma UNE-EN ISO 13606 ha sido desarrollada con el fin de facilitar el intercambio de información clínica entre las distintas organizaciones sanitarias. Esta norma especifica unas características comunes para la historia clínica electrónica con objeto de que puedan ser incorporadas a modelos de información y a mensajes interpretables por los sistemas clínicos, permitiendo intercambiar entre los diferentes sistemas extractos o historias clínicas completas de un paciente de una forma normalizada.

La meta global de la norma, según se indica, es definir una arquitectura de información rigurosa y estable para comunicar parte o toda la historia clínica electrónica de un único sujeto de la asistencia o paciente, con objeto de dar soporte a la interoperabilidad entre sistemas mediante la comunicación a través de mensajes electrónicos de los componentes de la historia clínica electrónica (UNE-EN ISO 13606-1, 2013).

La norma no se encarga de especificar una arquitectura interna para las bases de datos de historia clínica. Mantiene como propósito el intercambio de información para que la interoperabilidad pueda darse. Por esta razón, el modelo de información que propone se va a utilizar principalmente para la definición de mensajes, documentos, esquemas XML o para la definición de alguna interfaz.

La norma está basada en un modelo dual de arquitectura para la historia clínica electrónica, que

consiste en separar la información del conocimiento gestionado por los sistemas de información. El modelo alberga para ello dos esquemas perfectamente diferenciados, aunque relacionados entre sí, denominados: Modelo de Referencia y Modelo de Arquetipos.

El Modelo de Referencia se centra en las características comunes que se repiten en todas las historias clínicas. Este modelo define, a partir de una serie de clases, las estructuras para organizar cualquier tipo de información, incluyendo asimismo la información de contexto. Se trata de un modelo simple y flexible, ya que puede ser adaptado a cualquier situación o estructura de información. La norma nos viene a indicar que la información de una historia clínica electrónica es intrínsecamente jerárquica, de lo que pueden deducirse clases que van a poder ser completadas siguiendo una jerarquía preestablecida.

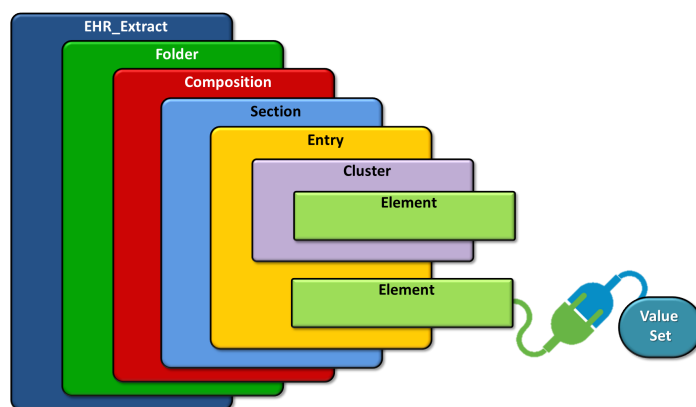


Figura 1. Representación de las clases del Modelo de Referencia de la norma UNE-EN 13606

El Modelo de Referencia se ocupará, por tanto, de representar la información a partir de las clases definidas, preservando el significado de las historias clínicas cuando se comunican entre sistemas heterogéneos. Para ello, subdivide formalmente en partes las diferentes jerarquías de la historia clínica electrónica proporcionando una correspondencia consistente con las formas de organización de los documentos o registros de la historia clínica.

Por otro lado, el Modelo de Arquetipos se va a ocupar de representar los conceptos clínicos de mayor nivel semántico, siendo definidos esencialmente por los profesionales de la salud. Estas definiciones van a requerir un alto grado de calidad dado que influirán en la forma en la que los datos de salud serán capturados, procesados, comunicados y visualizados. Con los arquetipos se pueden definir de una manera formal conceptos clínicos, como puede ser un informe de alta o una

prueba de laboratorio, basándose para su construcción en las clases del modelo de referencia y restringiéndolas a estructuras de datos precisas.

En este contexto, un arquetipo es una representación de una realidad clínica que se establece por medio de una serie de metadatos que van a ser utilizados para determinar las estructuras clínicas y las características particulares de las mismas, con objeto de cumplir con los requisitos de información de un grupo de profesionales sanitarios, de una especialidad, de una unidad o de un servicio. Para Kalra (2006) un arquetipo es la definición formal de las combinaciones establecidas a partir de las clases del Modelo de Referencia para determinados dominios u organizaciones clínicas. Se trata de una definición formal que utiliza un lenguaje procesable por ordenador, *Archetype Definition Language (ADL)* y contiene una semántica clara.

Un arquetipo también podría ser considerado desde una perspectiva más simple, como una definición conceptual de todos los elementos estructurales que conforman un formulario o plantilla para un registro de datos clínicos en un contexto determinado, que se establece a partir del estudio minucioso de unas necesidades asistenciales determinadas por un grupo de expertos de dominio.

Por lo tanto, el Modelo de Arquetipos nos permite representar el conocimiento del dominio, facilitando la representación de estructuras basadas en el conocimiento clínico utilizando las piezas del modelo de referencia. Cada arquetipo, por tanto, va a definir y limitar de forma efectiva las clases del Modelo de Referencia, mediante la especificación de los nombres de cada componente, tipos de datos, valores y rangos de esos valores. En cierto modo, los arquetipos van a servir de puente para unir dos orillas, la del conocimiento clínico y la de los sistemas, ya que permiten trasladar el conocimiento de los expertos a los sistemas sirviéndose de las clases del modelo de referencia (Romero y Marco, 2014).

Para una mejor comprensión de ambos modelos podríamos pensar en una sencilla analogía con los juegos de construcción de Lego, el Modelo de Referencia va a representar a las piezas que son necesarias para poder construir cualquier objeto y el Modelo de Arquetipos representará los planos o instrucciones que vienen a describir aquello que queremos hacer. Ambos modelos van a hacer posible que lo pensado pueda ser trasladado a la realidad.



Figura 2. Analogía Lego para la representación del Modelo de Referencia y el Modelo de Arquetipos de la norma UNE-EN-ISO 13606

El modelo dual parte de la idea de separar la información del conocimiento. El conocimiento es representado por medio de los arquetipos, cuya definición conceptual es especificada por los profesionales sanitarios, y la información necesaria para la representación de ese conocimiento en una estructura de datos se lleva a cabo a través de las clases del Modelo de Referencia, que son implementadas en los sistemas por profesionales informáticos. De este modo, la norma implica a diferentes roles de profesionales haciendo que sea imprescindible su participación y colaboración para conseguir los resultados más óptimos.

Otra característica lógica de estas definiciones de conceptos expresadas como arquetipos es que pueden evolucionar a lo largo del tiempo, sin afectar a los sistemas clínicos. El Modelo de Referencia se configura como una parte estable y no variable, mientras que los arquetipos constituyen aquella parte variable de las historias clínicas que puede adaptarse a las necesidades de cada dominio clínico. Tal y como indica Cáceres Tello y otros (2013) “*separando la información del conocimiento, se consigue que el sistema sea capaz de evolucionar y adaptarse de forma sencilla y automática a los cambios que se produzcan*”.

Asimismo, los arquetipos pueden enlazarse con terminologías para proporcionarles una capa semántica que pueda asegurar su comprensión y optimice la interoperabilidad semántica. Este proceso de enlace terminológico, durante el modelado de arquetipos, suele consistir en la asignación de conceptos aceptados, a modo de conjuntos de valores, para restringir los valores admitidos en un elemento determinado del modelo. Siendo habitual que un arquetipo limite el registro de información a un elemento, por medio de un nodo denominado *<term binding>*, usando un conjunto de valores controlados que forman parte, por lo general, de un recurso o terminología externa. Estos valores van a aportar también una importante capa de conocimiento, incluyendo conceptos previamente seleccionados y validados por los usuarios para el registro en un campo de información o variable concreta.

La implementación de estos enlaces permitirá al usuario la selección del significado de una forma



precisa y dependerá del implementador la decisión sobre la mejor forma de presentación de estos valores en la interfaz (por medio de listas desplegables, buscadores de texto predictivo, casillas de verificación, etc.). Esta implementación debe atender a criterios de usabilidad, ergonomía y ahorro de tiempo para el usuario.

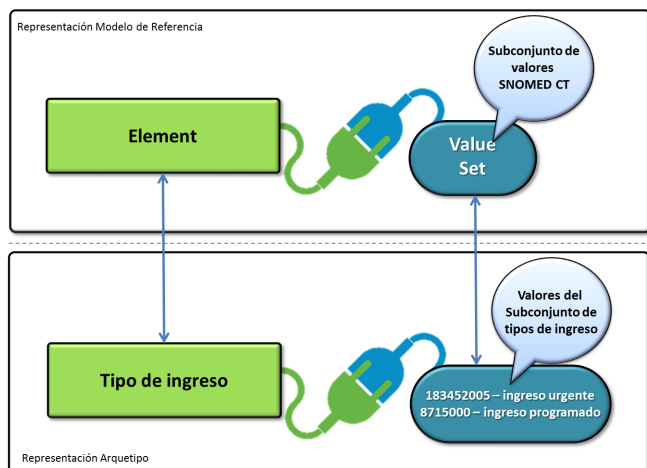


Figura 3. Ejemplo de representación del elemento <tipo de ingreso> por medio del binding o enlace terminológico

## 2.2. Recursos de modelado clínico

Esta estrategia se concreta con la publicación, por parte del Ministerio de Sanidad, de una serie de arquetipos de referencia. Cada uno de ellos consiste en un modelo básico de un componente comunicable de la historia clínica electrónica, que establece las directrices comunes aplicables dentro del SNS (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2014c). Estos arquetipos de referencia pueden ser refinados y extendidos por parte de los diferentes agentes del SNS, teniendo en cuenta las necesidades de cada territorio, centro, especialidad o proceso clínico.

Los recursos iniciales de este proyecto de modelado clínico se han centrado en arquetipos de referencia relacionados con el ya mencionado Real Decreto 1093/2010 y con el Real Decreto 1718/2010 sobre receta médica y órdenes de dispensación (2011). Estos arquetipos se encuentran en un fichero comprimido y disponible para su descarga. Se representan en lenguaje ADL e incluyen un catálogo y una representación esquemática por medio de mapas mentales, que facilitan su interpretación. Estos materiales se han elaborado siguiendo un diseño modular, con el fin de reutilizar de forma óptima los componentes más sencillos para la construcción de componentes más complejos (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2014d).

Clase	Nº Arquetipos
Composition	10
Section	7
Entry	38
Cluster	3
Total	58

Tabla IV. Arquetipos de referencia publicados por el Ministerio de Sanidad atendiendo a la clase del modelo de referencia

La intención es conseguir una base de conocimiento en la que la participación y la colaboración de expertos sea una fuente constante que pueda impulsar el progreso y la generación de nuevos arquetipos. Para que puedan evolucionar estos recursos, el Ministerio de Sanidad, en el contexto de la HCDSNS, ha desarrollado una plataforma de colaboración denominada Servidor de Objetos de Modelado para el SNS (somSNS), que consiste en una herramienta web que permite compartir arquetipos y otros recursos de modelado con los diferentes agentes del SNS. El servidor se encuentra aún en fase de desarrollo y permitirá, además del acceso y colaboración de expertos, la distribución de los diferentes recursos de modelado (Romero y Marco, 2014).

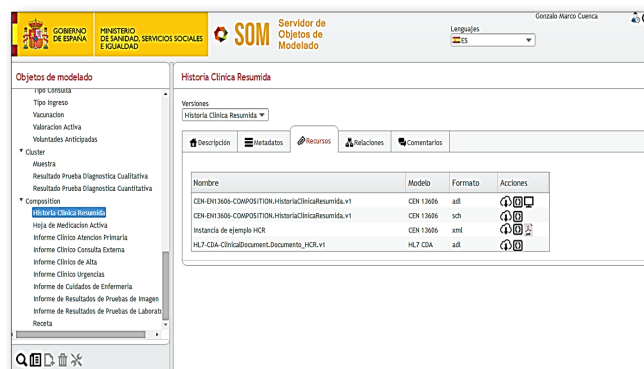


Figura 4. Interfaz del sistema somSNS (Romero y Marco, 2014)

## 2.2. Normalización terminológica

Tal y como indica Sempere (2001) “el uso apropiado y preciso de un vocabulario específico sobre un área de conocimiento es crucial para la comunicación entre los especialistas de ese campo, y la medicina no es una excepción”. La Medicina, como otras disciplinas del saber, es una ciencia y una profesión. Su lenguaje especializado, desde ambas perspectivas, científica y profesional, sirve de vehículo de transmisión del

conocimiento, lo que exige por un lado, una fidelidad absoluta a la realidad, y por otro, la comunicación con precisión de los conceptos propios de esa área de conocimiento.

Por lenguaje clínico vamos a interpretar aquel lenguaje que es usado para la comunicación del resultado de las observaciones y valoraciones que se producen durante cualquier episodio asistencial. Se trata de un lenguaje que es producto de la práctica clínica y que tiene como objetivo el registro veraz de la información sobre la salud de los pacientes. Hay que considerarlo como el lenguaje de trabajo del profesional de la salud, en su ámbito correspondiente, y queda expresado en el conjunto de documentos que conforman la historia clínica.

Es importante indicar que los lenguajes de especialidad tratan de diferenciarse del lenguaje común esencialmente por su capacidad de precisión y objetividad, manteniendo como objetivo común proporcionar una comunicación fluida y sin trabas entre los especialistas de una materia determinada. No obstante, tal y como precisaban los profesores López Piñero y Terrada (2005) el lenguaje médico no siempre se adapta a un criterio lógico uniforme, y además participa de las imprecisiones y ambigüedades propias del lenguaje natural. Estas imprecisiones y ambigüedades forman también parte del lenguaje clínico y, en muchos casos, se producen debidas a ciertos problemas y vicios lingüísticos y a una evidente falta de tiempo del profesional para una redacción más precisa.

Para Aleixandre y Amador (2001) *“en la comunicación médica se producen numerosas incorrecciones que provocan pérdida de precisión y claridad de los mensajes transmitidos y ponen en peligro la investigación y educación en ese campo”*. A estas afirmaciones habría que añadir, con relación a la interoperabilidad semántica de la historia clínica electrónica, que también se pone en peligro la calidad y seguridad de la atención de los pacientes, la salud pública y la gestión.

Representar el conocimiento clínico, evidentemente, no es una tarea sencilla. En el ámbito de la salud esta tarea es particularmente compleja por el uso de este lenguaje clínico sumamente especializado, que evoluciona con gran rapidez y que tiene una gran dependencia del contexto. Pero, además, el lenguaje clínico presenta una serie de problemas y vicios habituales, como son: la sinonimia y la polisemia, la abundancia de anglicismos, el excesivo abuso de abreviaturas, las siglas y acrónimos no normalizados, la elipsis, el excesivo uso de epónimos, los localismos, etc.

La necesidad de normalización terminológica surge, esencialmente, por las necesidades de información para la gestión y a partir de las propias necesidades individuales de los profesionales como mejora de su labor asistencial. En ambos casos estas necesidades de información van a encontrarse vinculadas a priori a demandas relacionadas con la recuperación de la información, el análisis estadístico con fines epidemiológicos y/o de investigación y con procesos relacionados con la gestión y la logística asistencial. La uniformidad terminológica se hace necesaria e imprescindible para alcanzar algunos objetivos científicos como, por ejemplo, comparar datos epidemiológicos extraídos por diferentes equipos, que solo va a ser posible utilizando un lenguaje común.

En un contexto interoperable, el mero intercambio de datos entre los sistemas de información usando el lenguaje natural no sería suficiente. El objetivo es conseguir que los sistemas logren interpretar los datos transmitidos de una misma manera y que las personas que deben recibirlos e interpretarlos lo puedan hacer de forma fehaciente. Por este motivo, un lenguaje controlado adquiere la condición de elemento clave para la interoperabilidad semántica en los sistemas de información sanitarios.

Una terminología es una colección de términos y conceptos organizados por su equivalencia de significado y por unas relaciones básicas previamente definidas. Las terminologías suelen incluir reglas de asociación entre conceptos, así como reglas editoriales relacionadas con la adición, combinación y evolución de sus componentes. Se constituyen como instrumentos que incluyen un gran número de conceptos y que admiten múltiples descripciones para cada uno de ellos. Los conceptos se definen, principalmente, por sus relaciones jerárquicas proporcionando una herramienta rica en datos, información y conocimiento. Además, cada concepto suele estar formado por una o varias descripciones, a modo de sinónimos, que se encargan de expresar su significado.

Por otro lado, suele ser común la identificación de estos conceptos por medio de códigos que aportan un identificador unívoco a cada elemento de la terminología. Todo ello con objeto, como vocabularios controlados, de ser precisos y evitar la ambigüedad y la redundancia.

En el caso específico de una terminología clínica podemos decir que es un conjunto de términos estructurados y normalizados que busca servir de instrumento para el registro de datos clínicos, como base para otras posibles investigaciones o



como medio de intercambio de información clínica entre profesionales para la atención de la salud de los pacientes.

En el dominio de las terminologías clínicas destaca de forma sobresaliente SNOMED CT. SNOMED CT es la terminología clínica integral, multilingüe y codificada de mayor amplitud, precisión e importancia desarrollada en el mundo, diseñada para un uso específico en registros clínicos y, fundamentalmente, para la historia clínica electrónica. SNOMED CT es una terminología cuyos derechos de autor son propiedad de una organización internacional sin ánimo de lucro, la International Health Terminology Standards Development Organisation (IHSTDO). España forma parte de esta organización desde el año 2009, lo que permite la distribución de este activo semántico dentro de nuestro territorio nacional de una forma gratuita a través del Área de descarga de SNOMED CT del Ministerio de Sanidad (s. f.).

Uno de los objetivos principales de SNOMED CT es facilitar un registro exacto de la información clínica y de otros temas relacionados con la salud, para proporcionar interoperabilidad semántica a los documentos clínicos. La terminología SNOMED CT provee un lenguaje común que permite una forma consistente de codificación, recuperación y análisis de datos clínicos. Su contenido busca cubrir la totalidad de la práctica clínica, incluyendo conceptos relacionados con la salud (hallazgos clínicos, patologías, organismos, sustancias, procedimientos, especímenes, etc.) y otros datos de interés en un registro clínico (datos administrativos, contextos sociales, eventos, calificadores, localizaciones ambientales, categorías profesionales, atributos, elementos de registro, dispositivos, escalas de evaluación, etc.).

Es considerada una nomenclatura integral, al ofrecer contenido clínico con suficiente expresividad y detalle para el registro y comunicación eficiente de la documentación clínica. Se trata también de un recurso validado científicamente por equipos de trabajo compuestos por clínicos, informáticos, investigadores y documentalistas. Es importante indicar que se trata de un recurso terminológico con una finalidad operativa y su misión es servir como herramienta útil para la práctica clínica.

En cuanto a su cobertura, su alcance va más allá de la Medicina y todas sus especialidades, pues incluye conceptos relacionados con el área de Enfermería, Farmacología, Biología, Laboratorio, Imagen clínica, etc. Cuenta además con conceptos que trascienden el uso humano, disponiendo de una extensión dedicada a la Veterinaria. SNOMED CT se compone de más de

311.000 conceptos activos con significados únicos y definiciones basadas en lógicas formales organizadas en jerarquías. Cuenta, además, con aproximadamente unas 900.000 descripciones y más de un 1.200.000 relaciones.

SNOMED CT es una terminología basada en conceptos y son, por tanto, los elementos clave de este recurso terminológico. En el contexto de SNOMED CT un concepto es un significado clínico identificado mediante un código o identificador numérico único, permanente, no modificable y no reutilizable. Este código de concepto, o identificador numérico único de concepto, se denomina *conceptID* y se presenta como una secuencia de dígitos que no transmite ninguna información acerca del significado o naturaleza del concepto. Su significado se representa de manera legible para las personas por medio de descripciones.

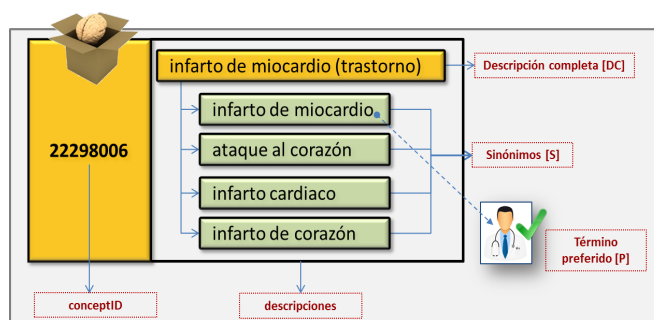


Figura 5. Ejemplo de representación del concepto "infarto de miocardio (trastorno)" en SNOMED CT

Cada concepto equivale a un significado clínico único. Los conceptos vienen a expresar unidades de conocimiento o pensamiento que se expresan mediante descripciones, que son las unidades del lenguaje que tienen como objetivo representar o describir un mismo concepto. Cada concepto, por tanto, puede ser representado por una o más descripciones. Lo que nos permite interpretar cada concepto como un pequeño contenedor de conocimiento que se identifica de forma unívoca por un código y que incluye diferentes descripciones que lo definen.

Los conceptos de SNOMED CT destacan por su alto nivel de detalle o de granularidad, aunque se pueden encontrar conceptos básicamente a dos niveles de profundidad: *nivel de agregación*, que constituye un nivel general que se encarga de definir y agrupar conceptos clínicos; y *nivel de detalle clínico*, un nivel detallado y específico de definición de conceptos clínicos. Los elementos al nivel de detalle clínico suelen ser usados como valores en los registros clínicos.

Los elementos que forman parte del *nivel de agregación* son considerados como conceptos de baja granularidad y los conceptos de *nivel de detalle clínico* son considerados de alta granularidad, al último concepto de este nivel se le considera como el de máxima granularidad.

Las descripciones de los conceptos son términos o frases vinculadas a cada concepto que sirven para describir su significado y, por lo general, se presentan a modo de sinónimos o variaciones léxicas de un mismo concepto. Todos los componentes de SNOMED CT utilizan una gramática para su composición y unas reglas editoriales.

Los conceptos se definen formalmente por sus relaciones con otros conceptos. Todos los conceptos activos en SNOMED CT (excepto el concepto raíz SNOMED CT) mantienen al menos una relación con un concepto *supertipo* o concepto *padre*. Las relaciones o “definiciones lógicas” otorgan una carga de significado explícito que viene a completar a cada concepto.

Las relaciones se establecen por medio de tripletas de componentes, basados en un mismo patrón semántico compuesto por: *un concepto origen + un atributo = un concepto de destino* (donde todos los componentes son representados por conceptos de SNOMED CT). La representación ontológica de estas tripletas básicas, que mantienen una misma interpretación en el contexto de SNOMED CT, se basan en el patrón: *nombre + atributo = predicado*. Por ejemplo, un concepto “*infarto de miocardio (trastorno)*” es descrito por medio del concepto “*es un[a] (atributo)*”, que establece una relación jerárquica con un concepto padre o ascendente como puede ser la “*cardiopatía isquémica (trastorno)*”. En otras palabras, por medio de estas relaciones podemos decir que un “*infarto de miocardio es una cardiopatía isquémica*”.

Las relaciones no son únicas, pudiendo establecerse una o varias relaciones de tipo jerárquico para un mismo concepto. Estas relaciones son conocidas como relaciones de “*supertipo-subtipo*” o “*relaciones padre-hijo*”. Además de este tipo de relaciones jerárquicas, existen otras relaciones, denominadas *calificadoras*, basadas en diferentes atributos. Estas relaciones siguen el mismo patrón de la tripleta y pueden relacionar conceptos de diferentes jerarquías de SNOMED CT. Siguiendo con el mismo ejemplo, el concepto “*infarto de miocardio (trastorno)*” puede tener un atributo calificador denominado “*morfología asociada (atributo)*” que lo relaciona con el concepto “*infarto (anomalía morfológica)*”.

El uso de formalismos basados en la lógica descriptiva utilizada en SNOMED-CT permite reali-

zar inferencias y extraer nuevas relaciones a partir de las ya existentes (Maldonado Segura, 2013). Cada concepto en SNOMED CT debe, por tanto, ser modelado atendiendo a sus relaciones para poder ser interpretado correctamente.

Actualmente, SNOMED CT dispone de más de 50 atributos definitorios para modelar las relaciones entre los conceptos. Cada atributo de SNOMED CT puede aplicarse a una jerarquía concreta de la nomenclatura y solamente unos pocos atributos pueden ser aplicados a más de una jerarquía, atendiendo siempre a un sentido lógico y con significado dentro del ámbito clínico. El conjunto de los componentes de SNOMED CT, incluyendo sus relaciones, hacen que la terminología se constituya como una ontología ligera basada en el conocimiento de los expertos. Tal es así, que el recurso, además de distribuirse en formatos de tablas de texto delimitado por tabuladores, ha comenzado su distribución en Web Ontology Language (OWL).

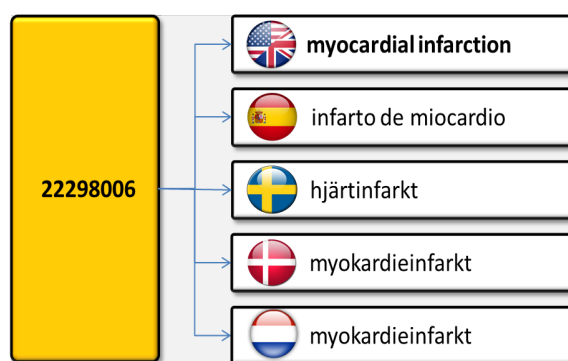


Figura 6. Representación del concepto “infarto de miocardio” por medio de descripciones en diferentes idiomas

Otra característica que hace de SNOMED CT una terminología única es ser un recurso multilingüe. SNOMED CT mantiene una metodología que permite integrar diferentes idiomas y dialectos. La edición internacional de SNOMED CT en inglés y español incluye un conjunto de conceptos, descripciones y relaciones que cada país miembro de la IHTSDO puede traducir a su propio idioma. Los códigos de concepto nunca se alteran, son permanentes, no reutilizables y unívocos, lo que es fundamental para poder conseguir la interoperabilidad semántica entre los sistemas de información y la captura de la información de forma consistente y trazable a lo largo del tiempo.

SNOMED CT incluye también un mecanismo de extensiones, basado en la asignación de espacios de nombre para organizaciones autorizadas. Este mecanismo facilita gran agilidad en el man-

tenimiento de la terminología, pues se pueden generar conceptos a nivel local y posteriormente ser elevados para su uso en un contexto más global, lo que beneficia al desarrollo de un conocimiento compartido y colaborativo a nivel internacional.

La terminología se plantea mejorar la comunicación de las historias clínicas electrónicas y aumentar la disponibilidad de información relevante de forma controlada en los sistemas. La información clínica puede conservarse de manera que permita su recuperación e intercambio basándose en el significado. Estos significados pueden ser seleccionados y adaptados dependiendo de cada contexto clínico, utilizando para ello lo que se denomina conjuntos de referencia (refsets) o subconjuntos.

### 2.3. Recursos terminológicos derivados de SNOMED CT

Un subconjunto es un derivado de SNOMED CT o de una extensión normalizada, que equivale a una fracción de la terminología que ha sido previamente seleccionada y agrupada para un uso o propósito concreto, lo que incluye un determinado número de conceptos, descripciones y relaciones que son apropiadas para un contexto en particular o para unos fines determinados (Marco Cuenca, 2013). Se trata, por tanto, de un recurso terminológico que procede de un activo mayor, SNOMED CT o una extensión nacional normalizada, y que se concreta con una serie de componentes previamente seleccionados, a partir de unos criterios definidos, con el objetivo de servir a un propósito específico.

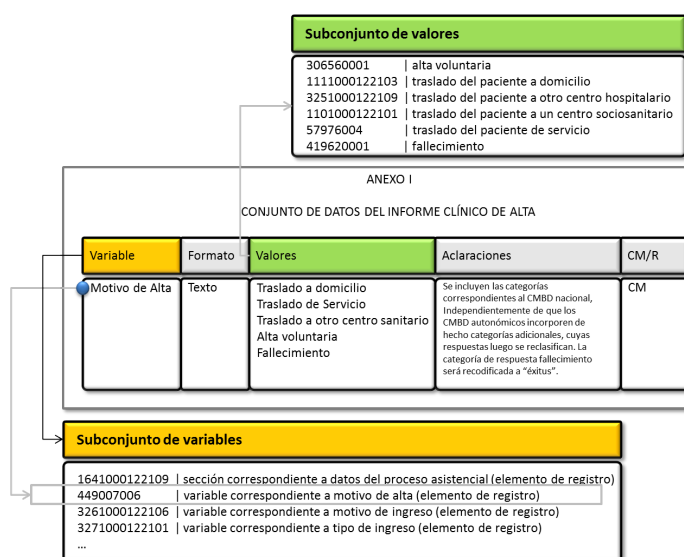


Figura 7. Ejemplo de asociación de subconjuntos de variables y valores para la variable motivo de alta del conjunto de datos del Informe Clínico de Alta

Esta terminología facilita la normalización semántica, tanto de los componentes de información que forman parte del modelo como de los valores que son vinculados a cada elemento o variable del modelo. De este modo, para la HCDSNS se han creado dos tipos de subconjuntos: subconjuntos de variables y subconjuntos de valores.

Los subconjuntos de variables equivalen a conceptos que sirven para representar de forma exacta el significado de las variables o etiquetas de un informe clínico definido en un modelo. Por otro lado, los subconjuntos de valores se componen de conceptos que representan los valores asociados o aceptados para una etiqueta o variable de un informe clínico definido en un modelo.

Algunos ejemplos de subconjuntos de variables son: *etiquetas del informe clínico de alta, etiquetas del informe clínico de consulta externa, etiquetas del informe de resultados de pruebas de laboratorio, informe de resultados de pruebas de imagen*, etc. Con respecto a los subconjuntos de valores podemos encontrar: *alergias, alertas, vacunaciones, diagnósticos, procedimientos, motivos de alta, recomendaciones terapéuticas, especialidades médicas y enfermeras, antecedentes quirúrgicos, antecedentes familiares y sociales, enfermedades de declaración obligatoria*, etc.

Tipo	Cantidad
Variables	9
Valores	58

Tabla V. Subconjuntos de variables y valores disponibles en el Área de descarga de SNOMED CT

Resulta importante destacar de entre estos recursos el *Nomenclátor de prescripción* de la Agencia Española del Medicamento y Productos Sanitarios (AEMPS). Este Nomenclátor de prescripción es una base de datos de medicamentos diseñada para proporcionar información básica de prescripción a los sistemas de información asistenciales, que incluye todos los medicamentos autorizados, suspendidos y revocados a partir de mayo de 2013, se encuentren financiados o no financiados. Los datos para la identificación clínica del medicamento a diferentes niveles se basan en la terminología SNOMED CT. Estos valores son aceptados para la representación del medicamento prescrito en la historia clínica electrónica.

Los subconjuntos disponibles para la HCDSNS se pueden descargar a través del Área de descarga de SNOMED CT, tanto de variables, como de valores aceptados. Los subconjuntos se publi-

can en diferentes formatos para facilitar su reutilización e implementación e incluyen una ficha técnica que facilita su interpretación. Estos activos semánticos se gestionan y versionan de forma continuada, pues son objeto de revisión e incorporación de nuevos conceptos cuando así son requeridos.

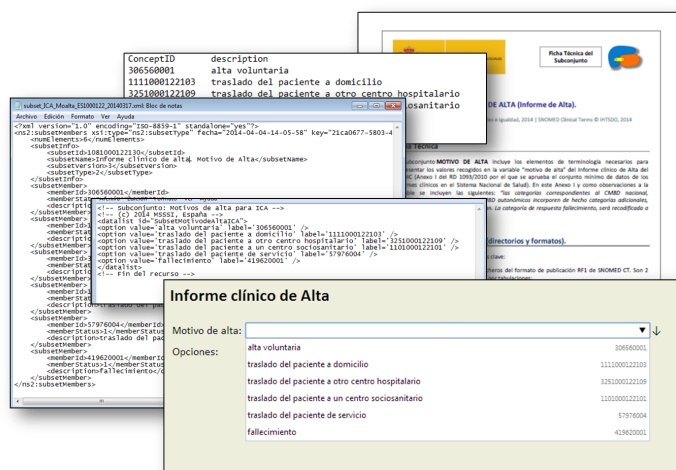


Figura 8. Diferentes formatos de publicación del subconjunto de motivo de alta del Informe Clínico de Alta

Los recursos terminológicos del SNS cuentan, al igual que los recursos de modelado clínico, con un servidor denominado Servidor de terminologías del SNS (stSNS). Un servidor de terminologías facilita el uso y la gestión de actualizaciones de los recursos terminológicos de referencia, de forma que estos procesos se realicen en un único sistema y la información pueda ser consultada o trasladada de forma sincronizada a otros sistemas desde un único punto de acceso, reduciendo de este modo las labores de mantenimiento.

El stSNS mantiene como objetivo principal mejorar la interoperabilidad semántica en las aplicaciones de historia clínica electrónica. Para ello cuenta con una serie de servicios que han sido definidos para poder compartir y consultar los diferentes recursos semánticos disponibles, como los subconjuntos, de una forma dinámica y eficiente. El stSNS actúa como servidor de referencia y permite actualizar a otros potenciales servidores que puedan estar alojados en las Comunidades Autónomas, sirviendo de nodo central de una red que comparte recursos terminológicos. Con este servidor se busca armonizar los recursos semánticos, de tal manera que puedan ser compartidos a partir de las mismas fuentes de referencia, evitando la duplicidad de esfuerzos por parte de todos los agentes del SNS.

### 3. Conclusiones

Los sistemas de historia clínica electrónica tienen, entre sus objetivos principales, facilitar el registro de la información clínica por parte de los profesionales sanitarios. Desde el punto de vista de la interoperabilidad semántica, es necesario contar con recursos que permitan la normalización de la información capturada y almacenada, y su comunicación e intercambio sin pérdida de significado. Por ello hay que normalizar la información, establecer estructuras comunes y vincular la actividad del registro a la introducción de datos controlados.

Si se consigue contextualizar la información y agregar datos de forma controlada, podremos disponer de herramientas que transformen estos datos y la información que contienen en conocimiento que los profesionales y las organizaciones podrán aprovechar para la mejora de la calidad asistencial. Además, la introducción de información de forma normalizada puede facilitar la relación de unos datos con otros y, por consiguiente, facilitar la implementación de reglas de inferencia en los sistemas, para que estos puedan dar soporte a una toma de decisiones basada en la evidencia.

Estos sistemas no pueden evolucionar y las organizaciones y profesionales no pueden adquirir conocimiento, sin dar prioridad al registro estructurado de la actividad asistencial. No es lo único que hay que hacer, pero es un primer paso necesario e imprescindible. Además, la interoperabilidad semántica no podrá producirse entre sistemas de forma precisa, si la información que un médico introduce en un sistema no se encuentra normalizada y no puede ser interpretada de forma correcta en otro sistema, en otro lugar y por otro profesional.

Para que esto sea posible hay que fortalecer, tal y como se está realizando en la HCDSNS, estos sistemas de recolección de datos con la producción de activos semánticos basados en estándares internacionales. Estos recursos semánticos deben ser compartidos, proporcionando mecanismos para su mejora continua, su mantenimiento y su puesta en producción de forma ágil.

La norma UNE-EN ISO 13606 favorece el consenso y la conexión entre dos mundos tan diferentes como son el clínico-asistencial y el de los sistemas de información. Deja en manos de los profesionales sanitarios la definición funcional de las estructuras y la propuesta de arquetipos y, por otro lado, la responsabilidad de los modelos de información recae en el personal informático. Por otro lado, la labor asociada a la normaliza-



ción de los datos y su significado debería ser asumida por profesionales especializados en la gestión de datos e información y, en concreto, por documentalistas. El documentalista, además, es un profesional con competencias para traducir las necesidades funcionales en requisitos técnicos, sirviendo de puente o pasarela entre el mundo asistencial y el de los sistemas de información.

Es importante también recalcar la flexibilidad que se consigue con el uso de estándares basados en un modelo dual, como la norma UNE-EN ISO 13606, ya que aportan un modelo de clases o de referencia estable y un modelo de arquetipos que puede evolucionar sin problemas de adaptación. Los arquetipos representan una propuesta para el control y normalización de la historia clínica electrónica, favoreciendo la interoperabilidad entre los sistemas de información clínicos. Son la base en la que se fundamenta el desarrollo de activos semánticos de calidad, relevantes, consistentes y sostenibles a lo largo del tiempo.

El uso de un lenguaje común y normalizado es una evolución necesaria de los sistemas de historia clínica, que deben incorporar herramientas que proporcionen un intercambio preciso de la información y un procesamiento eficaz, con objeto de facilitar el análisis del contenido con fines de evaluación clínica, gestión e investigación.

Los subconjuntos tienen que producirse atendiendo a un contexto asistencial determinado, pues no puede definirse un mismo producto para un médico de familia que para un cardiólogo. Su labor asistencial es diferente y los conceptos que necesitan para expresar esta labor también son diferentes. Utilizar una terminología clínica normalizada integral, como SNOMED CT, permite la producción de subconjuntos orientados a diversos usos, más generales o más específicos. Además, si utilizamos SNOMED CT como terminología de referencia, también podremos beneficiarnos de disponer de un lenguaje común y multilingüe para la comunicación entre diferentes sistemas, niveles asistenciales y diversos idiomas.

SNOMED CT ofrece otras ventajas adicionales, como son una metodología lógica y un procedimiento normalizado para extender el conocimiento clínico de forma colaborativa y compartida a nivel internacional. Para poder disponer de dicho conocimiento actualizado en todo momento y disponible para todos los agentes del SNS, es necesario crear una infraestructura organizativa que provea el mantenimiento y mejora continua de estos recursos, como es el caso de las herramientas somSNS y stSNS.

Todo parece indicar que el conocimiento clínico en nuestro SNS estará basado en el desarrollo

de ontologías y puede ser una buena forma de aproximarse a este objetivo la normalización y el uso de estándares de referencia.

## Referencias

- Aleixandre, R.; Amador, A. (2001). Problemas del lenguaje médico actual (I). Extranjerismos y falsos amigos. // *Papeles Médicos*. 10:3, 144-149.
- Cáceres Tello, J.; Sánchez de Madariaga, R.; Muñoz Carrero, A.; et.al (2013). Plataforma de Historia Clínica Electrónica conforme a la norma UNE-EN ISO 13606 como soporte a Grupos de Investigación. // XV Congreso Nacional de Informática Médica (Informed).
- Deloitte (2013). eHealth EIF eHealth European Interoperability Framework European Commission – ISA Work Programme. Study Report. Julio de 2013. [http://www.uems.eu/\\_data/assets/pdf\\_file/0010/1504/eHealth\\_European\\_Interoperability\\_Framework\\_-\\_Study\\_report\\_-\\_July\\_2013.pdf](http://www.uems.eu/_data/assets/pdf_file/0010/1504/eHealth_European_Interoperability_Framework_-_Study_report_-_July_2013.pdf) (2016-02-11).
- Directiva 2011/24/UE del Parlamento Europeo y del Consejo, de 9 de marzo de 2011, relativa a la aplicación de los derechos de los pacientes en la asistencia sanitaria transfronteriza. Diario Oficial de la Unión Europea. 4 de Abril de 2011. 88 (Serie L) 45-65.
- Etreros Huerta, J. J.; Marco Cuenca, G.; Abad Acebedo, I et.al (2009a). La interoperabilidad como base de la Historia Clínica Digital del Sistema Nacional de Salud. // *Todo Hospital*. 258 (Julio-Agosto) 467-474.
- Etreros Huerta, J. J.; Alonso Villar, C.; Marco Cuenca, G.; et.al (2009b). El sistema de Historia Clínica Digital en el Sistema Nacional de Salud: accesibilidad y protección de la información como elementos clave. // *DS: Derecho y Salud*. 18:1, 99-110.
- Indarte Galli, S.; Vero Suaya, A. (2014). Sistemas de apoyo a la toma de decisiones clínicas y de gestión en atención primaria de salud. // *Manual de salud electrónica para directivos de servicios y sistemas de salud*. Madrid: Sociedad Española de Informática de la Salud, 2014. 181-202.
- Kalra, D (2006). Electronic Health Records Standards. // *IMIA Yearbook of Medical Informatics*, 136-144.
- Ley 16/2003, de 28 de mayo, de cohesión y calidad del Sistema Nacional de Salud. // *Boletín Oficial del Estado*. (29 de Mayo de 2003) 128, 20567-20588.
- Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica. // *Boletín Oficial del Estado*. (15 de Noviembre de 2002) 274, 40126-40132.
- López Piñero, J. M.; Terrada Ferrandis, M. L. (2005). *Introducción a la terminología médica*. Madrid: Masson.
- Maldonado Segura, J. A. (2013). La frontera entre modelos de información y terminologías. // *Manual práctico de interoperabilidad semántica para entornos sanitarios basada en arquetipos*. Madrid: Instituto de Salud Carlos III. Unidad de Investigación en Telemedicina y e-Salud, 138-149.
- Marco Cuenca, G. (2013). Producción de subconjuntos con SNOMED CT. Un enfoque práctico. // *Manual práctico de interoperabilidad semántica para entornos sanitarios basada en arquetipos*. Madrid: Instituto de Salud Carlos III. Unidad de Investigación en Telemedicina y e-Salud, 124-137.
- Ministerio de Sanidad y Consumo (2004). Ley 14/1986, de 25 de abril, General de Sanidad. Colección Textos legales. Serie Sanidad; 1. Madrid : Ministerio de Sanidad y Consumo.

- Ministerio de Sanidad y Política Social (2010). El sistema de Historia Clínica Digital del SNS. Ministerio de Sanidad, Servicios Sociales e Igualdad. [http://www.msssi.gob.es/organizacion/sns/planCalidadSNS/docs/HCDNSNS\\_Castellano.pdf](http://www.msssi.gob.es/organizacion/sns/planCalidadSNS/docs/HCDNSNS_Castellano.pdf) (2016-02-11).
- Ministerio de Sanidad, Servicios Sociales e Igualdad. Consejo Asesor de Sanidad (2014a). La E-Salud. Prioridad estratégica para el Sistema Sanitario. [http://www.redaccion-medica.com/contenido/images/E\\_Salud.pdf](http://www.redaccion-medica.com/contenido/images/E_Salud.pdf) (2016-02-13).
- Ministerio de Sanidad, Servicios Sociales e Igualdad (2014b). Políticas de Normalización Terminológica. Principios y estrategias de Interoperabilidad Semántica para los Sistemas de Información Clínicos del SNS. Ministerio de Sanidad, Servicios Sociales e Igualdad. [http://www.msssi.gob.es/profesionales/hcdnsns/areaRecursosSem/FactoriaDocs/MSSSI\\_PNT\\_20140630.pdf](http://www.msssi.gob.es/profesionales/hcdnsns/areaRecursosSem/FactoriaDocs/MSSSI_PNT_20140630.pdf) (2016-02-15).
- Ministerio de Sanidad, Servicios Sociales e Igualdad (2014c). Preguntas más frecuentes sobre los Recursos de Modelado Clínico (arquetipos). Ministerio de Sanidad, Servicios Sociales e Igualdad. [http://www.msssi.gob.es/profesionales/hcdnsns/areaRecursosSem/preg\\_frec\\_rec\\_mod\\_clin.htm](http://www.msssi.gob.es/profesionales/hcdnsns/areaRecursosSem/preg_frec_rec_mod_clin.htm) (2016-02-18).
- Ministerio de Sanidad, Servicios Sociales e Igualdad (2014d). Lista de arquetipos (Recursos de Modelado Clínico). Ministerio de Sanidad, Servicios Sociales e Igualdad. [http://www.msssi.gob.es/profesionales/hcdnsns/areaRecursosSem/Lista\\_arquetipos\\_CMDIC\\_v1\\_20022014.pdf](http://www.msssi.gob.es/profesionales/hcdnsns/areaRecursosSem/Lista_arquetipos_CMDIC_v1_20022014.pdf) (2016-02-18).
- Ministerio de Sanidad, Servicios Sociales e Igualdad (2017). Proyecto HCDSNS. Historia Clínica Digital del Sistema Nacional de Salud. Informe de situación. Febrero 2017. Ministerio de Sanidad, Servicios Sociales e Igualdad. [http://www.msssi.es/profesionales/hcdnsns/contenido-Doc/Inf\\_Sit\\_HCDSNS\\_Febrero2017.pdf](http://www.msssi.es/profesionales/hcdnsns/contenido-Doc/Inf_Sit_HCDSNS_Febrero2017.pdf).
- Ministerio de Sanidad, Servicios Sociales e Igualdad (s. f.). Descarga de recursos terminológicos SNOMED CT. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad. <https://snomed-ct.msssi.es/snomed-ct>
- Muñoz Carrero, A (2013). Introducción. Manual práctico de interoperabilidad semántica para entornos sanitarios basada en arquetipos. Madrid: Instituto de Salud Carlos III, Unidad de Investigación en Telemedicina y e-Salud, 13-18.
- Orden de 6 de septiembre de 1984 por la que se regula la obligatoriedad del informe de alta. // Boletín Oficial del Estado. 221 (14 de septiembre de 1984) 26685-26686.
- Real Decreto 1093/2010, de 3 de septiembre, por el que se aprueba el conjunto mínimo de datos de los informes clínicos en el Sistema Nacional de Salud. Boletín Oficial del Estado. 225 (6 de septiembre de 2010) 78742-78767.
- Real Decreto 1718/2010, de 17 de diciembre, sobre receta médica y órdenes de dispensación. Boletín Oficial del Estado. 17 (20 de enero de 2011) 6306-6329.
- Romero Gutiérrez, A.; Marco Cuenca, G. (2014). Conocimiento abierto y modelos clínicos detallados. Avances en interoperabilidad en la Historia Clínica Digital del SNS (HCDSNS). // I+S. Informática y Salud. 107 (octubre) 16-28.
- Sempere J (2001). Problemas terminológicos en medicina: ¿alguna novedad? //Papeles Médicos, 10 (3): 136-143
- UNE-EN-ISO 13606-1 (2013). Informática sanitaria. Comunicación de la historia clínica electrónica. Parte 1: Modelo de referencia (ISO 13606-1:2008). Madrid: Aenor.

---

Enviado: 2016-03-02. Segunda versión: 2017-02-13.  
Aceptado: 2017-05-09.

---

---

# Método para la extracción masiva de canales de sindicación

*A method for the massive extraction of syndication channels*

---

**Manuel BLÁZQUEZ OCHANDO**

Departamento de Biblioteconomía y Documentación, Universidad Complutense de Madrid, España, manu-blaz@ucm.es

## Resumen

Uno de los problemas para la investigación de la producción informativa de canales de sindicación es contar con la cantidad de fuentes suficientes y del mismo dominio, temática o área de conocimiento, para componer una muestra. Esto se debe a la dispersión de las fuentes de información en la Web y por otra parte a la dificultad del investigador para conocer todos los recursos disponibles. A estos problemas se suma la dificultad de extraer y localizar los enlaces de los canales de sindicación en cada recurso o sitio web pertinente que se descubre. En este artículo se aborda el método para extraer los canales de sindicación mediante la composición de semillas para el análisis, con programas *web crawler*, identificando la configuración y posterior preparación.

**Palabras clave:** Minería de datos. Extracción de datos. Web crawler. Sindicación de contenidos. RSS. Canales de sindicación.

## 1. Introducción

La redifusión de información, sinónimo de sindicación de contenidos, corresponde a una tecnología clave para soportar diversas actividades de la información y documentación. De hecho, la sindicación se define como la transmisión de activos informativos y documentales, para su reutilización e integración en terceros recursos, a través de un archivo editado en lenguaje de marcado extensible XML, contenedor de la información, conforme a un formato de estructuración de datos RSS o Atom. A este recurso se le denomina canal de sindicación (Hammersley, 2003). A su vez, los canales de sindicación, normalizados de acuerdo a un formato o sintaxis determinada, pueden ser suscritos por los usuarios o destinatarios, mediante programas de agregación, especializados en la lectura y gestión de dichos canales. Ello ha permitido que las Ciencias de la Documentación aprovechen la tecnología para la redifusión de registros documentales y autoridades, el intercambio de contenidos entre bibliotecas digitales (Eito-Brun, 2015, p. 218), la alimentación de grandes proyectos europeos como Europeana (Houssos et al., 2011, p.2), la difusión selectiva de la información y su aplicación como sistema de vigilancia informacional y

## Abstract

One of the problems for investigating the informative production of syndication channels is counting on the sufficient number of sources from the same domain, subject or area of knowledge, to compile a sample. This is a consequence of the dispersion of information sources on the Web; the researcher's difficulty in knowing all the available resources; and the difficulty in extracting and locating the links of syndication channels in every relevant web site or Internet resource that is discovered. This article discusses the method to extract and compile syndication channels through the composition of seeds using a web crawler, and the configuration and subsequent processing of the obtained links.

**Keywords:** Data mining. Scraping. Web crawler. Content syndication. RSS. Feeds.

tecnológica (Peis et al., 2008, p. 522) o el desarrollo de estrategias de redifusión social y curación de contenidos (Nieto, 2015, p. 5). Desde un punto de vista informativo, la sindicación tiene un papel fundamental para el periodismo digital, como herramienta de difusión de los medios, como tecnología democratizadora del derecho de acceso a la información. Ello se debe a la libertad de elección del lector sobre las fuentes de información públicas, permitiéndole discriminar sus contenidos, comentarlos y sugerirlos, conformando así su propio menú de informaciones (McCownnet al., 2009, p. 1-2).

Como puede deducirse, la sindicación de contenidos está presente en los procesos de comunicación más importantes y ello ha propiciado una explosión demográfica en la creación de canales de sindicación, que está determinada por el uso de los sistemas de gestión de contenidos digitales. Si se asume que existen 348 millones de instalaciones de programas CMS registradas (BuiltWith, 2016), con sus correspondientes canales de sindicación, sin incluir sitios gubernamentales, instituciones académicas, bibliotecas y medios digitales. Se concluye que el principal problema es el descubrimiento de los canales de sindicación pertinentes. A ello se añade la dificultad de

recuperar masivamente la información, crear *big data*, clasificar y organizar los resultados, a fin de evaluar e integrar los canales de mayor valor en los flujos de comunicación. En esta investigación se plantean algunos interrogantes en relación a este problema: ¿Cómo detectar los canales de sindicación de un determinado dominio o área de conocimiento? ¿Qué métodos y herramientas existen para recuperar de forma masiva los canales de sindicación? ¿Cómo se discriminan los formatos de sindicación defectuosos?

El método que se propone combina programas web crawler y estrategias de búsqueda para dirigir el objetivo de recopilación. Si bien el uso de estos métodos de forma aislada es conocido, no lo es tanto a la hora de definir un método cualitativo para la extracción masiva de canales de sindicación, esto es, los enlaces a los recursos en formato RSS.

Las últimas investigaciones en la materia se refieren a técnicas de rastreo web para mejorar la calidad del contenido de los canales de sindicación mediante la implementación de contenidos extras relacionados (Hurtado, 2015, p. 2-3), buscar contenidos especializados en una fuente de información específica (Lubbers, 2015, p. 21-26), desarrollar agregadores de contenidos (Baporikar et al., 2015, p. 159-160), e incluso el uso de programas web crawler especializados con capacidades de consulta adaptativa (Lee et al., 2008, p. 220-223) destinados a la recopilación de noticias producidas en estos medios. En la misma línea, también es destacable la experiencia de buscador especializado en contenidos producidos por canales RSS, elaborado por Viseur (2012).

Sin embargo, las referencias citadas no abordan una técnica con la que puedan ser descubiertos nuevos canales de sindicación, desconocidos para el investigador en una materia o temática dada. Esto es, los enlaces de archivos RSS, con los que poder crear una colección lo suficientemente amplia como para ser considerada del dominio del Big Data.

La extracción masiva de canales de sindicación en un dominio concreto, permitiría estudios más exhaustivos y precisos, en relación a la producción informativa de los medios de comunicación de un país, de las revistas científicas en un área de conocimiento, de la universidad, la administración pública e incluso las redes sociales. De hecho, otros investigadores ya advirtieron la posibilidad de usar la información que los canales de sindicación proporcionan, para identificar debates públicos en los medios, determinar su relevancia, así como el grado de preocupación de la opinión pública (Thelwall et al., 2006, p. 1-3).

También cabe mencionar el papel de la sindicación de contenidos frente a herramientas de alerta de noticias como Google Alerts, cuyo objetivo es proporcionar una vigilancia informacional basada en las palabras clave. El resultado es el suministro de alertas de noticias procedentes de medios y fuentes específicas tales como vídeos, finanzas, libros y blogs. Sin embargo, no es posible conocer a priori, qué fuentes de información está utilizando y sobre todo la calidad de las mismas. Este hecho ya fue descubierto en investigaciones pretéritas (Prieto, Lloret, y Palomar, 2012) en las que se concluye que la efectividad de este tipo de herramientas dependía de los términos de consulta empleados y del dominio general de las alertas, que no siempre se adecua al área de conocimiento abordada. Este tipo de experiencias apoya la tesis de controlar y descubrir nuevas fuentes de información especializadas en la materia, frente al dominio global de la web y el desarrollo de estrategias de consulta basadas en lenguajes documentales con un cierto grado de normalización. Es en este apartado donde las técnicas de extracción y descubrimiento de canales de sindicación pueden ayudar al investigador a desarrollar sistemas de vigilancia informacional más certeros. Por todo ello, los canales de sindicación pueden tener aplicaciones clave en el seguimiento de información especializada, análisis de correlaciones entre contenidos, minería de datos, generación de big data, cálculo de impacto, periodismo de datos o el desarrollo de sistemas expertos para la asistencia en la toma de decisiones.

## 2. Metodología

El descubrimiento y recopilación masiva de nuevos canales de sindicación puede resultar una tarea complicada, si se tiene en consideración la dificultad de localizar fuentes de información fiables que proporcionen enlaces permanentes a los archivos de datos XML. Si se añade la necesidad de recuperar de forma masiva canales de sindicación para crear nuevas bases de conocimiento o colecciones de investigación, el problema se hace más evidente y obliga a desarrollar una metodología específica. El método propuesto, consta de los siguientes pasos: 1) delimitación del área de conocimiento y desarrollo de un vocabulario representativo, 2) diseño de estrategias de consulta para la recuperación de sitios web pertinentes, 3) creación de una semilla de enlaces para su análisis con herramientas web crawler, 4) análisis de enlaces con programas web crawler, y 5) preparación previa de los canales de sindicación para su procesamiento en agregadores de contenidos.



## 2.1. Área de conocimiento y vocabulario

La delimitación del área de conocimiento es fundamental para centrar el objetivo de rastreo, que se llevará a cabo posteriormente. Ello implica conformar un corpus textual, representativo del dominio temático en el que se pretenden descubrir nuevos canales de sindicación. El núcleo léxico de las consultas, pasa por la recopilación de descriptores, obtenidos a partir de la literatura del área, la identificación de los temas principales en las revistas científicas, el análisis estadístico de frecuencias de los términos de los textos especializados, así como la revisión de los términos

que componen los títulos más citados (Huang et al., 2015, p. 3). La importancia del trabajo de recuperación léxica, reside en la experiencia que el investigador adquiere para componer consultas con terminologías equivalentes o paralelas en la consecución de un objetivo de búsqueda concreto, demostrándose que ello supone una clara ventaja en la recuperación de información (Thatcher, 2008, p. 1309-1311). La organización del léxico no es rígida y basta con definir la categoría de cabecera, las subcategorías principales y sus términos específicos, recogidos a modo de bolsa de elementos, como se puede comprobar en la tabla I.

<i>Esquema de organización léxica</i>		
Término cabecera Categoría principal n1 (Término específico n1.1, Término específico n1.2, Término específico n1.3...  Término relacionado n1.1, Término relacionado n1.2, Término relacionado n1.3...)		
<i>Ejemplo de organización léxica</i>		
Comunicación Secciones de contenidos (Política, Parlamento, Política internacional, Cooperación política, Conflictos, Legislación, Finanzas, Función pública, Derecho civil, Derecho penal, Justicia, Derecho internacional, Libertades, Política económica, Análisis económico, Política comercial, Consumo, Instituciones financieras...)		
<i>Definición de calificadores</i>		
<i>Calificador</i>	<i>Descripción</i>	<i>Ejemplo</i>
Geográfico, Idioma	Localización o ubicación del recurso, país o región	Portugal, Alentejo, Algarve, Lisboa, portugués, Pt; Francia, Normandía, Bretaña, Aquitania, Alsacia, francés, Fr
Cronológico	Situación temporal en un año o intervalo específico	2010..2016
Tipo de recurso	Identificación de los tipos de recursos o fuentes de información que las que se desean obtener canales de sindicación	Scientific journal, Paper, E-print archive, Digital repositories, Digital newspaper, Magazines, Mass media, Media, Social network
Editor	Sistema de gestión y edición de contenidos que utiliza el recurso o fuente de información.	Blog, Blogger, Wordpress, Joomla, Drupal, Revist
Extensión y formato	Formato del tipo de archivo XML que se desea recuperar	RSS, Atom, MARC-XML, OAI-MARC, OPML, RDF, OWL
Texto denotativo de enlace	Cadenas de texto frecuentemente presentes en las direcciones URL permanentes de los canales de sindicación	feed, xml, rss, syndication, channel

Tabla I. Esquema de organización léxica y calificadores

Adicionalmente conviene identificar los calificadores, que podrán ser empleados en el diseño de las estrategias de consulta. Los calificadores geográficos permitirán recuperar sitios web de una determinada localización y tendrán implicaciones en el idioma de los recursos. El factor cronológico puede ayudar a diferenciar recursos o fuentes de información obsoletas, o que no reciben actualización desde una determinada fecha, evitando recopilar canales de sindicación muertos. El tipo de recurso es un calificador que ayuda a afinar los resultados en torno a un tipo de publicación web, como por ejemplo revistas científicas,

repositorios digitales, medios de comunicación social, etc. La extensión o formato, también ayuda a localizar aquellos sitios web que posean archivos de las características indicadas. El calificador "texto denotativo de enlace", se refiere a ciertas cadenas de caracteres que frecuentemente se encuentran presentes en los enlaces permanentes de algunos sitios web. Ello denota unívocamente la presencia de canales de sindicación. Por ejemplo, las palabras feed y rss se usan en consultas como (inanchor:"feed" OR inanchor:"rss") que recuperarían todos los sitios web, que contengan dichas cadenas de texto.

## 2.2. Diseño de estrategias de consulta

Una vez preparado el léxico y los calificadores, se define la estrategia de consulta. El objetivo es diseñar consultas optimizadas para buscadores, de los que se extraerán las páginas de resultados más relevantes. De los resultados se obtienen listas de enlaces, con las que conformar una relación más amplia, denominada semilla o matriz. La semilla representará el dominio temático de la web, de la que se desean extraer los canales de sindicación. Por ello las estrategias de consulta

(Blázquez, M. 2013a, p. 71-76), requieren operadores avanzados, que aseguren el filtrado efectivo de la web, para aumentar el grado de precisión. En la tabla II se muestra una selección de operadores y ejemplos de restricción de dominio, búsqueda en títulos, corpus textual y enlaces. También se exponen estrategias de consulta para recuperar enlaces de revistas científicas de acceso abierto, en editoriales científicas, medios de comunicación, instituciones científicas y repositorios especializados.

Operador	Función	Ejemplos
site:	Buscar entre los contenidos del sitio o dominio especificado.	<i>Buscar revistas científicas de Open Access especializadas en bibliotecas</i> [site:doaj.org "library"]  <i>Buscar contenidos de Elsevier especializados en medicina</i> [site:elsevier.com "clinical medicine"]
intitle:	Buscar en los títulos de las páginas web indexadas.	<i>Buscar secciones de noticias especializadas en economía española</i> [intitle:noticias intext:españa intext:economía]  <i>Buscar instituciones científicas de Estados Unidos</i> [intitle:"national institute" intext:"science" intext:"united states"]
intext:	Buscar en el corpus textual de las páginas indexadas.	<i>Buscar directorios de medios de comunicación especializados en prensa digital en el dominio de Reino Unido</i> [intext:"directory" intext:"journalism" intext:"newspapers" inurl:uk]
inurl:	Buscar cadena de texto en las direcciones URL indexadas.	<i>Buscar instituciones científicas de Estados Unidos de tipo gubernamental</i> [intitle:"national institute" intext:"science" intext:"united states" inurl:gov]
inanchor:	Buscar cadena de texto en los enlaces hipertextuales disponibles en el corpus textual de las páginas web.	<i>Buscar sitios web que contengan enlaces a repositorios especializados en Biblioteconomía y Documentación</i> [inanchor:"repository" intext:"library and information science"]

Tabla II. Principales operadores y ejemplos de estrategias de consulta

El diseño de las estrategias de consulta puede expresarse mediante combinaciones de términos, propios del léxico mencionado anteriormente. La consulta puede tener dos posibles orientaciones. La primera consiste en satisfacer una necesidad de información concreta y previamente razonada por el investigador. La segunda es capaz de resolver una necesidad concreta, para aplicar la matemática de combinaciones entre los términos del léxico, con la finalidad de obtener recursos y resultados de la Web que son parcial o completamente desconocidos para el investigador. Esto significa que las estrategias de consulta pueden combinar las necesidades de información puntuales del investigador y añadir la componente matemática de la minería de datos a las consultas.

En este sentido, se propone una fórmula que ayude a configurar una estrategia de consulta combinatoria, basada en la intersección del término cabecera, las categorías principales, los términos específicos y relacionados, así como los calificadores o complementos que ayuden a precisar y filtrar aún más los resultados.

$$Q_n = TC_{n(0,1)} \cap CP_{n(0,2)} \cap (TE \leftrightarrow TR)_{n(1,4)} \cap C_{\neq n(0,3)}$$

Tabla III. Propuesta de estrategia de consulta

En la tabla III, se identifica que "n" consultas serán el resultado de combinar de forma opcional el término cabecera mediante la intersección opcional de un máximo de dos categorías principales del mismo rango, añadiendo la intersección de uno a cuatro términos específicos o relacionados del mismo rango, y finalmente introduciendo hasta tres calificadores opcionales para el filtrado de las consultas de diferentes rangos

## 2.3. Creación de una semilla de enlaces

La estrategia de consulta en buscadores proporciona resultados de utilidad para configurar la semilla de enlaces. El problema se encuentra en la extracción de los mismos. Para resolverlo pueden realizarse copias sistemáticas de cada pá-

gina de resultados utilizando algunos complementos de los principales navegadores web (1), ya que los buscadores no posibilitan la descarga libre de sus contenidos, salvo excepciones (2). De hecho, la API del buscador Google para la recuperación de sus páginas de resultados, se encuentra fuera de servicio desde el año 2010 (Google, 2015) y la alternativa ofrecida sólo plantea una solución para crear un buscador de sitio web, que no responde a las necesidades que se están planteando. Otra posibilidad más efectiva, pero menos frecuente, es el empleo de servicios de descarga de enlaces en buscadores. Un ejemplo es la aplicación experimental Google2down (3), diseñada para la recuperación automática de enlaces de múltiples páginas de resultados, en los buscadores Google y Google Scholar.

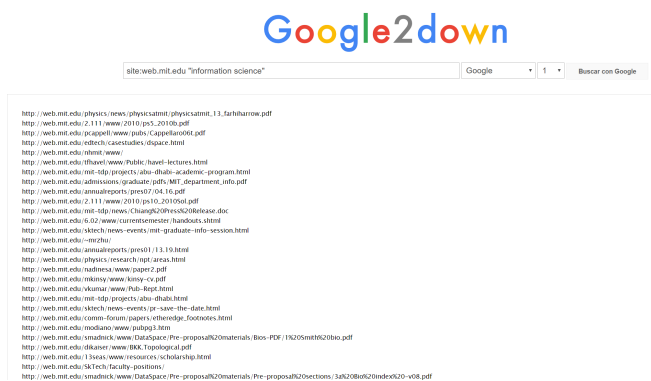


Figura 1. Servicio de descarga Google2down

## 2.4. Análisis de enlaces con web crawler

Una vez se dispone de la lista completa de enlaces, se procede a la configuración y ejecución de un programa web crawler que los analizará en profundidad a fin de extraer los canales de sindicación. Cualquier web crawler dispone de opciones de configuración que permiten programarlo con distintos niveles de profundidad de análisis y restricción de acceso a dominios y páginas web especificadas. Se aconseja que el número máximo de niveles de profundidad del análisis no exceda el valor 4, ya que está demostrado que más del 70% de los contenidos de la web son recuperables en los tres primeros niveles de análisis (Baeza-Yates, Castillo y Graells, 2008, p. 13). Por ejemplo, un sitio web expresado en la semilla con la designación “p1”, puede contener enlaces a otras páginas web del mismo dominio, designadas como “p1.1, p1.2, p1.n” que a su vez pueden contener enlaces a otras páginas designadas como “p1.1.1, p1.1.2, p1.1.n”. Cada salto de una página web a otra, dentro de un mismo dominio en nivel jerárquico descendente, se denomina nivel de enlazamiento o nivel de profundidad del análisis. La opción “restricción de sitio” debe ser

activada, para indicar al web crawler que no tiene que rastrear los enlaces salientes a páginas externas. De esta forma se evita que analice otros sitios web que no siempre tienen que ver con el objetivo planteado en la investigación. Para concluir la configuración, se debe activar la función “extracción de canales de sindicación” o bien definir los patrones de reconocimiento mediante expresiones regulares. Por ejemplo la siguiente expresión regular (<?xml version="..." encoding="..." | <rssversion="..." | <channel> | <item>) presenta fragmentos de código que identifican unívocamente cualquier archivo XML y canal de sindicación en formato RSS.

En cuanto al funcionamiento del web crawler, el proceso ya fue descrito en anteriores investigaciones (Blázquez, 2013b) de las que se pueden sintetizar los siguientes pasos: 1) Lectura de los enlaces disponibles en la semilla, 2) Selección del primer enlace, 3) Comprobación de errores de codificación del enlace, 4) Verificación por socket del enlace, mediante envío de cabeceras HTTP Head, vía puerto 80, 5) Verificación de respuesta del servidor, 6) Descarga del código fuente del sitio web enlazado, mediante técnica cURL, 7) Análisis del código fuente y extracción de todos los enlaces de la página, usando el método XPath de selección y expresiones regulares, tal como se muestra en la tabla IV, 8) almacenamiento de los enlaces recuperados en base de datos, 9) replicación del método con toda la semilla de enlaces, 10) cribado de enlaces almacenados, véase apartado 2.5.

Método	Ejemplo
XPath	<i>Extracción de canales RSS</i> <code>\$canalesrss = \$xpath -&gt;  query("/html/head/link[@type='application/rss+xml']");</code> <i>Extracción de enlaces</i> <code>\$enlaces = \$xpath -&gt;  query("/html/body//a");</code>
REGEXP (Expresiones Regulares)	<i>Extracción de enlaces</i> <code>[-a-zA-Z0-9:~#?&amp; =]{2,256}\\.([@\\  [a-z]{2,12})b(V[-a-zA-Z0-  9:%_+~#?&amp; =]*)?</code>

Tabla IV. Método de extracción de enlaces

## 2.5. Preparación de canales de sindicación

El resultado del análisis con web crawler suele ser un archivo con la lista de canales de sindicación procedentes del conjunto de sitios web analizados en la semilla. Es muy frecuente recuperar más canales de sindicación de los que realmente corresponden. Ello se debe a que no todos los archivos XML están bien contruidos, o bien corresponden a formatos de sindicación en los que

el investigador no está interesado, o bien presenten defectos sintácticos que hacen imposible su lectura (Tabla V). Por estos motivos es fundamental realizar un proceso de cribado.

Principales errores	Descripción
(^<!doctype ^<html>)	Confusión de códigos HTML con códigos XML propios del formato de sindicación
(^\s ^t ^R ^n)	La primera línea XML está precedida de un espacio en blanco que invalida la correcta formación del documento
(rssversion)	El nombre de la etiqueta y sus atributos aparecen juntos sin espacio separador

Tabla V. Errores de codificación más frecuente en canales de sindicación de contenidos

Sin embargo, cada vez es más frecuente la incorporación de métodos de detección de errores en los sistemas de agregación de contenidos (Petrova-Antonova y Simov, 2011, p. 642). Por lo tanto, una forma efectiva de eliminar todos aquellos canales de sindicación no válidos es la importación directa de los enlaces en un agregador. Por regla general los canales de sindicación defectuosos, serán desestimados borrando automáticamente su enlace de la lista. Los canales de sindicación correctamente formados y validados serán importados y constarán de alguna información básica como el título y la descripción que en todo caso el agregador pueda recopilar. No obstante, la importación de los canales de sindicación no implica que estén disponibles para su explotación. Para efectuar investigaciones sobre la producción informativa de los canales de sindicación es necesario realizar una tarea de categorización previa que permita indicar aspectos clave como la procedencia, tipo de fuente, factor geográfico, dominio temático, valor de la fuente de información, idioma y prioridad de recuperación.

## 2.6. Caso práctico: canales de sindicación de los medios digitales de Portugal

Recientemente se ha completado un estudio de la producción informativa de los medios de comunicación digitales de Portugal que se publicará en la revista *Transinformação* de la Universidade Estadual de Campinas (Blázquez Ochando, 2017) usando el método descrito en el presente artículo. Las semillas de enlaces utilizadas para la extracción de los canales de sindicación de Portugal pueden ser consultadas y descargadas desde las siguientes referencias:

- Semilla de sitios web de prensa de Portugal. <http://mblazquez.es/wp-content/uploads/semilla-prensa-pt.txt>
- Semilla de sitios web de radios digital de Portugal. <http://mblazquez.es/wp-content/uploads/semilla-radio-pt.txt>
- Semilla de sitios web de televisión de Portugal. <http://mblazquez.es/wp-content/uploads/semilla-television-pt.txt>

Después de aplicar el método de extracción masiva de canales de sindicación, se obtuvieron los siguientes canales de sindicación:

- Canales de sindicación de prensa de Portugal. <http://mblazquez.es/wp-content/uploads/feeds-prensa-pt.txt>
- Canales de sindicación de radio de Portugal. <http://mblazquez.es/wp-content/uploads/feeds-radio-pt.txt>
- Canales de sindicación de televisión de Portugal. <http://mblazquez.es/wp-content/uploads/feeds-television-pt.txt>

De esta forma se tienen todos los canales de sindicación de la prensa (889 feeds), radio (208 feeds) y televisión (231 feeds) de Portugal, con los que se pudo realizar un estudio exhaustivo de la producción informativa del país.

## 3. Conclusiones

A pesar de que la sindicación de contenidos ha sido abordada desde el punto de vista de la comunicación informativa y documental, no se ha profundizado en los métodos de recopilación masiva de canales de sindicación en dominios y áreas de conocimiento específicas.

El desarrollo de estudios informétricos, de producción de información periodística, de opinión, tendencias o incluso de producción científica, según el objeto de estudio, depende de una mayor exhaustividad en las fuentes de información utilizadas. Esto significa no restringir las investigaciones a las fuentes conocidas y abrir el campo de estudio a nuevas fuentes que están por descubrir.

La aplicación del método expuesto ayuda a localizar nuevas fuentes de información con las que complementar los estudios de la información. Con ello se determinan distintas fases lógicas en las que el investigador plantea un contexto sobre el que construye un vocabulario o léxico normalizado, especializado en el área de conocimiento sujeto al análisis. Tomando como referencia la terminología organizada, se proponen estrate-

gias de consulta que la combinan usando operadores de consulta avanzada en buscadores para obtener resultados más pertinentes posibles. De los resultados obtenidos se extraen los enlaces que serán procesados por herramientas web crawler para extraer los enlaces de los canales de sindicación. Finalmente se detectan errores en los canales de sindicación y se completa la información clasificatoria y descriptiva que caracteriza su contenido.

Aunque no se proporciona un análisis de los agregadores que permiten realizar estudios informáticos a partir de los canales de sindicación, se puede adelantar que originalmente no fueron diseñados con ese objetivo. Es por ello que los investigadores no cuentan con herramientas adecuadas para aprovechar todas las posibilidades que ofrecen estas fuentes de información. De este problema, pueden deducirse futuras líneas de investigación que aborden el desarrollo de agregadores de contenidos diseñados para realizar estudios sobre la producción informativa, capaces de clasificar la información automáticamente, categorizar los canales de sindicación y obtener las relaciones inherentes entre los contenidos recuperados, descubriendo la influencia y el impacto de la información publicada.

## Notas

- (1) Selección de pluggins para los navegadores Google Chrome y Mozilla Firefox especializados en la copia de enlaces:  
<http://chrome.google.com/webstore/search/copy+links>  
<http://addons.mozilla.org/firefox/search/?q=copy+links>
- (2) El buscador WauSearch permite exportar los resultados de las consultas en listas de enlaces de forma automática (<http://www.wausearch.com/>).
- (3) Google2down es un servicio de exportación masiva de resultados de Google y Google Scholar (<http://www.mblazquez.es/google2down/>).

## Referencias

- Baeza-Yates, R.; Castillo, C.; Graells, E. (2008). Características de la web chilena 2007. // Technical Report, Center for Web Research, University of Chile.
- Baporikar, M.; Salvi, S.; Sowany, V.; Sakhare, N. S. (2015). An approach towards news alert systems. // *International Journal of Computer Science and Mobile Computing*. 4:11 (noviembre 2015) 159-163.
- Blázquez Ochando, M. (2013a). Sistemas de recuperación e internet: Metadescripción, procesamiento, webcrawling, técnicas de consulta avanzada, hacking documental y posicionamiento web. // Madrid: mblazquez.es
- Blázquez Ochando, M. (2013b). Desarrollo tecnológico y documental del webcrawler Mbot: prueba de análisis web de la universidad española. // XIII Jornadas Españolas de Documentación Fesabid. (mayo 2013).
- Blázquez Ochando, M. (2017). Método para el estudio de la producción informativa: Medios digitales de Portugal. // *Transinformação*. 29:1.

- BuiltWith (2016). CMS technologies Web Usage Statistics. <http://trends.builtwith.com/cms/> (2016-02-01).
- Eito-Brun, R. (2015). Context-based aggregation of archival data: the role of authority records in the semantic landscape. // *Archival Science*. 15:3 (February 2014) 217-238.
- Google. (2015). Google Web Search API (Deprecated). <https://developers.google.com/web-search/docs/>
- Google. (2016). Google Alerts. <https://www.google.es/alerts>
- Hammersley, B. (2003). Content syndication with RSS. // Sebastopol: O'Reilly, 2003.
- Houssos, N.; Stamatis, K.; Banos, V.; Kapidakis, S.; Garoufllou, E.; Koulouris, A. (2011). Implementing enhanced OAI-PMH requirements for Europeana. // *Research and Advanced Technology for Digital Libraries*. Berlin: Springer, 2011.
- Huang, Y.; Schuehle, J.; Porter, A. L.; Youtie, J. (2015). A systematic method to create search strategies for emerging technologies based on the Web of Science: illustrated for Big Data. // *Scientometrics*. 105:3 (July 2015) 2005-2022.
- Hurtado, J. (2015). Automated System for Improving RSS Feeds Data Quality. // arXiv e-prints. <http://arxiv.org/pdf/1504.01433v1> (2016-01-14).
- Lee, B. S.; Im, J. W.; Hwang, B. Y.; Zhang, D. (2008). Design of an RSS Crawler with Adaptive Revisit Manager // *SEKE*. 219-222.
- Lubbers, M. (2015). Adaptable Crawler Specification Generation System for Leisure Activity RSS Feeds. // Nijmegen: Radboud Universiteit, 2015.
- McCown, F.; Nelson, M. L.; Van de Sompel, H. (2009). Everyone is a curator: human-assisted preservation for ore aggregations. // arXiv e-prints. (Consulta 2016-01-15) <http://arxiv.org/pdf/0901.4571v1>
- Nieto, J. Y. (2015). Las revistas sociales personalizadas a través de agregadores compiten con el resto de medios informativos digitales. // *Ambitos: Revista internacional de comunicación*. 28 (July 2015) 5-13.
- Peis, E.; Herrera-Viedma, E.; Morales-del-Castillo, J. M. (2008). Modelo de servicio semántico de difusión selectiva de información (DSI) para bibliotecas digitales. // *El profesional de la información*. 17:5 (diciembre 2007) 519-525.
- Petrova-Antonova, D.; Simov, R. (2011). jQuery RSS: a jQuery plugin for RSS and Atom feeds parsing. *Proceedings of the 12th International Conference on Computer Systems and Technologies*. // ACM. (June 2011) 641-646.
- Prient, C.; Lloret, E.; Palomar, M. (2012). Análisis de la calidad de la información recuperada por sistemas de alertas en el dominio Químico Textil. // II Congreso Español de Recuperación de Información CERI. (junio 2012).
- Thatcher, A. (2008). Web search strategies: The influence of Web experience and task type. // *Information Processing & Management*. 44:3 (September 2007) 1308-1329.
- Thelwall, M.; Prabowo, R.; Fairclough, R. (2006). Are raw RSS feeds suitable for broad issue scanning? A science concern case study. // *Journal of the American Society for Information Science and Technology*. 57:12 (agosto 2006) 1644-1654.
- Viseur, R. (2012). Create a Specialized Search Engine – The Case of an RSS Search Engine // *Proceedings of the International Conference on Data Technologies and Applications*. 245-248.

Enviado: 2016-02-05. Segunda versión: 2017-03-28.  
 Aceptado: 2017-05-09.



---

# Produção científica e inserção internacional da revista *Scire* no período de 2006 a 2014

*Producción científica e inserción internacional de la revista Scire en el periodo 2006-2014*  
*Scientific production and international insertion of the journal Scire in the 2006-2014 period*

---

**Ely Francina Tannuri de OLIVEIRA (1), Bruno Henrique ALVES (2),  
Marcos Rodrigues do PRADO (3), Maria Aparecida PAVANELLI (4)**

(1) Pesquisadora do Programa de Pós-Graduação em Ciência da Informação da UNESP-Marília-Brasil-Endereço: Rua Coronel José Brás, nº 155, apartamento, nº 502-Marília-SP-Brasil-CEP: 17501-570, etannuri@gmail.com; (2) Doutorando do Programa de Pós-Graduação em Ciência da Informação da UNESP-Marília-Brasil-Endereço: Rua Sargento Wilson Abel Oliveira, nº 1267 - Garça-SP-Brasil-CEP:17400-000, brhenriqueal-ves@gmail.com; (3) Doutorando do Programa de Pós-Graduação em Ciência da Informação da UNESP-Marília-Brasil-Endereço: Rua Abuassali Abujamra, nº 389-Ourinhos-SP-Brasil-CEP: 19907-170, marcos-prado@yahoo.com.br; (4) Doutoranda do Programa de Pós-Graduação em Ciência da Informação da UNESP-Marília-Brasil-Endereço: Avenida Sebas-tião Lacerda Corrêa, nº 928-Araraquara-SP-Brasil-CEP:14800-480, cidinhapavanelli@gmail.com

## Resumen

Se analiza bibliométricamente la producción científica de la revista *Scire: Representación y Organización del Conocimiento*, para comprobar su internacionalización, las instituciones más productivas en sus respectivos países y las coautorías internacionales y temáticas más relevantes en el periodo 2006-2014. Se evidenciaron 73 instituciones de las cuales 23 publicaron al menos un artículo. Se mapearon las temáticas de trabajo según la clasificación de Dahlberg. Las instituciones fueron agrupadas en clústeres según su similitud temática y representadas en un dendrograma. Se concluye que existe una tendencia, aunque frágil, a su internacionalización, especialmente en relación a los países de América Latina.

**Palabras clave:** Organización del conocimiento. Producción científica. Revistas científicas. *Scire*. Coautoría institucional. Sistema de clasificación para la literatura sobre Organización del Conocimiento. América Latina.

## Abstract

A bibliometric analysis of the scientific production of the journal *Scire: Representación y organización del Conocimiento* in the 2006-2014 period is offered, in order to show its internationalization. Specifically, the study aims to highlight institutions with higher productivity in their countries, as well as to examine institutional co-authorships, highlighting the international co-authorships and the most important themes. As total of 73 institutions were investigated, with 23 publishing at least two investigations. A network of institutional collaboration was created, a list of themes worked by the various collaborating institutions was developed, and the institutions were grouped according to Dahlberg's categories. The institutions and themes were grouped into clusters according to thematic similarity and presented through a dendrogram. As a conclusion, it was observed that the internationalization of the journal *Scire* is a trend, although fragile, especially in relation to Latin America.

**Keywords:** Knowledge organization. Scientific production. Internationalization. *Scire*. Institutional co-authorship. Classification System for Knowledge Organization Literature. Latin America.

## 4. Introdução

Visando à divulgação disseminação e visibilidade dos resultados das pesquisas realizadas, a revista *Scire: Representación y Organización del Conocimiento* é uma publicação semestral da Universidade de Zaragoza (Espanha), criada em 1995, de caráter interdisciplinar, que trata da representação, normalização, tratamento, recuperação e comunicação da informação e do conhecimento. Em seu editorial, do primeiro número da Revista *Scire: Representación y Organización del Conocimiento*, em 1995, vol. 1, n.1, Francisco Javier Garcia Marco, seu editor e diretor, explicita as razões do princípio da nova revista: "A revista

*Scire: Representación y Organización del Conocimiento* quer ser mais que um instrumento de difusão. Nossa revista, antes de tudo, é o ponto de encontro de um grupo de pessoas de distintos âmbitos científicos e profissionais preocupados com o mesmo problema: a representação e organização do conhecimento" (Garcia-Marco, 1995).

Seu índice h é 3 e se encontra indexada na base de dados *Scopus*, desde 2011, além das bases *DIALNET*, *Latindex*, *Fuete Académica Premier Lisa* e *Lista*. Recebeu o selo de qualidade *FECYT* e se encontra indexada na plataforma e-revistas (plataforma de acesso aberto de Revistas Científicas Eletrônicas Espanhola e Latino-Americanas). Atualmente é uma revista pertencente ao 3º

Quartil, observando-se que, inicialmente quando foi criada, pertencia ao 4º Quartil, que desde sua indexação na *Scopus* (2011) mostrou acentuada ascendência. Segundo o *Scimago Journal & Country Rank*, dando destaque para os anos de 2012, 2013 e 2014 se produziu 19, 44 e 62 documentos respectivamente, sinalizando assim um crescimento nos últimos três anos (1). Apesar desse visível crescimento, a revista *Scire* tem, Qualis CAPES B1, no Brasil, provavelmente devido aos diferentes critérios de avaliação utilizados pela *Scopus*.

A escolha deste periódico deve-se, inicialmente, à relevância do mesmo na área de Ciência da Informação, bem como à carência de estudos realizados nas poucas revistas já consolidadas na temática 'Organização e Representação do Conhecimento', além do aumento significativo de pesquisas na área. Observe-se que, considerando a subárea Biblioteconomia e Ciência da Informação, a Espanha publica 8 revistas, sendo que a revista *Scire: Representación y Organización del Conocimiento* se coloca na 4ª posição entre elas (3º Quartil). A revista *Cybermetrics* é a única do 1º Quartil, as revistas *Profesional de la Información* e *Revista Española de Documentación Científica* no 2º Quartil e *Revista General de Información y Documentación*, *Anales de Documentación*, *Ibersid* e *BiD* no 4º Quartil (*Scimago Journal & Country Rank*, 2015).

Na revisão de literatura sobre a análise ou estudos realizados sobre a revista *Scire: Representación y Organización del Conocimiento*, destacam-se recentes pesquisas em formato de artigo, a saber: "Análise bibliométrica da revista *Scire*: um estudo de colaboração científica institucional, no período de 1996 a 2010 (Alves, Oliveira e Gracio, 2012) e um segundo estudo denominado "Relações Teóricas da Organização do Conhecimento como com Abordagens de catalogação de Assunto, Indexação e Análise documental: uma Análise de domínio da revista *Scire* (1995-2010)" (Guimarães, Pinho e Ferreira, 2012). Este último artigo realiza uma análise de domínio da revista *Scire* (1995-2010), a fim de identificar a produção científica nas linhas teóricas mencionadas, bem como relações existentes entre elas. Ainda, Guimarães et al. (2014) estudam a dimensão temática da pesquisa em OC por meio da análise de domínio dos congressos da ISKO (Brasil, Espanha e América do Norte), com o objetivo de identificar os temas e as características da colaboração científica.

Nos últimos anos, os pesquisadores tem estudado questões relativas a internacionalização do conhecimento científico e de forma mais específica, à revista *Scire*. Assim, cita-se o trabalho intitulado "*International Comparative Domain*

*Analysis in Knowledge Organization Research Topics in Four Countries - Brazil, South Korea, Spain and the United States*" (Beak, Glover, Martínez-Ávila e Milani, 2013).

No sentido de apontar as justificativas para este estudo destaca-se que, com a globalização imposta ao mundo, acirra-se a exigência pela internacionalização da pesquisa e dos periódicos científicos de forma mais instigante. Assim, a grande utilidade dessa pesquisa é avaliar o alcance que os artigos publicados pela revista em questão, são absorvidos por outros países externos à Espanha. Ainda, considera-se relevante que os autores potenciais que queiram nela publicar, tenham parâmetros substanciais para realizar tal escolha. São estas contribuições que a subárea Estudos Métricos da Informação oferecem, no sentido de avaliação da ciência produzida no país, além do seu impacto, visibilidade e consequente inserção em âmbito local e internacional.

A 'Organização e Representação do Conhecimento (ORC)' foco principal da revista, é a área que trata do desenvolvimento de técnicas para construção, gestão, uso e avaliação de classificações científicas, taxonomias, nomenclatura e linguagens documentais. Trata, ainda, de metodologias de uso e recuperação por linguagem natural, bem como das atividades baseadas nas práticas de armazenagem e recuperação de documentos no decorrer de um processo de aperfeiçoamento da sistematização do conhecimento humano, tendo como objeto de estudo o "conhecimento socializado" (Esteban-Navarro e Garcia-Marco, 1995; Barité, 2001; Fujita, 2001).

Segundo Fujita (2008) destacam-se dois aspectos na representação do conhecimento, considerando a área de Ciência da Informação: o resultado da representação do conhecimento por meio da identificação dos conceitos e a representação da estrutura lógica do conhecimento, que resulta na 'Organização do Conhecimento (OC)'. Ainda, segundo Dahlberg (2006), a representação do conhecimento compreende a estrutura lógica da representação conceitual, e, também, o resultado da identificação de conceitos, em função da terminologia utilizada, por meio dos termos utilizados.

A pesquisa em ORC resulta de um contexto dinâmico de produção e reelaboração de conhecimentos. Estes se transformam continuamente em razão da sua aplicação concreta e da comunicação científica, pois as atividades de pesquisa são determinadas pelo contexto social e histórico, pelos marcos teóricos e pressupostos científicos e também pelas diferentes metodologias existentes.



Há diferentes procedimentos na pesquisa em ORC, tanto relativo à construção do conhecimento, como em relação ao contexto da produção, entre eles as avaliações da produção científica por meio dos estudos bibliométricos. Estes, associados às análises contextuais, evidenciam o referencial teórico-epistemológico dominante na área, as relações existentes, constituindo um dos instrumentos metodológicos que contribuem para a visualização do comportamento da ciência produzida em uma dada área. Gómez-Caridad *et al.* (2004) observa que dentro de produção científica destacam-se artigos, resumos de congressos, carta, revisão, material editorial, notícias de artigos, biografias, debates, entre outros documentos.

Os indicadores básicos de produção são constituídos pela contagem do número de publicações dos pesquisadores, instituições ou país, e objetivam refletir o impacto junto à comunidade científica à qual pertencem, dando visibilidade àqueles pesquisadores, instituições ou países mais produtivos, bem como às temáticas mais destacadas de uma área do conhecimento. A relevância de se conhecer as temáticas mais destacadas de determinada área do conhecimento, é o fato das temáticas mais estudadas sinalizar as questões em torno das quais os estudiosos desenvolvem suas pesquisas, além de se constituírem focos de interesse daquela área do conhecimento e traduzirem, muitas vezes, os interesses econômicos, sociais e políticos do contexto local.

A colaboração científica entre autores, instituições ou países supõe uma consociação de hipóteses e objetivos centrais de um projeto, o estabelecimento de uma divisão de trabalho, a interação entre os investigadores e o compartilhamento de informação e conhecimento (Olmeda Gómez, Perianes-Rodriguez e Ovalle-Perandones, 2008). Katz e Martin (1997) apontam a coautoria como indicador da atividade de colaboração científica, que possibilita identificar a cooperação nacional ou internacional, bem como descrever a estrutura social e científica de um grupo, que pode ser representada por uma rede social. Neste contexto, a análise de coautoria reflete um rol possível de intercâmbios e trocas, sendo medida pelo número de publicações em colaboração entre autores, instituições ou países.

Considerando o exposto, esta pesquisa objetiva analisar, por meio de procedimentos bibliométricos, a produção científica da revista *Scire*, a fim de evidenciar a produção científica e inserção internacional da mesma no período de 2006 a 2014. De forma mais específica, objetiva-se destacar as instituições mais produtivas e seus respectivos países de origem, bem como analisar as coautorias institucionais e explicitar as temáticas

que mais têm contribuído para os debates e desenvolvimento do saber no tema, no período estudado, temas estes classificados segundo as categorias de Dahlberg (1993). E ainda, estabelecer a relação instituições e temáticas por meio de representações gráficas e analíticas.

## 5. Metodologia

Para levantamento de dados, tomaram-se os 18 volumes da revista *Scire*, de 2006 até 2014, que constam no endereço eletrônico: <http://www.iber-sid.eu/ojs/index.php/scire/issue/archive> e, manualmente constituiu-se o rol do conjunto de 154 artigos publicados no período. Inicialmente, levantaram-se as 73 instituições que mais se fizeram visíveis na revista, bem como os países a que estas pertenciam com a finalidade de se destacar os países participantes da internacionalização. A Tabela I (na página seguinte) apresenta o rol das 23 instituições que produziram pelo menos 2 artigos no período de 2006 a 2014.

Considerando o objetivo de se analisar as coautorias institucionais, levantou-se o total de 41 instituições que fizeram coautorias e a partir delas, gerou-se a rede de colaboração institucional, por meio do *software Pajek*, relativa a todo período em estudo. Utilizou-se o indicador de densidade, que permite avaliar a coesão e a estrutura da rede. Em relação aos indicadores de centralidade, calculou-se a centralidade de grau e a centralidade de intermediação, por serem os indicadores com maior significação no caso de coautorias. A primeira avalia o número de ligações que uma instituição tem, com as demais da rede (Otte e Rousseau, 2002), mostrando assim aquelas que mais fizeram parcerias. A centralidade de intermediação pode ser definida como o número de vezes que um nó (instituição) necessita para alcançar um outro nó, isto é, o número de caminhos mais curtos, que passam por determinado nó, colocando em evidência as instituições mais importantes da rede. Avaliou-se as instituições de maior inserção internacional que publicaram na revista *Scire*, no período em estudo, bem como as tendências de ampliação das subredes formadas.

Para explicitar as temáticas que mais têm contribuído para os debates e desenvolvimento do saber no tema, somente foram utilizados os artigos cujos autores trabalharam em coautoria, em um total de 83 artigos (~54%), do total de 154 artigos. Estas foram sistematizadas segundo o instrumento de classificação temática das categorias de Dahlberg (1993), o *Classification Scheme for Knowledge Organization Literature* (CSKOL). Para a classificação utilizou-se o título do artigo, resumo e as palavras-chave. Em alguns casos foi

necessário a análise do artigo na íntegra. Utilizou-se o *cluster* via apresentação do dendograma, utilizando-se o método *Ward* com distância euclidiana, sendo as instituições agrupadas por similaridade temática.

## 6. Apresentação e análise dos resultados

Em relação ao total de 73 instituições, encontraram-se 23 instituições mais produtivas, com pelo menos 2 artigos publicados. Assim, do total de instituições, 50 (68,5%) foram autoras de um

único artigo no período. Apresenta-se a Tabela I com o rol das 23 instituições que produziram pelo menos 2 artigos no período de 2006 a 2014.

Destacando as origens das 23 instituições, registra-se que 13 delas são da Espanha (56,5%), 7 do Brasil, uma do México, uma do Reino Unido e uma do Uruguai. Assim, 39% das publicações da revista *Scire* são originárias da América Latina, valor representativo, porém há baixa inserção em outros continentes, justificado pela indexação, em 2012, portanto recentemente da revista em bases internacionais.

Instituições	2006	2007	2008	2009	2010	2011	2012	2013	2014	Total
Universidade Estadual Paulista-FFC-Marília (Brasil)	3	1	0	2	6	2	6	3	5	28
Universidade de Salamanca (Espanha)	5	0	3	0	1	2	3	1	0	15
Universidade de Murcia (Espanha)	1	1	2	1	2	1	0	1	2	11
Universidade de Extremadura (Espanha)	1	4	1	0	1	0	1	0	0	8
Universidade Rey Juan Carlos (Espanha)	1	3	0	0	2	0	0	0	0	6
Universidade Carlos III de Madrid (Espanha)	1	0	0	2	0	2	0	1	0	6
Universidade Complutense de Madrid (Espanha)	1	0	2	0	0	0	1	1	0	5
Universidade de Zaragoza (Espanha)	0	0	0	1	0	1	0	2	1	5
Universidade de São Paulo (Brasil)	1	1	0	0	0	0	0	1	1	4
Universidade de Granada (Espanha)	0	1	1	0	1	1	0	0	0	4
Universidade Nacional Autónoma de México	0	3	1	0	0	0	0	0	0	4
Universidade Federal de São Carlos (Brasil)	0	0	0	0	0	1	1	1	0	3
Universidade de Alcalá de Henares (Espanha)	1	0	0	1	0	0	0	0	1	3
Universidade Federal de Santa Catarina (Brasil)	3	0	0	0	0	0	0	0	0	3
Universidade de Valladolid (Espanha)	0	0	0	1	0	0	2	0	0	3
Universidade de León (Espanha)	0	0	0	1	1	0	1	0	0	3
Instituto Brasileiro de Inf. e Tecnologia (Brasil)	1	1	0	0	0	1	0	0	0	3
Universidade Autónoma de Barcelona (Espanha)	0	0	0	0	0	2	1	0	0	3
Universidade Federal Fluminense (Brasil)	1	0	0	0	0	0	0	0	1	2
Universidade Federal de Pernambuco (Brasil)	0	0	0	0	0	0	1	0	1	2
Universidade de la República (Uruguai)	0	0	0	1	0	0	0	0	1	2
Cura Consortium and TFPL Ltd (Reino Unido)	1	0	1	0	0	0	0	0	0	2
Cindoc-Csic (Espanha)	1	0	0	0	1	0	0	0	0	2
<b>Total</b>	<b>22</b>	<b>15</b>	<b>11</b>	<b>10</b>	<b>15</b>	<b>13</b>	<b>17</b>	<b>11</b>	<b>13</b>	<b>127</b>

Tabela I. Instituições e países que produziram pelo menos 2 artigos

Portanto, a partir da análise dos dados, o Brasil é o país com maior destaque e inserção na revista em questão. É relevante destacar a estreita relação, quase que histórica entre Espanha e Brasil. Fazendo uma restropectiva, considera-se o Tratado Geral de Cooperação e Amizade entre a República Federativa do Brasil e o Reino da Espanha, em seu 12º artigo deixa explicitado a necessidade de facilitar o intercâmbio acadêmico entre

representantes das respectivas universidades, instituições de pesquisa, educação superior e cultura, bibliotecas, arquivos e outras entidades. Para esse fim, ambas as partes, Brasil e Espanha, acordam estabelecer um sistema de concessão de bolsas de estudo visando a comunicação entre os diferentes países.

Cite-se ainda, que em março de 2001, firma-se o convênio de Cooperação entre o Ministério de

*Educación, Cultura y Deporte* (MECD) da Espanha e o Ministério da Educação do Brasil, para o desenvolvimento de programas de formação, aperfeiçoamento e atualização de conhecimentos para pós-graduados e professores universitários (Capes, 2001). Há outros convênios de comparação entre estes países, que fugiriam ao escopo dessa pesquisa. Justifica-se a destacada presença dos pesquisadores brasileiros, em periódicos espanhóis. Cite-se como exemplo, especialmente na Universidade Estadual Paulista-FFC-Marília a presença de Victor Herrero-Solana, da Universidade de Granada, e com continuidade em 2005, com a visita do prof. Felix de

Moya Anégon da mesma universidade, culminando com a vinda do professor visitante Antonio Pulgarín Guerrero-Espanha, da Universidade de Extremadura, em 2010, com auxílio do Conselho Nacional de Desenvolvimento Científico e Tecnológico-CNPq.

A Tabela I apresenta ainda, a produção de artigos por ano, nas diferentes instituições e países. Destaca-se assim, na linha dos totais, a produção de 17 artigos, coincidindo com o ano de indexação da revista. Depois de 2012, a revista volta a ter número de artigos próximo ao que havia antes da indexação.

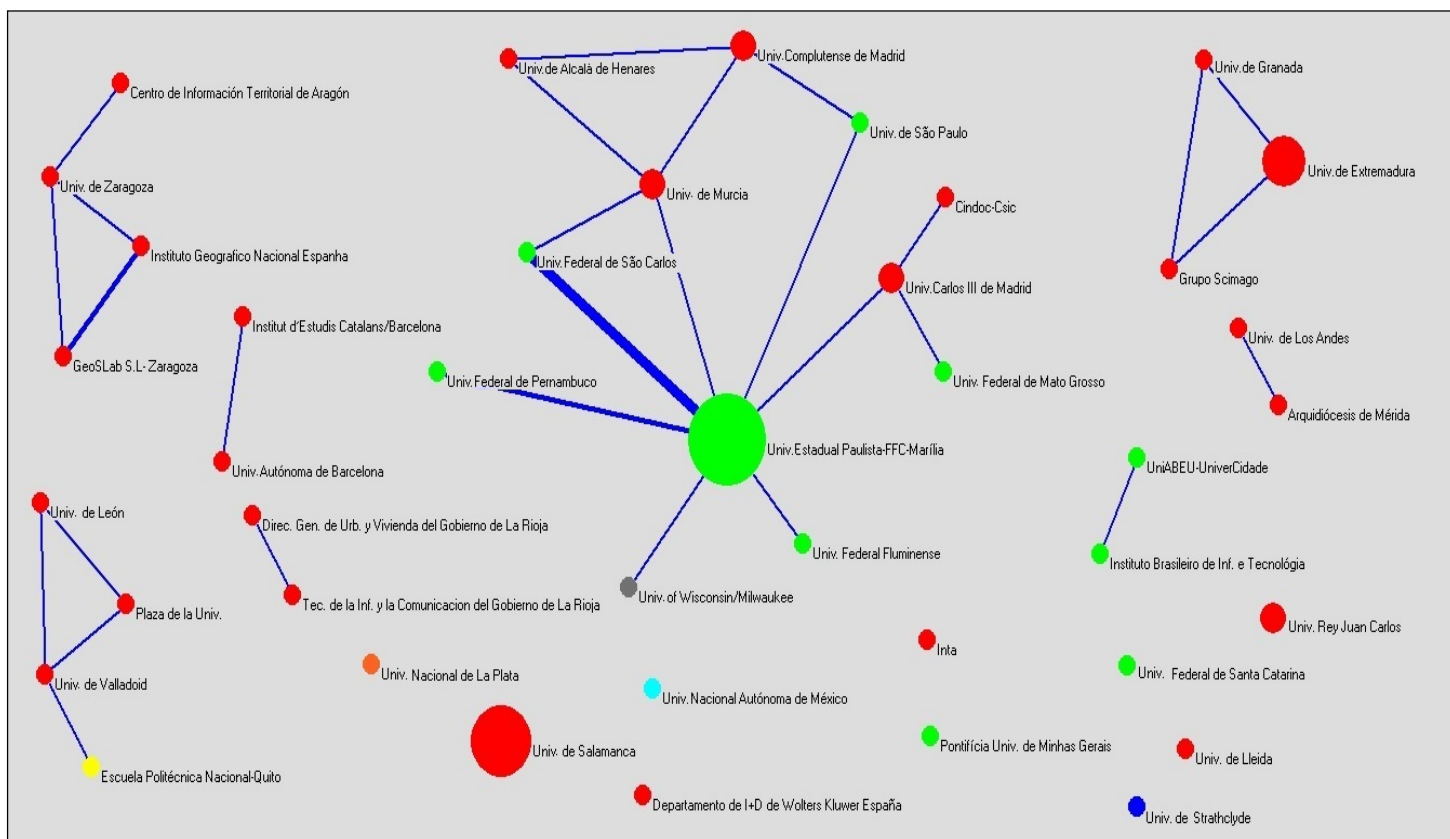


Figura 1. Rede de colaboração institucional (2006 a 2014)

Destaque-se que em 2006, há a maior produção da revista *Scire* com 22 artigos. Tem-se por hipótese que, em vista da revista ter dedicado um número especial ao tema OC em diferentes fases do ciclo da informação documental, acabou por incentivar e incrementar a pesquisa na referida temática (Garcia-Marco, 2006).

Em síntese, a Tabela I mostra por meio do estudo diacrônico, entre 2006 e 2014, uma produção equilibrada e constante de artigos, com pequenas variações, no momento de sua criação e de

sua indexação em base de dados, apesar da referida tabela apresentar somente instituições que publicaram até dois artigos.

Em relação às coautorias, apresenta-se a Figura 1, com a rede de colaboração institucional construída considerando o período total em estudo, ou seja, todas as coautorias que foram realizadas no período, para que se possa analisar a dinâmica da rede nestes 9 anos.

Os círculos em vermelho são de instituições espanholas, em verde de instituições brasileiras, o azul, amarelo e cinza de outras instituições. Os

círculos sem ligação alguma, significam instituições, cujos pesquisadores fazem somente coautoria interna; a espessura do segmento significa as frequências de coautorias. Destaca-se a subrede centrada na Universidade Estadual Paulista-FFC-Marília, com o maior número de relações, constituída de sete instituições coautoras. Ainda nesta subrede, ressalta-se o grande número de coautorias institucionais com a Universidade Federal de São Carlos, explicitada pelo segmento mais forte e um pouco mais fraco com a Universidade Federal de Pernambuco, além de parcerias feitas com instituições estrangeiras, como Universidade de Murcia, Universidad Carlos III de Madrid e *University of Wisconsin-Milwaukee* e de forma indireta com o *Cindoc-Sic*, Universidade Complutense de Madrid e Universidade Alcalá de Henares.

Esses dados mostram que a subrede centrada na Universidade Estadual Paulista-FFC-Marília tende à fazer coautorias com outras instituições estrangeiras, consignando o início da internacionalização, sem deixar de fazer suas coautorias com universidades nacionais, tais como a Universidade de São Paulo e Universidade Federal Fluminense.

Observa-se ainda, que a subrede formada pelas instituições espanholas – Universidade de León, Universidade Valadolid, Plaza de la Universidad – adquire características de um grupo internacionalizado, com a coautoria com a *Escuela Politécnica Nacional de Quito*. Ainda, destacam-se outros grupos espanhóis, a saber: *Centro de Información Territorial de Aragón*, Universidade de Zaragoza, Instituto Geográfico Nacional Espanha e GeoSLab S. L. Zaragoza; Universidade de Granada, Universidade de Extremadura e Grupo SCImago.

Identifica-se na rede de colaboração institucional quatro subredes formadas por duas instituições, sendo três delas espanholas e uma brasileira. Outras instituições espanholas, brasileiras, inglesa, mexicana e argentina, aparecem isoladas somente com coautoria interna. A densidade da rede apresenta-se, aproximadamente, com 4%, justificado pelo número das instituições relativamente alto. A maior centralidade de grau é sete, da Universidade Estadual Paulista-FFC-Marília, caracterizada pelo maior número de ligações, com centralidades menores, tais como, centralidade quatro da Universidade de Murcia e Universidade Carlos III de Madrid e Universidade Complutense de Madrid, com centralidade três. Quanto aos indicadores de intermediação, destacam-se a Universidade Estadual Paulista-FFC-Marília com 43,50 de intermediação, a Universidade Carlos III de Madrid, com intermediação 19

com intermediação em um total de 7 universidades com valor de intermediação maior que zero, é o número de caminhos mais curtos, que passam por determinado nó (instituição).

Em síntese, a rede de colaboração institucional da revista *Scire*, nesse período em estudo, apresenta-se constituída por instituições espanholas, com destacada presença de instituições brasileiras, mostrando uma consolidação de parcerias entre Brasil e Espanha, advindas de diferentes convênios institucionais firmados entre estes dois países, que proporcionam pesquisa comuns, intercâmbios de pesquisadores, como também de outras instituições estrangeiras, porém não tão significativas.

Para explicitar as temáticas que mais têm contribuído para os debates e desenvolvimento do saber no tema, no período estudado, temas estes classificados segundo as categorias de Dahlberg, apresenta-se a Tabela II, que se constituiu a partir da construção do dendograma, de forma a agrupar as instituições no entorno de temáticas similares ou próximas.

Buscou-se aporte na concepção filosófica da teoria de Dahlberg (1993), que oferecem fundamentos no sentido de se constituírem em procedimentos instrumentais utilizados nos processos de organização, interpretação e estruturação do conhecimento humano. Para Dahlberg (1993), o Sistema de Classificação da literatura em OC assim se apresenta:

- Categoria 0: Este se refere à divisão de forma dos documentos, na literatura em OC. No entanto, este não é o objeto desta pesquisa, sendo assim, nenhuma pesquisa foi neste grupo classificada.
- Categoria 1: Fundamentos teóricos e problemas gerais da OC, com os seguintes conteúdos: Ordenação e OC, Conceptologia em OC, Matemática e OC, Teoria de sistemas e OC, Psicologia e OC, Ciência e OC, Problemas em OC, Pesquisa da Classificação, História da OC;
- Categoria 2: Sistemas de Classificação e Tesouros, estrutura e construção. Apresentam-se os seguintes conteúdos: Estruturas e elementos de Sistemas de Classificação e Tesouros, Construção de Sistema de Classificação, Relações entre conceitos, Taxonomias, Notações e códigos, Manutenção, atualização e armazenagem em C&T, Concordância entre linguagens de indexação, Avaliação de Sistemas de Classificação e Tesouros;
- Categoria 3: Metodologia de Classificação e Indexação, subentendendo os seguintes conteúdos: Teoria da Classificação e Indexação,

Análise de Assunto, Técnicas de Classificação e Indexação, Classificação e Indexação Automática, Ordenação Manual e Automática,

Geração de Índices e Programas, Reclassificação, Avaliação de Classificação e Indexação, Codificação e decodificação.

Gr.	Instituições	Total	C1	C2	C3	C4	C7	C8	C9
G1	Grupo Scimago	1					1		
	Univ. Federal de Mato Grosso	1					1		
	Univ. de Lleida	1					1		
	Escuela Politécnica Nacional-Quito	1					1		
	GeoSLab S.L. - Zaragoza	1					1		
	Institut d' Estudis Catalans/Barcelona	1					1		
	Instituto Geografico Nacional Espanha	1					1		
	Tec. de la Inf. y la Com. Del Gov. de La Rioja	1					1		
	Univ. Autónoma de Barcelona	1					1		
	Direc. Gen. de Urb. y Viv. del Gob. de La Rioja	1					1		
G2	Univ. de León	2		1			1		
	Plaza de la Univ.	1		1			1		
	Univ. de Valladolid	3		2			3		
	Univ. Rey Juan Carlos	1					2		
	Dep. I+D de Wolters Kluwer Espanha	1		1			1		
	Univ. de Zaragoza	2					2		
G3	Instituto Nacional de Técnica Aeroespacial	1					1	1	
	Pontifícia Univ. de Minas Gerais	1	1						
	Univ. de Alcalá de Henares	2	1		1		1	1	
	Univ. de São Paulo	4	1	1			1	1	1
	Univ. Federal de Santa Catarina	2	1				1	1	1
	Univ. Federal Fluminense	2	1				1		
	Univ. Nacional de La Plata	1		1				1	
	Univ. of Wisconsin/Milwaukee	1	1						
	Univ. Federal de Pernambuco	2	2				1		
G4	Univ. de Granada	3					2		2
	Univ. Nacional Autónoma de México	2							2
	UniABEU-UniverCidade	1					1		1
	Instituto Brasileiro de Inf. e Tecnologia	2					1		1
	Univ. de Los Andes	1							1
	Arquidiócesis de Mérida	1							1
	Univ. of Strathclyde	1				1			1
	Univ. de Aragon	1							1
G5	Cindoc-Csic	3					3	2	
	Univ. Carlos III de Madrid	5	1				4	1	1
	Univ. Complutense de Madrid	4					3	1	1
G6	Univ. de Extremadura	6			1	1	2	1	3
	Univ. Federal de São Carlos	3	1		3		1		
	Univ. de Murcia	6	2		2		3	2	2
G7	Univ. Estadual Paulista-FFC-Marília	20	10	2	4		10	2	
G8	Univ. de Salamanca	15					14	1	2

Tabela II. Clusters das instituições articuladas com as categorias temáticas de Dahlbeg

- Categoria 4: Sistemas de Classificação e Tesouros, com os seguintes conteúdos: Dewey Decimal Classification, Library of Congress Classification, Library of Congress Subject Headings, Colon Classification; este Grupo é mais específico do que o Grupo 2.
- Categoria 5: Classificação orientada à objetos (Taxonomias), nas diferentes áreas do conhecimento.
- Categoria 6: Classificações e Tesouros de assuntos específicos.
- Categoria 7: Representação do Conhecimento por Linguagens e Terminologia, com os seguintes conteúdos: Problemas gerais de linguagem natural em relação à OC, Processamento automático da linguagem, Sistemas de recuperação *on-line* e tecnologias, Problemas de terminologia, entre outros.
- Categoria 8: Indexação e Classificação aplicadas, compreendendo os seguintes conteúdos: Problemas gerais, Catálogos, Diretrizes, Regras, Índices, Indexação e Classificação de dados, títulos, literatura primária, de livros, entre outros.
- Categoria 9: Ambiente da OC, com os seguintes conteúdos: Organização profissional e espacial (nível nacional e internacional) bem como itens de Educação e treinamento, Legislação, Econômicos, Uso de serviços, Normalização do trabalho de OC, entre outros.

Apresenta-se a Tabela II, com as diferentes instituições formando *clusters*, segundo as temáticas comuns ou similares.

Para análise dos *clusters*, destaca-se inicialmente que não foram encontrados artigos nas categorias (C) - C5 e C6, sugerindo assim que a revista *Scire* não possui como foco de pesquisa as temáticas 'Taxonomias, nas diferentes áreas do conhecimento' ou 'Classificações e Tesouros para assuntos específicos', considerando-se que as diferentes categorias de Dahlberg, não foram mutuamente exclusivas, a soma de cada linha não resulta o total registrado, devido ao fato de um artigo pertencer a uma ou mais categorias.

Encontraram-se oito agrupamentos. O primeiro deles denominado grupo G1, o maior grupo constituído de 10 instituições, apresentou cada uma delas um único trabalho em coautoria, na C7, denominada "Representação do conhecimento por linguagens e Terminologia", na qual aborda aspectos do processamento automático da linguagem, terminologia e recuperação da informação. Este grupo tem como foco de investigação os temas em C7, mas não significativa, devido à sua baixa representatividade na frequência.

O segundo grupo G2, trata das categorias C2 "Sistemas de Classificação e Tesouros, estrutura e construção" e C7 "Representação do Conhecimento por Linguagens e Terminologia". Das seis instituições todas apresentaram artigos em C7 e quatro delas em C2. Destaque-se que essas duas categorias se relacionam por apresentarem sub-temáticas similares, tais como: Relações entre conceitos, Problemas de terminologia, trabalho de terminologia orientada ao assunto, concordância entre linguagens de indexação, processamento automático da linguagem, entre outros, ampliando assim seu foco de pesquisa.

Constituído por nove universidades, o grupo G3, concentra pesquisas especialmente em C1 denominado "Fundamentos teóricos e problemas gerais da OC", C7 "Representação do Conhecimento por Linguagens e Terminologia" e C8 "Indexação e Classificação aplicadas", com pouca presença nas demais categorias. Estas instituições se preocupam com os fundamentos teóricos e conceituais, bem como os aspectos aplicados da OC, dispersando seus focos de pesquisas e por outro lado concentrando em C1, C7 e C8.

Todas as oito instituições do grupo G4, apresentam pesquisas em C9 "Ambiente da OC, com menor presença em C7 "Representação do Conhecimento por Linguagens e Terminologia" e C4 "Sistemas de Classificação e Tesouros". O C9 focaliza, especialmente, em organização profissional, bem como itens de educação e treinamento, legislação, estudos de usuários, entre outros.

O grupo G5, formado por três instituições todas elas espanholas, concentra suas pesquisas em C7 "Representação do Conhecimento por Linguagens e Terminologia" e C8 "Indexação e Classificação aplicadas", portanto para os aspectos mais técnicos da Ciência da Informação e de forma mais dispersa em C1 e C9.

As três instituições do grupo G6 contemplam pesquisas em todas as categorias de Dahlberg aqui presentes, exceto em C2, com maior ênfase em C3 e C7, que trabalham metodologias de classificação e indexação.

Ainda, o grupo G7 formado por uma única universidade, a Universidade Estadual Paulista-FFC-Marília, a mais produtiva, com destaque para C1 e C7, categorias que se articulam na medida que tratam de Fundamentos teóricos e questões de OC.

Por fim, o grupo G8 formado pela Universidade de Salamanca com prevalência para pesquisas em C7, com frequência 14 nesta categoria e ainda pesquisas em C8 e C9.

Comparando os dados do *cluster*, com a rede de colaboração institucional (Figura 1) e com as instituições mais produtivas (Tabela I), destacam-se nas três representações a Universidade Estadual Paulista-FFC-Marília, que no *cluster*, por ser a instituição mais produtiva aparece em destaque e sozinha; na rede de colaboração institucional, faz coautorias com países, como a Espanha, Estados Unidos da América-EUA e instituições do Brasil, com destaque para a Universidade de São Paulo e a Universidade Federal de São Carlos com maior frequência e centra a maior sub-rede. Analisando a rede de colaboração esta instituição é a que mais possui coautores, dentro do Brasil e com instituições espanholas.

Em relação à Universidade de Salamanca, comparando os dados do *cluster*, com a rede de co-autoria (Figura 1) e com as instituições mais produtivas, destaca-se nas três representações a Universidade de Salamanca, que no *cluster* e na rede, aparece em destaque e sozinha, por ser a instituição mais produtiva, aparecendo na Tabela I com 15 artigos e na segunda posição, logo após a Universidade Estadual Paulista-FFC-Marília. Não faz interlocução com outras instituições, porém possui grande número de coautorias interna.

#### 4. Considerações finais

Este estudo procurou caracterizar os principais momentos que determinaram a criação da revista *Scire*, desde o editorial do seu primeiro número, o ano em que foi feito um número monográfico da mesma, o ano de sua indexação em bases internacionais, bem como sua produção nas diferentes instituições, por meio de uma análise diacrônica. Identificou-se as instituições e respectivos países que publicaram na revista, para avaliar sua inserção internacional, destacando-se sua elite institucional. Por meio das palavras-chave, foram explicitadas as temáticas que contribuíram para os debates e desenvolvimento do saber no tema. Apresentou-se a rede de colaboração institucional em destaque no tema e utilizou-se a classificação de Dahlberg, agrupando-se, as instituições por similaridade temática, via representação do dendograma.

A formação das redes no período de 2006 a 2014, mostra a tendência, embora frágil, à internacionalização da revista *Scire*, especialmente, em relação aos países da América Latina. Observou-se uma rede de colaboração institucional com poucas subredes, onde as instituições são originárias de regiões tais como: América Latina, América Central, especialmente, Cuba e México, ainda América do Norte e Europa.

Há uma maior subrede que sugere a integração e expansão, para uma rede com maior consistência internacional com instituições advindas de diferentes países, bem como instituições sem interlocução com as demais. Ainda, pequenas tríades e díades formadas com instituições advindas do mesmo país, mostrando que a interlocução pouco extrapola os limites geográficos dos seus próprios países. A recente indexação da revista em bases de dados internacionais, explica a pouca coautoria internacional, apesar da relevância temática da revista.

Em relação às temáticas publicadas pela revista, elas têm como foco os fundamentos teóricos da área, sem deixar de lado os procedimentos técnicos da ORC.

No entanto, os agrupamentos das instituições, via dendograma, mostra que a revista *Scire* tem dois focos principais de pesquisa: C7 'Representação do Conhecimento por Linguagens e Terminologia' e C1 'Fundamentos teóricos e problemas gerais da OC', nas quais apresenta 70 e 22 pesquisas respectivamente, neste período em estudo dentro dos 83 artigos em coautoria (2).

A análise dos dados procurou destacar a articulação existente, entre instituições mais produtivas, presença na rede de colaboração institucional e proximidade temática entre as instituições.

Recomenda-se que este estudo possa servir de fonte para futuros trabalhos que estudam questões de internacionalização de periódicos científicos que tratam deste tema nuclear na área de Ciência da Informação.

#### Agradecimentos

Agradecemos, especialmente, ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-Brasil)

#### Notas

- (1) O Quartil é um indicador ou medida de uma posição do periódico em relação a totalidade de sua área. Se dividirmos em 4 partes iguais uma lista de revistas ordenadas do maior para o fator de impacto menor, cada uma dessas partes denomina-se Quartil. Revistas com maior Fator de Impacto (FI) será do primeiro Quartil (pertencem ao grupo dos primeiros 25%), os Quartis médios são o segundo e o terceiro Quartil é o de menor FI, está nos últimos 25% do conjunto todo (Barbetta, 1994).
- (2) Justifica-se a construção do dendograma feito somente a partir das coautorias, para assim melhor articular dendograma e rede entre si. Ratifica-se ainda que, devido as categorias serem mutuamente excludentes, a quantidade de artigos por categoria não corresponde ao total de artigos.

## Referências

- Alves, Bruno Henrique; Oliveira, Ely Francina Tannuri de; Gracio, Maria Cláudia Cabrini (2012). Análise bibliométrica da revista Scire: um estudo de colaboração científica institucional. // Scire. 18:2 (Julho 2012) 43-48.
- Barité, Mario (2001). Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación. // Carrara, Kester (Org.). Educação, universidade e pesquisa. Marília: Unesp-Marília-Publicações, 2001. 35-50.
- Beak, Jihee; Glover, Jeanette; Martínez-Ávila, Daniel; Milani, Suellen Milani (2013). International Comparative Domain Analysis in Knowledge Organization Research Topics in Four Countries - Brazil, South Korea, Spain and the United States. // Proceedings from North American Symposium on Knowledge Organization. 4:1, 30-39. <http://journals.lib.washington.edu/index.php/nasko/article/view/14643> (2017-02-19).
- Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (2001). Avaliação da pós-graduação. // CAPES.(2001).<http://capes.gov.br/acesoainformacao/perguntas-frequentes/avaliacao-da-pos-graduacao> (2016-06-16).
- Dahlberg, Ingetraut (1993). Knowledge Organization: Its Scope and Possibilities. // Knowledge Organization. 20:4, 211-222.
- Dahlberg, Ingetraut (2006). Knowledge organization: a new science?. // Knowledge Organization. 33:1, 11-19.
- Esteban Navarro, Miguel Angel; García Marco, Francisco Javier (1995). Las "Primeras Jornadas sobre Organización del Conocimiento: Organización del Conocimiento e Información Científica". // Scire. 1:1 (Enero-Junio 1995) 149-157.
- Fujita, Mariângela Spotti Lopes (2001). Organização do conhecimento: algumas considerações para o tratamento temático da informação. // Carrara, Kester. (Org.). Educação, universidade e pesquisa. Marília: Unesp-Marília-Publicações, 2001. 29-34.
- Fujita, Mariângela Spotti Lopes (2008). Organização e Representação do Conhecimento no Brasil: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. // Tendência da Pesquisa Brasileira em Ciência da Informação. 1:1, 1-32.
- García Marco, Francisco Javier (1995). Editorial. // Scire. 1 (Julio-Diciembre 1995) 1-2.
- García Marco, Francisco Javier (2006). Editorial. // Scire.1 (Enero-Junio 2006) 1-2.
- Gómez Caridad, Isabel; Fernández Muñoz, María Teresa ; Bordons, María ; Morillo Ariza, Fernanda (2004). Proyecto de obtención de indicadores de producción científica y tecnológicas de España. // Centro de Información y Documentación Científica (CINDOC) - Consejo Superior de Investigaciones Científicas (CSIC), Madrid. (2004). <http://digital.csic.es/bitstream/10261/12260/1/informe1.pdf> (2016-06-16).
- Guimarães, José Augusto Chaves; Pinho, Fábio Assis; Ferreira, Gustavo Meletti. (2012). Relações Teóricas da Organização do Conhecimento como com Abordagens de catalogação de Assunto, Indexação e Análise documental: uma análise de domínio da revista Scire (1995-2010). // Scire. 18:2 (Agosto 2012) 31-41.
- Guimarães, José Augusto Chaves et al. (2014). A dimensão temática da pesquisa em organização do conhecimento: uma análise de domínio dos congressos nacionais e regionais da ISKO (Brasil, Espanha e América do Norte). // Scire. 20:2 (Jul.- Dic. 2014) 19-25.
- Katz, J. Sylvan; Martin, Ben R. (1997). What is research collaboration?. // Research Policy. 26 (January 1995) 1-18.
- Olmeda Gómez, Carlos; Perianez-Rodríguez, Antonio; Ovalle-Perandones, María Antonia (2008). Estructura de las redes de colaboración científica entre las universidades españolas. // Ibersid: Revista Internacional de Sistemas de Información y Documentación, Zaragoza, 2, 129-140.
- Otte, E.; Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the Information sciences. // Journal of Information Science. 28:6 (December 2002), 441-453.
- SCImago Journal & Country Rank (2015). Science analysis. // SJR. (2015). <http://www.scimagojr.com/> (2016-06-16).
- Scopus (2011). What does it cover?. // Elsevier. (2011). <https://www.elsevier.com/solutions/scopus> (2016-06-16)
- Tratado Geral de Cooperação e Amizade entre a República Federativa do Brasil e o Reino da Espanha (1994). Decreto nº 1.211. // Presidência da República. (1994). [http://www.planalto.gov.br/ccivil\\_03/decreto/1990-1994/D1211.htm](http://www.planalto.gov.br/ccivil_03/decreto/1990-1994/D1211.htm) (2016-06-16).

Enviado: 2016-08-16. Segunda versão: 2017-02-22.  
Aceptado: 2017-05-09.



---

# A catalogação de assunto e a indexação em bibliotecas universitárias: um estudo comparativo com protocolo verbal

*La catalogación de materias y la indización en bibliotecas universitarias: un estudio comparativo con protocolo verbal*

*Subject cataloguing and indexing in university libraries: a verbal-protocol comparative study*

---

**Roberta Cristina Dal'Evedove TARTAROTTI y Mariângela Spotti Lopes FUJITA**

Universidade Estadual Paulista, Av. Hygino Muzzi Filho, 737, Campus Universitário, CEP. 17525900, Marília, SP, Brasil, roberta\_tartarotti@yahoo.com.br, mariangelaslf57@gmail.com

## Resumen

Se investigan las diferencias en el tratamiento temático de libros y tesis doctorales realizadas por los bibliotecarios en bibliotecas universitarias a través de la indización en la base de datos y de la catalogación de materias en el catálogo colectivo en línea. Mediante la metodología cualitativa de recogida de datos denominada Protocolo Verbal Individual (PVI), se estudiaron cinco sujetos (2 catalogadores, 2 indizadores y 1 catalogador-indizador) en tres bibliotecas universitarias públicas de la Universidad de São Paulo, Brasil. Se constata que los profesionales realizan la catalogación de materias y la indización de tesis de manera distinta, ya que en las bases de datos el proceso de indización se realiza a un nivel más alto de especificidad que la catalogación de materias en el catálogo colectivo en línea. Parece que esta actitud diferencial de los profesionales está directamente relacionada con la concepción del análisis documental que adquirieron durante su formación universitaria y con la política de indización del sistema donde actúan, especialmente con los diferentes vocabularios controlados utilizados en los sistemas de recuperación de la información.

**Palabras clave:** Indización. Catalogación de materias. Protocolo verbal individual. Catálogos colectivos en línea. Bases de datos. Tesis doctorales. Bibliotecas universitarias. Brasil.

## 1. Introdução

No âmbito da Ciência da Informação, a área de Organização e Representação do Conhecimento dispõe do arcabouço teórico-prático necessário no que tange à produção, tratamento e recuperação da informação/conhecimento na sociedade, tendo o tratamento da informação como elo complementar entre a produção e o uso de informações. Embora presente em outras unidades de informação como arquivos e museus, a abordagem teórica da catalogação de assunto, conforme Fujita, Rubi e Boccato (2009, p. 39) está essencialmente ligada à construção de catálogos

## Abstract

The differences in the subject cataloguing and indexing of books and doctoral theses carried out by university librarians were analyzed. Using the qualitative data collection methodology called Individual Verbal Protocol (PVI), five subjects (2 catalogers, 2 indexers and 1 cataloger-indexer) were studied from three public university libraries of the University of São Paulo, Brazil. It was verified that the professionals perform the indexing of theses in the databases at a higher level of specificity than the cataloging of library materials for the online catalog. It seems that this differential attitude of the professionals is directly related to the conception of document analysis that they acquired during their university education and to the indexing policy of the system where they act, specially with the different controlled vocabularies used in the information retrieval systems.

**Keywords:** Indexing. Subject cataloguing. Verbal Protocol. Online public catalogues. Databases. University libraries. Doctoral thesis. Brazil.

de bibliotecas, enquanto que a indexação está ligada à construção de índices de bibliografias em serviços de informação bibliográficos que produzem bases de dados. Ainda para as autoras, a atuação dos catálogos como verdadeiras bases de dados é uma tendência devido a dois principais fatores:

a dimensão que a internet deu aos catálogos das bibliotecas, uma vez que agora eles estão disponíveis sem fronteiras espaciais e temporais, permitindo ao usuário acessá-los de qualquer lugar a qualquer hora; e a exigência cada vez maior do usuário em aspirar a que os catálogos atuem como verdadeiras bases de dados, oferecendo especificidade, rapidez e hiperlinks a textos completos.

Como responsáveis pelo armazenamento, localização, busca e seleção, sob demanda, de dados pertinentes a um determinado assunto, os catálogos coletivos online e as bases de dados de bibliotecas universitárias são considerados sistemas de recuperação da informação. Contudo, se a qualidade da catalogação de assunto e da indexação inexistir no tratamento/organização documental, conseqüentemente se refletirá na qualidade dos próprios produtos. Desta forma, nestes sistemas de recuperação da informação a qualidade da representação temática está diretamente relacionada à qualidade da análise temática, considerando sua complexidade inerente ao fazer profissional.

Partindo-se da relevância de investigações teórico-metodológicas em torno do profissional que realiza a análise de assunto, visando contribuir para a otimização do tratamento temático da informação em contexto de bibliotecas universitárias, o objetivo do artigo é investigar a atuação bibliotecária no tratamento temático da informação de bibliotecas universitárias, por meio da indexação em base de dados e da catalogação de assunto em catálogo coletivo online, utilizando-se como metodologia qualitativa de coleta de dados a técnica introspectiva do Protocolo Verbal (PV) sistematizada por Ericsson e Simon (1993), na modalidade Protocolo Verbal Individual (PVI).

## **2. Os processos de catalogação de assunto e de indexação em bibliotecas universitárias**

As bibliotecas universitárias foram uma das primeiras instituições em prover o uso de sistemas de recuperação da informação, em duas principais frentes: provendo acesso tanto para os catálogos da própria biblioteca como para bases de dados comerciais. Na contemporaneidade, diversos fatores contribuíram para a utilização dos sistemas de recuperação da informação em bibliotecas universitárias de forma mais efetiva.

Os catálogos coletivos online e as bases de dados de bibliotecas universitárias são sistemas de recuperação da informação fundamentais, na medida em que possuem âmbito informacional definido, informação especializada e fornecimento de produtos informacionais específicos.

O catálogo é um produto da catalogação, atividade de natureza descritiva da área de tratamento da informação. Como um sistema de recuperação da informação, tanto a catalogação (processo), como o catálogo (produto) permitem a mediação entre documentos e usuários. Por realizarem um processo de comunicação, são determinados historicamente de acordo com o contexto em que estão inseridos.

A catalogação de assunto é a dimensão do processo de catalogação responsável pela análise de assunto de itens informacionais, bem como o fornecimento de notações de classificação e cabeçalhos de assunto que representem o assunto dos documentos. É definida como “representação, nos catálogos de biblioteca, dos assuntos contidos no acervo” (Fiuza, 1985, p. 257) e uma operação intelectual do exame do documento para extração e representação de assuntos nele contidos, que servem de ponto de acesso para busca e recuperação do documento original no acervo, bem como para agrupamento das obras por área de abrangência temática dos documentos relacionados entre si (Shoham y Kedar, 2001).

Por outro lado, as bases de dados são um conjunto de dados estruturados de forma sistemática, geralmente específicos sobre algum campo do conhecimento. Atualmente, as bases de dados em bibliotecas universitárias promovem a divulgação e comunicação de informações especializadas e criação de conhecimento, visto seu papel em fornecer informação atualizada, precisa e confiável, de acordo com a demanda dos usuários (Tartarotti, 2014, p. 112).

No processo de atribuição de assuntos em bases de dados ou indexação, os assuntos utilizados para descrever os documentos dependem de cada área científica especializada. Estes assuntos podem ser descritos por meio de informações localizadas em determinadas partes do documento, como título, resumo ou em um vocabulário controlado. A maioria das bases de dados utilizam termos de indexação em um campo descritor, geralmente oriundos de um tesouro específico, ou seja, um vocabulário de indexação formalmente estruturado (Rasmussen, 2011, p. 699).

Conceitualmente, a indexação é um processo formado por subprocessos ou etapas que tem como objetivo identificar o conteúdo de um documento, por meio de uma metalinguagem construída – a linguagem documental – com o intuito de promover a efetiva recuperação da informação (Tartarotti, 2014, p. 21). Suas principais etapas são leitura documental, análise de assunto/identificação de conceitos, seleção de conceitos e tradução de conceitos.

Sintetizadas por Gil Leiva (2008, p. 62), as diferentes abordagens na prática da indexação de documentos por indexadores podem ser reunidas em quatro vertentes presentes na literatura: indexação orientada para o documento, indexação orientada para o conteúdo, indexação orientada para o usuário e indexação orientada para o

domínio, sendo esta última considerada a concepção ideal, ao considerar outros elementos além do documento ou o usuário.

O objetivo da prática profissional tanto do indexador como do catalogador de assunto é tornar conhecido o assunto de um documento, ou seja, representar seu conteúdo por meio de conceitos significativos. Mesmo considerados como processos de tratamento temático da informação distintos, a qualidade da indexação e da catalogação de assunto está relacionada à capacidade de reconstruir o assunto tratado em um documento em conceitos para recuperação posterior pelo usuário do sistema de informação.

### 3. Procedimentos metodológicos

No intuito de realizar a coleta de dados com os catalogadores com a técnica introspectiva do PVI para observação comparativa do processo de análise de assunto na catalogação de assunto e na indexação de teses na área de Saúde, elaborou-se o plano metodológico da pesquisa.

O universo de pesquisa constitui-se de três bibliotecas universitárias públicas da Universidade de São Paulo: Biblioteca da Faculdade de Odontologia (FO), Biblioteca da Faculdade de Saúde Pública (FSP) e Biblioteca do Instituto de Psicologia (IP). A amostra foi definida após pesquisa nos sistemas de recuperação da informação e contato com as bibliotecas participantes, verificando-se que realizam, com as mesmas teses, tanto a catalogação de assunto no catálogo coletivo online DEDALUS(1) do Sistema de Bibliotecas da Universidade de São Paulo quanto a indexação na base de dados Literatura Latino-americana e do Caribe em Ciências da Saúde (LILACS) (2).

Cabe esclarecer que, inicialmente, a proposta da pesquisa seria realizar um estudo comparativo no contexto do Conselho de Reitores das Universidades Estaduais Paulistas (CRUESP) (3), ou seja, tendo como universo de pesquisa uma biblioteca da USP, uma biblioteca da UNESP e uma biblioteca da UNICAMP com o documento tipo livro. Entretanto, verificou-se que nenhuma biblioteca da UNICAMP realiza a indexação de livros na LILACS, justificando-se, assim, a escolha de teses para este estudo.

No tocante à infraestrutura material para a realização da aplicação do PVI, as bibliotecas atenderam os quesitos necessários. Para tanto, foi utilizado o próprio ambiente de trabalho do bibliotecário catalogador e indexador, ou seja, com os sujeitos em situação natural, para que assim a coleta de dados pudesse refletir a realidade das atividades de catalogação e indexação.

Salienta-se que as bibliotecas possuíam ambientes adequados para a aplicação do PVI, pois as atividades a serem investigadas (catalogação e indexação) fazem parte da rotina dos bibliotecários que nelas atuam. A coleta de dados foi realizada uma única vez em cada Biblioteca universitária, gerando 6 protocolos verbais, com 5 sujeitos participantes. No caso da Biblioteca da Faculdade de Saúde Pública (FSP), considerando-se que o mesmo profissional realiza tanto a catalogação de assunto no DEDALUS quanto a indexação de teses na LILACS, foram aplicados dois (2) protocolos com o mesmo bibliotecário. A aplicação do PVI pautou-se nos seguintes procedimentos metodológicos (adaptados de Nardi, 1999).

#### 3.1. Procedimentos anteriores à aplicação do PVI

*Definição do universo da pesquisa:* Após verificação de quais bibliotecas realizavam tanto a catalogação de assunto em catálogo coletivo online como a indexação em base de dados, as bibliotecas universitárias foram selecionadas.

*Seleção dos sujeitos:* A aplicação da técnica do PVI ocorreu com 5 bibliotecários que atuam no tratamento temático da informação, mais especificamente nas atividades de catalogação de assunto e indexação das bibliotecas supracitadas. As coletas propiciaram a geração de 6 protocolos verbais. Ressalta-se a discrepância entre os números apresentados, considerando-se que em uma das bibliotecas pesquisadas, o mesmo profissional realizou tanto a catalogação de assunto como a indexação de tese.

*Definição da tarefa:* Realizar a catalogação de assunto/indexação de uma tese no DEDALUS e da mesma tese na LILACS, sendo 1 em cada biblioteca, para observação por meio da técnica do PVI, que consiste na gravação de voz, enquanto o profissional realiza a atividade proposta.

*Seleção do texto-base:* O material selecionado para o tratamento temático da informação foram três (3) teses de doutorado pertencentes à área científica especializada das bibliotecas pesquisadas:

Gomes, N. F. (2010). A conduta moral na administração pública: um estudo com ocupantes de cargos comissionados. // 202 f. Tese (Doutorado em Psicologia) – Instituto de Psicologia, Universidade de São Paulo.

Silva, B. M. H. (2010). Efeito do tratamento térmico sobre a sorção e solubilidade, tenacidade à fratura e grau de conversão de compósitos restauradores diretos. // 88 f. Tese (Doutorado em Ciências Odontológicas) – Faculdade de Odontologia, Universidade de São Paulo.

Bocchiglieri, M. M. (2010). O lixiviado dos aterros sanitários em estações de tratamento dos sistemas públicos de esgotos. // 257 f. Tese (Doutorado em Saúde Pública) – Faculdade de Saúde Pública, Universidade de São Paulo.

*Conversa informal com os sujeitos:* Realizou-se uma conversa informal com a direção e com os sujeitos catalogadores e indexadores de cada biblioteca universitária por telefone e e-mail, resultando na aceitação e definição das datas para a realização da coleta de dados. Antes da aplicação do PVI, foram mencionados os objetivos da pesquisa e a importância da mesma para o desenvolvimento da área de tratamento temático da informação em bibliotecas universitárias. Foi entregue para cada sujeito o Termo de Consentimento Livre e Esclarecido (TCLE), assinado tanto pela pesquisadora como pela orientadora. Solicitou-se aos sujeitos a assinatura do documento e aceitação formal da participação na pesquisa, ficando uma cópia para os sujeitos. Ressaltou-se, ainda, que a identidade de cada um dos sujeitos permaneceria oculta, com o propósito de não comprometer os dados e deixá-los à vontade durante a realização da tarefa de modo mais natural possível, seguindo sua rotina diária de atividade.

*Familiarização dos sujeitos com a técnica do PVI:* Os sujeitos participantes foram apresentados à técnica introspectiva do PVI e seus respectivos procedimentos, sendo-lhes instruídos no modo como deveriam proceder durante a aplicação da mesma. Foi esclarecido que é preciso “pensar alto” durante a leitura e exteriorizar seus processos mentais, procurando esquecer a presença da pesquisadora que esteve presente apenas com o intuito de lembrar que é preciso “pensar alto” e controlar o gravador.

### 3.2. Procedimentos durante a aplicação do PVI

*Gravação do “pensar alto” dos sujeitos participantes:* Foi realizada a gravação do “pensar alto” dos bibliotecários durante a catalogação de tese no catálogo coletivo online e a indexação na base de dados, por meio de um gravador digital. A conversa informal, familiarização com a tarefa, instruções aos sujeitos e a aplicação do PVI foram aplicadas individualmente com cada profissional, respeitando sua própria individualidade e necessidades.

### 3.3. Procedimentos após a aplicação do PVI

*Transcrição literal das gravações das falas dos sujeitos participantes:* Foram realizadas as transcrições literais das falas dos sujeitos participantes, preservando a identidade dos mesmos por meio de siglas específicas de acordo com a atividade realizada – indexador/catalogador de assunto, sendo que para as transcrições utilizou-se como apoio o software *Express Scribe Transcription*. No intuito de obter uma melhor visualização dos processos adotados pelos sujeitos, nas

transcrições foram utilizadas notações específicas (adaptadas de Cavalcanti, 1989), por exemplo, as apresentadas no Quadro 1:

Notação	Significado
....	pausas curtas
...~~	pausas longas
((RM))	tom de ironia do sujeito
((RI))	riso do sujeito
((FR))	vocalização e riso ao mesmo tempo do sujeito
[...]	trecho do texto-base vocalizado pelo sujeito
< >	termos identificados pelo sujeito
(< >)	termos finais atribuídos ao documento pelo sujeito
(->)	trecho do texto-base “saltado” (ignorado) na leitura
“...”	trecho de instruções do sistema de recuperação da informação
{ }	trecho de material de apoio (obra de referência) utilizado para compreensão do conteúdo documental pelo sujeito ou definição do termo no sistema de recuperação da informação
(...)	omissão de trecho não relevante na transcrição

Quadro 1. Notações específicas para as transcrições

*Leitura detalhada dos dados em busca de fenômenos significativos e recorrentes para elaboração de categorias de análise:* Posteriormente, foi realizada leitura detalhada das transcrições com o intuito de buscar fenômenos significativos para a construção das categorias de análise dos dados.

*Construção das categorias de análise:* Após leitura detalhada dos dados, foram construídas as unidades de análise, eixos temáticos e categorias de análise baseadas tanto na literatura como na prática profissional, por meio das quais as informações obtidas pela aplicação do PVI foram estruturadas.

*Retorno aos dados para retirar trechos da discussão que exemplifiquem cada categoria de análise:* Realizou-se a releitura das transcrições visando à retirada de trechos da discussão que melhor exemplificassem cada categoria de análise, mediante a elaboração de quadros e síntese dos principais aspectos observados.

## 4. Resultados e discussão

Para atender aos objetivos da pesquisa, com os dados coletados dos protocolos verbais individu-

ais foi realizada a transcrição integral dos mesmos, em busca de fenômenos significativos como subsídios à elaboração de categorias de análise, pautadas tanto nos aportes teóricos como nas declarações realizadas pelos sujeitos participantes. Por sua vez, cada eixo temático com suas respectivas categorias de análise é apresentado em quadros, com o intuito de evidenciar trechos das declarações dos sujeitos participantes durante a realização da catalogação de assunto e indexação por meio da abordagem qualitativa do PVI. Salienta-se que cada trecho foi identificado por meio das seguintes siglas (Quadro II).

<i>Profissional</i>	<i>Sigla</i>
Catalogador da Faculdade de Odontologia	(C-FO)
Indexador da Faculdade de Odontologia	(I-FO)
Catalogador da Biblioteca da Faculdade de Saúde Pública	(C-FSP)
Indexador da Biblioteca da Faculdade de Saúde Pública	(I-FSP)
Catalogador da Biblioteca do Instituto de Psicologia	(C-IP)
Indexador da Biblioteca do Instituto de Psicologia	(I-IP)

*Quadro II. Siglas dos sujeitos participantes das bibliotecas pesquisadas utilizadas nas transcrições*

Desse modo, são apresentadas a seguir as unidades de análise, os eixos temáticos e as categorias de análise elaboradas como subsídios para a análise qualitativa (Quadro III).

#### 4.1. Aspectos relacionados ao processo documental

##### 4.1.1. Eixo temático: Tratamento descritivo da informação para catalogação (catálogo coletivo online)

*Manual de catalogação descritiva:* O catalogador da FO pautou-se tanto em um manual elaborado pelo próprio bibliotecário como na ajuda de campo do software ALEPH do DEDALUS, que apresenta ainda exemplos para a catalogação descritiva de forma padronizada no SIBiUSP.

e tudo o que eu for catalogar eu tenho que seguir isso aqui ó ... "Orientações de Catalogação", tá vendo? (...) Então eu vou ter que preencher campo a campo ... todos os dados (C-FO)

Os catalogadores da FSP e do IP também utilizaram a ajuda de campo do sistema. Além disso, para cada tipo de material é apresentada uma planilha *default* com os campos de preenchimento obrigatório e ao enviar o registro para o

sistema, caso tenha alguma inconsistência, o mesmo acusa, facilitando a atividade.

<i>Unidade de análise</i>	<i>Eixos temáticos</i>	<i>Categorias de análise</i>
Aspectos relacionados ao processo documental	Tratamento descritivo da informação para catalogação (catálogo coletivo online)	· Manual de catalogação descritiva
	Tratamento descritivo da informação para indexação (base de dados)	· Manual de descrição documental
	Tratamento temático da informação (catalogação de assunto/indexação)	· Análise de assunto/ identificação de conceitos · Seleção de conceitos · Representação/ tradução de conceitos
	Linguagem documental	· Adequação/atualização da linguagem documental para catalogação de assunto/indexação · Desempenho da linguagem documental para catalogação de assunto/indexação · Pertinência da linguagem documental com a linguagem da comunidade usuária
Aspectos relacionados ao catalogador de assunto/ indexador	Atuação profissional do catalogador/ indexador	· Concepção sobre catalogação de assunto · Concepção sobre indexação · Fatores interferentes · Formação do catalogador de assunto · Formação do indexador
	Política de tratamento temático da informação documental	· Manual de catalogação de assunto · Manual de indexação · Nível de especificidade/ exaustividade · Quantidade de termos
Aspectos relacionados ao contexto de bibliotecas universitárias	O usuário do sistema de recuperação da informação	· Recuperação por assunto em catálogo coletivo online e em bases de dados na biblioteca · Preocupação com a perspectiva do usuário

*Quadro III. Elaboração das categorias e subcategorias de análise a partir dos referenciais teóricos apresentados nas seções anteriores e dos objetivos da pesquisa*

##### 4.1.2. Eixo temático: Tratamento descritivo da informação para indexação (base de dados)

###### *Manual de descrição documental*

Quando surge alguma, às vezes na hora que você tá preenchendo surge umas dúvidas, tal, então você tem no próprio aplicativo você tem como você buscar a informação ... pra te ajudar, uma informação auxiliar ... tá? (I-FO)

O indexador da FO pautou-se na ajuda de campo do software LILDBI-Web da Metodologia LILACS para descrição documental, o que denominou de "informação auxiliar ou assistente". Este manual, denominado de *Manual de Descrição Bibliográfica* orienta o preenchimento dos campos de dados definidos no formato LILACS, descreve cada campo de dados e apresenta exemplos de preenchimento. A padronização de alguns elementos segue a AACR2 (Código de Catalogação Anglo-Americano) e a ISO (International Standard Organization), possibilitando o controle de qualidade e cooperação no tratamento das informações bibliográficas. O profissional ressalta ser positivo o fato de o sistema acusar alguma inconsistência por meio de alertas coloridos.

O indexador da FSP também utilizou o assistente para realizar a descrição documental da tese, inclusive explicitando durante o protocolo algumas instruções, sanando suas dúvidas. O indexador do IP compara a descrição documental da LILACS com a catalogação do MARC no DEDALUS, pois os campos obrigatórios na planilha *default* na LILACS aparecem em destaque, na cor vermelha, proporcionando uma descrição mais rápida, enquanto que no DEDALUS isso não ocorre. Além disso, sugere que o campo de preenchimento do *Currículo Lattes* do autor da tese, implantado mais recentemente, também possa ser validado pelo sistema.

#### 4.1.3. Eixo temático: Tratamento temático da informação (catalogação de assunto/indexação)

##### Análise de assunto/identificação de conceitos

O que que a gente costuma fazer, a gente vai, eu já abro quando pega a tese, já vou e já abro direto no ... no resumo ... certo? Aí eu dou uma ... dou uma lida assim, né? Dou uma lida no resumo, tal. É ... tento ver também com relação ao título né e no resumo (...) Como às vezes é ... muito, eu tento dentro do texto já no olhar clínico já tento tirar algumas palavras ... né? Algumas palavras, isso dentro de uma experiência já adquirida de indexação então eu já vou direto e já olho, já vejo aqui se é possível tirar algumas palavras ... aí ... olhei né, aí o que que eu faço? Olho na ficha catalográfica que tem também a ... já tem aqui a indexação. (I-FO)

Deixa eu olhar nas palavras-chave <Estações de tratamento>, <Sistemas públicos de esgoto> ... Tô abrindo no SIBIX (...) enquanto isso eu vou aqui pro resumo da tese ... vou começar indexando pelos termos principais que aparecem no título depois eu vou pro objetivo ... no resumo ... aqui (C-FSP)

Todos os profissionais, sejam catalogadores de assunto ou indexadores, realizaram a análise do conteúdo do documento, entretanto, em diferentes níveis de profundidade. A leitura documental

é considerada como a fase mais importante do processo de análise de documentos, sendo o sucesso nesta fase inicial que garantirá o restante do processo.

##### Seleção de conceitos

Estes foram os termos que ficaram. São cinco termos que foram dados pra essa tese e eu vou agora (...) voltar no formulário da catalogação pra acrescentar mais campos aqui, só tenho um. ... Então eu vou criar mais campos que faltam pra eu colocar todos os assuntos (C-IP)

##### Representação/tradução de conceitos

É ... <Tenacidade>. Tenho aqui o descritor exato (...) palavra ou termo. Então eu vou colocar <Tenacidade>. <Tenacidade> não tem. Volta ... vou pra palavra. Também não. Então... sem chance. Daí, bom, eu tento... buscar alguma ... tento esmiuçar o máximo que eu posso, mas (...) Aí eu vou lá e coloco as palavras do autor, porque aí é o autor que tá colocando ... <Tenacidade à fratura>, vamos ver aqui, <Grau de conversão> ... <Grau de conversão>, também não ... Termo exato também não, acho que não vai ter ... bom... não consegui ir mais ... dentro do que tá aqui (I-FO)

então é só transportar do vocabulário pro formulário de catalogação do DEDALUS ... essa parte da indexação então ... já tá ... pronta. (...) Bom, então essa tese ficou com os termos (<Ética>) ... (<Moral>) ... (<Trabalho>), (<Administração pública>) e (<Teoria crítica>). (C-IP)

<Individualização> tem no DeCS, mas <Individualização - Psicologia> não tem. (...) Mas eu ficaria numa dúvida aqui, colocar ou não. Na dúvida ... regra ... geral que a gente dá ... embora tem que ter ... pelo menos 50, 70% de certeza. ((FR)) (...) Por exemplo, aqui <Psicologia do indivíduo> é um ... termo que tá parecendo aqui é um campo totalmente ... que vai em outra linha ... não é ... esse eu não usaria ... <Formação do indivíduo> ... tá aqui ó. ... Você tá vendo como a nossa terminologia tão específica que é né? (...) Aqui já foram quatro termos hein? O que eu faço também é depois dar uma olhadinha, pelo menos na parte em português e já ... reavaliar se eu não to colocando um muito geral ou um concorrente que não necessitaria ... (I-IP)

O catalogador da FO explicita que o SIBIX não alcança o nível de especificidade abordado na tese. Desse modo, termos que poderiam ser atribuídos ao documento não puderam ser utilizados devido à restrição do vocabulário controlado utilizado na etapa de representação e tradução de conceitos. Entretanto, a mesma situação ocorre com o indexador da mesma instituição, pois ao tentar traduzir na linguagem de indexação termos identificados e selecionados durante a análise de assunto, percebe que no DeCS os termos não são localizados.

O catalogador da FSP denota que não será utilizada, na etapa de tradução, a mesma quantidade

de termos identificados e selecionados. O catalogador do IP, ao analisar a ficha catalográfica da tese, expõe que, pela sua experiência profissional, é possível verificar se os termos ali contidos fazem ou não parte do vocabulário controlado, ou seja, do SIBIX. O indexador do IP, ao mesmo tempo em que identifica e seleciona os conceitos significativos do conteúdo da tese, também realiza a pesquisa dos descritores no vocabulário controlado específico da área de Psicologia. O vocabulário controlado também auxilia o profissional na compreensão do conteúdo do documento.

Fujita, Rubi e Boccatto (2009, p. 24) explicam que durante o processo de indexação, as etapas não precisam ser, necessariamente, realizadas de maneira sequencial, visto que o indexador profissional já familiarizado com o processo pode realizá-las simultaneamente. Do mesmo modo, quanto mais familiarizado com a área científica especializada no qual atua, mais estas etapas são superpostas durante os processos de catalogação de assunto/indexação em bibliotecas universitárias.

#### 4.1.4. Eixo temático: Linguagem documental

##### *Adequação/atualização da linguagem documental para catalogação de assunto/indexação*

ai as palavras-chaves é do autor, e tá aqui, posso colocar só, mas aí o que que acontece? É ... como eu faço a indexação pelo DeCS ... que é um aplicativo da BIREME ... tá? Daí ele tem a palavra-chave, tal. Então geralmente eu uso isso aí só depois se caso eu não conseguir nada pelo DeCS aí eu coloco esse aqui ... coloco as palavras-chave do autor. (I-FO)

palavra-chave do autor ... esse a gente tá usando ainda, que já é justamente aqueles termos não ... inclusos (...) ... então eu coloco aqui e coloco em um arquivo word depois e trimestralmente eu passo pra (...) ... porque eu não posso pré-julgar e dizer que aquele termo ... que o autor colocou pra tese dele ... não estando no DeCS, não estando na ... terminologia nossa ... não seja válido. Eu não posso pré-julgar, dizer não, esse termo não se aplica ... Aí vem em seguida os nossos termos ... né? ... Então eu venho aqui na terminologia ... aí que começa a brincadeira da indexação. (I-IP)

O catalogador da FO explica que quando o registro da tese é carregado no ALEPH (oriundo da Biblioteca Digital), no campo 950 (termo livre) são carregados os assuntos atribuídos pelo autor da tese (palavras-chave do autor quando da elaboração de sua ficha catalográfica).

Este mesmo campo é utilizado pelo catalogador para inserir termos que julga relevantes para a recuperação da tese, mas cujo vocabulário controlado (VocaUSP) não contempla. Ressalta a importância de acréscimo destes termos em lin-

guagem natural no campo livre, para futura análise de sua inserção ou elaboração de remissivas no vocabulário controlado. O indexador da FO salienta que, caso as palavras-chave não sejam contempladas na linguagem controlada da LI-LACS – O DeCS, as mesmas são utilizadas no campo “palavras-chave do autor” no registro do documento. Além disso, compara o DeCS com o SIBIX, considerando que neste último, devido às atualizações, é mais provável encontrar termos para a tradução na indexação.

O indexador do IP também se vale do mesmo campo para inserir os termos potenciais que poderiam fazer parte da terminologia em Psicologia. Desse modo, o bibliotecário que atua no tratamento temático da informação desempenha um importante papel ao realizar o levantamento dos termos relevantes/pertinentes que representem o conteúdo do documento e que possam ser buscados pelos usuários do sistema de recuperação da informação, mas que não fazem parte da linguagem documental, ou seja, do vocabulário controlado utilizado pela instituição. É importante destacar que tanto o catálogo coletivo online como a base de dados utilizam os vocabulários controlados, VocaUSP para a catalogação de assunto e DeCS para a indexação, embora em diferentes níveis de especificidade e abrangência.

## 4.2. Aspectos relacionados ao indexador

### 4.2.1. Eixo temático: Atuação profissional do catalogador/indexador

#### *Concepção sobre indexação*

É lógico, como diz o ... o Lancaster qualquer indexação por, melhor que ... mais racional que ela tem que ser ela é ... subjetiva ... e ... no caso das outras ... é ... Unidades ... às vezes ... o ... o profissional bibliotecário deixa passar né? ... Termos ... que é o próximo que eu vou colocar é ... <Percepção do sujeito> (I-IP)

Somente o IP explicita sua concepção sobre indexação, pois percebe a subjetividade do processo de indexação e a relevância do profissional em selecionar termos que representam o conteúdo do documento.

#### *Formação do indexador*

Uma coisa que facilita em termos de indexação lá na frente, até pro professor a gente sabe a linha ... teórica que ele escreve ... no caso eu como são mais de 80 professores aqui da Escola eu busco ainda ... entender ... por mais que ... eu tenha experiência nunca é ... o pensamento nunca para ... tem que se atualizar (I-IP)

Somente o indexador do IP aponta a necessidade de atualização profissional, visto a complexidade

inerente à área científica especializada, considerando-se que o profissional em geral não é um especialista na área em que atua. De acordo com Guinchat e Menou (1994, p. 497), “a formação contínua é indispensável em uma profissão em que as técnicas e os conhecimentos evoluem rapidamente e na qual a demanda e as novas necessidades estão em crescimento contínuo”.

#### 4.3. Aspectos relacionados ao contexto

##### 4.3.1. Eixo temático: Política de tratamento temático da informação documental

###### Nível de especificidade/exaustividade

Então ele tá sempre falando dessa ... as <Propriedades> desse <Material> ...~~ Aí ele fala muito olha, ele fala de alguns termos ... que ele não vai ter, eu não tenho esse tal de <Calorimetria> também aqui, então tem termos que eu mais ou menos que eu já sei que não vou achar ... é muito específico ... esse banco, ele não ... ah, métodos ... ele não tem termos tão específicos (C-FO)

O catalogador da FO salienta que os termos da tese são muito específicos, e pela sua experiência profissional não poderão ser inseridos no campo 650 como assuntos, pois o vocabulário controlado não contempla tão nível de especificidade.

Durante a indexação, o indexador da FSP explicita que a atribuição dos descritores na LILACS é pautada nos termos atribuídos durante a catalogação de assunto da mesma tese no DECS. Entretanto, é pertinente destacar que, de modo análogo, a catalogação de assunto em catálogos coletivos online deveria alcançar o mesmo nível de especificidade da indexação em bases de dados, ou seja, que os assuntos realmente representem o conteúdo dos documentos do tipo tese.

Para tanto, a linguagem documental tem um relevante papel, pois de nada vale o profissional realizar uma concepção de análise pautada no conteúdo documental se a linguagem documental utilizada (o vocabulário controlado) não acompanhar tal nível de especificidade. Desse modo, a constante atualização da linguagem com a participação de bibliotecários, dirigentes, especialistas e usuários permite maior qualidade na etapa de representação/tradução de conceitos durante o processo de catalogação de assunto tanto de teses como de livros no contexto das bibliotecas universitárias – elemento importante da política de indexação. A política de indexação refere-se à forma de realizar a indexação de uma determinada instituição para concretizar, sistematizar e reunir em manuais os processos da indexação, isto é, um guia para tomada de decisões (Gil Leiva, 2008; Carneiro, 1985). Esta reunião dos

procedimentos em manuais está diretamente relacionada com o número de profissionais que se dedicam a esta atividade de indexação e ao nível de qualidade que o sistema de informação pretende alcançar.

Cada biblioteca universitária dispõe de elementos de política de indexação onde a interação destes visa, por meio de seus produtos e serviços, à efetiva recuperação da informação por seus usuários, pois cada instituição se beneficia de uma estrutura específica e de uma forma singular de atuar, considerando-se suas próprias especificidades e características. Nesse sentido, pressupõe-se que a atuação profissional do bibliotecário do tratamento temático da informação ocorra de maneira diferente em cada sistema de recuperação da informação, considerando-se que a atividade é influenciada direta ou indiretamente pela política de indexação adotada. Assim, torna-se necessário que a política de indexação em bibliotecas universitárias esteja descrita de maneira clara, objetiva e ao alcance de todos os atores que contribuem para o desenvolvimento desse processo dentro do sistema de informação – inclusive os próprios usuários.

###### Quantidade de termos

aí o que que eu faço? Nossa, só <Materiais dentários>? Aí eu vou colocar outro assunto ... vamos ver (C-FO)

Outra questão relevante, referente à quantidade de termos, é explicitada também pelo catalogador da FO, que explica que a quantidade limite de termos dos assuntos no registro do ALEPH é de no máximo 5 termos, o que denota este aspecto da política de indexação do sistema de recuperação da informação. Por outro lado, para o indexador da mesma instituição, quanto mais termos forem indexados, melhor para o pesquisador, proporcionando maior possibilidade de recuperação do documento. No entanto, em outro trecho, o profissional sinaliza que a política de indexação aponta quatro termos, para a representação dos assuntos contidos no material do tipo tese. Caso não seja possível este número de termos, três assuntos são aceitáveis, independentemente de se tratar de descritores primários ou descritores secundários.

Outro ponto a ser destacado é que o indexador da FSP explicita a quantidade limite de qualificadores iguais, sendo até três, durante a atribuição dos descritores na indexação.

O processo de indexação de materiais na LILACS pauta-se no *Manual de Indexação de Documentos*, que descreve os procedimentos que deverão ser seguidos para a análise do conteúdo de documentos a serem incorporados na mesma.



Observa-se a inexistência de um manual na instituição que norteie a catalogação de assunto no DEDALUS, explicando, assim, a ausência desta temática nos protocolos verbais individuais dos catalogadores.

Desse modo, apreende-se a necessidade de uma política de indexação formalizada em um manual de indexação, que deve ser um instrumento real de trabalho e norteador dos princípios do processo de tratamento documental pela biblioteca universitária, a fim de que seja garantida a consistência tanto na catalogação de assunto como na indexação.

#### 4.3.2. Eixo temático: O usuário do sistema de recuperação da informação

##### *Recuperação por assunto em catálogo coletivo online e em bases de dados na biblioteca*

então é sempre assim, alguém vai procurar direto essa ... <Compósitos dentários>, vai achar essa tese, mas ele também pode procurar, ele quer saber quais são os <Materiais dentários pra restauração>, aí nesse, ele pode achar esse <Compósitos> (C-FO)

O catalogador da FO demonstra preocupação com a recuperação da tese pelo usuário, não apenas pelo assunto como por outros campos inseridos no registro do MARC.

##### *Preocupação com a perspectiva do usuário*

então eu sempre vejo outra perspectiva que o ... o usuário também pode estar procurando, né? (C-FO)

dar o link pro ... Currículo Lattes do autor é super importante, principalmente pra saber quais outros pesquisadores da minha própria Escola estão escrevendo e publicando sobre aquele assunto ... então ... é ... pra auxiliar até na rede de relacionamento que ele vai ter ... seja como psicólogo lá no futuro ou seja como pesquisador, mesmo, nato, né? (I-IP)

O mesmo profissional tenta prever a forma com que o usuário realizará a busca no sistema de recuperação da informação, proporcionando outros termos que o mesmo poderia utilizar para recuperar a tese. O catalogador reforça a categoria de recuperação por assunto, salientando que no campo 650 procura colocar os termos do ponto de vista de sua experiência profissional na área científica especializada, no caso, a Odontologia. Prioriza os termos em um nível de especificidade maior, ou seja, mais geral e menos específico, colocado pelo autor. O indexador do IP considera relevante a descrição do *Currículo Lattes* dos autores, auxiliando na recuperação dos mesmos.

Os usuários são um componente ativo e determinante no contexto de bibliotecas universitárias, pois são quem utilizam a informação. Enquanto

os usuários reais referem-se a pessoas/instituições que fazem o uso efetivo dos recursos e serviços de uma unidade de informação, os usuários potenciais são vistos como aqueles que reúnem as características pelas quais foi elaborado o sistema de recuperação da informação, mas que por diferentes razões, não utilizam tais recursos. Assim, ao analisar um documento, o catalogador de assunto/ indexador não deve se limitar a representar ou resumir apenas a informação explícita nele contida, devendo ir além, indagando-se sobre as possibilidades de tornar o conteúdo do documento, ou parte dele, visível tanto para o usuário real como para o potencial.

De forma mais sistemática, verificou-se que algumas categorias propostas inicialmente para a análise dos protocolos não apareceram nas declarações dos sujeitos, o que denota que os protocolos verbais coletam uma parcela do processo em si, e não necessariamente todos os elementos que compõem a complexidade do processo de análise documental, materializado no processo de catalogação de assunto e de indexação no contexto das bibliotecas analisadas.

## 5. Considerações finais

A aplicação do PVI como metodologia qualitativa permitiu a observação comparativa dos processos de catalogação de assunto e de indexação realizados pelos profissionais em bibliotecas universitárias e do contexto de trabalho do catalogador e do indexador, levantando alguns pontos como as facilidades e dificuldades da realização dos processos em si.

Quanto ao *catalogador de assunto/indexador*, a questão da subjetividade, presente nas discussões da área e que acompanha o bibliotecário ao longo de sua trajetória profissional na determinação do(s) assunto(s), pode ser minimizada tanto pela realização de todas as etapas dos processos de catalogação de assunto/indexação como pela adoção de uma política de indexação bem definida em cada sistema de recuperação da informação, norteador sua atuação profissional.

Quanto aos *processos de catalogação de assunto/indexação*, verificou-se que o profissional realiza a catalogação de assunto e a indexação de forma distinta, visto que nas bases de dados o processo de indexação é realizado com mais especificidade do que na catalogação de assunto em catálogo coletivo online. Acredita-se que esta diferença esteja relacionada não apenas com o contexto de trabalho do bibliotecário, de um lado, em sua atuação como catalogador de assunto em catálogo coletivo online e, de outro, em sua atuação como indexador em base de dados de área

científica especializada, mas também com os diferentes vocabulários controlados adotados. A atuação profissional é adaptada de acordo com cada tipo de sistema recuperação da informação, onde o principal diferencial reside no nível de especificidade alcançado pelo vocabulário controlado.

Enquanto no catálogo coletivo online utiliza-se um vocabulário mais geral que abarca diversos campos do conhecimento da Universidade, nas bases de dados o vocabulário é específico da área de Saúde, o que possibilita ao bibliotecário maior amplitude e possibilidade de escolha de determinados descritores para indexação de teses, embora nos dois sistemas de recuperação da informação exista a abertura para incorporação de novos termos aos vocabulários, de acordo com a demanda. Desse modo, cada vocabulário controlado - definido pela política de indexação de cada sistema de recuperação da informação - reflete as características e a filosofia de cada contexto. O vocabulário controlado adotado pelo sistema de recuperação da informação é um instrumento da política de indexação que visa à recuperação da informação documental. Como elemento mediador entre a linguagem do conteúdo do documento e a linguagem do usuário do sistema de recuperação da informação, o vocabulário controlado tem papel fundamental tanto em catálogos coletivos online como em bases de dados. Assim, ressalta-se a necessidade de elaboração/atualização de um vocabulário controlado próprio para cada contexto, conferindo maior credibilidade da representatividade e a visibilidade em bibliotecas universitárias.

Em relação ao *contexto* do processo de catalogação de assunto/indexação, recomenda-se a necessidade de uma política de indexação formalizada em um manual de indexação que norteie a catalogação de assunto no catálogo coletivo online e a indexação em bases de dados científicas especializadas, cujos registros são inseridos pelos profissionais que neles atuam. O estabelecimento de uma política de indexação do tratamento documental pressupõe maior qualidade da informação recuperada, pois as diretrizes e procedimentos de catalogação de assunto/indexação bem estabelecidos e seguidos pelos profissionais podem contribuir para uma maior consistência nos resultados da indexação.

## Notas

- (1) O DEDALUS é o catálogo coletivo online do Sistema de Bibliotecas da Universidade de São Paulo (SibiUSP), tendo o VocaUSP (Vocabulário Controlado da USP) como linguagem de representação temática.

- (2) A base de dados LILACS (Literatura Latino-americana e do Caribe em Ciências da Saúde), criada em 1985 constitui atualmente o principal índice e repositório da produção científica e técnica em saúde nos países da América Latina e Caribe. Coordenada em âmbito regional pela BI-REME/OPAS/OMS, é resultado de um esforço cooperativo de centenas de Centros Cooperantes de 37 países da América Latina e Caribe. Complementa índices internacionais, como MEDLINE e Web of Science e está disponível em três idiomas: português, espanhol e inglês.
- (3) O CRUESP/BIBLIOTECAS iniciou suas atividades em 1999, como Grupo de Estudos instituído pela Resolução do Conselho de Reitores das Universidades Estaduais Paulistas (CRUESP) 149/99, tendo por objetivo a integração dos Sistemas de Bibliotecas da Universidade de São Paulo (USP), Universidade Estadual Paulista (UNESP) e Universidade Estadual de Campinas (UNICAMP) do Brasil.

## Annexes in a separate file / Apêndices en un fichero adjunto

### Referências

- Carneiro, M. V. (1985). Diretrizes para uma política de indexação. // Revista da Escola de Biblioteconomia da UFMG. 14: 2 (set. 1985) 221-241.
- Cavalcanti, M. C. (1989). I-n-t-e-r-a-ç-ã-o leitor-texto: aspectos de interpretação pragmática. // Campinas: UNICAMP, 1989.
- Ericsson, K. A.; Simon, H. A. (1993). Protocol analysis: verbal reports as data. // Cambridge: MIT Press, 1993.
- Fiuza, M. M. (1985). O ensino da catalogação de assunto. // Revista da Escola de Biblioteconomia da UFMG. 14:2 (set. 1985) 257-269.
- Fujita, M. S. L.; Rubi, M. P.; Boccatto, V. R. C. (2009). As diferentes perspectivas teóricas e metodológicas sobre indexação e catalogação de assuntos. // Fujita, M. S. L. (Org.). A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias. São Paulo: Cultura Acadêmica, 2009. 19-42.
- Gil Leiva, I. (2008). Manual de indización: teoría y práctica. // Gijón: Trea, 2008.
- Guinchat, C; Menou, M. (1994). Introdução geral às ciências e técnicas da informação e documentação. Brasília: IBICT, 1994. p. 133-165.
- Nardi, M. I. A. (1999). A metáfora e a prática de leitura como evento social: instrumentos do pensar a Biblioteconomia do futuro. 272 f. Tese (Doutorado em Linguística Aplicada e Estudos da Linguagem) – Pontifícia Universidade Católica, São Paulo.
- Rasmussen, E. (2011). Library systems. // Baeza-Yates, R.; Ribeiro-Neto, B. Modern information retrieval: the concepts and technology behind search. 2. ed. Harlow: Addison Wesley, 2011.
- Shoham. S.; Kedar, R. (2001). The subject cataloging of monographs with the use of keywords. // Cataloging & Classification Quarterly. 33:2 (2001) 29-54.
- Tartarotti, R. C. D. E. (2014). Atuação bibliotecária no tratamento temático da informação em unidades informacionais: um estudo comparativo qualitativo-quantitativo. // 277 f. Dissertação (Mestrado em Ciência, Tecnologia e Sociedade) – Universidade Federal de São Carlos.

Enviado: 2016-047-02. Segunda versão: 2016-07-11.  
Aceptado: 2017-09-13.

---

# Pedro Garhel y su implicación en “Espacio P”: propuesta de un cuadro de clasificación para el archivo personal de un artista polivalente

*Pedro Garhel and its involvement in "P Space": proposing a classification scheme for a versatile artist*

---

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

Instituto Agustín Millares, Universidad Carlos III de Madrid, Calle Madrid 128 (28903),  
castromo@hum.uc3m.es, amorales@bib.uc3m.es, fatimag@bib.uc3m.es, k.ohlenschlaeger@gmail.com

## Resumen

Se propone un cuadro de clasificación del archivo personal del artista Pedro Gharel, configurado como una sección del fondo del Archivo Espacio P, que permita organizar y recuperar la documentación generada por el artista en el contexto de un archivo no institucional con voluntad de actuar como centro de documentación de prácticas artísticas contemporáneas internacionales, pionero en España. Al tratarse de un fondo complejo, constituido por una documentación heterogénea y diversa, se ha estructurado mediante un cuadro de clasificación de amplia temática que contempla todas las actividades polifacéticas e interconectadas a las que se dedicaba el artista entre las artes plásticas, visuales, escénicas y sonoras.

**Palabras clave:** Archivos personales. Espacio P. Pedro Garhel. Cuadros de clasificación de archivo. Arte contemporáneo. Performance. Video. Multimedia. Espacio artístico independiente. Autogestión. España.

## 1. Introducción

Los archivos personales han pasado de ser los grandes olvidados a suscitar, actualmente, gran interés por parte de los investigadores, por lo que no es de extrañar que las distintas instituciones culturales deseen albergar este tipo de fondos entre sus colecciones (Rucio Zamorano, 2011).

La bibliografía sobre el tema, relativamente reciente, subraya los problemas que plantea el tratamiento técnico de estas colecciones y la necesidad de compartir experiencias concretas entre los profesionales de la información que se dedican a su organización (Seminario de Archivos Personales, 2006; Asociación de Archiveros de Navarra, 2012).

El Archivo de Espacio P y el Archivo Personal de Pedro Garhel, depositados en el Archivo Histórico Provincial de Santa Cruz de Tenerife, constituyen dos subfondos de un mismo legado, superpuestos y entrelazados.

Espacio P(erformance), también llamado Espacio P(olivalente) (1981-1997), fue una parte

## Abstract

A classification scheme is proposed for the Personal Archive of the artist Pedro Gharel, an archival section of the Archivo Espacio P. It will enable the organization and retrieval of the documentation preserved in a non-institutional archive that was planned as a Documentation Center for International Contemporary Art Practices, pioneer in Spain. Since it is a complex archive holding, composed by very diverse materials, it has been organized following a wide-scope classification scheme. In this way, this device embraces all the versatile and interdisciplinary activities of an artist outstanding by his combinations of a wide range plastic, visual, scenic and sound arts.

**Keywords:** Personal archives. Espacio P. Pedro Garhel. Archival classification schemes. Contemporary art. Performance. Video. Multimedia. Independent art space. Self-management. Spain.

esencial del proyecto artístico de Pedro Garhel (Puerto de la Cruz, 1952 - La Guancha, 2005), un centro independiente, autogestionado y autofinanciado, pionero en el ámbito de la performance, el videoarte y las nuevas prácticas multimedia en España, capaz de generar recursos propios y crear redes de proyección nacional e internacional. Aunque fue un proyecto colectivo y una asociación, muchas iniciativas se sufragaron con el esfuerzo económico personal del propio artista y otros integrantes del centro.

La presente propuesta se enmarca dentro del proyecto de investigación *Archivo de Espacio P: Propuesta Metodológica para su Continuidad Digital* (2014-2017) (MINECO, Ref.HAR2013-44726-R). En la primera fase del proyecto de investigación se elaboró un modelo o cuadro de clasificación del fondo generado por Espacio P, que ha permitido organizar y recuperar la documentación que conserva este archivo.

En el presente trabajo se propone un cuadro de clasificación para la sección: Fondo Personal Pedro Garhel, que corresponde a los documentos

personales generados y/o reunidos por el artista a lo largo de su trayectoria; y da continuidad al estudio *Archivo de Espacio P: propuesta de un cuadro de clasificación para su continuidad digital*, publicado en esta misma revista (Castro Morales, García López, Morales García y Ohlenschläger, 2015).

Al tratarse de un archivo personal, presenta una documentación compleja y diversa que ha precisado de un estudio exhaustivo e individualizado del fondo. Para la organización de este fondo ha sido necesario contemplar un cuadro de clasificación amplio para categorizar una documentación tan heterogénea. Para ello, se ha tenido en cuenta el cuadro clasificación propuesto por Galiana Chacón de la Biblioteca de Valencia (Galiana Chacón 2006; p. 25-28), así como las propuestas de la Base de Datos de Archivos Personales de la Biblioteca Nacional (Rucio Zamorano, 2011). El principal reto que ha planteado este cuadro de clasificación ha sido articular una estructura de ordenación que permitiera representar las actividades tan polifacéticas e interconectadas a las que se dedicaba el artista entre las artes plásticas, visuales, escénicas y sonoras.

En este sentido, destaca la amplia serie de correspondencia personal, cuyo estudio permite trazar una cartografía emocional y de sus relaciones profesionales y personales. De gran valor son los apuntes (anotaciones sobre proyectos y obras) y el material docente, pues reflejan las relaciones entre la docencia no formal y las incipientes enseñanzas regladas de arte y nuevos medios en facultades de Bellas Artes. Y, finalmente, refleja la diversidad de intereses y prácticas artísticas: *mail art*, *copy art*, *fax art*, arte electrónico, *performance*, ente otras.

## 2. Propuesta de cuadro de clasificación

El cuadro de clasificación para el Archivo Personal de Pedro Garhel constituye un reto como consecuencia del carácter polifacético del artista, así como por los diversos ámbitos de actividad en los que desarrolló su trayectoria. Por ello, ha sido necesario estructurar el cuadro de clasificación en ocho grandes secciones:

- Sección 1. Bibliografía relacionada con el autor y/o su obra
- Sección 2. Correspondencia
- Sección 3. Documentación gráfica
- Sección 4. Documentación legal, judicial, fiscal y mercantil
- Sección 5. Documentación personal y familiar
- Sección 6. Documentación profesional
- Sección 7. Objetos personales

- Sección 8. Obra propia

Señalamos sintéticamente su contenido.

### 2.1. Sección 1: Bibliografía relacionada con el autor y/o su obra

En esta primera sección abundan las referencias de prensa impresa, reseñas y críticas recibidas por sus primeras exposiciones en el ámbito de la expresión plástica; también los ensayos y tesis doctorales sobre su trayectoria y sus más complejas propuestas vinculadas a las artes de la acción. A ellas se suman monografías nacionales e internacionales, folletos y catálogos, que permiten acercarnos a sus diferentes identidades artísticas (Pedro Garhel, Schinca, Depósito Dental, Corps), así como a muestras colectivas y programas de eventos nacionales e internacionales, especialmente de vídeo y performance, en los que participó. Debido a que la producción bibliográfica no ha cesado, a pesar de su prematuro fallecimiento en 2005, se ha añadido una nueva subsección "Guías y listados bibliográficos", donde se actualiza la recepción de su legado con el propósito de facilitar a los investigadores la aproximación al análisis sobre la fortuna crítica e interpretación historiográfica del artista y su obra (Castro Morales, 2011). Las cada vez más frecuentes referencias al artista en soporte electrónico (webs y blogs profesionales, así como revistas *on line* y trabajos académicos en repositorios institucionales) se recogen además en la web del proyecto de investigación en curso (ArchivoEspacio P, 2014)

### 2.2. Sección 2: Correspondencia

Contempla dos grandes series: correspondencia personal (incluye la familiar y la sostenida con los coleccionistas Otto y Barbara Dobermann) y correspondencia profesional (artística, docente universitaria y relacionada con la gestión de Espacio P), en cada caso, ordenadas cronológicamente.

La correspondencia profesional constituye un epistolario compuesto por 3.608 cartas (las primeras registradas en 1983), reflejo de los vehículos sobre los que se sustenta la proyección internacional experimentada por el artista y sus proyectos: los viajes para participar en exposiciones y festivales que van articulando paso a paso las relaciones con instituciones y distribuidoras (Vídeo 235 Colonia, Time Based Arts Amsterdam, Infermental Colonia, entre otros) de vídeos y performances. Permite documentar la actividad tanto individual de Garhel como la colectiva de Espacio P, aunque presenta algunas lagunas, explicable por el hecho de que los artistas del ámbito cercano no solían recurrir al intercambio epistolar para organizar su participación.

En España el modo de gestionar esta distribución, incluso en el ámbito institucional era muy poco profesional, debido a la práctica ausencia de instituciones culturales interesadas en estos ámbitos de creación. Piénsese que el Ministerio de Cultura se había fundado apenas tres años antes, en 1978; la feria de arte contemporáneo, ARCO, iniciaba su andadura en 1982; y el Centro Nacional de Exposiciones del Ministerio de Cultura comienza su programa en 1983. Entre las galerías españolas casi es una excepción Helga de Alvear y su galería de Madrid, que acogió la muestra de Darío Corbeira, producida y coordinada por Espacio P en 1988.

En otros países circunstancias análogas habían llevado al surgimiento de numerosas distribuidoras —gran parte de las cuales habían surgido en la década de los 70 y principios de los 80— por iniciativa de los propios artistas para articular nuevos canales de difusión y exposición de sus obras audiovisuales. De ahí, la importante decisión de crear en 1985 su propia distribuidora de vídeo (Alliance Video Art y AVA); aunque, a diferencia de las anteriores, fue concebida desde un principio como un proyecto de colaboración entre el Espacio P y la galería Grita Insam de Viena, con el doble propósito de facilitar la proyección internacional de videocreadores españoles y la presentación en España de importantes figuras internacionales. La correspondencia entre AVA y estas distribuidoras nos permite observar el modo en el que se promocionó a los artistas españoles en el extranjero y los cauces para difundir a los artistas internacionales que AVA representaba en España. De ahí, la abundante correspondencia con las distribuidoras europeas y americanas de videocreación. Pero también refleja el modo en el que Pedro Garhel y otros integrantes de Espacio P contactaban directamente con creadores extranjeros para invitarles a presentar sus proyectos en Espacio P. De este *modus operandi* surgen, por ejemplo, los encuentros con Michelangelo Pistoletto, Meredith Monk o Phillip Corner.

Asimismo, especial interés encierra la información que ofrecen sobre el rico entramado de relaciones en el ámbito de la performance, el vídeo y otras prácticas artísticas muy marginalizadas en la década de los 80 en el Estado español. Las cartas desvelan además los procesos de internacionalización de las mismas, articulados a partir de la conexión de Espacio P con las redes independientes extranjeras y también con algunas de las instituciones más comprometidas con las prácticas artísticas de aquel momento dentro del país. El epistolario (cartas, tarjetas, faxes y telegramas) nos permite conocer como gestionaban sus proyectos y producían sus obras Pedro Garhel y los

artistas que participaban en Espacio P. También permite cifrar el efecto de la participación en eventos de proyección nacional e internacional, como Ars Electronica, la documenta de Kassel (Ohlenschläger, 2007), o el programa Amsterdam Capital Cultural Europea, todo ello en 1987, que propició el encuentro interprofesional, el acceso a publicaciones que por entonces tenían escasa circulación en España y ampliar la red de contactos del artista. Además, dan testimonio de un conjunto de proyectos formativos no reglados de producción de obras multimedia y performativas en las que se exploraba formal y conceptualmente los nuevos medios de creación colectiva con participación ciudadana y de artistas de los más diversos campos de la creación visual, escénica y sonora.

Otro aspecto reseñable sobre la correspondencia es la terminología utilizada, detectándose una interesante hibridación de léxicos provenientes de profesiones hasta entonces escasamente vinculadas a las artes plásticas. De ahí que de su lectura se extraigan abundantes ideas, conceptos y términos que resultan de gran utilidad para construir el Tesoro de Prácticas Artísticas Contemporáneas (TEPAC) en el que nos encontramos empeñados.

### 2.3. Sección 3: Documentación gráfica

La tercera sección recoge la documentación gráfica. Se trata de una selección amplia y rica en soportes y tipologías que ilustran el tránsito de la era analógica a la digital. En ocasiones se trata de documentos gráficos y audiovisuales, de sus propuestas prácticas performativas, multimedia y audiovisuales que se producían en Espacio P. El afán por documentar procesos y presentaciones revela la fuerte conciencia que tenía Pedro Garhel acerca del carácter innovador de su trabajo. Con los equipos audiovisuales de vídeo del Instituto Alemán ya en el año 1979, dos años antes de la inauguración del Espacio P, el artista había realizado su primera performance para vídeo, grabada en la sede del Goethe Institute de Madrid. Posteriormente, documentaría los procesos de la actividad en Espacio P.

Una primera subsección del material gráfico está compuesta por hojas de contacto y fotografías (personales, familiares y de las obras); la segunda por diapositivas y negativos (personales y de la obra); la tercera recoge grabaciones sonoras y finalmente la cuarta recoge los soportes audiovisuales (Beta, Film Super 8, 1.3.4.3. Hi8, MINI DV, U-MATIC, VHS). Las cintas en Super 8 han sido telecinadas y editadas para su difusión a través del canal de Vimeo de Archivo de Espacio P (Espacio P, 2015).

Sin embargo, queda pendiente la digitalización de las cintas en soporte U-MATIC, grabaciones sonoras y las obras pioneras del arte electrónico en España, realizadas por Pedro Garhel. Este cuadro de clasificación y los inventarios asociados al mismo serán de gran utilidad para cumplir con la directriz europea *Conservar la memoria del mañana – Conservar los contenidos digitales para las generaciones futuras* (2002), a la espera de la implantación de una política de conservación del patrimonio digital acorde con las directrices de UNESCO (2012).

#### 2.4. Sección 4: Documentación legal, judicial, fiscal y mercantil

Destacamos en esta sección la serie de ámbito laboral, que incluye altas y bajas en la Seguridad Social, cotizaciones y el expediente de reconocimiento de invalidez, así como la documentación fiscal, que incluye declaraciones de la renta, trámites relativos a la licencia fiscal, así como IVA, impuestos y tributos, porque ambas permiten acotar el ámbito de precariedad en el que el artista desarrolló su proyecto artístico y vital. Las dificultades económicas o la incompreensión ante sus propuestas creativas, queda reflejada en la documentación judicial, así como en reclamaciones, denuncias, procesos y sentencias. Para comprender los mecanismos de producción de las obras propias y las producciones que sufraga a artistas y colectivos vinculados a Espacio P resulta fundamental la documentación mercantil, compuesta por contratos, facturas, albaranes, recibos, entre otros.

Los avatares del local, primero alquilado, luego adquirido y finalmente vendido, forma parte de la documentación patrimonial (contratos de alquiler, tasaciones, escrituras de compra-venta, comunidad de propietarios, obras y reformas), a ella hemos asociado también los inventarios de bienes y la documentación sobre el traslado de los mismos a Tenerife. Asimismo, forma parte de esta sección la serie sobre propiedad intelectual relativa a su obra.

Especialmente relevante para estudiar los mecanismos de la autogestión artística es la documentación bancaria (préstamos, cuentas, extractos, movimientos) que permite cruzar datos con la documentación integrante de las dos series anteriores.

#### 2.5. Sección 5: Documentación personal y familiar

Destaca junto a la documentación civil y de identidad, diversos carnets, documentación académica que acredita la formación recibida, al igual que los diplomas y premios recibidos, así como

la participación en diversos concursos y convocatorias. Debido al delicado estado de salud del artista, la documentación médica presenta abundantes informes y pruebas diagnósticas. También se recoge aquí la documentación necrológica.

#### 2.6. Sección 6: Documentación profesional

Comprende la documentación más interesante, especialmente la Serie dedicada a la “Docencia”, tarea a la que se dedicó Garhel en diversos ámbitos de la enseñanza no formal, adquiriendo una experiencia que luego quiso transmitir a las aulas universitarias, como profesor de la asignatura de Arte y Nuevos Medios en la Universidad de Salamanca. Se conserva documentación laboral (contratos, nombramientos, seguros, licencias, bajas, nóminas, entre otras), documentación administrativa relativa a la organización docente (horarios, planes de estudio, fichas de estudiantes, entre otras) y material docente en papel (apuntes, anotaciones, fotocopias) y visual o audiovisual, utilizado para la impartición de sus clases. Especial interés tienen los trabajos y dossiers de estudiantes, así como la documentación generada durante la preparación y presentación de las actividades públicas realizadas con los estudiantes, planteadas como intersecciones entre la ciudad, la institución universitaria y otros profesionales invitados para favorecer la interacción con los estudiantes.

Convencido de los beneficios que tendría para la formación de los estudiantes en el ámbito de la gestión de programas públicos, entre 1994 y 1996, organizó junto con sus estudiantes los programas colectivos de arte público “Abierto” (1994 y 1996) y “+ o - Yak el Performador” y “Vuelta y Vuelta” (1996). Destacamos la presencia de Darío Corbeira, Nacho Criado y Marisa González. Primero, en 1992, el artista acogió a los estudiantes en “Transformación interactiva por Fax”, un taller que tuvo lugar en el Círculo de BBAA de Madrid. Posteriormente, en 1995 impartiría un segundo taller, “Estación Fax”, en el contexto de *Bibliograph '95*, contando con la intervención de los alumnos Iñigo Zubikoa, Gorkoa y Elisa Espina. También coordinó la participación de sus estudiantes en “El País de los Tuertos” (1994, 1995, 1996), un certamen nacional que convocaba a los estudiantes de todas las facultades de Bellas Artes del estado español en torno a la videocreación.

En esta sección también se ha incluido una serie dedicada a la documentación relativa a “Colaboraciones” en el ámbito de la enseñanza superior, en otras universidades y en el Instituto de Estética, invitado por José Jiménez.

Asimismo, se ha identificado una serie relativa a la "Formación de formadores", un ámbito de influencia necesario para diseminar la sensibilidad hacia las prácticas artísticas contemporáneas en los niveles primario y secundario. Con igual motivación, aunque abierto a cualquier público, destacan los cursos, seminarios, conferencias y talleres impartidos.

El reconocimiento a su labor docente (formal y no formal) y a su trayectoria artística individual o colectiva queda plasmada en la serie dedicada a la "Participación en Comités Científicos y Organizadores y Jurados".

También se ha agrupado en una serie la documentación relativa a "Proyectos de comisariado y de exposiciones", así como la de "Dirección de videoforum o ciclos de videocreación".

La conciencia profesional y su convicción acerca de la necesidad de estructurar el sector de arte contemporáneo, le condujo a formar parte de asociaciones y colegios profesiones. La "Documentación sobre estas organizaciones" (AMAVI, SGAE, AMAP, entre otras), integra otra serie, al igual que la "Sanación espiritual", ámbito al que se dedicó profesionalmente los últimos años de su vida, llegando a formar parte de la directiva de la Asociación Española de Sanación espiritual.

Significativo también para conocer las políticas de apoyo a la creación es la serie formada por "Becas, ayudas y subvenciones", solicitadas y/o concedidas para producir y exhibir las propuestas de Pedro Garhel, reflejo a su vez de la escasa consideración que tenían en la estructura ministerial las prácticas interdisciplinarias (Albarrán, 2012).

Esta sección de "Documentación profesional" se cierra con la serie que recoge los *curricula* profesionales de Pedro Garhel, de los que se conservan numerosas versiones y copias, necesarias en muchas ocasiones para completar la documentación de solicitud de apoyos a la creación o para ofrecer servicios profesionales.

## 2.7. Sección 7: Objetos personales

Entre los objetos personales del artista se encuentran los trajes, con zapatos a juego y todo tipo de ropa que solía utilizar en sus actuaciones públicas. Alguna de esta indumentaria ha sido especialmente confeccionada para la ocasión, incluso por diseñadores de renombre de la época, en los años ochenta.

También se guardan los enseres personales, los distintos elementos utilizados en sus performances y otros objetos que empleó en sus clases, en su taller personal y en su mesa de trabajo.

## 2.8. Sección 8: Obra propia

Esta sección está dedicada a la documentación sobre la obra propia de Pedro Garhel en los ámbitos de la performances, videos, instalaciones y proyectos multimedia; pero no solo aquella que el artista expuso nacional e internacionalmente, sino también los proyectos que no llegaron a realizarse.

Cada uno de los ochenta y seis proyectos llevados a cabo dispone de una carpeta propia en la que se guardan los apuntes previos de ideas y primeros bocetos, la descripción conceptual y tecnológica de las obras, las sinopsis finales, los presupuestos, folletos y tarjetas de invitaciones de los lugares donde la obra fue expuesta y otra documentación complementaria sobre papel que el artista utilizó para el desarrollo del proyecto en cuestión. Esta documentación se complementa con el listado de material proyectable de la obra, que se encuentra depositado en la sección 1.3, así como una relación bibliográfica y hemerográfica en la que se menciona el proyecto del artista.

Entre los cuarenta y un proyectos no realizados figuran no solo videos, instalaciones y obras multimedia, sino también propuestas para la realización de exposiciones o la organización de festivales y otros eventos. Cada uno de estos proyectos están igualmente documentados con sus respectivos apuntes, bocetos, presupuestos, descripciones y materiales complementarios sobre papel que el artista utilizó al concebir su propuesta.

Otra serie contiene los dossieres y *curricula* de los grupos Schinca, Corps, Depósito Dental o Galindo=Garhel, de los que Pedro Garhel había formado parte a lo largo de su vida, mayoritariamente en la dirección de los mismos.

Por último, la serie final contiene otros textos escritos por el artista, fragmentos de textos desconexos, así como hojas con apuntes, tareas y quehaceres de toda índole.

En cambio, la obra pictórica, dibujos y diseños sobre papel de Pedro Garhel, están en depósito, por tiempo indefinido, en el Museo de Arte Contemporáneo TEA, Tenerife Espacio de las Artes, con sede en Santa Cruz, integrados en su colección permanente.

Todas estas obras anticipaban distintas maneras y nuevos modos de investigación, producción, autogestión y difusión, como consecuencia de la desmaterialización y la performatividad. Se trataba de proyectos que requerían nuevos canales de producción y distribución audiovisual y la participación activa del artista en el proceso (Aramburu, 2010).

Autogestión y performance centrarían precisamente la intervención de Pedro Garhel en los *Encuentros de Arte Actual*, de la *Red Arte y Colectivos Independientes en el Estado Español*, celebrados en Vitoria-Gasteiz en 1997, justo cuando Espacio P cerraba sus puertas y Pedro Garhel se planteaba retornar a su isla natal (Garhel, 1997).

### 3. Consideraciones finales

El cuadro de clasificación del archivo personal del artista Pedro Garhel, como parte del fondo del Archivo Espacio P, ha permitido organizar y recuperar la documentación generada por el artista en el marco de un archivo no institucional de gran valor para el conocimiento de las prácticas artísticas contemporáneas lo que le categoriza como una fuente de información de primordial importancia para abordar investigaciones sobre esta temática. El principal reto planteado ha sido articular una estructura de ordenación que permita representar las actividades tan polifacéticas e interconectadas a las que se dedicaba el artista entre las artes plásticas, visuales, escénicas y sonoras, por lo que ha sido necesario contemplar un cuadro de clasificación de amplia temática para categorizar una documentación tan heterogénea y diversa.

Un aspecto importante que queremos resaltar finalmente es que la organización del archivo ha permitido la identificación y establecimiento de prioridades de intervenciones para la conservación y preservación de un fondo que, pese a su juventud, se muestra muy vulnerable. Esto afecta, especialmente, a los soportes de algunas piezas de fax-art, mail-art y copy-art, muy inestables, como también los soportes de numerosas creaciones audiovisuales de los inicios del arte electrónico y el Media Art, en el que Pedro Garhel fue un pionero en España. En este sentido, queremos resaltar la aportación que supone esta investigación a la organización de los archivos de creadores que se adentran en los territorios de la creación con nuevos medios, así como a los instrumentos habilitados para compartir en la red los fondos documentales de naturaleza audiovisual.

Este cuadro de clasificación tiene además valor como guía topográfica para la ordenación física del fondo e identificación de tipologías documentales, soportes y dispositivos, facilitando la planificación de las medidas de conservación y el ensayo de las metodologías para la continuidad digital. El glosario de términos empleados para identificar el archivo personal y profesional de Pedro Garhel constituye un importante aporte para el inicio del Tesoro de Prácticas Artísticas Contemporáneas (TEPAC) con el que se avanzará en un nivel más profundo de catalogación

del archivo.

### Apéndice electrónico

El cuadro de clasificación está disponible junto con el artículo en el sitio web de la revista: <http://www.iberid.eu/ojs/index.php/scire/editor/submission/4351>

### Referencias

- Albarrán, Juan (ed.) (2012). *Arte y Transición*. Madrid: Bruñaria, 2012.
- Aramburu, Nekane (2010). *Historia y situación actual de los colectivos de artistas y espacios independientes en el Estado español: la otra historia*. Madrid: Bubok, 2010. ISBN 978-8461492572.
- Asociación de Archiveros de Navarra (2012). *Los fondos personales y/o familiares, ¿son fondos de segunda o tercera clase?* // Asociación de Archiveros de Navarra. <http://www.archiverosdenavarra.org/?p=286> (2017-02-11).
- ArchivoEspacio P (2014). *Proyecto de investigación "Archivo de Espacio P: Propuesta Metodológica para su Continuidad Digital"*. // *Archivoespaciop*. <http://www.archivoespaciop.es/> (2017-02-11).
- Castro Morales, Federico, García López, Fátima, Morales García, Ana M<sup>a</sup>, Ohlenschläger, Karin (2015). *Archivo de Espacio P: propuesta de un cuadro de clasificación para su continuidad digital*. // *Scire*. 21:2 (jul.-dic. 2015) 42-52. ISSN 1135-3716.
- Castro Morales, Federico (2011). *Sobre archivos y museos: a propósito de la retrospectiva de Pedro Garhel*. // *Revista de Museología*. 51 (2011) 42-52.
- Galiana Chacón, Juan. P. (2006). *De los archivos personales, sus características y su tratamiento técnico*. // *Seminario de Archivos personales*. Madrid: Biblioteca Nacional, 2006. 12-28. ISBN 84-88699-X.
- Espacio P (2015). <https://vimeo.com/archivoespaciop> (2017-02-11).
- Garhel, Pedro (1997). "P de Polivalente". // Aramburu, N. (ed.). *Encuentros de Arte Actual, Red Arte y Colectivos Independientes en el Estado Español*. Vitoria-Gasteiz: *Trasforma*, 1997. 21-24.
- Martín Abad, Julián; Romero Tobar, Leonardo; Iglesias Nieves (coord.) (2006). // *Seminario de Archivos Personales*. Madrid: Biblioteca Nacional, 2006. ISBN 84-88699-X.
- Ohlenschläger, Karin (2007). *Pedro Garhel: entre la acción analógica y la proyección digital, 20 aniversario del arte multimedia español en la documenta de Kassel*. // *a.mimima, new media art now*. 21 (2007), 160-175.
- Unión Europea (2002). *Resolución 25 de junio de 2002 del Consejo Europeo*. // *Diario Oficial de las Comunidades Europeas C*. 162, 45º año (6 de julio de 2002).
- Rucio Zamorano, María José (2011). *Los archivos personales, fondos visibles en la web de la Biblioteca Nacional de España*. // *Quintas Jornadas de Archivo y Memoria. Extraordinarios y fuera de serie: formación, conservación y gestión de archivos personales*. Madrid, 17-18 febrero. [http://www.docutren.com/ArchivoyMemoria/ArchivoyMemoria2011/pdf/5J\\_Com\\_30\\_Rucio\\_web.pdf](http://www.docutren.com/ArchivoyMemoria/ArchivoyMemoria2011/pdf/5J_Com_30_Rucio_web.pdf) (2017-02-11).
- UNESCO (2017). *El patrimonio digital*. // *Unesco: comunicación e información*. [http://www.unesco.org/new/es/communication-and-information/access-to-knowledge/preservation-of-documentary-heritage/digital-heritage/\(2017-02-11\)](http://www.unesco.org/new/es/communication-and-information/access-to-knowledge/preservation-of-documentary-heritage/digital-heritage/(2017-02-11)).

Enviado: 2016-04-04. Segunda versión: 2017-05-09.  
Aceptado: 2017-05-09.



---

# 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](http://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/multi doc/multidoc/revista/num9/cine/sagredo.htm> (2001-01-22).

### 8. Apéndice I. El resumen

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

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

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

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

#### 9.1. Sistemática

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

#### 9.2. Claridad y concisión

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

#### 9.3. Pertinencia y parsimonia

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

#### 9.4. 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*

---

Alves, Bruno Henrique, 47	Morales García, Ana MARÍA, 67
Blázquez Ochando, Manuel, 39	Ohlenschläger, Karin, 67
Castro Morales, Federico, 67	Oliveira, Ely Francina Tannuri de, 47
Dal'Evedove, Paula Regina, 13	Pavanelli, Maria Aparecida, 47
Fujita, Mariângela Spotti Lopes, 13, 57	Prado, Marcos Aparecido Rodrigues do, 47
García López, Fátima, 67	Salvador Oliván, José Antonio, 25
Marco Cuenca, Gonzalo, 25	Tartarotti, Roberta Cristina Dal'Evedove, 57

---

# Índice de materias en español

*Subject index in Spanish*

---

América Latina, 47	Multimedia, 67
Aproximación sociocultural, 13	Normalización, 25
Archivos personales, 67	Ontologías, 25
Autogestión, 67	Organización del conocimiento, 47
Bases de datos, 57	Pedro Garhel, 67
Bibliotecas universitarias, 13, 57	Performance, 67
Brasil, 13, 57	Políticas de indización, 13
Canales de sindicación, 39	Producción científica, 47
Catalogación de materias, 13, 57	Protocolo verbal, 13
Catálogos colectivos en línea, 57	Protocolo verbal individual, 57
Coautoría institucional, 47	Revistas científicas, 47
Cuadros de clasificación de archivo, 67	RSS, 39
Doctoral thesis, 57	Scire, 47
Espacio P, 67	Sindicación de contenidos, 39
España, 25, 67	Sistema de clasificación para la literatura sobre Organización del Conocimiento, 47
Extracción de datos, 39	Sistema Nacional de Salud (España), 25
HCDSNS, 25	SNOMED CT, 25
Historia Clínica Digital, 25	Terminologías clínicas, 25
Historias clínicas electrónicas, 25	Tratamiento temático de la información, 13
Indización, 57	Video, 67
Interoperabilidad semántica, 25	Web crawler, 39
Minería de datos, 39	
Modelos clínicos, 25	

---

# Índice de materias en inglés

*Subject index in English*

---

Archival classification, 67	Clinical models, 25
Archival classification schemes, 67	Clinical terminologies, 25
Brazil, 13, 57	Contemporary art, 67
Classification System for Knowledge Organization Literature, 47	Content syndication, 39
	Data mining, 39

Databases, 57  
Doctoral thesis, 57  
Electronic Health Records, 25  
Espacio P, 67  
Feeds, 39  
HCDSNS, 25  
Independent art space, 67  
Indexing, 57  
Indexing policies, 13  
Institutional co-authorship, 47  
Internationalization, 47  
Knowledge organization, 47  
Latin America, 47  
Multimedia, 67  
National Health System (Spain), 25  
Online public catalogues, 57  
Ontologies, 25  
Pedro Garhel, 67  
Performance, 67  
Personal archives, 67  
RSS, 39  
*Scire*, 47  
Scraping, 39  
Self-management, 67  
Semantic interoperability, 25  
SNOMED CT, 25  
Sociocultural approach, 13  
Spain, 25, 67  
Standardization, 25  
Subject cataloguing, 13, 57  
University libraries, 13, 57  
Verbal protocol, 13, 57  
Video, 67  
Web crawler, 39