



Universidade Estadual Paulista “Júlio de Mesquita Filho”  
Faculdade de Filosofia e Ciências  
Campus de Marília

Luiza de Menezes Romanetto

## **O CONTROLE DE AUTORIDADE NO CONSÓRCIO VIAF**

Marília  
2017



Universidade Estadual Paulista “Júlio de Mesquita Filho”  
Faculdade de Filosofia e Ciências  
Campus de Marília

Luiza de Menezes Romanetto

## **O CONTROLE DE AUTORIDADE NO CONSÓRCIO VIAF**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Informação.

Linha de Pesquisa: Informação e Tecnologia.

Orientadora: Profa. Dra. Plácida L. V. Amorim da Costa Santos.

Marília

2017

Romanetto, Luiza de Menezes.

R758c O controle de autoridade no consórcio VIAF / Luiza de Menezes Romanetto. – Marília, 2017.  
104 f. ; 30 cm.

Orientador: Plácida Leopoldina V. A. da Costa Santos.  
Dissertação (Mestrado em Ciência da Informação) –  
Universidade Estadual Paulista (Unesp), Faculdade de  
Filosofia e Ciências, 2017.  
Bibliografia: f. 95-104

1. Catálogo de autoridade (Recuperação da  
informação). 2. Metadados. 3. Web semântica. 4. Dados  
ligados. I. Título.

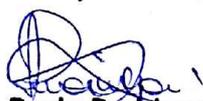
CDD 005.72

**ATESTADO DE APROVAÇÃO - DEFESA**

Atestamos que **LUIZA DE MENEZES ROMANETTO**, RA nº: 186737-1, RG nº 42.915.211-5, expedido pela SSPSP/SP, defendeu, no dia 24/01/2017, a dissertação intitulada **O controle de autoridade no consórcio VIAF**, junto ao Programa de Pós Graduação em CIÊNCIA DA INFORMAÇÃO, Curso de Mestrado Acadêmico, tendo sido 'APROVADA'.

Atestamos ainda que a obtenção do título dependerá de homologação pelo Órgão Colegiado competente.

Marília, 24 de janeiro de 2017

  
**Ana Paula Rodrigues**  
**Supervisor Técnico de Seção**

Luiza de Menezes Romanetto

## **O CONTROLE DE AUTORIDADE NO CONSÓRCIO VIAF**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Informação.

**Linha de Pesquisa:** Informação e Tecnologia.

**Data da defesa:** 24/01/2017

**Banca examinadora:**

Dra. Plácida L. V. Amorim da Costa Santos  
(UNESP/Marília)  
Livre-docente em Catalogação (orientadora)

---

Dra. Zaira Regina Zafalon (UFSCar/São Carlos)  
Doutora em Ciência da Informação

---

Dra. Rachel Cristina Vesu Alves (UNESP/Marília)  
Doutora em Ciência da Informação

---

## **AGRADECIMENTOS**

À minha orientadora Profa. Plácida, pela oportunidade, confiança e por me ter possibilitado cursar o mestrado e desenvolver esta pesquisa. Por todas as orientações e conselhos compartilhados, que contribuíram de forma ímpar em minha aprendizagem.

Às professoras Rachel Cristina Vesu Alves e Zaíra Regina Zafalon, por todas as recomendações apresentadas à pesquisa e, principalmente, por me permitir acompanhá-las nos últimos anos.

Aos professores suplentes, Ana Carolina Simionato e José Eduardo Santarém Segundo, e demais professores do Programa de Pós-Graduação em Ciência da Informação da UNESP, com quem tive a oportunidade de aprender durante as disciplinas cursadas no mestrado, por todo o conhecimento compartilhado, que contribuiu direta e indiretamente para o desenvolvimento da pesquisa.

Aos colegas do Grupo de Pesquisa Novas Tecnologias em Informação (GPNTI) e aos do PPGCI/UNESP, pela troca de conhecimentos e pelo apoio nos momentos de incerteza e dúvida.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo fomento cedido para o desenvolvimento da pesquisa.

Aos meus familiares e amigos, que, apesar da distância, se fizeram presentes, pelo apoio e companheirismo em todos os momentos. Em especial, à minha mãe Lourdes, meu maior exemplo de vida.

Se o ser humano sempre procurou padrões que unem ou diferenciam tudo o que existe, é porque essa tem sido sua ferramenta para lidar com uma inquietação atávica: a necessidade de dar ordem ao caos.

Umberto Eco

ROMANETTO, L.M. **O controle de autoridade no consórcio VIAF**. 2017. 104f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, 2017.

## RESUMO

O controle de autoridade garante consistência aos sistemas de informação. Adotado na catalogação, o tema recebeu maior visibilidade durante a década de 1980 com a automação dos catálogos. Desde então, surgiram projetos direcionados ao intercâmbio e compartilhamento de dados de autoridade. O *Virtual International Authority File* (VIAF) é um consórcio de cooperação internacional entre bibliotecas e agências nacionais que disponibiliza arquivos de autoridade no *Linking Open Data*. Este estudo foi desenvolvido com o intuito de responder ao problema de pesquisa: como o controle de autoridade no VIAF contribui para a efetivação da Web Semântica de modo a proporcionar maior qualidade aos sistemas de busca e recuperação da informação? Para isso, foram definidos os objetivos: 1) descrever sobre os princípios, técnicas e padrões que proporcionam o controle de autoridade na catalogação; 2) analisar o controle de autoridade de nomes para pessoas, instituições e geográfico; 3) apresentar o VIAF, sua origem, as tecnologias envolvidas em sua estrutura e sua perspectiva de contribuição para a Web Semântica. A metodologia compreende um estudo qualitativo, de natureza aplicada, com objetivos exploratórios. Os resultados obtidos apresentam a abrangência internacional do VIAF e a caracterização sobre as tecnologias e a forma pela qual são estabelecidos os registros de autoridade no consórcio. Os registros de autoridade estabelecidos no VIAF compreendem a agregação de vocabulários de valor estabelecidos pelas principais agências catalogadoras no mundo, que têm sido adotados em *Linking Open Data*.

**Palavras-chave:** Controle de autoridade. Padrões de metadados. *Virtual International Authority File*. *Linked Open Data*. Web Semântica.

ROMANETTO, L.M. **The authority control in the consortium VIAF**. 2017. 104f. Dissertation (Master of Information Science) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, 2017.

## **ABSTRACT**

Authority control ensures consistency in information systems. Adopted in the cataloging, the theme received greater visibility from the automation of the catalogs, during the decade of 1980 with automation of catalogs. Since then emerged projects cooperation targeted to exchange and cooperation of authority data. The Virtual International Authority File (VIAF) is consortium of international cooperation between library and national agencies what provides authority file in the linking open data. This study was developed with intent to respond to research problem: how does the authority control in VIAF contributes to the effectiveness from the semantic web so provide higher quality to the search systems and information retrieval? For this were defined the objectives: 1) to describe about the principles, techniques and standards what provide authority control in cataloging; 2) to analyze authority control of names for people, institutions and geographical; 3) present the VIAF, your origin, the technologies involved in your structure and its perspective of contribution to the Semantic Web. The methodology comprises a qualitative study of an applied nature with exploratory objectives. The obtained results present the international scope of VIAF and characterization about the technologies and the way in which are established authority records in the consortium. The authority established the VIAF records comprise the vocabulary of aggregate value established by the main catalogers agencies in the world. What stands out to the importance relevance of the consortium to the international community.

**Keywords:** Authority control. Metadata standards. Virtual International Authority File. Linked Open Data. Semantic Web.

## LISTA DE FIGURAS

Figura 1 - Padrões para implementação do controle bibliográfico.....	22
Figura 2 - Produção de artigos indexados na Scopus e Web of Science.....	41
Figura 3 – Produção anual nas bases de dados Scopus e Web of Science..	42
Figura 4 – Produção por país entre Scopus e Web of Science.....	43
Figura 5 - Produtividade por periódico na Web of Science e Scopus.....	44
Figura 6 - Mapa da participação internacional no VIAF.....	59
Figura 7 – As camadas da Web Semântica.....	65
Figura 8 – Diagrama <i>Linking Open Data</i> .....	71
Figura 9 - Cluster VIAF de nome geográfico Brasil.....	76
Figura 10 - Cluster VIAF para nome geográfico São Paulo.....	77
Figura 11 - Cluster VIAF entrada autorizada de nome geográfico Brasil.....	78
Figura 12 - Cluster VIAF entrada autorizada de nome geográfico São Paulo.....	79
Figura 13 - Registro VIAF para nome de entidade coletiva UNESP.....	86
Figura 14 - Cluster VIAF entrada autorizada de nome para entidade coletiva.....	87
Figura 15 - Registro VIAF para pessoa Machado de Assis.....	89
Figura 16 - Cluster VIAF entrada autorizada de nome pessoal Machado de Assis.....	90

## LISTA DE QUADROS

Quadro 1 - Descrição dos cinco formatos MARC.....	18
Quadro 2 - Campos e etiquetas MARC autoridade.....	19
Quadro 3 – Descrição das entidades do modelo FRBR.....	23
Quadro 4 – Descrição de entidades e atributos FRAD.....	24
Quadro 5 – Elementos e subelementos MADS.....	37
Quadro 6 - Tipologia dos pontos de acesso.....	47
Quadro 7 – ISO 25964 parte 1.....	51
Quadro 8 – Conteúdo ISO 25964 parte 2.....	52
Quadro 9 - Elementos descritivos da norma ISO 26964.....	53
Quadro 10 - Código da norma ISO 26964.....	53
Quadro 11 - Relações da norma ISO 26964.....	53
Quadro 12 – Registro de autoridade estruturado por tesauro.....	55
Quadro 13 - Vocabulário de classes SKOS.....	57
Quadro 14 – Trecho da correspondência entre SKOS/XL e MADS/RDF.....	58
Quadro 15 - Descrição dos dados fornecidos por unidade cooperante ao VIAF.....	60
Quadro 16 - Mapeamento entre elementos do MARC Autoridade e MADS 2.0.....	69
Quadro 17 - Campos de apresentação de metadados adotados no VIAF....	75
Quadro 18 - Trecho registro MARC para nome geográfico.....	80
Quadro 19 - Trecho registro MARC-21 em XML.....	81
Quadro 20 - Trecho de classes e subclasses no registro MARC-21 em XML.....	82
Quadro 21 – Elementos de metadados e valores no registro MARC-21 em XML.....	83
Quadro 22 - Trecho de registro em estrutura RDF.....	84
Quadro 23 - Mapeamento de ontologias de domínio nos registros RDF.....	85
Quadro 24 - Cluster VIAF.....	86
Quadro 25 – Trecho registro MARC de nome para entidade coletiva.....	88
Quadro 26 - Trecho registro MARC para nome pessoal.....	91

## LISTA DE SIGLAS

AACR	Anglo-American Cataloguing Rules
ALA	American Library Association
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
DND	Deutsche Nationalbibliothek
FOAF	Friend of a friend
FRAD	Functional Requirements for Authority Data
FRBR	Functional Requirements for Bibliographic Records
FRSAD	Functional Requirements for Subject Authority Data
GARE	Guidelines for Authority and Reference Entries
GARR	Guidelines for Authority Records and References
IFLA	International Federation of Library Associations and Institutions
ISADN	International Standard Authority Data Number
ISBD	International Standard Bibliographic Description
ISO	International Organization for Standardization
KOS	Knowledge Organization Systems
LC	Library of Congress
LOD	Linking Open Data
MADS	Metadata Authority Description Schema
MARC	Machine Readable Cataloging
NISO	National Information Standards Organization
OCLC	Online Computer Library Center
OPAC	Online public access catalog
OWL	Online Writing Lab
RDA	Resource Description and Access
RDF	Resource Description Framework
SKOS	Simple Knowledge Organization System
SRU	Search/Retrieve via URL
TDI	Tratamento descritivo da Informação
TIC	Tecnologias da Informação e Comunicação
UBC	Universal Bibliographic Control
UNESCO	United Nations Educational, Scientific and Cultural Organization

UNIMARC	Universal Machine Readable Cataloging
URI	Universal Resource Identifier
VIAF	Virtual International Authority File
W3C	World Wide Web Consortium
WWW	World Wide Web
XML	eXtensible Markup Language

## SUMÁRIO

1	INTRODUÇÃO.....	14
1.1	Problema de pesquisa .....	28
1.2	Justificativa.....	29
1.3	Objetivo geral .....	30
1.4	Objetivos específicos .....	31
1.5	Materiais e métodos .....	31
1.6	Estrutura do trabalho.....	33
2	CONTROLE DE AUTORIDADE .....	34
2.1	Produção científica sobre controle de autoridade .....	39
2.2	A representação de autoridade no tratamento descritivo da informação ...	45
2.3	Relações hierárquicas, associativas e de equivalência no controle de autoridade.....	49
3	<i>VIRTUAL INTERNATIONAL AUTHORITY FILE</i> .....	59
3.1	Infraestrutura de dados em Linked Data .....	63
3.2	Web Semântica e Linked Open Data .....	64
3.3	Publicação de dados de autoridade em Linked Open Data.....	70
4	CARACTERIZAÇÃO DE ARQUIVOS DE AUTORIDADE NO VIAF .....	73
5	CONSIDERAÇÕES FINAIS.....	92
	REFERENCIAS.....	95

## 1 INTRODUÇÃO

Em um contexto global, no qual o acesso e o uso da informação são considerados determinantes para a sociedade da informação, assim como estabelece o segundo paradigma de Manuel Castells, as tecnologias têm sido desenvolvidas com o intuito de promover o acesso e o uso da informação e proporcionam avanços individuais e coletivos na sociedade (WERTHEIM, 2007). Essa afirmação, corroborada por Capurro e Hjørland (2007), considera que a produção, distribuição e o acesso à informação tornaram-se essenciais para o desenvolvimento econômico na atualidade.

A década de 1960 marca o surgimento da Ciência da Informação, período em que os governos passam a manifestar interesse e preocupação com a recuperação, acesso e uso da informação científica e tecnológica (SIQUEIRA, 2010), e impactou no desenvolvimento e implementação de tecnologias, como as adotadas nos processos de representação direcionados à recuperação e acesso à informação.

Neste cenário, bibliotecas e instituições do patrimônio cultural possuem tradição e experiência no desenvolvimento de técnicas para a representação (ZENG; QIN, 2008), como as adotadas na catalogação, que podem ser entendidas como "[...] o estudo, a preparação e a organização de mensagens, com base em registros do conhecimento, reais ou ciberespaciais, existentes ou passíveis de inclusão em um ou vários acervos [...]" (MEY, SILVEIRA, 2010, p.126).

Os avanços tecnológicos obtidos desde a implementação, evolução e popularização das tecnologias da informação e comunicação (TIC) potencializaram o aprimoramento de técnicas adotadas no tratamento descritivo de informação (TDI), representando para a catalogação uma mudança de paradigmas relacionada aos aspectos tecnológicos e representacionais (ALVES; SANTOS, 2013).

O tratamento descritivo da informação estabelecido pela catalogação pode ser entendido como o processo de descrever os elementos formais do objeto físico ou digital para a definição de pontos de acesso que, por meio da descrição, atuam na mediação entre documento e a necessidade de informação dos usuários, expressada na recuperação da informação, uma vez que possibilita o acesso às características da expressão. Desta forma, o produto da catalogação é o registro bibliográfico (CATARINO; SOUZA, 2012).

Durante a década de 1980, o controle de autoridade foi implementado como atividade fundamental para a automação de catálogos e consiste na adoção de uma forma padronizada na descrição de pontos de acesso e de estabelecer relações entre diferentes pontos de acesso (HERRERO PASCUAL, 1999). Deste modo, compreende-se que o conceito de controle de autoridade pode ser entendido a partir do termo autoridade, que remete a um recurso ou conceito com rótulo controlado.

Por meio da identificação de formas variantes para nomear pessoas, entidades, obra, tema ou conceito, elegem-se termos ou rótulos, autorizados e não autorizados, a fim de realizar a unificação de pontos de acesso e a devida associação com os respectivos termos variantes. Dentre os benefícios de sua implementação, Sandín Enriquez (2011) destaca a maior precisão conferida aos sistemas de busca e economia nos serviços de catalogação cooperativa.

Os pontos de acesso podem ser classificados como não controlados e controlados (SANDÍN ENRIQUEZ, 2011). O catálogo de autoridade é produto dos pontos de acesso controlados.

Segundo Albás (2003), as principais funções do controle de autoridade são:

- Unificar pontos de acesso idênticos, que apresentem diferentes denominações, pela ocorrência de variações de nomes e sinonímias;
- Diferenciar pontos de acesso distintos, que apresentem a mesma forma de denominação, pela ocorrência de homonímia e polissemia;
- Estabelecer relação entre o termo autorizado de ponto de acesso e suas variantes;
- Estabelecer relações associativas e hierárquicas entre diferentes pontos de acesso autorizados.

Embora o conceito esteja bastante associado à automação dos catálogos, o controle de autoridade passou a ser considerado fundamental desde o surgimento do conceito de Controle Bibliográfico Universal, durante a década de 1970, quando foi definido que cada país deveria desenvolver, por meio de agências de bibliografias nacionais, uma forma de padronização de autoridades, e que, ao realizar a classificação de obras estrangeiras, deveriam ser utilizados os padrões específicos da nacionalidade de origem da obra (TILLET, 2007).

Neste período, houve “[...] o esforço da biblioteconomia em direção ao aperfeiçoamento do acesso à produção bibliográfica mundial e representação pelo

desenvolvimento das redes de informação, resultantes da aplicação da informática aos processos de organização bibliográfica” (CAMPELLO, 2006, p.2).

A ênfase no compartilhamento de dados bibliográfico e de autoridade tem por objetivo aproveitar registros elaborados por outros e reduzir a necessidade de retrabalho, tendo em vista que a partilha internacional reduz os custos com a catalogação (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2001).

O projeto Controle Bibliográfico Universal foi desenvolvido durante o período de 1974 a 1990 com o objetivo de construir e a difundir um sistema normativo para a descrição e o intercâmbio de dados bibliográficos. Contudo, ao longo da década de 1990, a falta de recursos financeiros causou o fim das atividades desenvolvidas no projeto (ANGENJO BULLON; HERNANDEZ CARRASCAL, 2004).

O projeto de Controle Bibliográfico Universal foi estabelecido a partir de dois princípios: 1) cada país é responsável pelo controle bibliográfico de publicações realizadas em seu próprio país; 2) cada país é responsável por elaborar e disponibilizar aos demais países do mundo os registros bibliográficos dessas publicações. Por meio do compartilhamento de registros bibliográficos, é possível organizar a produção de dados bibliográficos e de controle de qualidade em uma escala universal para a partilha de responsabilidades e de benefícios (ANGJELI; MAC EWAN; BOULET, 2014).

O intercâmbio e o compartilhamento de dados foram idealizados desde os primeiros códigos de catalogação, mesmo quando as tecnologias ainda não possibilitavam sua concretização, com destaque para o trabalho de catalogação cooperativa idealizado por Charles Coffin Jewett, uma atividade fundamental para a compreensão do conceito de Controle Bibliográfico Universal.

A partir das transformações potencializadas pelas tecnologias da informação e da comunicação (TIC) sobre as técnicas de tratamento descritivo da informação, cria-se infraestrutura para “[...] a catalogação ajustar-se à mecanização a fim de possibilitar que um livro seja catalogado uma única vez em seu país de origem, proporcionando, entre outras facilidades, uma rápida troca de informações” (MACHADO, 2003, p. 49).

Santos e Pereira (2014) indicam que, a partir do desenvolvimento e implementação do formato *Machine Readable Cataloging* (MARC) pela Library of

Congress, durante a década de 1960, as instituições bibliotecárias passaram a cooperar e a compartilhar informações catalográficas de forma autorizada.

O formato de intercâmbio MARC é um conjunto de códigos e designações de conteúdo cuja estrutura é formada pelo *Information Interchange Format* (ANSI Z39.2), pelo *Format for Information Exchange* (ISO 2709) (LIBRARY OF CONGRESS, 1996), e compreende, ainda na atualidade, um dos padrões de metadados de maior adesão entre as bibliotecas. Em 2012, correspondia a 86% de adesão entre bibliotecas nacionais sul-americanas (CARLOS; ZAFALON, 2013).

Um registro MARC é composto por três elementos: a estrutura do registro, a designação de conteúdo e o conteúdo de dados dos registros.

A estrutura do registro é formada pelo padrão internacional – *Format for Information Exchange* (ISO 2709) – e seu correspondente americano – *Bibliographic Information Interchange* (ANSI / NISO Z39.2). A designação de conteúdo compreende códigos e convenções adotados para identificar e caracterizar os elementos de dados estabelecidos em cada um dos cinco formatos MARC. O conteúdo dos elementos compreende os valores estabelecidos por meio da descrição, que, por sua vez, é definida a partir de códigos, diretrizes e vocabulários de catalogação (LIBRARY OF CONGRESS, 2006a), códigos como a segunda edição do *Anglo-American Cataloguing Rules* (AACR), e seu sucessor o *Resource Description and Access* (RDA), e vocabulários como o *Library of Congress Subject Headings* (LCSH), entre outros.

O MARC é formado por um conjunto de cinco padrões, que cobrem informações específicas e podem ser adotados de forma conjunta ou separadamente, apresentados no quadro 1.

### Quadro 1 – Descrição dos cinco formatos MARC

 <p>Formato MARC 21 Conciso para <b>DATOS BIBLIOGRÁFICOS</b> Biblioteca del Congreso Oficina de Desarrollo de Redes y Normas MARC</p>	<p>O formato MARC 21 para dados bibliográficos registra informações bibliográficas sobre materiais textuais impressos e manuscritos, arquivos de computador, mapas, músicas, recursos contínuos, materiais visuais e mistos. O que inclui título, nomes, tópicos, notas, dados de publicação e informações sobre a descrição física do item.</p>
 <p>MARC 21 Format for <b>HOLDINGS DATA</b> Library of Congress Network Development and MARC Standards Office</p>	<p>O formato MARC 21 para dados de coleção estabelece códigos e convenções para a criação e manutenção de informações sobre dados de coleção em sistemas de biblioteca.</p>
 <p>MARC 21 Format for <b>AUTHORITY DATA</b> Library of Congress Network Development and MARC Standards Office</p>	<p>O formato MARC 21 para dados de autoridade estabelece formas autorizadas adotadas como pontos de acesso nos registros MARC e cobre nomes, assuntos e subdivisões e as relações entre essas entradas.</p>
 <p>MARC 21 Format for <b>CLASSIFICATION DATA</b> Library of Congress Network Development and MARC Standards Office</p>	<p>O formato MARC 21 para dados de classificação foi desenvolvido para a elaboração e manutenção de sistemas para o intercâmbio e processamento de registros de classificação.</p>
 <p>MARC 21 Format for <b>COMMUNITY INFORMATION</b> Library of Congress Network Development and MARC Standards Office</p>	<p>O formato MARC 21 para dados de informação comunitária foi desenvolvido para a criação e manutenção de sistemas para o intercâmbio e processamento de registros sobre informação comunitária</p>

Fonte: Library of Congress (1999; 2006; 2016a; 2016b; 2016c).

Dentre os padrões desenvolvidos, destaca-se, nesta pesquisa, o formato MARC autoridade, que foi desenvolvido para ser utilizado na construção de pontos de acesso para nomes, assuntos e subdivisões de assunto como referência a

formas autorizadas de representação e para estabelecer relações com formas variantes (LIBRARY OF CONGRESS, 2009).

Os pontos de acesso no registro MARC autoridade são estabelecidos a partir de etiquetas numéricas definidas por conteúdo nos campos 1XX, 4XX e 5XX, com o intuito de promover a indexação, recuperação e organização de dados em um registro (quadro 2). Os campos 1XX são definidos a partir da convenção entre os códigos de catalogação e as normas de construção de tesouros (LIBRARY OF CONGRESS, 2006).

**Quadro 2 – Campos e etiquetas MARC autoridade**

<b>Campos para nome</b>	
Nome de pessoa	100
	400
	500
Nome de jurisdição	151
	451
	551
Nome de entidade coletiva	110
	410
	510
Títulos uniformes	130
	430
	530
Nome de evento	111
	411
	511
<b>Campos para assunto</b>	
Termo tópico	150
	450
	550
Termo tópico para nome geográfico	151
	451
	551
Termo gênero ou forma	155
	455
	555
Termos cronológicos	148
	448
	548
Título uniforme com subdivisões temáticas	130
	430
	530

Continua.

Continuação.

<b>Campos para subdivisão de assunto</b>	
Termos gerais de compartimentação	180
	480
	580
Nomes de subdivisão geográfica	181
	481
	581
Termo de subdivisão cronológico	182
	482
	582
Termo forma de compartimentação	185
	485
	585

Fonte: Adaptado de Library of Congress (2009).

O desenvolvimento dos formatos MARC deu origem a outras iniciativas relacionadas ao desenvolvimento e implementação de padrões de metadados, pois os “[...] metadados são considerados uma solução fundamental para promover a representação e o acesso dos recursos informacionais digitais” (ALVES; SANTOS, 2013, p.99).

Segundo Foulonneau e Riley (2008 *apud* CARLOS; ZAFALON, 2013), os padrões de metadados são classificados em padrões de estrutura de metadados descritivos e padrões de conteúdo: os padrões de estrutura definem listas de elementos considerados importantes para descrever o recurso, sua repetibilidade e ordem; e os padrões de conteúdo estabelecem regras de entrada em um campo de metadados pela sintaxe.

Neste contexto, Alves (2005) e Alves e Santos (2013) classificam os padrões de metadados em três níveis: simples, estruturados e ricos.

Segundo as autoras, os padrões de metadados simples são construídos por meio da definição de metadados não estruturados, que apresentam limitações relacionadas à sua estrutura e à descrição, mas proporcionam a localização e a recuperação de recursos informacionais em ambientes como a Web. Os padrões de metadados estruturados são estabelecidos a partir de normas emergentes, com o intuito de promover a descrição simples e proporcionar a identificação, localização e recuperação de recursos informacionais. Os padrões de metadados ricos são estabelecidos a partir de normas e códigos especializados e têm como principal

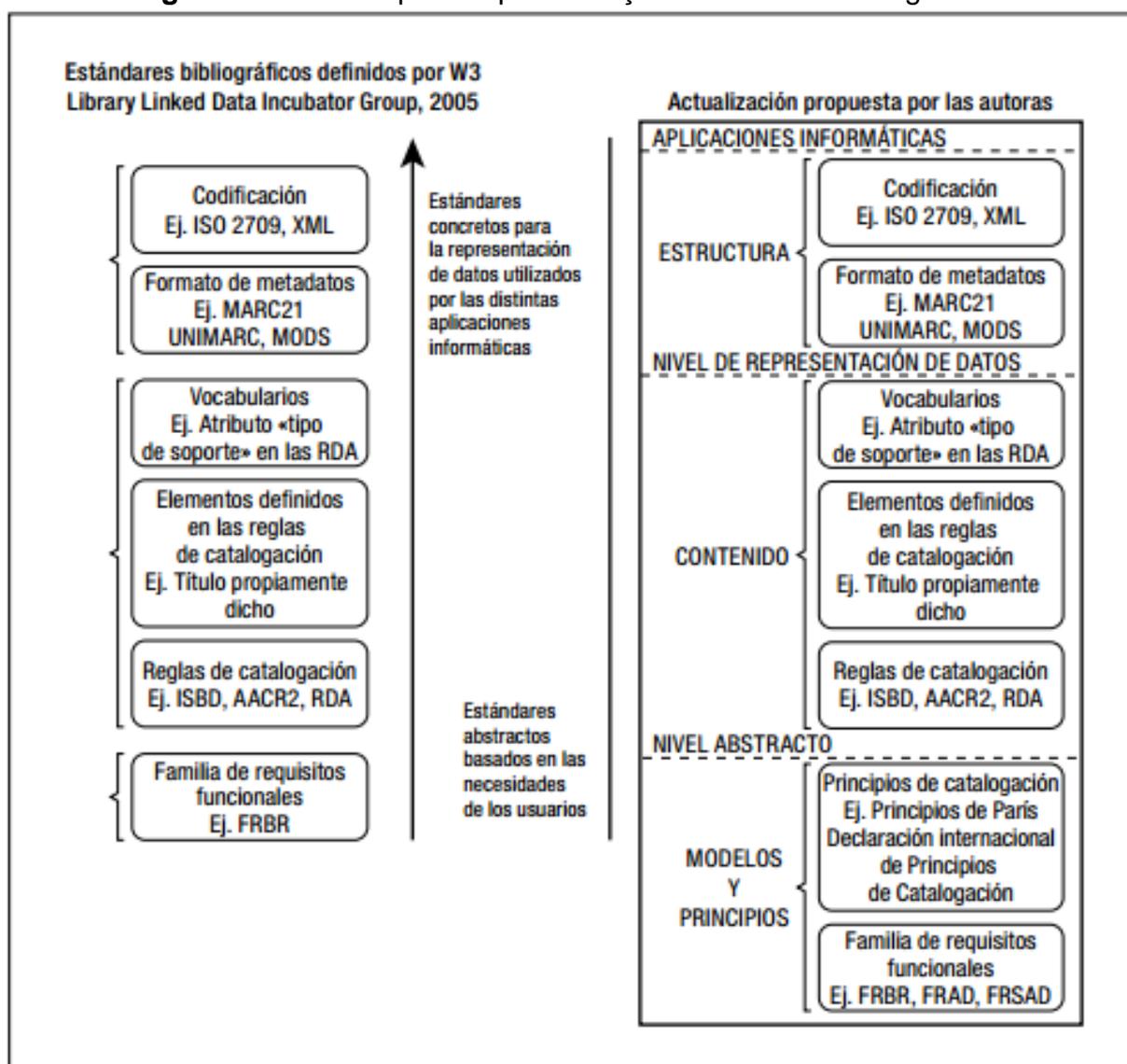
característica um alto grau de estruturação, descrição mais formal e detalhada (ALVES, 2005; ALVES; SANTOS, 2013).

Segundo Alves (2005, p. 94), a padronização dos metadados no domínio bibliográfico envolve aspectos tecnológicos e representacionais, por meio do desenvolvimento ou integração entre:

Modelagem Conceitual do sistema, determinação dos requisitos funcionais, mapeamento dos elementos essenciais, determinação da estrutura de descrição do esquema, determinação dos espaços do valor e determinação dos esquemas de codificação para padronização dos espaços do valor.

Para a correta definição de pontos de acesso em um sistema de informação, torna-se imprescindível estabelecer entidades, atributos e relações que compõem os catálogos automatizados, sustentados por padrões de metadados, códigos de catalogação e de modelos conceituais, com o intuito de promover estrutura, descrição, modelos e princípios aos registros (PICCO; ORTIZ REPISO, 2012). Os padrões de estrutura dão suporte para a definição de entidades, e os padrões de conteúdo fornecem critérios a serem seguidos durante a descrição dos espaços de valor, que são definidos pelos atributos e infraestrutura, conforme ilustra a figura 1.

**Figura 1 - Padrões para implementação do controle bibliográfico**



Fonte: Picco e Ortiz Repiso (2012, p. 149).

O modelo conceitual Requisitos Funcionais para Registros Bibliográficos (FRBR), publicado em 1998, foi desenvolvido pela *International Federation of Library Associations* (IFLA) por meio do Grupo de Trabalho dos FRBR, Seção de Catalogação, dirigida por Patrick LeBoeuf, entre os anos de 1992 e 1995 (TILLET, 2007). O modelo compreende a especificação dos requisitos funcionais de modelagem de banco de dados direcionados às funções dos catálogos.

Catálogo é um meio de comunicação, que veicula mensagens sobre os registros do conhecimento, de um ou vários acervos, reais ou ciberespaciais, apresentando-as com sintaxes e semânticas próprias e reunindo os registros do conhecimento por semelhanças, para os usuários desses acervos. O catálogo explicita por meio das

mensagens, os atributos das entidades e os relacionamentos entre elas (MEY; SILVEIRA, 2009, p.12).

Os FRBR realizam a descrição do modelo conceitual Entidade-Relacionamento e seus atributos, por meio dos quais é estabelecido mapeamento do sistema, sua estrutura atual, entidades, atributos e os relacionamentos presentes nos registros bibliográficos e de autoridade (TILLET, 2007). As entidades do modelo FRBR são identificadas em três grupos, descritos no quadro 3.

**Quadro 3 – Descrição das entidades do modelo FRBR**

	<b>Entidade</b>	<b>Descrição</b>
<b>Grupo 1</b> "Produto de trabalho intelectual ou artístico" (FUSCO, 2011, p. 88)	Obra	"[...] entidade abstrata, uma criação intelectual ou artística distinta [...]" (FUSCO, 2011, p.88).
	Expressão	"[...] realização intelectual ou artística específica que assume uma Obra ao ser realizada [...]" (FUSCO, 2011, p.88).
	Manifestação	"[...] materialização de uma expressão de uma Obra, ou seja seu suporte físico [...]" (FUSCO, 2011, p.89).
	Item	"[...] Item pode exemplificar uma e apenas uma Manifestação" (FUSCO, 2011, p.89).
<b>Grupo 2</b> "Responsáveis pelo conteúdo intelectual ou artístico, pela produção física e disseminação ou pela guarda das entidades do primeiro grupo" (FUSCO, 2011, p. 88).	Pessoa	"[...] entidade individual, pessoa física [...]" (FUSCO, 2011, p.90).
	Entidade coletiva	"[...] entidade coletiva, pessoa jurídica [...]" (FUSCO, 2011, p.90).
<b>Grupo 3</b> "Conjunto adicional de entidades que servem como assunto de Obras [...]" (FUSCO, 2011, p. 88).	Conceito	"[...] abrange uma variedade de abstrações que podem ser objeto de uma Obra [...]" (FUSCO, 2011, p.91).
	Objeto	"[...] abrange uma variedade de coisas materiais que podem ser objeto de uma Obra [...]" (FUSCO, 2011, p.92).
	Evento	"[...] abrange ações e ocorrências que podem ser objeto de uma Obra [...]" (FUSCO, 2011, p.92).
	Lugar	"[...] abrange uma gama completa de localidades [...]" (FUSCO, 2011, p.92).

Fonte: Adaptado de Fusco (2011, p. 88-92).

No FRBR, as entidades compreendem tabelas de identificação dos dados e possuem, cada qual, um conjunto de atributos específicos. Os atributos são a descrição das características de cada entidade e compreende os pontos de acesso.

O FRBR prevê que uma obra pode ser produzida por uma ou mais entidades, *pessoa, entidade coletiva e/ou família*, e expressar, por meio de diferentes formatos (manuscrita, sonora, imagem, entre outras), o conteúdo intelectual ou artístico da obra; a *expressão* dá origem à *manifestação*, que compreende o documento físico e publicado, que, dentre outras possibilidades, poderá se tornar um *item* em um acervo bibliográfico.

O modelo conceitual Requisitos Funcionais para Dados de Autoridade (FRAD), publicado em 2009, compreende uma extensão do modelo estabelecido no grupo 2 do FRBR, portanto, uma especificação para dados de autoridade. O FRAD abrange as entidades *pessoa* e *entidade coletiva* do modelo FRBR, além de incluir no grupo a entidade *família* e uma extensão com as entidades *nome, identificador, ponto de acesso controlado, regra, agência catalogadora*, totalizando oito entidades (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2009), conforme descritas no quadro 4.

**Quadro 4** – Descrição de entidades e atributos FRAD

<b>Entidade</b>	<b>Descrição</b>	<b>Atributos</b>
Pessoa	Indivíduo ou uma personalidade ou identidade estabelecida ou adotada por um indivíduo ou grupo	Data associada à pessoa (nascimento – morte), Título. Outras informações associadas: Gênero, Local de nascimento, Local de morte, País, Afiliação, Endereço, Idioma, Campo de atividade, Profissão, Biografia / história.
Entidade coletiva	Organização ou grupo de pessoas e / ou organizações identificadas por um nome específico agindo como uma unidade.	Local associado à entidade, Data associada à entidade, Idioma, Endereço, Campo de atividade, História.

Continua.

## Continuação.

Família	Duas ou mais pessoas relacionadas por nascimento, casamento, adoção, união civil, ou do estatuto jurídico semelhante, ou que de outra forma se apresentam como uma família.	Tipo de família, Data associada à família, Local associado à família, Campo de atividade, História da família.
Nome	Carácter ou grupo de palavras e / ou caracteres por uma entidade que é conhecida no mundo real.	Tipo de nome, Cadeia de nome, Escopo de uso, Data de uso, Idioma de uso, Escrita do nome, Esquema de transliteração do nome.
Identificador	Número, código, palavra, frase, logotipo, dispositivo, etc., que estão associados a uma entidade e servem para diferenciar essa entidade de outras entidades dentro do domínio em que o identificador é atribuído.	Tipo de identificador.
Ponto de acesso controlado	Nome, termo, código, etc., em que um registro bibliográfico e autoridade de referência serão encontrados.	Tipo, Status, Uso designado, Ponto de acesso indiferenciado, Idioma, Roteiro, Esquema de transliteração, Fonte, Ponto de acesso de base, Adição.
Regras	Conjunto de instruções relativas à formulação e / ou gravação de pontos de acesso controlados (formas autorizadas, formas variantes ou referências, etc.).	Citações Identificador.
Agência	Organização responsável pela criação ou modificação de um ponto de acesso controlado	Nome, Identificador, Localização.

Fonte: Adaptado de Jin (2012, p. 17-53).

Embora a evolução tecnológica tenha proporcionado uma série de avanços no desenvolvimento de sistemas e na forma como a representação da informação tem sido estabelecida na catalogação, existem disparidades entre os sistemas dos catálogos com a Web, o que tem dificultado a integração dos registros no ambiente da Web.

Diferente das funções estabelecidas nos catálogos, em que os recursos informacionais são representados de maneira exaustiva, a *World Wide Web*, conhecida por Web, é um sistema que proporciona o acesso a recursos informacionais interligados e executados na Internet. Entretanto a forma como é estabelecida sua infraestrutura, baseada em *HyperText Markup Language* (HTML), apresenta limitações quando adotada na descrição de conteúdos, o que pode prejudicar a manutenção e a recuperação da informação. A linguagem HTML promove a descrição para a apresentação, e não a estrutura, dos conteúdos.

Atrelado a esse cenário, soma-se a necessidade de se integrar as bibliotecas com outras instituições do patrimônio cultural, em seus registros bibliográficos e de autoridade na Web, com o intuito promover maior alcance na disseminação de recursos informacionais e aumentar a visibilidade dos serviços realizados por essas instituições.

Com o objetivo de potencializar uma melhor estrutura para a Web, foi estabelecido o conceito de Web Semântica, que, diferente da Web atual, em que conteúdos são estruturados para serem exibidos e lidos por pessoas, propõe que os conteúdos sejam estruturados de maneira semântica e possam ser "inteligíveis" por programas chamados de "agentes de software" (MARCONDES, 2005).

Na Web Semântica, os padrões de metadados são adotados como recurso de garantia para a interoperabilidade entre sistemas heterogêneos e para favorecer a ligação entre os conteúdos, formando, assim, a denominada teia de dados (COYLE, 2010).

A interoperabilidade pode ser entendida como "A capacidade de um sistema de hardware ou de software de se comunicar e trabalhar efetivamente no intercâmbio de dados com outro sistema, geralmente de tipo diferente, projetado e produzido por um fornecedor diferente" (SAYÃO; MARCONDES, 2008, p.136). Na Web Semântica existem diferentes níveis de interoperabilidade que são alcançadas por meio da integração entre as tecnologias XML, RDF e ontologias, permitindo estruturar, representar e interpretar os dados.

A infraestrutura estabelecida por meio dessas tecnologias compreende a base para a concretização da Web Semântica, que requer uma vasta quantidade de dados disponíveis em um formato padrão e relacionáveis (WORLD WIDE WEB CONSORTIUM, 2015). Neste contexto, surge o conceito de *Linked Data*, que "se

refere ao conjunto de melhores práticas para a publicação de dados estruturados na Web" (WORLD WIDE WEB CONSORTIUM, 2014, não paginado, tradução nossa).

Embora o conceito de *Linked Data* vislumbre a vinculação de dados, ele pode ser adotado para dados de acesso restrito, o que limita e restringe o potencial de reuso de dados e motivou a criação do *Linked Open Data*, que adota os princípios de vinculação de dados estabelecidos pelo W3C, mas se aplica aos dados abertos. Segundo Marcondes (2012), essa é uma oportunidade para que os registros que compõem os acervos de arquivos, bibliotecas e museus sejam interligados por meio das tecnologias da Web Semântica.

Isso tem motivado as instituições bibliotecárias no sentido de atuarem no estudo e na conversão de registros bibliográficos e de autoridade, tendo como base os princípios *Linked Data* estabelecido pela W3C e as tecnologias da Web Semântica, também denominada Web de dados, na publicação de dados vinculados.

O *Virtual International Authority File (VIAF)* é um consórcio internacional que segue os princípios *Linked Data*, por meio do qual bibliotecas e agências nacionais contribuem de forma mútua na conversão e disponibilização de arquivos de autoridade em âmbito mundial.

O VIAF teve seu início no ano de 1998 por meio de um projeto de cooperação entre a *Library of Congress (LC)*, a *Deutsche Nationalbibliothek (DNB)* e *Online Computer Library Center (OCLC)*; foi formalizado como consórcio em 2003, e, a partir de 2012, passou a ser oferecido pelo *Online Computer Library Center (OCLC)* (ONLINE COMPUTER LIBRARY CENTER, 2015). Atualmente, o VIAF conta com a participação de 52 bibliotecas e agências nacionais localizadas em 38 países.

NO VIAF, os arquivos são estruturados a partir dos princípios *Linked Data* nos formatos XML e RDF e agregados entre si, o que estabelece a formação de *Clusters* de valores por entidade. São estabelecidas ligações entre os registros de autoridade elaborados em cada unidade cooperante em seu país de origem. Neste ponto, o conceito de Controle Bibliográfico Universal é expandido, ao permitir variações nacionais e regionais na forma autorizada a coexistir e apoiar as necessidades para variações de idioma, roteiro e ortografia (ONLINE COMPUTER LIBRARY CENTER, 2015). Esses relacionamentos são estabelecidos a partir dos elementos de metadados determinados pelo formato de intercâmbio MARC autoridades e UNIMARC, padrões amplamente adotados por bibliotecas nacionais.

### **1.1 Problema de pesquisa**

A falta de padronização na descrição de autoridade para nomes de autores, de instituições e geográficos é uma das dificuldades existentes em sistemas de recuperação de informação disponibilizados na Web. A ausência de padrões no registro de dados pode repercutir em variações na descrição de nomes dificultando o acesso ao recurso desejado. Dentre os motivos para o não uso regular dos padrões, destacam-se: 1) desconhecimento sobre padrões e regras de descrição de autoridade ou das vantagens em adotá-los, por parte dos desenvolvedores e responsáveis pelos sistemas; 2) limitações de recursos humanos e tecnológicos para implantação; 3) a instituição que elabora e disponibiliza os registros de dados não adota uma política de padronização para controle de autoridade; 4) diferentes instituições adotam diferentes formas de padronização; 5) a descrição fica a critério de terceiros, como ocorre no autodepósito de artigos científicos, em que o próprio autor descreve seus dados; 6) tentativas de tradução de nomes de instituições; 7) abreviação de nomes de instituições; 8) descrição genérica (ex.: Instituto de química), que pode ser atribuída a diversas instituições.

A dificuldade de recuperação é intensificada no atual contexto tecnológico informacional (de aumento exponencial de registros do conhecimento, disponibilizados de modo disperso no ambiente da Web) sem que sejam estabelecidos relacionamentos entre eles por meio de pontos de acesso controlados e por conta da baixa interoperabilidade entre sistemas. Deste modo, os resultados da busca tornam-se condicionados ao conhecimento prévio do usuário sobre as diferentes fontes de busca disponíveis e, principalmente, na elaboração intuitiva do usuário para definir termos ou expressão de busca que correspondam aos termos registrados no sistema.

Nesse contexto, o problema deste estudo pode ser assim apresentado:

Como o controle de autoridade no consórcio VIAF contribui para a efetivação da Web Semântica de modo a proporcionar maior qualidade aos sistemas de busca e recuperação da informação?

Diante do exposto, tem-se, como hipótese de estudo, que o controle de autoridade realizado em bibliotecas e disponibilizados por meio dos arquivos de autoridade no consórcio VIAF pode proporcionar maior qualidade na representação e vinculação de dados para a concretização da Web Semântica. Também permite

que os arquivos sejam reutilizados por outras instituições bibliotecárias na definição de pontos de acesso, durante o tratamento descritivo da informação.

## **1.2 Justificativa**

O aumento expressivo de recursos informacionais dispersos em ambiente Web tem gerado grandes dificuldades para a busca e recuperação da informação. As tecnologias da Web 2.0 funcionam a partir de sistemas que, em sua maioria, disponibilizam registros documentais sem o prévio tratamento descritivo de informação (TDI), devido às limitações tecnológicas e pela grande quantidade de registros indexados periodicamente.

Embora existam técnicas ideais de TDI, desenvolvidas principalmente no âmbito das atividades de catalogação em bibliotecas e unidades de informação, esse serviço torna-se local e muitas vezes não chega a ser integrado a outros sistemas como da Web, prática que poderia contribuir para uma melhor qualidade dos registros na Web e proporcionar maior visibilidade aos recursos informacionais.

Com o intuito de possibilitar o compartilhamento e reuso de arquivos de autoridade, bibliotecas e agências nacionais têm atuado por meio do consórcio VIAF, ao estruturar e disponibilizar arquivos de autoridade seguindo os princípios *Linked Data* estabelecidos pelo *World Wide Web Consortium (W3C)* na conversão de registros em formatos como XML, RDF, entre outros.

O estudo realizado nesta pesquisa contribui no plano teórico e social. Sua relevância teórica relaciona-se à abordagem de aspectos e princípios de catalogação e sua contextualização sobre o aprimoramento das técnicas de tratamento descritivo da informação, que mostra como o avanço tecnológico tem possibilitado que o controle de autoridade obtido por meio da catalogação pode ser aproveitado por outras áreas de atividades no ambiente Web, a partir da interoperabilidade entre padrões de metadados.

A relevância social se atribui à apresentação e à divulgação do consórcio VIAF para as comunidades que atuam na representação de dados de autoridade em diferentes domínios para que possam tomar conhecimento da estrutura e perspectiva de aplicação.

No contexto nacional, verifica-se que, na catalogação, o controle de dados de autoridade no ambiente Web tem sido pouco abordado na literatura, apesar de a representação de nomes para pessoas, instituições e nomes geográficos ser considerada pontos de acesso básico na representação de documentos, como livros, artigos publicados em revistas científicas, entre outros.

No Brasil, destacam-se: os estudos de Assumpção (2012), de ordem teórica, que abordam definições, processos e componentes do controle de autoridade; e Salgado (2015), que investigou sobre o controle de autoridade estabelecido pela diretriz RDA.

Internacionalmente, destacam-se: os estudos de Sandín Enriquez (2011), em que o autor realizou uma investigação sobre o estado da arte no desenvolvimento do consórcio e sobre a recuperação de informações de autoridade no VIAF; e Garzon Farinos (2014), que descreve sobre o controle de autoridade no contexto da Espanha, a partir de um projeto de construção de uma base de autores da Espanha.

O VIAF disponibiliza, por meio de arquivos de autoridade, um conjunto de vocabulários construído pelas principais agências catalogadoras no mundo, proporcionando maior qualidade a representação da informação e melhorando os sistemas de busca e recuperação da informação, sem se limitar às funções de catálogo e ao domínio bibliográfico. Entretanto, no Brasil, o VIAF e sua perspectiva de aplicação são pouco conhecidos e pouco abordados, provavelmente em virtude de o Brasil não ser um contribuinte, o que pode aumentar a importância deste estudo em nível nacional, a fim de apresentar a relevância do VIAF e suas possibilidades de aplicação e proporcionar uma maior qualidade dos serviços de catalogação e visibilidade aos serviços desenvolvidos no país.

### **1.3 Objetivo geral**

O objetivo geral refere-se à análise dos aspectos teóricos de controle de autoridade para nomes de pessoa, instituições e nomes geográficos no consórcio *Virtual International Authority File (VIAF)*, na perspectiva de aplicação na Web Semântica.

#### **1.4 Objetivos específicos**

- Descrever os princípios, técnicas e padrões que proporcionam o controle de autoridade na catalogação, destacando sua evolução, características e conceitos;
- Analisar o controle de autoridade de nomes para pessoas, instituições e nomes geográficos;
- Apresentar o consórcio *Virtual International Authority File* (VIAF), sua origem, as tecnologias envolvidas em sua estrutura e sua perspectiva de contribuição para a Web Semântica.

#### **1.5 Materiais e métodos**

A abordagem adotada nesta pesquisa é de ordem qualitativa, de natureza aplicada e com objetivos exploratórios. Para Neves (1996, p. 1), a pesquisa qualitativa “compreende um conjunto de diferentes técnicas interpretativas que visam descrever e decodificar os componentes de um sistema complexo de significados”. Desta forma, esta pesquisa é qualitativa, pois envolve a compreensão e a interpretação de princípios da catalogação descritiva para controle de autoridade, que deram origem ao serviço VIAF, e pela interpretação das tecnologias envolvidas no seu desenvolvimento para o atual contexto sociocultural e tecnológico. Caracteriza-se como sendo de cunho exploratório, pois visa apresentar o serviço VIAF, seus conceitos e tecnologias, que, em uma abordagem preliminar, representam um tema relevante, mas pouco abordado na literatura científica em âmbito nacional.

Quanto ao procedimento, destaca-se a pesquisa bibliográfica, documental e levantamento de dados, para fundamentação teórica da pesquisa, a fim de identificar pesquisas desenvolvidas que apresentem relação com o tema pesquisado, as metodologias, os resultados e a contribuição para a área de Ciência da Informação.

Para atender aos objetivos propostos, serão adotados como procedimentos da pesquisa:

- 1 Levantamento bibliográfico;
- 2 Elaboração de referencial teórico;

3 Caracterização sobre catalogação, tratamento descritivo da informação, controle de autoridade, padrões de metadados, interoperabilidade, *Linked Data* e Web Semântica;

4 Apresentação do consórcio VIAF, seus conceitos e tecnologias;

5 Descrição da perspectiva de aplicação do consórcio VIAF no ambiente Web.

Para o levantamento bibliográfico, foram realizadas buscas em livros, relatórios, manuais, periódicos da área de Ciência da Informação e tecnologias, anais de eventos nacionais e internacionais, teses, dissertações, bases de dados textuais e referenciais, por meio do Portal de Periódicos CAPES, documentos disponíveis em meio eletrônico.

Com o intuito de caracterizar a produção bibliográfica pertinente ao tema, foi elaborada a análise bibliométrica nas bases de dados Web of Science e Scopus, apresentada no item 2.1.

A Web of Science, mantida pela *Thomson Reuters*, é uma base de dados que fornece acesso às principais fontes de informação, desde 1900, por meio da integração do *Science Citation Index Expanded*, *Social Sciences Citation Index* e *Arts & Humanities Citation Index* (THOMSON REUTERS, 2016).

A Scopus, mantida pela Elsevier, disponibiliza acesso à literatura científica por meio de textos completos, resumos e citações nas áreas de ciência, matemática, engenharia, tecnologia, saúde, medicina, ciências sociais, artes e humanidades. O conteúdo disponibilizado na base abrange conteúdos *peer-reviewed* desde o ano de 1960 e é considerada uma das maiores bases de dados do conhecimento (SCOPUS, 2015).

As buscas nas bases foram realizadas, considerado o período de início de abrangência da base até o ano de 2015. Assim, a busca por artigos indexados realizada na *Scopus* abrange o período de 1960 a 2015 e na *Web of Science* o período de 1900 a 2015.

O *corpus* da pesquisa compreende artigos recuperados por meio da expressão de busca "*Cataloging AND ("Authority work" OR "Authority control" OR "Authority data" OR "Metadata authority" OR "Authority Files")*" nos campos de título, resumo e palavras-chave.

Os dados recuperados foram extraídos e estruturados a partir do Microsoft Excel Office 2013, para formatação e correção de dados que possibilitaram a

elaboração de indicadores de produção por meio de *Rankings* de publicação por ano, países e periódicos.

### **1.6 Estrutura do trabalho**

A dissertação foi desenvolvida seguindo uma estrutura de apresentação por capítulos descritos a seguir.

Capítulo 1 - caracteriza o contexto da pesquisa, apresenta a justificativa, os objetivos geral e específicos e os materiais e métodos adotados no desenvolvimento da pesquisa.

Capítulo 2 - contextualiza os conceitos e fundamentos que deram origem ao controle de autoridade, aborda os padrões, códigos e diretrizes internacionais que foram criados e iniciativas internacionais relacionadas ao compartilhamento de dados de autoridade. O capítulo apresenta, ainda, uma breve análise da produção científica nas bases Web of Science e Scopus e expõe o modo pelo qual a representação de dados de autoridade tem sido abordada no tratamento descritivo da informação e sua relação com os sistemas de organização do conhecimento.

Capítulo 3 - contextualiza as tecnologias adotadas no compartilhamento de dados vinculados, por meio dos conceitos de Web Semântica e *Linked Data*, e discorre sobre a atuação da comunidade internacional na publicação de dados em Linked Open Data.

Capítulo 4 – apresenta como resultados da pesquisa a infraestrutura e a abrangência internacional do VIAF, por meio do detalhamento dos registros de autoridade estabelecidos pelo consórcio, a partir dos formatos XML e RDF.

Capítulo 5 - apresenta as considerações da pesquisa a partir da descrição dos objetivos estabelecidos e alcançados.

## 2 CONTROLE DE AUTORIDADE

O controle de autoridade é um tema adotado no tratamento descritivo da informação e passou a ser reconhecido como atividade fundamental para a automação dos catálogos desde a década de 1980. O mesmo é realizado a partir do trabalho de autoridade, que compreende a investigação e o esforço intelectual aplicado na criação e manutenção de registros de autoridade (JIMENES PELAYO; GARCÍA BLANCO, 2002).

O controle de autoridade compreende o resultado do processo denominado trabalho de autoridade, que Maxwell (2002, *apud* ASSUMPÇÃO; SANTOS, 2013, p.5) define como "[...] pontos de acesso uniformes, na manutenção do registro de autoridade e no registro das decisões tomadas em um arquivo de autoridade". Assim, se estabelecem a descrição de pontos de acesso aos catálogos e as relações entre os diferentes pontos de acesso, por meio da identificação de formas variantes para denominar pessoas, entidades, obra, tema ou conceito, e se realiza a eleição de termos autorizados e não autorizados (HERRERO PASCUAL, 1999).

Segundo Jimenes Pelayo e García Blanco (2002), o controle de autoridade compreende três atividades: 1) a compilação, o registro e a manutenção dos dados de autoridade; 2) a verificação destes dados; 3) o uso de formas estabelecidas e autorizadas como pontos de acesso no catálogo da biblioteca.

O trabalho de autoridade pode ser realizado em um nível de maior ou menor exaustividade e de modo cooperativo ou individual. Burke e Shorten (2010), em estudo sobre os serviços realizados em bibliotecas de diferentes portes, apontam que o trabalho de autoridade realizado por essas instituições são distintos: bibliotecas de pequeno porte são menos propensas a adquirir serviços de fornecedores externos, de modo que realizam trabalho em nível local; bibliotecas de grande porte apresentam maior incidência de adesão sobre serviços externos, como redes de catalogação cooperativas e consórcios.

Uma das primeiras iniciativas para a implementação do controle de autoridade em nível internacional foi estabelecida pela *United Nations Educational, Scientific and Cultural Organization* (UNESCO), em 1974, que apresentou recomendações e atribuiu a responsabilidade sobre as agências bibliográficas nacionais em estabelecer o controle de autoridade por meio da eleição de formas autorizadas para descrição de nomes de pessoas e entidades coletivas em seu país de origem

(INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2001).

O desenvolvimento e a viabilidade de um sistema internacional de dados de autoridade foram estabelecidos em 1978, em uma Conferência realizada pela IFLA, em Strbske Pleso, na antiga Checoslováquia, concretizados por meio da publicação da *Guidelines for Authority and Reference Entries* (GARE), em 1984, e, posteriormente, da *Guidelines for Subject Authority and Reference Entries*, em 1993, cuja atualização foi denominada por *Guidelines for Authority Records and References* (GARR) (PLASSARD, 2001).

De acordo com as recomendações da *International Federation of Library Associations and Institutions* (2001), os pontos de acesso de autoridade devem ser compostos por uma estrutura formada por três entradas: 1) entrada de autoridade, que compreende o registro de autoridade completo ou apenas a entrada autorizada que é exibida ao usuário; 2) entrada de referência, que fornece informações para que o leitor possa estabelecer relações e associações entre entidades; 3) entrada explicativa geral, um tipo especial de entrada de referência que apresenta entradas não aceitas nas categorias anteriores.

A catalogação tem como característica a adoção de acordos internacionais que, quando estabelecidos para o controle de autoridade, visem garantir que pontos de acesso controlados sejam registrados de forma consistente. Plassard (2001) afirma que por meio do avanço tecnológico tornou-se possível integrar registros de autoridade criados em diferentes países sem a necessidade de determinar formas padronizadas de descrição, preservando tradições culturais e linguísticas.

Diante do esforço internacional no desenvolvimento de projetos em busca do desenvolvimento de um serviço internacional que cubra a agregação de registros de autoridade, diversas são as iniciativas de países que atuam em projetos de harmonização de códigos de catalogação e padrões de metadados (TILLET, 2000).

O Projeto AUTHOR, iniciado em 1995, consiste na cooperação entre França, Reino Unido, Espanha, Portugal e Bélgica, e tem com foco o desenvolvimento de vocabulários multilíngues para dados de autoridade. Esse projeto mostra-se relevante, pois, de acordo com Tillet (2000, p.168, tradução nossa), "O controle de acesso do futuro deve proporcionar diferenças sobre a adoção de distintos códigos de catalogação, formas de transliteração e diferenças culturais".

Buttlar e Garcha (1998) pontuam que a catalogação tem passado por

mudanças significativas, podendo expandir o uso da catalogação em áreas não convencionais. Essa mudança requer que o profissional catalogador desenvolva novas habilidades que o capacitem para atuar em funções como analistas de acesso à informação, design de sistemas, gestores de recursos técnicos, entre outros. Essa ideia corrobora os resultados apresentados por Munnich (2000), segundo o qual o reuso de registros bibliográficos elaborados pela Library of Congress (LC) e British National Bibliography em bibliotecas alemãs demandou a intervenção intelectual e manual, sobretudo em redes de arquivos em *Linked Data*.

Angenjo Bullon e Hernandez Carrascal (2004) indicam que a comunidade internacional tem atuado na reformulação de conceitos e fundamentos de catalogação que proporcionem o intercâmbio e a integração entre sistemas, o que está diretamente relacionado com mudanças e atualizações nas tecnologias adotadas na catalogação.

Essa reformulação de conceitos tem sido motivada por três razões: 1) para que seja realizada a manutenção, promoção e harmonização de códigos de acordo com os conceitos relacionados ao controle bibliográfico e aos recursos em meio eletrônico e digital; 2) para o desenvolvimento de estratégias que assegurem a promoção de novas regras internacionais e o controle bibliográfico; 3) com o intuito de obter avanços relacionados aos aspectos que permeiam os arquivos e a conservação de recursos em meio eletrônico (ANGENJO BULLON; HERNANDEZ CARRASCAL, 2004).

Esses objetivos têm sido implementados por meio de duas iniciativas: 1) o consórcio VIAF, que tem como principal característica seu caráter universal de cooperação e reuso de dados e sua integração no projeto *Linking Open Data*; 2) o estudo dos requisitos de metadados para a descrição bibliográfica, baseados em XML, que vem sendo realizado pela comunidade internacional Dublin Core Metadata Initiative e pela Library of Congress por meio dos padrões de metadados: *Metadata Encoding and Transmission Standards* (METS); *Metadata Object Description Schema* (MODS); *Metadata Object Description Schema* (MIX); e *Metadata Authority Description Schema* (MADS). O intuito é promover o acesso, utilizando diferentes estruturas de metadados para a transmissão da informação na Web (ANGENJO BULLON; HERNANDEZ CARRASCAL, 2004).

Complementando essas iniciativas, a IFLA tem atuado por meio da conversão de vocabulários de valor dos principais padrões, códigos, diretrizes e modelos

conceituais MARC 21, UNIMARC, ISBD, RDA e FRBR e suas extensões, entre outras, em formatos XML e RDF, com o intuito de realizar a unificação de valores e promover a interoperabilidade desses vocabulários com outros vocabulários, o que tem sido publicado por meio do sistema de gerenciamento *Open Metada Registry*<sup>1</sup> e permite a recuperação da informação por meio da linguagem de recuperação SPARQL (WILLER; DUNSIREE, 2013).

O padrão de metadados MADS, publicado em 2005, compreende um esquema XML desenvolvido como complemento para o padrão MODS, para fornecer metadados sobre um conjunto de elementos de autoridade sobre pessoas, organizações, eventos e termos (LIBRARY OF CONGRESS, 2016).

Os registros no MADS são estabelecidos pelo elemento raiz, que pode ser individual, por meio do elemento *mads*, ou uma coleção, pelo elemento *madsCollection* (LIBRARY OF CONGRESS, 2016b). Um elemento *madsCollection* não possui atributos e compreende um conjunto de subelementos *mads*. Cada elemento *mads* conta como atributo ID e versão, e subelementos obrigatórios e opcionais, apresentados no quadro 5.

**Quadro 5 – Elementos e subelementos MADS**

<pre> &lt;authority&gt;   &lt;name&gt;     &lt;namePart&gt;     &lt;description&gt;   &lt;titleInfo&gt;     &lt;title&gt;     &lt;subTitle&gt;     &lt;partNumber&gt;     &lt;partName&gt;     &lt;nonSort&gt;     &lt;topic&gt;   &lt;temporal&gt;   &lt;genre&gt;   &lt;geographic&gt;   &lt;hierarchicalGeographic&gt;     &lt;country&gt;     &lt;province&gt;     &lt;region&gt;     &lt;state&gt;     &lt;territory&gt; </pre>
--

Continua.

<sup>1</sup> Fonte: Open Metadata Registry. Disponível em: <<http://metadataregistry.org/>>.

## Continuação.

<county>
<province>
<region>
<state>
<territory>
<county>
<city>
<island>
<area>
<extraterrestrialArea>
<citySection>
<occupation>
<related> - same as <authority>
<variant> - same as <authority>
<affiliation>
<position>
<organization>
<address>
<street>
<city>
<state>
<country>
<postcode>
<email>
<phone>
<fax>
<hours>
<dateValid>
<classification>
<fieldOfActivity>
<identifier>
<language>
<languageTerm>
<scriptTerm>
<note>
<url>
<extension>
<recordInfo>
<recordContentSource>
<recordCreationDate>
<recordChangeDate>
<recordIdentifier>
<recordOrigin>
<languageOfCataloging>
<languageTerm>
<languageScript>
<descriptionStandard>

Fonte: Library of Congress (2016b).

As relações entre elementos e subelementos no MADS estabelecem que um elemento mads <authority> esteja vinculado a diversos subelementos, como <name>, <titleInfo>, <temporal>, <genre>, entre outros, o que permite caracterizar uma entidade a partir de suas formas variantes, associativas e suas características, também conhecidos por atributos.

O MADS em formato XML foi desenvolvido para ser compatível com o formato MARC autoridades por meio dos rótulos de cada elemento de metadados. O formato é adotado como base para o modelo de dados MADS/RDF, um vocabulário estabelecido como um sistema de organização do conhecimento (KOS) projetado em OWL, como uma ontologia mapeada em SKOS. Tem sido adotado na conversão de dados de autoridade estabelecidos pelo formato MARC autoridades em formato RDF, para uso em projetos de dados vinculados.

Neste ponto, verifica-se que o controle de autoridade tem passado por profundas transformações diretamente relacionadas à tecnologia adotada na implementação de padrões de estrutura e dos elementos de descrição, sobretudo diante de sua perspectiva de ampliação para outros ambientes e domínios de aplicação. Considera-se relevante contextualizar como o tema tem sido desenvolvido e abordado na literatura científica, apresentando-se, a seguir, uma análise bibliométrica sobre a temática em estudo.

## **2.1 Produção científica sobre controle de autoridade**

A literatura científica reflete as mudanças e as relações que deram origem ao conhecimento, permitindo entender a evolução do conhecimento e as relações estabelecidas pela comunidade científica ao longo dos anos.

Neste contexto, a comunicação científica passa a ser fundamental para a transferência e o compartilhamento do conhecimento gerado pela e entre a comunidade científica. Para isso, são estabelecidos os canais de comunicação formais, por meio dos quais é composta a literatura científica, que abrange documentos publicados, como livros, artigos, entre outros, e os informais, resultados de pesquisa de mestrado, doutorado e eventos científicos.

A elaboração de indicadores científicos tornou-se ao longo dos anos um requisito fundamental para o monitoramento das atividades envolvidas no

desenvolvimento científico entre as áreas da Ciência. O “[...] reconhecimento de que a atividade científica pode ser recuperada, estudada e avaliada a partir de sua literatura sustenta a base teórica para a aplicação de métodos que visam à construção de indicadores de produção e de desempenho científico” (SILVA; HAYASHI; HAYASHI, 2011, p.111).

Os indicadores são estabelecidos a partir dos canais formais de comunicação, que compõem a literatura científica. Rostaing (1996 *apud* SILVA; HAYASHI; HAYASHI, 2011, p.112) afirma que “[...] um escrito científico é produto objetivo da atividade de um pensamento, o que significa no contexto científico que a publicação é uma representação da atividade de pesquisa de seu autor [...]”. Desse modo, concorda-se com Guedes e Borshiver (2006), ou seja, por meio da avaliação de pesquisadores, instituições, entidades geográficas, entre outros, é possível caracterizar o grau de desenvolvimento estabelecido entre as áreas do conhecimento, para fins da gestão da informação, do conhecimento e no planejamento científico e tecnológico.

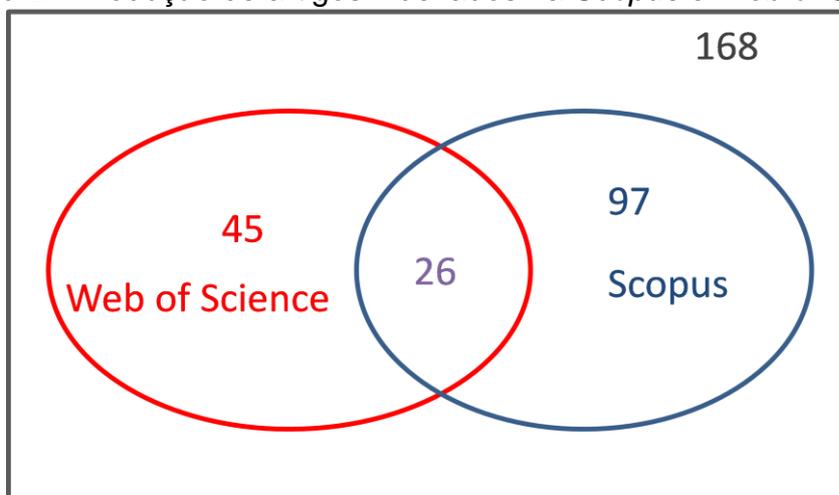
Segundo Okubo (1997), os indicadores podem ser subdivididos em três categorias: 1) indicadores de produção; 2) indicadores de citação; 3) indicadores de ligação. Lima, Velho e Faria (2007) afirmam que os indicadores de produção apresentam listas de frequência de publicação, de modo que são frequentemente adotados na elaboração de *rankings* para pesquisadores, instituições, países, entre outros. Grácio e Oliveira (2011, p.252), por sua vez, afirmam que os indicadores de produção têm por objetivo "refletir seu impacto junto à comunidade científica a qual pertencem, dando visibilidade àqueles mais produtivos, bem como às temáticas mais destacadas de uma área do conhecimento".

Com o intuito de identificar os estudos realizados sobre a temática da pesquisa, apresentam-se a seguir os resultados de uma análise bibliométrica que permitiu apontar aspectos relacionados à evolução do conhecimento e abrangência dos estudos sobre controle de autoridade indexados nas bases de dados *Scopus* e *Web of Science*, por meio das quais são apresentados indicadores de produção, destacando a abrangência das bases, a produção anual por países e os periódicos de maior representatividade.

Foram recuperados 194 artigos: 71 artigos na *Web of Science* e 123 artigos na *Scopus*. Na análise sobre os campos título, autor, periódico e data de publicação dos artigos indexados, identificaram-se 26 artigos indexados por ambas as bases de

dados, conforme apresentado na figura 2, configurando um total de 168 artigos distintos nas bases.

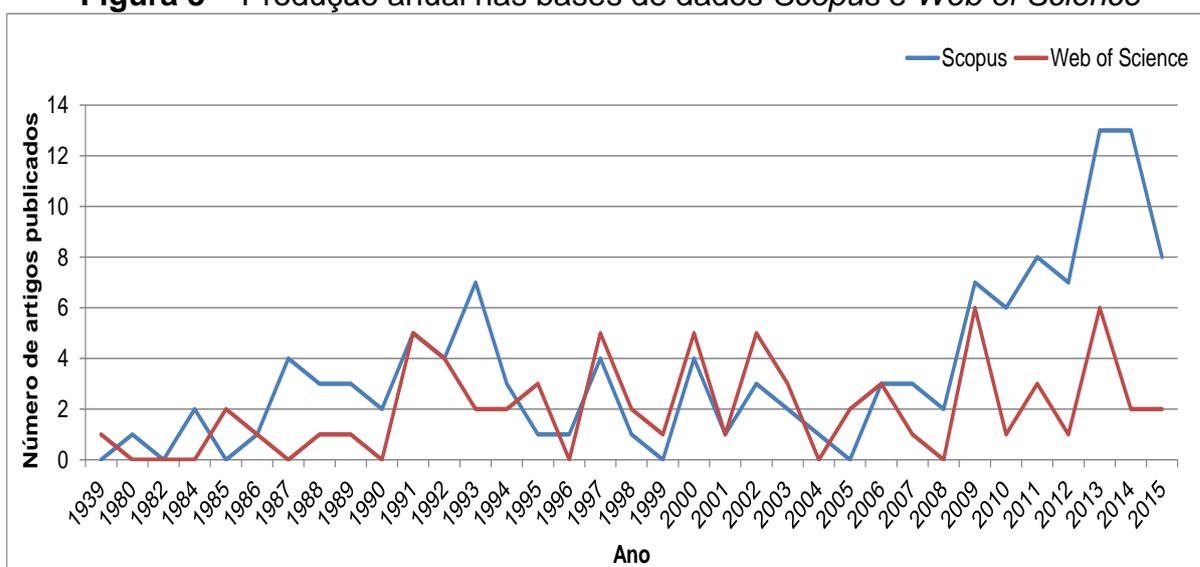
**Figura 2** - Produção de artigos indexados na *Scopus* e *Web of Science*



Fonte: Elaborado pela autora a partir de dados coletados da Web of Science e Scopus.

A análise da produção de artigos por autores e instituições apontou que o desenvolvimento da temática vem sendo realizada de maneira dispersa, visto que muitos autores e instituições apresentam um número reduzido de publicações. Identificou-se que 218 autores apresentaram produção científica entre um e dois artigos, com artigos publicados individualmente ou em coautoria.

Na figura 3, apresenta-se a análise da produção científica anual, que indica um número reduzido de artigos indexados nas bases ano a ano.

**Figura 3 – Produção anual nas bases de dados Scopus e Web of Science**

Fonte: Elaborado pela autora a partir de dados coletados da Web of Science e Scopus.

Embora o primeiro artigo indexado na *Web of Science* tenha sido publicado em 1939, verifica-se que somente a partir da década de 1980 o tema começou a ser abordado com maior frequência ano a ano. Outro comportamento a ser destacado é o crescimento exponencial de artigos na *Scopus*, a partir do ano 2000.

Durante a década de 1980, os estudos tinham como principal abordagem a implementação do controle de autoridade nos catálogos em linha, com destaque para os serviços oferecidos pela *Library of Congress* (LC), *American Library Association* (ALA) e *Online Computer Library Center* (OCLC), bem como o padrão de intercâmbio de dados *Machine Readable Cataloging* (MARC) e código de catalogação *Anglo-American Cataloguing Rules* (AACR).

Na década de 1990, o tema tornou-se um pouco mais abrangente, visto que os estudos enfocavam a avaliação do controle de autoridade implementado pelas bibliotecas, a discussão sobre os procedimentos da catalogação e a catalogação cooperativa. Esse período marca o surgimento de estudos sobre entradas multilíngues para catálogos de acesso Online, também conhecidos por *Online Public Access Catalog* (OPAC).

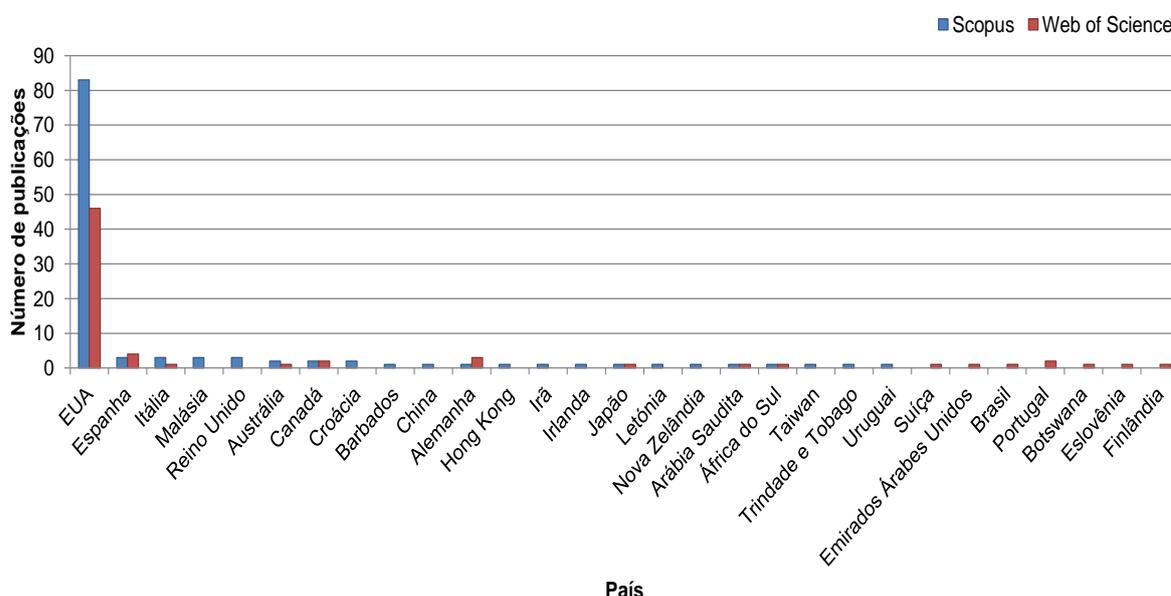
A década de 2000, período em que os padrões de metadados, os padrões de intercâmbio de dados e os códigos de catalogação eram amplamente adotados por bibliotecas, foi marcada pelo avanço de questões apresentadas anteriormente e pelo surgimento de estudos direcionados à interoperabilidade e à harmonização entre padrões de metadados, assim como aos modelos conceituais FRBR e ao FRAD,

que representam uma mudança de paradigmas da catalogação diante das tecnologias da informação e da comunicação (TIC).

Os artigos publicados entre os anos de 2010 e 2015 têm como característica os catálogos automatizados. Os estudos abordam os modelos conceituais FRBR e FRAD e as mudanças do *Anglo-American Cataloguing Rules*, segunda edição, (AACR2) para a diretriz *Resource Description and Access* (RDA). Os estudos apresentam, ainda, uma mudança nas tecnologias adotadas na estruturação de padrões de metadados e sua interoperabilidade, a partir de formatos como o *eXtensible Markup Language* (XML) e *Resource Description Framework* (RDF).

Na figura 4, apresenta-se a distribuição de publicações por país e observa-se que a literatura indexada nas bases *Scopus* e *Web of Science*, em sua maior parte, origina-se dos Estados Unidos da América (EUA): 69% na *Scopus* e 64% na *Web of Science*.

**Figura 4 – Produção por país entre *Scopus* e *Web of Science***

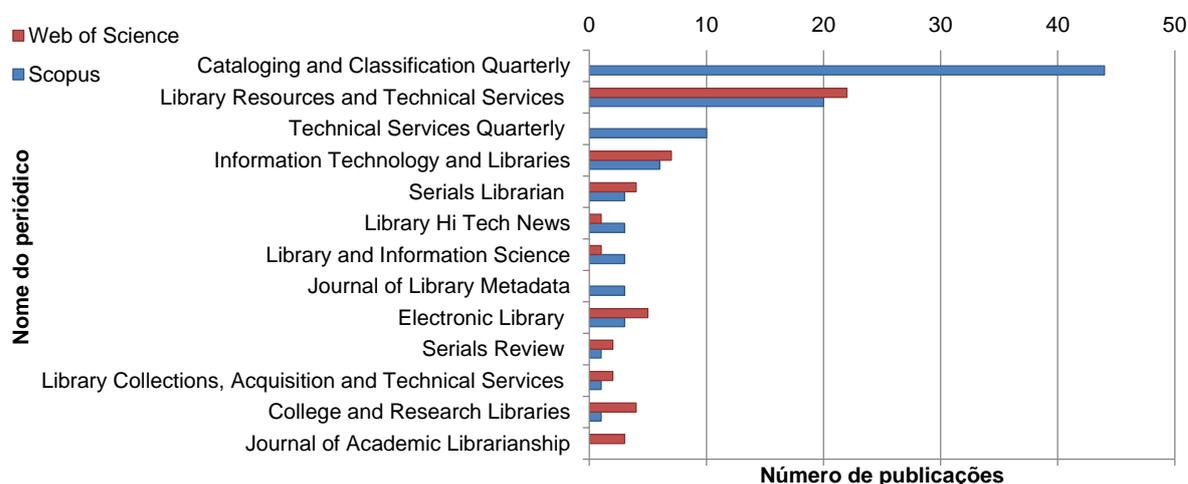


Fonte: Elaborado pela autora a partir de dados coletados da Web of Science e Scopus.

Existem diversos fatores que influenciam a internacionalização da ciência, dentre eles os fatores linguísticos. Este resultado pode estar relacionado ao número reduzido de artigos publicados pelos demais países, que em sua maior parte não adotam a língua inglesa como língua oficial.

A figura 5 apresenta os periódicos de maior produtividade relacionados ao tema nas bases.

**Figura 5 - Produtividade por periódico na Web of Science e Scopus**



Fonte: Elaborado pela autora a partir de dados coletados da Web of Science e Scopus.

Os dados apresentados na figura 5 indicam que os periódicos de maior representatividade na área são indexados na base *Scopus*. Os números confirmam a representatividade e relevância do periódico *Cataloging and Classification Quarterly* com a maior parcela de artigos publicados, seguido dos periódicos *Library Resources and Technical Services* e *Technical Services Quarterly*, entre os de maior produtividade. Considera-se, também, que a *Scopus* é mais abrangente para a busca de artigos na temática se comparada à *Web of Science*.

Por meio dos indicadores apresentados, verifica-se que as transformações relacionadas ao tema controle de autoridade foram ocorrendo gradualmente ao longo dos últimos 35 anos.

Considera-se que seria ideal se as bases analisadas fossem mais abrangentes, cobrindo um maior número de artigos originários de outros países, o que permitiria identificar o grau de desenvolvimento do tema por país e a adesão de produtos e serviços como padrões de metadados, códigos de catalogação, serviços internacionais de cooperação, entre outros. No cenário internacional, a baixa participação de países como o Brasil desperta a preocupação sobre como tem sido a atuação das instituições bibliotecárias, indicando a necessidade de um maior posicionamento por meio do desenvolvimento de estudos e publicação dos resultados.

## **2.2 A representação de autoridade no tratamento descritivo da informação**

Desde os primeiros códigos de catalogação, são adotados critérios na definição de pontos de acesso. Anthony Genesio Maria Panizzi, reconhecido como um dos percursores da catalogação com as 91 regras para a compilação de catálogos, definiu a descrição de título, autor e assunto (LEHNUS, 1972). Contudo, ao longo dos anos, surgiram novas tecnologias e necessidades que provocaram mudanças na representação da informação com o intuito de promover maior eficácia e rapidez na descrição, recuperação e disseminação da informação (SANTOS; PEREIRA, 2014).

Na década de 1960, ocorreram grandes mudanças na catalogação, quando instituições norte-americanas passaram a atuar de modo decisivo no desenvolvimento de técnicas e de ferramentas que culminaram no estabelecimento de padrões e códigos de catalogação assim como são conhecidos na atualidade.

No ano de 1961, na Conferência Internacional sobre Princípios de Catalogação, realizada em Paris, foram apresentados os resultados de um estudo da *International Federation of Library Associations and Institutions* (IFLA) sobre a teoria da catalogação e a normalização, de modo que estabelecesse sua prática em nível internacional, o que foi acordado e tornou-se amplamente conhecido por Princípios de Paris (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2009a).

Posteriormente, em 1969, durante a Reunião Internacional de Peritos em Catalogação, realizada em Copenhague, foi estabelecida uma resolução em prol dos padrões internacionais de descrição de forma e conteúdo, que culminou em 1971 na publicação da *International Standard Bibliographic Description for Monographic Publications* (ISBD), pela IFLA (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2009a).

Ainda, durante a década de 1960, foi desenvolvido e implementado o formato de intercâmbio de dados *Machine Readable Cataloging* (MARC), pela Library of Congress que, conforme contextualizado por Santos e Pereira (2014, p. 126), possibilitou “A cooperação e o compartilhamento de informações catalográficas e bibliográficas entre diversas instituições no ambiente automatizado [...]”.

Por meio da catalogação, elaborou-se o registro bibliográfico, que, de acordo com Gorman (2004), é composto por três partes: 1) ponto de acesso; 2) descrição

bibliográfica; 3) localização ou o próprio documento. A base da catalogação consiste na adoção de pontos de acesso com o intuito de permitir ao usuário localizar durante o processo de busca o registro, cuja descrição lhe permite decidir se o item desejado é aquele requerido, e a localização leva o usuário para o documento desejado (GORMAN, 2004).

Os pontos de acesso para indicação de responsabilidade foram definidos desde os primeiros códigos de catalogação. Nas 91 regras de Panizzi, definiu-se que as entradas para autor deveriam ser adotadas como ponto de acesso, no qual o nome do primeiro autor era considerado como ponto de acesso principal, e, em caso de múltiplos autores, os mesmos eram indicados como pontos de acesso secundários. Para isso, eram criadas entradas monolíngues seguindo como critérios para a descrição: 1) conforme fonte de referência; 2) a forma em que o autor é identificado em suas obras; 3) o último nome utilizado pelo autor (LEHNUS, 1972). Naquele momento, apenas uma forma de descrição era permitida.

Embora tenham sido estabelecidos os pontos de acesso por autoria, Panizzi não apresentou uma definição sobre autoria, dando margem para inúmeros questionamentos sobre o que deveria ser considerado na indicação de autoria (WAJENBERG, 1989). Svenonius (1989) esclarece que a descrição dos pontos de acesso de autoria estabelecida no tratamento descritivo da informação não é realizada tendo como base o conceito de autoria, embora reconheça que o conceito seja fundamental para a teoria da catalogação.

Quanto aos pontos de acesso por entidade coletiva, estes passaram a ser considerados a partir de 1853, quando Charles Coffin Jewet publica o código "*On the Construction of Catalogues of Libraries, and their Publication by Means of Separate Stereotyped Titles: With Rules and Examples*" (LEHNUS, 1972).

A primeira diretriz sobre a elaboração de pontos de acesso para indicação de responsabilidade foi publicada pela IFLA, em 1963, intitulada por *Name of persons* (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 1996). Este documento foi atualizado nos anos de 1967, 1977, 1980 e 1996.

O código AACR2, parte II, apresenta a definição de pontos de acesso, títulos, uniformes e remissivas. No código (2004, p.21-6), a indicação de responsabilidade é realizada identificando "[...] a pessoa fundamentalmente responsável pela criação do conteúdo intelectual ou artístico de uma obra". Os cabeçalhos para pessoas e os

critérios que devem ser seguidos são apresentados no capítulo 22, tais como a criação do ponto de acesso principal para o responsável principal ou primeiro responsável; em casos de múltiplos responsáveis, a criação de pontos de acesso secundários para até três responsáveis; e para casos de quatro ou mais responsáveis, o ponto de acesso principal é estabelecido pelo título (CÓDIGO..., 2004).

Compreende-se, portanto, que a função de autoria pode ser entendida como umas das funções passíveis de indicação de responsabilidade sobre o conteúdo intelectual ou artístico da obra tendo como base códigos, padrões e diretrizes.

Ao contrário do código AACR2, a diretriz *Resource Description and Access* (RDA) não apresenta restrições sobre a quantidade de pontos de acesso para indicação de responsabilidade e passa a considerar inclusive as relações entre as entidades *Obra, Pessoa, Família e Entidade Coletiva*, previstas nos modelos conceituais FRBR e FRAD (AMERICAN LIBRARY ASSOCIATION, 2013).

O novo código de catalogação *Resource Description and Access* (RDA), que substitui o *Anglo-American Cataloguing Rules* (AACR2), fornece instruções sobre como registrar dados a partir do modelo conceitual *Functional Requirements for Authority Data* (FRAD) (SALGADO; SILVA, 2013), sendo projetado para

[...] criar um quadro de referência claramente definido, estruturado, de modo a relacionar os dados dos registros de autoridade às necessidades dos usuários de tais dados; assim como auxiliar na avaliação do potencial de compartilhamento internacional de dados de autoridade, não restrito ao universo bibliográfico (MEY; SILVEIRA 2009, p.38).

Os registros de autoridade são estabelecidos por meio do controle de autoridade (forma padronizada na descrição de pontos de acesso autorizados) e nas relações entre diferentes pontos de acesso, a fim de realizar a unificação de pontos de acesso e a devida associação desses com seus respectivos termos variantes (HERRERO PASCUAL, 1999). Os pontos de acesso podem ser classificados como não controlados e controlados, apresentados no quadro 6.

**Quadro 6** - Tipologia dos pontos de acesso

Pontos de acesso não controlados	Pontos de acesso controlados
<ul style="list-style-type: none"> <li>• Títulos próprios</li> <li>• Palavras-chave</li> </ul>	<ul style="list-style-type: none"> <li>• Nomes de pessoas</li> <li>• Nomes de entidades</li> <li>• Nomes geográficos</li> <li>• Títulos uniformes</li> <li>• Série</li> <li>• Materiais</li> </ul>

Fonte: Adaptado de Sandín Enriquez (2011, p.8).

Uma das questões fundamentais para o controle bibliográfico nacional e internacional é o controle de autoridade, que estabelece os pontos de acesso controlados, a identificação de nomes de pessoas, nomes de entidades, nomes geográficos, títulos uniformes, série e materiais (SANDÍN ENRIQUEZ, 2011). De acordo com a Declaração de Princípios (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2009a), registros de autoridade são adotados a fim de permitir o controle de formas autorizadas e variações na identificação de pontos de acesso. Formas autorizadas devem coexistir por meio de relacionamentos que estabeleçam uniformidade ao acesso, e a representação de pontos de acesso deve considerar variações linguísticas e culturais (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2004).

Para dar suporte à representação de nomes, foram criados os elementos e padrões de metadados, que fornecem estrutura de descrição e de intercâmbio de dados aos registros. Na representação de dados de autoridade, adotam-se formatos de metadados ricos, que, segundo Alves (2005, p.140), "[...] proporcionam uma representação completa e detalhada de um recurso informacional".

Embora o termo metadado tenha tido maior visibilidade no contexto da Web (BREITMAN, 2006), os profissionais da informação adotam o termo como referência aos dados que compõem registros elaborados durante a catalogação e indexação, a fim de organizar e tornar os registros do conhecimento acessíveis (BACA, 1999).

Os padrões de metadados especificam os elementos para a descrição de recursos, e são definidos por Alves (2010, p.47) como “[...] estruturas de descrição constituídas por um conjunto predeterminado de metadados (atributos codificados ou identificadores de uma entidade) metodologicamente construídos e padronizados”. Segundo Willer e Dunsiree (2013), eles fornecem consistência à estrutura, pois determinam quais aspectos de um recurso devem ser descritos e

como os dados devem ser armazenados, enquanto as regras de catalogação oferecem consistência e coerência ao conteúdo dos metadados.

A partir dos padrões de metadados é estabelecida a estrutura do registro, o que pode envolver relações hierárquicas, associativas e de equivalência entre elementos de metadados, que serão descritas a seguir.

### **2.3 Relações hierárquicas, associativas e de equivalência no controle de autoridade**

A Organização do Conhecimento pode ser entendida como uma disciplina científica e aplicada, que tem por objetivo melhorar a circulação da informação, contribuindo para a produção de novos conhecimentos e facilitando o acesso aos já existentes (GARCIA MARCO, 1995). Para isso, são estabelecidos os Sistemas de Organização do Conhecimento (KOS), tais como linguagens documentárias, tesouros, taxonomias, ontologias, entre outros. Segundo Moreiro Gonzalez (2011), esses sistemas são adotados na indexação de recursos nos ambientes da Web com uso de diferentes vocabulários, estando diretamente relacionados ao tratamento temático da informação, que, de acordo com Guimarães (2008), compreende três correntes teóricas: catalogação de assunto, indexação e análise documental.

Ao longo dos anos, o desenvolvimento de tesouros tem sido adotado pela área de estudos Organização do Conhecimento como Sistema de Organização do Conhecimento (KOS), cuja sigla remete ao termo em língua inglesa *Knowledge Organization Systems*. Para Tristão, Fachin e Alarcon (2004, p. 161), os KOS “incluem a variedade de esquemas que organizam, gerenciam e recuperam a informação”.

Os tesouros são entendidos como “[...] sistemas artificiais de signos normalizados que permitem representação mais fácil e efetiva do conteúdo documental, com o objetivo de recuperar manual ou automaticamente a informação que o usuário solicita” (TRISTÃO; FACHIN; ALARCON, 2004, p.162). Ideia complementada pela afirmação de Tálamo, Lara e Kobashi (1992, p.198) de que tesouro é um

[...] objeto cultural que registra e representa o conhecimento segundo parâmetros estáveis e previamente determinados. Tais parâmetros, que se materializam sob a forma de redes de relações entre descritores, determinam não só o modo de organização e disseminação da informação, mas também a análise dos textos.

Os tesouros podem ser considerados vocabulários controlados, estabelecidos a partir de uma estrutura rígida que estabelece a ligação entre termos e conceitos. Sales e Café (2009, p.108-109) consideram que "tesouros atribui aos termos a função de evitar ou diminuir a flexibilidade da linguagem e descrever um conceito de maneira unívoca em um sistema de informação".

Dentro de sua estrutura são estabelecidas relações hierárquicas de gênero-espécie e todo-parte, e associativas entre os termos-descritores relacionados. Para as relações associativas, são possíveis três tipos de relações, descritas por Sales e Café (2009, p.109):

- a) Equivalência - quando um termo apresenta uma relação de sinonímia com outro, e, neste caso, o termo adotado pelo tesouro (termo preferido), também conhecido como descritor, é determinado na elaboração do tesouros lançando mão da sigla UP (Usado Para). O termo preterido é marcado pela sigla USE (que remete para o descritor correspondente).
- b) Associativa - apresenta relação semântica não hierárquica.
- c) Nota Explicativa - orientação que elucida o emprego de determinado termo, fornecendo informações como a definição do termo e sua relação com outros termos.

Ao estabelecer relações entre conceitos-termos, os tesouros atuam como importante instrumento na recuperação da informação, principalmente devido a sua interoperabilidade com outros vocabulários. A partir da norma ISO 25964 Informação e documentação – *Tesouro e interoperabilidade com outros vocabulários, Parte 1: Tesouro para recuperação da informação e Parte 2: Interoperabilidade com outros vocabulários*, eles fornecem recomendações para promover a interoperabilidade com modelos de dados e esquemas de metadados, entre outras.

A ISO 25964 foi elaborada pelo Comitê Técnico ISO / TC 46, informação e documentação, Subcomitê SC 9, identificação e descrição, da *International Organization for Standardization (ISO)*, instituição internacional responsável pela elaboração de normas técnicas. A norma cancela e substitui a ISO 2788/1986, que previa o desenvolvimento de tesouros monolíngues, e a ISO 5964/1985 para

tesauros multilíngues (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, p.vii).

A primeira parte da norma abrange o conteúdo para a construção de tesauros, e a recuperação da informação é dividida em 18 capítulos, apresentados no quadro 7. Os capítulos 1 a 13 abrangem o conteúdo contido nas normas ISO 5964/1985 e ISO 2788/1986, que foram substituídas.

**Quadro 7 – ISO 25964 parte 1**

<b>Capítulo</b>	<b>Descrição da estrutura da norma</b>
1	Escopo
2	Termos e definições
3	Simbolos, abreviações de termos e outras convenções
4	Visão geral e os objetivos dos tesauros
5	Conceitos e o escopo dos tesauros
6	Termos do tesouro
7	Conceitos complexos
8	Relações de equivalência no contexto monolíngue
9	Equivalência entre idiomas
10	Relações entre conceitos
11	Análise de faceta
12	Apresentação e layout
13	Gerenciamento de construção e manutenção do dicionário de sinônimos
14	Diretrizes para software de gerenciamento de tesouro
15	Modelos de dados
16	Integração entre tesauros com aplicações
17	Formatos de troca
18	Protocolos
Anexo A	Exemplos de displays encontrados em tesauros publicados
Anexo B	Esquema XML para troca de dados

Fonte: Adaptado de *International Organization for Standardization* (2011, p.iii-iv).

A partir do capítulo 14 da primeira parte e a segunda parte da norma, há informações para orientar o desenvolvimento e a manutenção de tesouros em ambiente Web.

A segunda parte da norma, que se refere à interoperabilidade com outros vocabulários, é dividida em 24 capítulos e um anexo, apresentados no quadro 8. A norma descreve, compara e contrasta os elementos e características entre esquemas de classificação, taxonomias, esquemas de cabeçalhos de assuntos e listas de autoridade de nomes.

**Quadro 8 – Conteúdo ISO 25964 parte 2**

<b>Capítulo</b>	<b>Descrição da estrutura da norma</b>
1	Escopo
2	Referências normativas
3	Definições
4	Simbolos, abreviações e outras convenções
5	Objetivos e identificação
6	Modelo estrutural por meio do mapeamento entre vocabulários
7	Tipos de mapeamento
8	Equivalência de mapeamentos
9	Mapeamento hierárquico
10	Mapeamento associativo
11	Mapeamento exato, inexato e parcial
12	Uso de mapeamento na recuperação da informação
13	Manipulação de pré-coordenação
14	Gerenciamento e troca de dados
15	Técnicas para identificar mapeamentos
16	Exibição de vocabulários mapeados
17	Esquemas de classificação
18	Sistemas de classificação utilizados para a gestão de registros
19	Taxonomias
20	Esquemas de rotulação de assuntos

Continua.

Continuação.

21	Ontologias
22	Terminologias
23	Listas de autoridade para nomes
24	Anéis de sinônimos
Anexo A	Gestão de dados terminológicos em apoio à interoperabilidade

Fonte: Adaptado de *International Organization for Standardization* (2011, p.iii-vi).

O mapeamento elementos descritivos, códigos e relacionamentos é estabelecido por meio de símbolos e abreviaturas. Os elementos descritivos são apresentados no quadro 9, os códigos no quadro 10, e os relacionamentos no quadro 11.

**Quadro 9 - Elementos descritivos da norma ISO 26964**

<b>Significado</b>	<b>TAG</b>
Nota de escopo	SN
Definição	DEF
Nota histórica	HN

Fonte: Adaptado de *International Organization for Standardization* (2011, p.13).

**Quadro 10 - Código da norma ISO 26964**

<b>Significado</b>	<b>TAG</b>
Categoria de assunto	SC
Código de conceito ou notação	CC

Fonte: Adaptado de *International Organization for Standardization* (2011, p. 13).

**Quadro 11 - Relações da norma ISO 26964**

<b>Significado</b>	<b>TAG</b>
Use	USE
"Usado por" ou "Usado para"	UF
Dois ou mais termos preferenciais seguintes às etiquetas devem ser usados em conjunto para representar o conceito indicado.	USE+...
Os termos não preferidos seguintes às tags devem ser representados por uma combinação de termos preferenciais, incluindo o termo preferido que precede a tag.	UF+

Continua.

Continuação.

Termo topo	TT
Termo amplo	BT
Termo amplo (genérico)	BTG
Termo amplo (instância)	BTI
Termo amplo (partitivo)	BTP
Termo restrito	NT
Termo restrito (genérico)	NTG
Termo restrito (instância)	NTI
Termo restrito (partitivo)	NTP
Termo relacionado	RT

Fonte: Adaptado de *International Organization for Standardization* (2011, p. 13).

Os princípios estabelecidos na ISO 25964 permitem definir relações de equivalência, hierárquicas e associativas entre conceitos de diferentes vocabulários. Para o mapeamento de equivalência é possível a aplicação de marcadores que representam o nível de relacionamento exato "=" e inexato "~" entre termos; as relações hierárquicas são possíveis apenas para mapeamento exato; e as relações associativas não permitem a designação de equivalência por abranger diferentes tipos de relacionamento.

O mapeamento de arquivos de autoridade estabelecido pela ISO 25964 tem por objetivos: 1) assegurar a consistência de entradas em catálogos, índices ou bases de dados; 2) de navegação e busca selecionando termos durante a recuperação de informação; 3) extração automática de nomes de entidades (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011).

As listas de autoridade oferecem pontos de acesso para a recuperação por meio de denominações conhecidas de uma entidade particular, podendo incluir siglas, pseudônimos, diferentes formas de tradução ou transliteradas, formas invertidas, formulários, entre outras formas. De modo que podem ser elaboradas como vocabulários estruturados e não estruturados.

A relação de equivalência é adotada para relacionar entradas autorizadas entre o nome preferido, e suas variações alternativas podem ser indicadas por meio da etiqueta USE/UF. A relação associativa de vínculo é indicada pela etiqueta RT de "Ver também". Além das três relações principais, pode ocorrer, ainda, a relação cronológica, quando ocorrem mudanças ao longo do tempo. Por exemplo: mudanças de nomes para pessoa e razão social para entidade coletiva; e relações compostas, para quando uma entidade é dividida em duas ou mais e mantém o vínculo.

As relações hierárquicas são exemplificadas por meio do registro de autoridade estruturado, apresentado no quadro 12.

**Quadro 12 – Registro de autoridade estruturado por tesouro**

<b>Brasil</b>	
<b>Usado por:</b>	República Federativa del Brasil
<b>Término genérico:</b>	Países lusófonos
<b>Término específico:</b>	Acre (Brasil, Estado) Alagoas Amapá (Brasil) Amazonas (Brasil, Estado) Bahía (Estado) Brasilia Ceará (Estado) Espírito Santo (Estado) Goiás (Estado) Maranhão Mato Grosso Mato Grosso do Sul Minas Gerais Pará (Estado) Paraíba (Estado) Paraná (Brasil, Estado) Pernambuco Piauí Río de Janeiro (Estado) Rio Grande do Norte Rio Grande do Sul Roraima Santa Catarina (Estado) São Paulo (Estado) Sergipe
<b>Término relacionado:</b>	Brasileños
<b>Fuentes:</b>	Gran Atlas Aguilar, 1969; (Brasil) LCSH; (Brazil) RAMEAU; (Brésil) Otro identificador normalizado: <a href="http://id.loc.gov/authorities/names/n79128015">http://id.loc.gov/authorities/names/n79128015</a> lcnaf

Continua.

Continuação.

<b>Registro en datos.bne.es:</b>	<a href="http://datos.bne.es/resource/XX4575422">http://datos.bne.es/resource/XX4575422</a>
<b>Nº Registro:</b>	XX4575422

Fonte: Biblioteca Nacional da Espanha (2015, não paginado).

O registro no quadro apresenta o exemplo de como são estabelecidas as relações hierárquicas, associativas e de equivalência por tesouros no registro de autoridade elaborado pela biblioteca nacional da Espanha. Embora tenha sido atualizado em 2015, foi elaborado a partir de informações de 1969, o que indica para a necessidade de atualização dos registros de autoridade. No quadro é possível identificar a ausência de informações sobre os Estados de Tocantins e Rondonia, assim como o Distrito Federal, que é representado por sua capital Brasília.

As contribuições da norma ISO 25964/2011 permeiam uma interligação entre os sistemas de organização estabelecidos por tesouros e os de rotulação por metadados. A interoperabilidade entre classes de aplicações depende de que os dados estejam estabelecidos por um modelo comum no qual

Sistemas de organização do conhecimento, tais como tesouros e esquemas de classificação são aplicadas em uma ampla gama de situações, e um sistema de organização do conhecimento individual pode ser usado em muitos sistemas de informação diferentes. Ao definir o modelo de dados SKOS como uma ontologia OWL completa, a Web Semântica pode então ser usado como um meio para a publicação, trocar, partilhar e ligação de dados envolvendo estes sistemas de organização do conhecimento (MILES; BECHHOFER, 2009, não paginado, tradução nossa).

O Simple Knowledge Organization System (SKOS) pode ser considerado um padrão de compartilhamento de dados, que estabelece por meio de rótulos o mapeamento de e entre sistemas de organização do conhecimento (KOS) como tesouros, taxonomias, esquemas de classificação, por apresentarem alguns níveis de semelhança. Esses sistemas de organização do conhecimento possuem características próprias na disposição dos dados. Assim, para que o conhecimento estabelecido por eles possa ser legível por máquina, torna-se necessário transformar sua estrutura e seu conteúdo em um modelo de dados formal, também denominado ontologia.

SKOS estabelece rótulos por meio de 32 classes apresentadas no quadro 13, e cada classe possui propriedades próprias.

**Quadro 13** - Vocabulário de classes SKOS

skos:Collection
skos:Concept
skos:ConceptScheme
skos:OrderedCollection
skos:altLabel
skos:broadMatch
skos:broader
skos:broaderTransitive
skos:changeNote
skos:closeMatch
skos:definition
skos:editorialNote
skos:exactMatch
skos:example
skos:hasTopConcept
skos:hiddenLabel
skos:historyNote
skos:inScheme
skos:mappingRelation
skos:member
skos:memberList
skos:narrowMatch
skos:narrower
skos:narrowerTransitive
skos:notation
skos:note
skos:prefLabel
skos:related
skos:relatedMatch
skos:scopeNote
skos:semanticRelation
skos:topConceptO

Fonte: Miles e Bechofer (2009, não paginado).

As classes SKOS possuem correspondência com outros vocabulários, como a ontologia MADS/RDF, apresentada no quadro 14. O MADS/RDF deriva do esquema MADS em XML, desenvolvido pela Library of Congress na conversão do formato MARC autoridade em formato XML e RDF para a publicação de conjuntos de dados estruturados de arquivos de autoridade.

**Quadro 14** – Trecho da correspondência entre SKOS/XL e MADS/RDF

SKOS/SKOS-XL	MADS/RDF
skos:ConceptScheme	madsrdf:MADSScheme
skos:Concept	madsrdf:Authority
skos:inScheme	madsrdf:isMemberOfMADSScheme
skos:topConceptOf (or its inverse skos:hasTopConcept)	madsrdf:isTopMemberOfMADSScheme
inverse of skos:inScheme	(inverse of madsrdf:hasTopMemberOfMADSScheme)
skos:notation	madsrdf:hasMADSSchemeMember
skos-xl:Label	madsrdf:code
skos:altLabel	possible sub-class: mads:Variant
OR skos-xl:altLabel	madsrdf:hasVariant

Fonte: Adaptado de ISAAC (2013, não paginado).

O processo de conversão e publicação de dados vinculados permeia diversas etapas: primeiramente, na conversão dos registros em MARC para o formato MADS em XML, realizado por meio de extração; posteriormente, é realizado o mapeamento entre os elementos do MADS em XML para a ontologia MADS/RDF, que finalmente é mapeado em SKOS. Tal processo está relacionado com limitações na captura dos dados, tendo em vista que SKOS estabelece apenas a ligação entre os KOS na Web (*LIBRARY OF CONGRESS*, 2012).

Na vinculação de dados, são definidas correspondências entre as classes e ou propriedades SKOS para os elementos de metadados MADS, conforme ilustrado no quadro, em que o elemento *madsrdf:Authority* corresponde a um conceito que foi mapeado a um *skos:Concept*. Essa relação permite a vinculação do elemento com outros vocabulários de domínio que tenham sido mapeados ao conceito SKOS. O mesmo tipo de correspondência pode ser estabelecido para os elementos MARC autoridade com SKOS; e vocabulários definidos a partir de elementos MARC e MADS podem ser vinculados a partir desse mapeamento.

Essa relação permite que a conversão de registros em XML, inicialmente definidos em elementos MARC, seja feita com uso dos elementos MADS, assim como é possível mapear vocabulários já estruturados em RDF com a ontologia MADS/RDF. Isso vem sendo realizado por instituições bibliotecárias na conversão de registros que posteriormente são fornecidos ao *Virtual International Authority File*, que será apresentado no próximo capítulo.



vocabulários por meio do projeto *Linking Open Data*.

Entre os vocabulários integrados ao SKOS, os arquivos de autoridade são os que apresentam maior frequência, nos quais os vocabulários de maior volume são o *Virtual International Authority File* (VIAF) e a *Faceted Application of Subject Terminology* (FAST) derivada da *Library of Congress Subject Headings*, o que corresponde a 73% dos dados (LCSH) (PASTOR SANCHEZ; MARTINEZ MENDEZ; RODRIGUES MUNOZ, 2009). A principal diferença entre esses vocabulários: o VIAF tem sido construído por meio da cooperação internacional enquanto o FAST é um vocabulário elaborado pela Library of Congress.

No VIAF, as instituições fornecem diferentes tipos de dados de autoridade, descrição apresentada no quadro 15.

**Quadro 15** – Descrição dos dados fornecidos por unidade cooperante ao VIAF

<b>Nome</b>	<b>País</b>	<b>Descrição e quantidade de registros no VIAF</b>
Library of Congress/NACO	Estados Unidos da América	Registros de pessoas: 6.310.631 Eventos: 183.495 Entidades coletivas: 1.444.944 Nomes geográficos: 130.208 Títulos: 526.623 Total de registros de autoridade: 9.008.128
German National Library	Alemanha	Pessoas: 189.198 Nomes de pessoas: 4.569.069 Entidades coletivas: 1.161.175 Conferências: 581.274 Nomes geográficos: 290.172 Assunto: 201.145 Títulos de obras: 87.232 Títulos de obras musicais: 103.127
National library of France	França	Pessoas: 1.435.797 Entidades coletivas: 327.409 Nomes geográficos: 113.463 Títulos: 149.930 Registros de autoridade Rameau: 169.992 Registros de autoridade Dewey: 100.626
National Library of Sweden	Suécia	Informação não disponível

Continua.

Continuação.

National Library of Australia	Austrália	Pessoas: 793.045 Entidades coletivas: 193.648 Títulos: 93.359
National Library of Spain	Espanha	Pessoas: 369.736 Entidades coletivas: 62.893 Títulos: 57.818
National Library of Portugal	Portugal	Pessoas: 505.736 Entidades coletivas: 117.863 Nomes geográficos: 11.815 Nomes familiares: 512 Títulos: 3.253 Tópico de assunto: 65.546 Registros de autoridade UDC: 275.497
Central Institute for the Union Catalogue of Italian Libraries and for Bibliographic Information (ICCU)	Itália	Nomes de pessoais: 43.506 Nomes de entidades coletivas: 56 Registros bibliográficos: 43.506
National Library of the Czech Republic	República Checa	Pessoas: 519.046 Entidades coletivas: 122.805
National Library of Israel	Israel	Informação não disponível
Bibliotheca Alexandrina	Egito	Pessoas: 45.596 Entidades coletivas: 1.394 Reuniões: 548
Vatican Library	Vaticano	Pessoas: 293.437 Entidades coletivas: 56.273
Swiss National Library	Suíça	Informação não disponível
Library and Archives Canada	Canadá	Nomes pessoais: quantidade não disponível Nomes de entidades coletivas: quantidade não disponível Nomes geográficos: quantidade não disponível Títulos: quantidade não disponível
Getty Research Institute	Estados Unidos da America	Informação não disponível
NUKAT Center of the University of Warsaw Library	Polonia	Registros de autoridade para pessoas e entidades coletivas: 1.204.267 Registros bibliográficos: 2.537.935
National Széchényi Library	Hungria	Informação não disponível
Library Network of Western Switzerland	Suíça	Informação não disponível
Bibliographic Agency for Higher Education	França	Informação não disponível

Continua.

Continuação.

VLACC: Flemish Central Catalogue	Bélgica	Nomes pessoais: 6.244 Nomes de entidades coletivas: 1.935 Total de registros de autoridade: 6.274.964
National Library of Russia	Russia	Informação não disponível
National Library of the Netherlands	Holanda	Informação não disponível
National Library of Norway	Noruega	Informação não disponível
DBC (Danish Bibliographic Centre)	Dinamarca	Informação não disponível
National Diet Library	Japão	Pessoas: 793.540 Famílias: 2.192 Entidades coletivas: 185.815 Nomes geográficos: 27.574 Títulos uniformes: 4.020 Total de registros: 1.013.478
National Library Board Singapore	Singapura	Informação não disponível
National Library of Latvia	Letônia	Informação não disponível
National Library of Poland	Polônia	Registros de autoridade para pessoas e entidades coletivas: 1.157.032 Registros bibliográficos: 2.493.490
National Library of Catalonia	Espanha	Informação não disponível
Lebanese National Library	Líbano	Informação não disponível
Perseus Digital Library	Estados Unidos da América	Informação não disponível
Syriac Reference Portal	Estados Unidos da América	Informação não disponível
Wikipedia	Estados Unidos da América	VIAF adiciona URLs da Wikipédia aos clusters.
International Standard Name Identifier	Reino Unido	VIAF adiciona identificadores ISNI aos clusters.
National Library of Ireland	Irlanda	Informação não disponível
National and University Library in Zagreb	Croácia	Informação não disponível
National Central Library, Taiwan	Taiwan	Nomes pessoais: 3.122
National Library and Archives of Québec	Canadá	Informação não disponível
National Library of Korea	Coréia do Sul	Registro de autoridade de pessoas Registros bibliográficos relacionados a pessoas.

Continua.

Continuação.

National Library of Luxembourg	Luxemburgo	
National Library of Chile	China	Informação não disponível
National Library of Morocco	Marrocos	Informação não disponível

Fonte: Virtual International Authority File (2017, não paginado).

A partir dos dados apresentados, é possível identificar que as instituições contribuem com dados de autoridade específicos, esses registros têm sido estruturados por meio de elementos dos padrões MARC para dados de autoridade, MARC para dados bibliográficos e UNIMARC. Os registros são convertidos e estruturados de acordo com os princípios *Linked Data*, que foram definidos pelo W3C, apresentados a seguir.

### **3.1 Infraestrutura de dados em *Linked Data***

Os avanços tecnológicos obtidos nas últimas décadas proporcionaram uma série de facilidades na publicação de dados e recursos informacionais no ambiente Web e repercutiu no aumento expressivo de conteúdos criados e gerenciados por comunidades heterogêneas.

A diversidade de dados, recursos informacionais e de técnicas adotadas na publicação de dados favoreceu a identificação de falhas e inconsistências, que repercutiram em dificuldades e, em alguns casos, impediam a recuperação da informação; por isso a comunidade internacional procurou definir melhores práticas e desenvolver tecnologias a serem adotadas na publicação de dados no ambiente da Web.

Dentre essas comunidades, está o Consórcio *World Wide Web Consortium*, conhecido por W3C, que é uma comunidade internacional fundada em 1994 por Timothy John Berners-Le, conhecido por Tim Berners-Lee, no Massachusetts Institute of Technology (MIT) (WORLD WIDE WEB CONSORTIUM, 2005). Liderado por Tim Berners-Lee e Jeffrey Jaffe, o consórcio é formado por especialistas que atuam no desenvolvimento de padrões internacionais a serem adotados na publicação de dados no ambiente da Web.

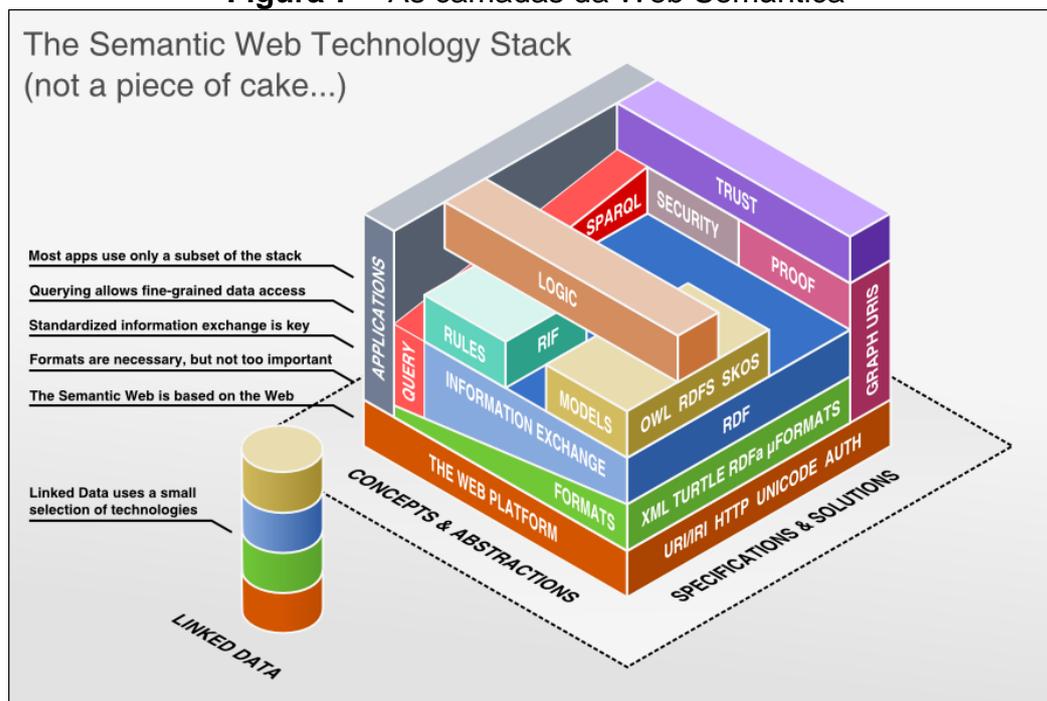
O conceito de *Linked Data* foi criado pelo W3C e consiste em um "conjunto de melhores práticas para a publicação de dados estruturados na Web" (WORLD WIDE WEB CONSORTIUM, 2014, não paginado, tradução nossa), com o intuito de contribuir na implementação das tecnologias da Web Semântica, sendo elas: 1) use URIs como nomes para as recursos; 2) use URIs HTTP como localização dessas recursos; 3) adote padrões na descrição e recuperação desses recursos, como RDF e SPARQL para melhorar a recuperação da informação; 4) estabeleça links entre os URIs para que outros recursos sejam recuperáveis.

A concretização da Web Semântica requer uma vasta quantidade de dados disponíveis em um formato padrão e relacionáveis (WORLD WIDE WEB CONSORTIUM, 2015), de modo que projetos de dados vinculados se situem em investigações em grande escala, o que tem resultado em grandes volumes de metadados e vocabulários controlados estruturados a partir dos princípios *Linked Data* do W3C e convertidos em RDF (VAN HOOLAND; VERBORGH, 2014).

### **3.2 Web Semântica e Linked Open Data**

A arquitetura da Web Semântica é composta por uma estrutura de sete camadas de aplicação (figura 7), dentre as quais as quatro primeiras camadas fornecem estrutura sintática e semântica, e as demais fornecem suporte nas relações, na execução e confiabilidade das camadas anteriores.

**Figura 7 – As camadas da Web Semântica**



Fonte: Hoppe (2015, não paginado).

A primeira camada – denominada característica internacional – é composta pelos elementos Unicode, um padrão ISO que permite o acesso de computadores em qualquer idioma (THE UNICODE CONSORTIUM, 2015); e URI, que identifica um recurso por meio da representação de seu mecanismo de acesso primário (WORLD WIDE WEB CONSORTIUM, 2001). O URI é um identificador único usado para referenciar diferentes entidades, sejam pessoas, instituições, locais, entre outras, fornecendo consistência aos dados.

Na segunda camada – denominada por sintática –, a linguagem XML + NS + xmlschema estabelece a estrutura de descrição do conteúdo que suporta a terceira camada de dados, o modelo formal de dados e sintaxe *Resource Description Framework* – RDF, que codifica metadados de forma legível por máquinas (BREITMAN, 2006).

O formato *Extensible Markup Language* (XML) foi desenvolvido em 1996 pelo Working Group do *World Wide Web Consortium* (W3C) e compreende um sistema gramatical elaborado para criar linguagens de marcação, que, de maneira similar ao *HyperText Markup Language* (HTML), foi projetado para armazenar e transportar dados (WORLD WIDE WEB CONSORTIUM, 2005). Derivado do *Standard Generalized Markup Language* (SGML), estabelecido pela Norma *ISO 8879:1986*

*Information processing--Text and office systems--Standard Generalized Markup Language.*

Um arquivo XML é formado por uma raiz e um conjunto de elementos, formando uma estrutura hierárquica por camadas. Para evitar ambiguidades e promover a identificação dos vocabulários de dados, são definidos *namespaces*, potencializando o reuso de dados (PETIYA, 2014).

O modelo de dados *Resource Description Framework* (RDF), por sua vez, compreende uma aplicação da linguagem XML, que proporciona representação de recursos informacionais no ambiente da Web por meio do processamento de metadados, com o objetivo de criar redes de informação a partir de dados distribuídos (LASSILA; SWICK, 1999; SANTAREM SEGUNDO, 2014). O RDF permite codificar, representar e transmitir metadados, favorecendo o processamento automático e promovendo a integração desses vocabulários em sistemas de informação disponíveis na Web (LASSILA; SWICK, 1999, 1999).

Desenvolvido durante a década de 1990 pelo *World Wide Web Consortium*, o RDF tem como principal influência as linguagens HTML, XML e SGML, metadados de catalogação, ontologias, programação orientada a objetos, entre outros (DZIEKANIAK; KIRINUS, 2004, *apud* FERREIRA; SANTOS, 2013).

A estrutura RDF é estabelecida por meio do modelo matemático conhecido por grafo, que é adotado em conjunto com o modelo de construção de triplas, compostas pelos elementos recurso, propriedade e valor (WORLD WIDE WEB CONSORTIUM, 2014; SANTAREM SEGUNDO, 2014). Um recurso é uma informação ou objeto que contém a informação, e a esse recurso pode ser atribuído um URI, o que possibilita distingui-lo de outros recursos. A propriedade compreende as características ou atributos que qualificam o recurso. O valor compreende pontos de acesso para o recurso e abrange dados descritivos (SANTAREM SEGUNDO, 2014).

A representação visual dos grafos RDF é descrita por meio de diagramas, favorecendo a aprendizagem e a leitura do modelo por humanos, enquanto para o processamento de máquinas a representação é realizada por meio de cadeias de caracteres (FERREIRA; SANTOS, 2013).

Na camada ontológica, são mapeados os relacionamentos possíveis entre entidade e suas interações, estabelecidos por meio de um vocabulário e de regras de inferência, que Campos (2007) denomina de níveis taxonômico e lógico. O nível

taxonômico compreende "[...] um conjunto comum de termos que são usados para descrever e representar um domínio [...]" (LIMA; CARVALHO, 2005, p.2); e o lógico provê maior expressividade na descrição de um domínio, estabelecendo restrições de propriedades, disjunção de classes, combinações entre classes, características de propriedades, entre outras (WAGNER FILHO; LÓSCIO, 2009).

Jacob (2003) considera que as ontologias promovem o entendimento comum e compartilhado do conhecimento entre pessoas e programas de aplicação, e isso se estabelece por meio da categorização e do relacionamento entre "coisas".

A estrutura complexa da Web Semântica tem como principal motivação o fato de ser diferente da Web atual, denominada Web 2.0, em que conteúdos são estruturados para ser exibidos e lidos por pessoas, e propõe uma maneira para que os conteúdos sejam estruturados de maneira semântica e possam ser "inteligíveis" por programas denominados "agentes de software" (MARCONDES, 2005). Santarém Segundo (2014, p.3.865) considera ser essa uma nova estrutura de armazenamento de dados: "O ponto principal está na separação da apresentação do conteúdo e do conteúdo da estrutura, tratando as unidades atômicas de uma informação como componentes independentes".

Berners-Lee, Hendler e Lassila (2001) definem Web Semântica como "uma extensão da Web atual, na qual à informação atribui-se um significado bem definido, permitindo que computadores e pessoas trabalhem em cooperação" (BERNERS-LEE; HENDLER; LASSILA, 2001, não paginado). Isso requer a "[...] padronização de tecnologias, de linguagens e de metadados descritivos, de forma que todos os usuários da Web obedeçam a determinadas regras comuns e compartilhadas sobre como armazenar dados e descrever a informação armazenada [...]" (SOUZA; ALVARENGA, 2003, p.134).

Neste contexto, a interoperabilidade entre metadados e ontologias passa a ser fundamental para o uso apropriado de vocabulários de valor na Web Semântica, pois fornece a definição de contextos e restringe o uso desses vocabulários (DUNSIRE, 2010). De acordo com Van Hooland e Verborgh (2014), a interoperabilidade é realizada por meio do desenvolvimento de modelos de abstração da realidade, estruturados por modelos de dados simples como os formatos RDF e OWL, que permitem a interpretação e o reuso dos dados.

Segundo Silva (2013, p.28),

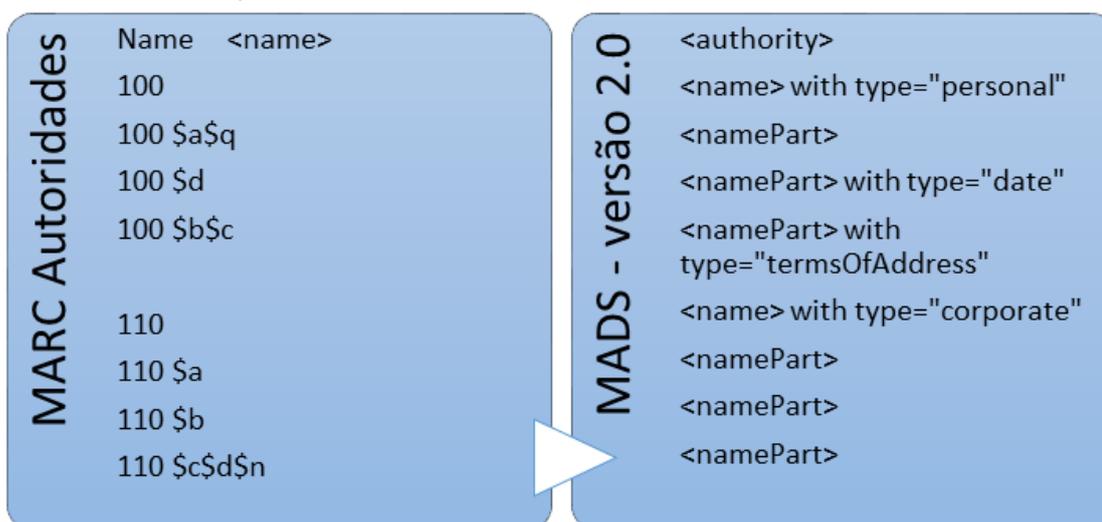
O emprego de tecnologias capazes de atribuir semântica aos recursos informacionais possibilita uma recuperação mais eficiente no contexto digital, permitindo, então, que as questões dirigidas a um sistema por meio de um mecanismo de busca obtenham respostas mais relevantes ao interesse de seu usuário, levando em conta não somente a agilidade da pesquisa e a quantidade dos recursos recuperados, mas sim a qualidade informacional desses recursos e possível maior relevância a quem necessita deles.

A estrutura da Web Semântica representa, ainda, uma maneira para que metadados autoridade, que foram elaborados por bibliotecas ao longo dos anos e estabelecidos a partir de códigos internacionais de catalogação, sejam integrados à Web Semântica e passem a contribuir para além das funções de catálogo. Isto oferece maior visibilidade aos produtos e serviços oferecidos por bibliotecas, favorece a descoberta e o acesso aos registros do conhecimento e aumenta a qualidade da descrição de conteúdos na Web.

Neste contexto, registros de autoridade elaborados no âmbito de bibliotecas e centros de documentação, caracterizados por seu alto grau de padronização, fornecem vocabulários de valor que garantem maior uniformidade na descrição de atributos dos registros disponibilizados na Web. Estes, quando estruturados em formato *Resource Description Framework* (RDF), podem agregar representações com diferenças culturais e linguísticas por meio da formação de *Clusters* entre elementos de metadados e dados, promovendo o relacionamento entre esses vocabulários e os registros do conhecimento. Isto tem sido motivado por elementos de metadados, e vocabulários de valor apresentaram grandes benefícios e um alto potencial na descoberta de recursos, reuso de dados, compartilhamento de dados e criação de novos conteúdos a partir de dados vinculados (SUBIRATS; ZENG, 2015).

Os padrões de metadados de maior adesão entre bibliotecas com o formato MARC não foram desenvolvidos, inicialmente, a partir dos formatos XML e RDF, gerando a necessidade de conversão de formatos. Dentre as formas de conversão adotadas por bibliotecas está o mapeamento de correspondência entre os elementos de metadados do formato de intercâmbio MARC autoridade para os do MADS, ilustrado no quadro 16.

**Quadro 16** - Mapeamento entre elementos do MARC Autoridade e MADS 2.0



Fonte: Adaptado de *Library of Congress* (2011, não paginado).

A correspondência apresentada permite que os elementos de metadados em MARC autoridades sejam vinculados aos elementos do MADS XML e as classes e propriedades da ontologia MADS/RDF. Para que os metadados possam contribuir na organização, é preciso que eles sejam compatíveis e mapeados a outros sistemas de organização do conhecimento, pois a integridade dos dados depende da definição formal estabelecida no modelo de dados. O modelo de dados *Simple Knowledge Organization System* (SKOS) compreende um modelo para representar sistemas de organização do conhecimento, usando o RDF e OWL.

*Simple Knowledge Organization System* (SKOS) é uma iniciativa do Consórcio *World Wide Web* sob a forma de uma aplicação RDF fornecendo um modelo para representar a estrutura básica e conteúdo dos esquemas conceituais como listas de cabeçalho de assunto, taxonomias, esquemas de classificação, tesouros e qualquer tipo de vocabulário controlado (PASTOR SANCHEZ; MARTINEZ MENDEZ; RODRÍGUEZ MUNOZ, 2009, não paginado, tradução nossa).

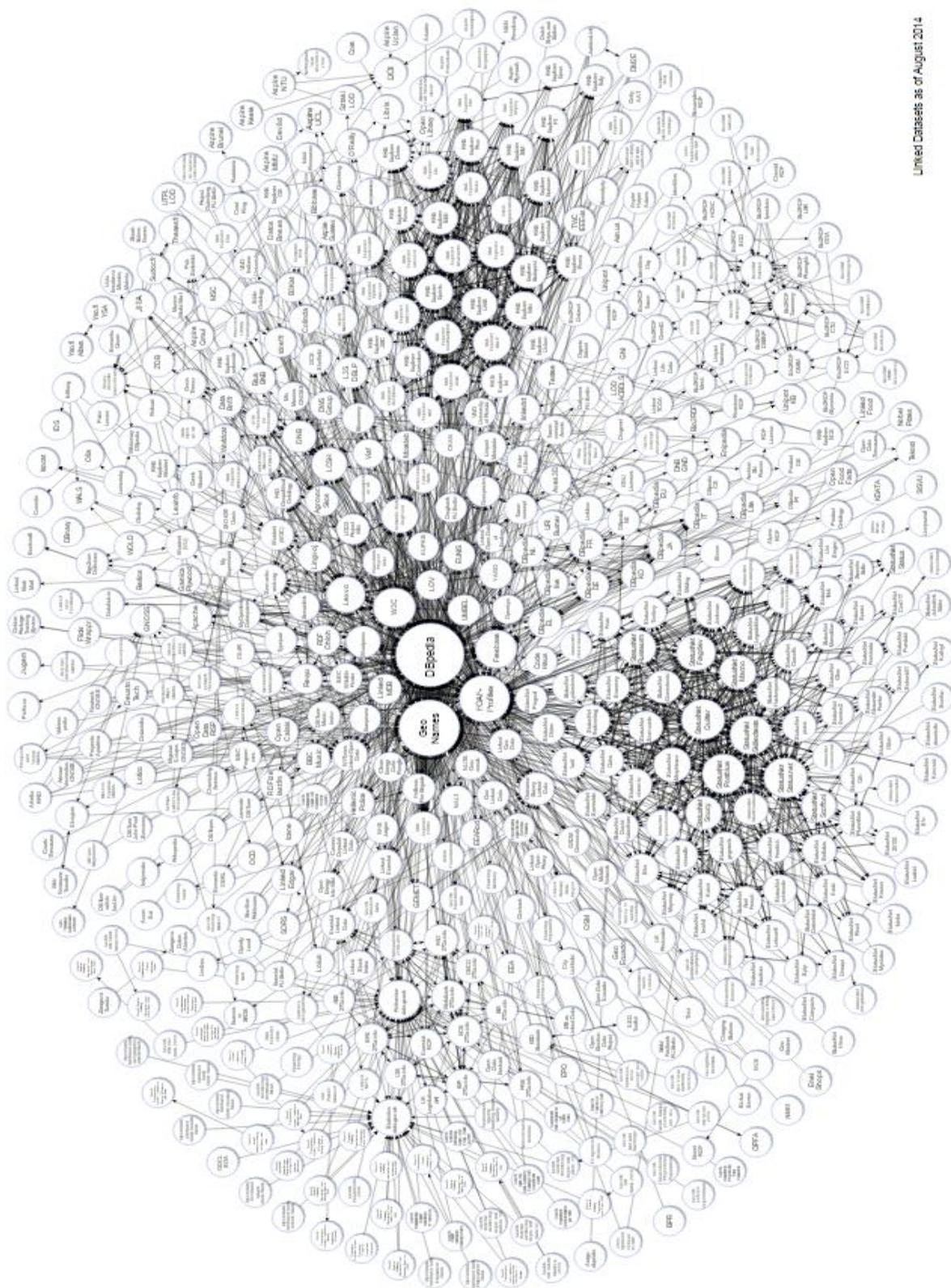
De acordo com as recomendações do *World Wide Web Consortium* (W3C) apresentadas por Miles e Bechhofer (2009), os conceitos SKOS podem ser mapeados para outros conceitos SKOS em diferentes esquemas conceituais, e permite quatro tipos de ligação: o mapeamento hierárquico, associativas, equivalentes aproximados e equivalente exato.

### **3.3 Publicação de dados de autoridade em *Linked Open Data***

*Linked Open Data* consiste na adoção dos princípios *Linked Data* para a publicação de dados abertos, no qual a Web Semântica tem sido adotada de duas formas: 1) para aplicações sofisticadas, com o intuito de permitir regras de inferência; 2) para a publicação e integração de dados estruturados, formando assim a teia de dados (BROOKS, 2008).

Daí a origem do projeto *Linking Open Data* (LOD), fundado em 2007, que tem por objetivo a publicação de dados abertos vinculados. Desde então, tem crescido de maneira expressiva, somando, em setembro de 2011, cerca de 31 bilhões de triplas em RDF e 504 milhões de ligações em RDF (BAUER; KALTENBOCK, 2011). A figura 8 apresenta o diagrama LOD.

Figura 8 – Diagrama *Linking Open Data*



Fonte: Cyganiak e Jentzsch (2014, não paginado).

Bauer e Kaltenbock (2011) citam os passos a serem seguidos para a publicação na nuvem LOD, recomendações que são baseadas nos quatro princípios da W3C:

- i. Escolha vocabulários estabelecidos passíveis de conversão em RDF;
- ii. Crie identificadores URI's;
- iii. Analise vocabulários já publicados para reuso;
- iv. Especifique licença de uso;
- v. Converta seus dados em RDF, de acordo com as recomendações da W3C;
- vi. Vincule seus dados a outros dados que sejam relevantes aos seus;
- vii. Publique os dados.

Para Bauer e Kaltenbock (2011), o LOD possibilita integrar e fornecer informações de alta qualidade e pode ser utilizado pelas organizações na tomada de decisão, gestão de desastres, gestão do conhecimento e soluções de inteligência de mercado. Tem influenciado em uma nova perspectiva de infraestrutura para o intercâmbio de registros bibliográficos e para o controle bibliográfico, assim como cria condições para que metadados elaborados por bibliotecas possam ser utilizados por outras comunidades da Web.

Esse novo ambiente tecnológico de aplicação trouxe maior complexidade sobre os princípios de intercâmbio e interoperabilidade de registros bibliográficos, pois “Um padrão de metadados pode apresentar um bom esquema de descrição, mas pode proporcionar uma representação insuficiente se não forem utilizados esquemas de codificação para a unificação dos dados e valores [...]” (ALVES; SANTOS, 2013, p. 126).

Willer e Dunsirre (2013) consideram que as técnicas adotadas para controle de vocabulário de valor na descrição dos elementos de metadados fornecem precisão, consistência, coerência, completude aos registros e uma maior qualidade se comparados à descrição elaborada por comunidades que não possuem domínio sobre técnicas de tratamento descritivo da informação. Essas comunidades são caracterizadas pelos autores como usuários leigos.

Nesse cenário, bibliotecas e agências nacionais possuem domínio das técnicas adotadas no tratamento descritivo da informação e têm acumulado, ao longo dos anos, registros que caracterizam o patrimônio cultural em cada país, integrando o consórcio VIAF ao *Linking Open Data*.

#### 4 CARACTERIZAÇÃO DE ARQUIVOS DE AUTORIDADE NO VIAF

Bibliotecas e agências nacionais são responsáveis pela preservação e divulgação do patrimônio cultural em cada país e detêm o registro do patrimônio cultural e intelectual acumulado por um longo período na história da humanidade. Por meio do consórcio VIAF, são convertidos e disponibilizados registros do conhecimento, que fornecem a representação do patrimônio cultural e intelectual internacional, que, por sua vez, foram elaborados pelas principais agências catalogadoras no mundo.

Os registros de bibliotecas consistem em declarações sobre as obras. Cada declaração é composta por um elemento da entidade e seu respectivo valor, abrangendo padrões de estrutura, metadados descritivos e padrões de conteúdo, e requerendo o uso de vocabulários de valor (ISAAC et al., 2011).

No VIAF, os registros são estabelecidos a partir de elementos de autoridade dos padrões MARC para dados bibliográficos, MARC para dados de autoridade e UNIMARC, padrões esses que são tradicionalmente e amplamente adotados pelas agências catalogadoras e apresentam correspondência de elementos entre si.

O padrão MADS em XML compreende uma iniciativa da *Library of Congress*, que adota MADS em XML na extração de elementos do formato MARC autoridades durante a conversão dos registros em XML. Esses registros são fornecidos pela *Library of Congress* para uso de outras instituições e ao conjunto de dados do VIAF; entretanto a base dos elementos de metadados fornecidos pela *Library of Congress* no VIAF permanece em formato MARC autoridades.

Os registros de autoridade formados no VIAF são agrupados automaticamente por meio de um algoritmo de comparação entre os registros de cada instituição cooperante e a base *WorldCat*, estabelecidos a partir do modelo conceitual FRBR (GARZON FARINOS, 2014). Por meio deste são formados *Cluster's*<sup>2</sup> de elementos de autoridade, que fornecem o registro dos vocabulários de valor definidos em cada agência catalogadora e criadas as relações entre as entidades nos registros de autoridade.

Vocabulários de valor são adotados na elaboração de metadados por

---

<sup>2</sup> Termo adotado por estatísticos para referir-se ao agrupamento por similaridade. No VIAF, é adotado para referir-se ao agrupamento de elementos correspondentes a uma mesma entidade a partir dos elementos de metadados contidos nos registros fornecidos em cada agência catalogadora.

proporcionar maior consistência na descrição do conteúdo de sistemas de recuperação da informação (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2005), em que um valor vocabulário representa uma lista controlada de valores permitidos para um elemento (ISAAC et al., 2011).

Na definição dos vocabulários de valor, são adotados padrões de conteúdo, diretrizes e vocabulários controlados, que fornecem modelos de descrição para a definição dos pontos de acesso, refletindo em fatores culturais, linguísticos e sociais de grande importância para a representação de informações do patrimônio cultural internacional. Como exemplo, pode-se citar o código AACR2, que apresenta recomendações para que o registro seja estabelecido conforme consta na fonte principal de informação e reflete a forma pela qual os nomes de pessoas, entidades coletivas e nomes geográficos são registrados nos recursos informacionais no país de origem.

No VIAF, a interface de busca de registros do sistema permite que as buscas sejam realizadas por campos:

- Todos os campos,
- Todos os títulos,
- Nomes corporativos,
- Nomes geográficos,
- Nomes pessoais,
- Obras,
- Expressões,
- Títulos preferenciais,
- Títulos bibliográficos.

O sistema permite que a busca por registros seja realizada selecionando todas ou apenas uma das instituições cooperantes.

O sistema recupera os registros que apresentem a expressão de busca correspondente e identifica os pontos de acesso, o tipo de entidade e exemplo de obras que ajudem a identificar a entidade com maior exatidão.

Ao selecionar um registro de uma entidade recuperada, são apresentados os metadados por tipo de informação por campos, conforme o quadro 17.

**Quadro 17** - Campos de apresentação de metadados no VIAF

<b>Campo</b>	<b>Descrição</b>
Forma preferida	Entrada autorizada adotada em cada unidade cooperante.
Forma de nome alternativo	Identifica variações e formas alternativas de descrição adotada em cada unidade cooperante.
Nome relacionado	Apresenta nomes relacionados ao registro de autoridade no registro bibliográfico. Exemplo: coautores; instituições de filiação; entre outros.
Obra	Identifica os tipos do recurso representado no registro bibliográfico relacionado ao registro de autoridade.
Coautores selecionados	Apresenta nomes de coautores.
País de publicação	Identifica o local de publicação conforme consta no registro bibliográfico.
Estatísticas de Publicações	Apresenta a representação gráfica de estatísticas de publicações da entidade.
Editores selecionados	Apresenta informações sobre o editor.
Sobre	Descreve informações pessoais por nacionalidade e idioma, assim como apresenta <i>links</i> externos.
Registros	Disponibiliza acesso aos registros em formato MARC 21/XML, VIAF <i>Cluster</i> XML, RDFS, JSON.
Histórico do ID no VIAF	Apresenta data e hora em que cada instituição realizou a inclusão ou exclusão de qualquer informação ou registros no sistema.

Fonte: Adaptado de Virtual International Authority File (2016a, não paginado).

A agregação dos elementos de autoridade é estabelecida a partir do formato XML com uso do protocolo *Search/Retrieve via URL* de comunicação cliente-servidor desenvolvido pela *Library of Congress*, que identifica a requisição, realiza a busca e retorna os resultados encontrados de acordo com a requisição do cliente (MOREIRA; RIBEIRO, 2008) e, uma Interface de Programação de Aplicações (API) determina como o sistema deve estabelecer as relações entre elementos de metadados, permitindo a ocorrência de variações nacionais e regionais na forma autorizada a coexistir e apoio às necessidades para variações de idioma, roteiro e ortografia, campos de nomes de autores, instituições e nomes geográficos (ONLINE COMPUTER LIBRARY CENTER, 2015).

A cada *Cluster* gerado é atribuído um Identificador Único e um link permanente (URI), que oferece maior consistência e integridade aos registros, apresentados na figura 9. Segundo Marcondes (2012), as URIs podem identificar recursos informacionais disponíveis em formato digital e de objetos físicos, e quando estes não estão em formato digital são substituídos pela representação do seu

conteúdo.

**Figura 9 - Cluster VIAF de nome geográfico Brasil**



Fonte: Virtual International Authority File (2016a, não paginado).

A identificação das entidades pessoas, entidade coletiva e nome geográfico é estabelecida por meio da representação de conteúdo. As entidades são identificadas por um identificador único nomeado VIAF ID e localizados por um link URL que identifica a localização do registro no VIAF. Por meio desses identificadores são estabelecidas as relações entre entrada autoridades de cada agência catalogadora.

Por meio da formação de *Cluster's*, são estabelecidas relações entre as entidades. As figuras 9 e 10 correspondem aos nomes geográficos Brasil (país) e São Paulo (município), cada registro possui um *Cluster* específico e ambos estão relacionados pelo *Cluster* São Paulo (estado), como forma de nome relacionado.

Figura 10 - Cluster VIAF para nome geográfico São Paulo

**São Paulo** 🇨🇭 🇩🇪 🇮🇹 🇪🇸 🇸🇪  
**São Paulo (Brazil)** 🇺🇸 🇨🇦 🇮🇱 🇦🇺  
**סאו פאולו (ברזיל)** 🇮🇱  
**São Paulo (Brésil)** 🇨🇦 🇩🇪 🇫🇷 🇵🇹 🇩🇪  
**Brazil São Paulo** 🇫🇦 🇹  
**São Paulo (Brazylia ; stan). Prefeitura.** 🇵🇹  
 VIAF ID: 126223335 (Geographic)  
 Permalink: <http://viaf.org/viaf/126223335>  
 ISNI: [0000 0001 2108 298X](https://www.isni.org/0000-0001-2108-298X)

Fonte: Virtual International Authority File (2016b, não paginado).

Os registros estabelecidos no VIAF apresentam os vocabulários formados pelos *Cluster's*. Na visualização dos dados, é possível identificar os valores definidos em cada agência catalogadora, o que abrange diferentes formas de transcrever o nome das entidades e a tradução para outros idiomas.

A figura 11 apresenta o *Cluster* de nome geográfico para Brasil, a partir do qual é possível identificar o valor definido em cada agência catalogadora, o metadado correspondente e o vocabulário de valor. Assim, verificam-se semelhanças e diferenças na forma descrição da entrada autorizada.

**Figura 11 - Cluster VIAF entrada autorizada de nome geográfico Brasil**

-  151 \_\_ [ta Brasil](#)
-  151 \_\_ [ta Brasil](#)
-  151 \_\_ [ta Brasilien](#)
-  151 \_\_ [ta Brasilien](#)
-  151 \_\_ [ta Brasilien](#)
-  151 \_\_ [ta Brazil](#)
-  151 \_\_ [ta Brazil](#)
-  110 2 \_ [ta Brésil](#)
-  210 || [ta Brésil](#)
-  110 2 \_ [ta Brésil](#)
-  210 0 1 [ta República Federativa do Brasil.](#)
-  151 \_\_ [ta ブラジル](#)

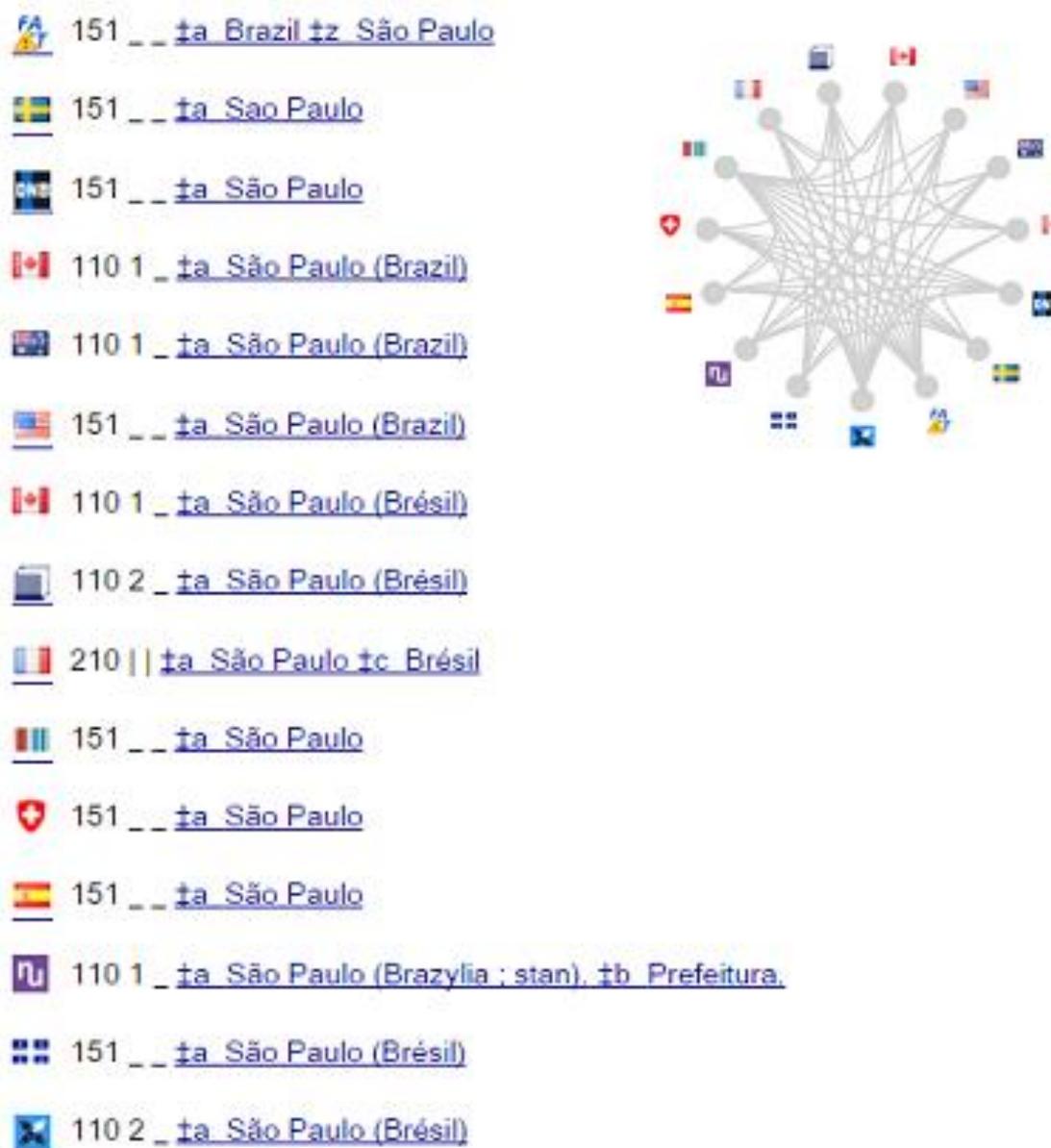


Fonte: Virtual International Authority File (2016a, não paginado).

No registro, é possível identificar que os metadados foram estabelecidos identificando a entidade Brasil como nome geográfico e como entidade coletiva pelas agências catalogadoras, por meio dos elementos de metadados 151 de nome geográfico e 110 de nome corporativo do formato MARC, e 210 do UNIMARC de nome corporativo. Os valores para nome foram estabelecidos por diferentes idiomas como: Português, Inglês, Frances, Espanhol e Esperanto.

A figura 12 apresenta as entradas autorizadas para nome geográfico São Paulo (cidade).

**Figura 12 - Cluster VIAF entrada autorizada de nome geográfico São Paulo**



Fonte: Virtual International Authority File (2016b, não paginado).

O Estado de São Paulo tem sido representado como nome geográfico e como entidade coletiva, de maneira similar ao Brasil (figura 11), por meio dos elementos de metadados 151 de nome geográfico e 110 de nome corporativo do formato MARC, e 210 de nome corporativo do formato UNIMARC. A associação entre diferentes entidades é estabelecida pelas entradas autorizadas, formas variantes da entidade e entradas relacionadas.

O sistema permite selecionar individualmente as entradas autorizadas definidas em cada agência catalogadora. Ao selecionar a entrada autorizada, o

sistema direciona o acesso ao registro no VIAF, conforme apresentado no quadro 18, ou ao link que permite o acesso ao registro no catalogo da agência catalogadora.

**Quadro 18 – Trecho registro MARC para nome geográfico**

Leader	00000nz a2200037n 45 0
001	WKP Q155 (VIAF cluster) (Authority/Source Record)
003	WKP
005	20161121160839.0
008	161121nneanz  abbn n and d
035	‡a (WKP)Q155
034	‡d -053.000000 ‡e -053.000000 ‡f -011.000000 ‡g -011.000000
035	‡a (OCoLC)Q155
035	‡a (GeoNames)3469034
151	‡a Brasil †9 xh †9 ang †9 pms †9 vi †9 ca †9 an †9 ia †9 cy †9 pt- br †9 tpi †9 eu †9 gl †9 ie †9 id †9 es †9 pt †9 min †9 nb †9 tl †9 cbk-zam †9 war †9 mwl †9 ceb †9 yo †9 ast †9 lad †9 jv †9 lmo †9 fur †9 fo †9 bcl †9 qu †9 nn †9 ace †9 eml †9 oc †9 su †9 pam †9 ext †9 kw †9 ilo †9 nah
368	‡b χώρα της Νότιας Αμερικής †9 el
368	‡b lando en Sudameriko †9 eo
368	‡b federal republic in South America †9 en
368	‡b 南美洲国家 †9 zh
368	‡b држава у Јужној Америци †9 sr-ec
368	‡b federació d'estats de l'Amèrica del Sud †9 ca
368	‡b Stato dell'America Meridionale †9 it
368	‡b 南美洲國家 †9 zh-hk
368	‡b الجنوبية أمريكية كافي دولة †9 ar
368	‡b país da América do Sul †9 pt-br
368	‡b valtio Etelä-Amerikassa †9 fi
368	‡b país de América do Sur †9 gl
368	‡b país de América del Sur †9 es
368	‡b country in South America †9 en-gb
368	‡b страна в Южной Америке †9 ru
368	‡b land in Zuid-Amerika †9 nl
368	‡b civitas in America Australis †9 la
368	‡b land i Sør-Amerika †9 nb
368	‡b दक्षिण अमेरिकाको देश †9 ne
368	‡b bansa sa Timog Amerika †9 tl
368	‡b bangsa king Mauling Amerika †9 pam

Continua.

## Continuação.

368	‡b państwo federacyjne w Ameryce Południowej ‡9 pl
368	‡b தென் அமெரிக்க நாடு ‡9 ta
368	‡b pays d'Amérique du Sud ‡9 fr
368	‡b lân yn Súd-Amearika ‡9 fy
368	‡b Staat in Südamerika ‡9 de
368	‡b állam Dél-Amerikában ‡9 hu
368	‡b Land a Südamerika ‡9 lb
368	‡b 南アメリカの国 ‡9 ja
368	‡b המדינה הדרומית אמריקה ‡9 he

Fonte: Virtual International Authority File (2016a, não paginado).

Os registros permitem acesso a outros campos disponíveis por meio dos elementos de metadados, permitindo interligar pessoas, instituições e localidades a outras informações bibliográficas.

A interface de consulta aos registros disponibiliza arquivos em quatro formatos: MARC 21 *record* XML, VIAF *Cluster* in XML, RDF *record* e *Just Links* em JSON, que estabelecem as relações sintáticas e semânticas.

Seguindo a estrutura dos registros em formato MARC-21, os registros estabelecidos no VIAF em formato XML apresentam em sua estrutura um cabeçalho contendo campos de controle do registro no VIAF, conforme apresenta o quadro 19.

### Quadro 19 – Trecho registro MARC-21 em XML

```
<mx:leader>00000nz##a22.....n..4500</mx:leader>
<mx:controlfield tag="001">viaf142985214</mx:controlfield>
<mx:controlfield tag="003">OCoLC</mx:controlfield>
<mx:controlfield tag="008">170205n||aznbnbn#####|a#aaa####|#</mx:controlfield>
- <mx:datafield tag="024" ind2=" " ind1="7">
  <mx:subfield code="a">http://viaf.org/viaf/142985214</mx:subfield>
  <mx:subfield code="2">uri</mx:subfield>
</mx:datafield>
- <mx:datafield tag="040" ind2=" " ind1=" ">
  <mx:subfield code="a">VIAF</mx:subfield>
  <mx:subfield code="c">VIAF</mx:subfield>
</mx:datafield>
```

Fonte: Virtual International Authority File (2016a, não paginado).

O registro no VIAF em MARC-21 em formato XML é formado a partir de uma estrutura de classes e subclasses. As classes são definidas a partir dos elementos de metadados do formato MARC, enquanto as subclasses consistem em identificar as agências catalogadoras e os metadados fornecidos em cada agência.

Quadro 20 - Trecho de classes e subclasses no registro MARC-21 em XML

```

- <mx:datafield tag="751" ind2=" " ind1=" " >
  <mx:subfield code="a">Brasil</mx:subfield>
  <mx:subfield code="9">xh</mx:subfield>
  <mx:subfield code="9">ang</mx:subfield>
  <mx:subfield code="9">pms</mx:subfield>
  <mx:subfield code="9">vi</mx:subfield>
  <mx:subfield code="9">ca</mx:subfield>
  <mx:subfield code="9">an</mx:subfield>
  <mx:subfield code="9">ia</mx:subfield>
  <mx:subfield code="9">cy</mx:subfield>
  <mx:subfield code="9">pt-br</mx:subfield>
  <mx:subfield code="9">tpi</mx:subfield>
  <mx:subfield code="9">eu</mx:subfield>
  <mx:subfield code="9">gl</mx:subfield>
  <mx:subfield code="9">ie</mx:subfield>
  <mx:subfield code="9">id</mx:subfield>
  <mx:subfield code="9">es</mx:subfield>
  <mx:subfield code="9">pt</mx:subfield>
  <mx:subfield code="9">min</mx:subfield>
  <mx:subfield code="9">nb</mx:subfield>
  <mx:subfield code="9">tl</mx:subfield>
  <mx:subfield code="9">cbk-zam</mx:subfield>
  <mx:subfield code="9">war</mx:subfield>
  <mx:subfield code="9">mwI</mx:subfield>
  <mx:subfield code="9">ceb</mx:subfield>
  <mx:subfield code="9">yo</mx:subfield>
  <mx:subfield code="9">ast</mx:subfield>
  <mx:subfield code="9">lad</mx:subfield>
  <mx:subfield code="9">jv</mx:subfield>
  <mx:subfield code="9">lmo</mx:subfield>
  <mx:subfield code="9">fur</mx:subfield>
  <mx:subfield code="9">fo</mx:subfield>
  <mx:subfield code="9">bcl</mx:subfield>
  <mx:subfield code="9">qu</mx:subfield>
  <mx:subfield code="9">nn</mx:subfield>
  <mx:subfield code="9">ace</mx:subfield>
  <mx:subfield code="9">eml</mx:subfield>
  <mx:subfield code="9">oc</mx:subfield>
  <mx:subfield code="9">su</mx:subfield>
  <mx:subfield code="9">pam</mx:subfield>
  <mx:subfield code="9">ext</mx:subfield>
  <mx:subfield code="9">kw</mx:subfield>
  <mx:subfield code="9">ilo</mx:subfield>
  <mx:subfield code="9">map-bms</mx:subfield>
  <mx:subfield code="9">nah</mx:subfield>
  <mx:subfield code="0">(WKP)Q155</mx:subfield>

```

Fonte: Virtual International Authority File (2016a, não paginado).

A partir da estrutura de classes e subclasses, é possível identificar e preservar os elementos de metadados e a descrição de valor definidos em cada agência catalogadora na representação do nome geográfico Brasil, apresentado no quadro 21.

### Quadro 21 – Elementos de metadados e valores no registro MARC-21 em XML

<pre>- &lt;mx:datafield tag="751" ind2=" " ind1=" " &gt;   &lt;mx:subfield code="a"&gt;Brasilien&lt;/mx:subfield&gt;   &lt;mx:subfield code="0"&gt;(DNB)04008003X&lt;/mx:subfield&gt; &lt;/mx:datafield&gt;</pre>	<pre>&lt;mx:subfield code="a"&gt;Brasilien&lt;/mx:subfield&gt; &lt;mx:subfield code="0"&gt;(DNB)04008003X&lt;/mx:subfield&gt;</pre>	<p>← Classe</p> <p>← Subclasse</p>
<pre>- &lt;mx:datafield tag="751" ind2=" " ind1=" " &gt;   &lt;mx:subfield code="a"&gt;Brasilien&lt;/mx:subfield&gt;   &lt;mx:subfield code="0"&gt;(SELIBR)141619&lt;/mx:subfield&gt; &lt;/mx:datafield&gt;</pre>		
<pre>- &lt;mx:datafield tag="751" ind2=" " ind1=" " &gt;   &lt;mx:subfield code="a"&gt;Brasilien&lt;/mx:subfield&gt;   &lt;mx:subfield code="0"&gt;(SWNL)vtls005299101&lt;/mx:subfield&gt; &lt;/mx:datafield&gt;</pre>		
<pre>- &lt;mx:datafield tag="751" ind2=" " ind1=" " &gt;   &lt;mx:subfield code="a"&gt;Brazil&lt;/mx:subfield&gt;   &lt;mx:subfield code="0"&gt;(LC)n 79128015&lt;/mx:subfield&gt; &lt;/mx:datafield&gt;</pre>		
<pre>- &lt;mx:datafield tag="751" ind2=" " ind1=" " &gt;   &lt;mx:subfield code="a"&gt;Brazil&lt;/mx:subfield&gt;   &lt;mx:subfield code="0"&gt;(FAST)fst01206830&lt;/mx:subfield&gt; &lt;/mx:datafield&gt;</pre>		
<pre>- &lt;mx:datafield tag="751" ind2=" " ind1=" " &gt;   &lt;mx:subfield code="a"&gt;Brazil&lt;/mx:subfield&gt;   &lt;mx:subfield code="9"&gt;lat&lt;/mx:subfield&gt;   &lt;mx:subfield code="0"&gt;(NLI)000977286&lt;/mx:subfield&gt; &lt;/mx:datafield&gt;</pre>		
<pre>- &lt;mx:datafield tag="751" ind2=" " ind1=" " &gt;   &lt;mx:subfield code="a"&gt;Brazil&lt;/mx:subfield&gt;   &lt;mx:subfield code="0"&gt;(NSK)000193398&lt;/mx:subfield&gt; &lt;/mx:datafield&gt;</pre>		
<pre>- &lt;mx:datafield tag="710" ind2=" " ind1="1"&gt;   &lt;mx:subfield code="a"&gt;Brazil&lt;/mx:subfield&gt;   &lt;mx:subfield code="0"&gt;(LAC)0000B8344E&lt;/mx:subfield&gt; &lt;/mx:datafield&gt;</pre>		
<pre>- &lt;mx:datafield tag="710" ind2=" " ind1="2"&gt;   &lt;mx:subfield code="a"&gt;Brazil&lt;/mx:subfield&gt;   &lt;mx:subfield code="0"&gt;(NII)DA15339025&lt;/mx:subfield&gt; &lt;/mx:datafield&gt;</pre>		

Fonte: Virtual International Authority File (2016a, não paginado).

No quadro, as classes identificam o elemento de metadados, e as subclasses indicam, a partir do indicador “a”, o valor de “o”, que identifica a agência catalogadora.

No modelo RDF, os elementos são codificados a outros vocabulários de domínio, como as ontologias SKOS, *Geonames*, *Genont*, *Schema.org*, *Friend of a friend* (FOAF), entre outras.

- *Geonames* consiste em uma base de dados geográfica que contém mais de 10 milhões de nomes geográficos que descrevem características de cada localidade e é adotado para vincular nomes geográficos (GEONAMES, 2016).
- *Schema.org* compreende um vocabulário de entidades, relações e ações entre entidades, dividido em 583 tipos, associado a 846 propriedades e 114 valores, a partir de formatos como RDFa, Microdados e JSON-LD. Pode ser vinculado a



### Quadro 23 – Mapeamento de ontologias de domínio nos registros RDF

```

- <rdf:Description rdf:about="http://viaf.org/viaf/sourceID/NSK%7C000193398#skos:Concept">
  <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
  <skos:inScheme rdf:resource="http://viaf.org/authorityScheme/NSK" xmlns:skos="http://www.w3.org/2004/02/skos/core#"/>
  <skos:prefLabel xmlns:skos="http://www.w3.org/2004/02/skos/core#">Brazil</skos:prefLabel>
  <skos:altLabel xmlns:skos="http://www.w3.org/2004/02/skos/core#">Brasil</skos:altLabel>
  <skos:altLabel xmlns:skos="http://www.w3.org/2004/02/skos/core#">Republica Federativa do Brasil</skos:altLabel>
  <foaf:focus rdf:resource="http://viaf.org/viaf/142985214" xmlns:foaf="http://xmlns.com/foaf/0.1/">
</rdf:Description>
- <rdf:Description rdf:about="http://viaf.org/viaf/sourceID/LAC%7C000088344E#skos:Concept">
  <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
  <skos:inScheme rdf:resource="http://viaf.org/authorityScheme/LAC" xmlns:skos="http://www.w3.org/2004/02/skos/core#"/>
  <skos:prefLabel xmlns:skos="http://www.w3.org/2004/02/skos/core#">Brazil</skos:prefLabel>
  <skos:prefLabel xmlns:skos="http://www.w3.org/2004/02/skos/core#">Brésil</skos:prefLabel>
  <skos:altLabel xmlns:skos="http://www.w3.org/2004/02/skos/core#">Etats-Unis du Brésil</skos:altLabel>
  <foaf:focus rdf:resource="http://viaf.org/viaf/142985214" xmlns:foaf="http://xmlns.com/foaf/0.1/">
</rdf:Description>
- <rdf:Description rdf:about="http://viaf.org/viaf/sourceID/NII%7CDA15339025#skos:Concept">
  <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
  <skos:inScheme rdf:resource="http://viaf.org/authorityScheme/NII" xmlns:skos="http://www.w3.org/2004/02/skos/core#"/>
  <skos:prefLabel xmlns:skos="http://www.w3.org/2004/02/skos/core#">Brazil</skos:prefLabel>
  <skos:altLabel xmlns:skos="http://www.w3.org/2004/02/skos/core#">Brasil</skos:altLabel>
  <skos:altLabel xmlns:skos="http://www.w3.org/2004/02/skos/core#">República Federativa do Brasil</skos:altLabel>
  <foaf:focus rdf:resource="http://viaf.org/viaf/142985214" xmlns:foaf="http://xmlns.com/foaf/0.1/">
</rdf:Description>
- <rdf:Description rdf:about="http://viaf.org/viaf/sourceID/SUDOC%7C02643895X#skos:Concept">
  <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
  <skos:inScheme rdf:resource="http://viaf.org/authorityScheme/SUDOC" xmlns:skos="http://www.w3.org/2004/02/skos/core#"/>
  <skos:prefLabel xmlns:skos="http://www.w3.org/2004/02/skos/core#">Brésil</skos:prefLabel>
  <skos:altLabel xmlns:skos="http://www.w3.org/2004/02/skos/core#">Brazil</skos:altLabel>
  <skos:altLabel xmlns:skos="http://www.w3.org/2004/02/skos/core#">República federativa do Brasil</skos:altLabel>
  <foaf:focus rdf:resource="http://viaf.org/viaf/142985214" xmlns:foaf="http://xmlns.com/foaf/0.1/">
</rdf:Description>

```

Fonte: Virtual International Authority File (2016a, não paginado).

O registro estabelece o mapeamento entre as classes e propriedades das ontologias *skos* e *foaf* com os valores definidos em cada agência catalogadora. No registro, é possível identificar as etiquetas *skos:preflabel* para identificar a entrada autorizada e *skos:altlabel* para identificar as entradas variantes. O *foaf:focus* fornece o identificador do registro no VIAF.

Outro registro disponibilizado é o *VIAF Cluster* em formato XML. Para a formação de cada *Cluster*, é atribuído a cada agência catalogadora um identificador único. O identificador é adotado em toda a estrutura do registro para identificar a agência como fonte do valor definido, apresentado no quadro 24.

### Quadro 24 - Cluster VIAF

```

▼<ns1:sources>
  <ns1:source nsid="n79128015">LC|n 79128015</ns1:source>
  <ns1:source nsid="http://d-nb.info/gnd/4008003-1">DNB|04008003X</ns1:source>
  <ns1:source nsid="fst01206830">FAST|fst01206830</ns1:source>
  <ns1:source nsid="02643895X">SUDOC|02643895X</ns1:source>
  <ns1:source nsid="141619">SELIBR|141619</ns1:source>
  <ns1:source nsid="00561061">NDL|00561061</ns1:source>
  <ns1:source nsid="vtls005299101">SWNL|vtls005299101</ns1:source>
  <ns1:source nsid="0000000121972691">ISNI|0000000121972691</ns1:source>
  <ns1:source nsid="Q155">WKP|Q155</ns1:source>
  <ns1:source nsid="http://catalogue.bnf.fr/ark:/12148/cb11868637v">BNF|11868637</ns1:source>
  <ns1:source nsid="XX4575422">BNE|XX4575422</ns1:source>
  <ns1:source nsid="562432">PTBNP|562432</ns1:source>
  <ns1:source nsid="0000B8344F">LAC|0000B8344F</ns1:source>
  <ns1:source nsid="000977286">NLI|000977286</ns1:source>
  <ns1:source nsid="000193398">NSK|000193398</ns1:source>
  <ns1:source nsid="0000B8344E">LAC|0000B8344E</ns1:source>
  <ns1:source nsid="DA15339025">NII|DA15339025</ns1:source>
</ns1:sources>

```

Fonte: Virtual International Authority File (2016a, não paginado).

A estrutura de identificação da fonte e dos valores permite identificar e preservar no registro as entradas definidas em cada agência catalogadora, conforme apresentado na interface de apresentação dos vocabulários de valor, que identificam a forma de descrição de autoridade definida pela agência catalogadora como entrada autorizada.

Os registros fornecem vocabulários que agregam diferentes formas de descrição entre a entrada autorizada, forma variante e forma relacionada. A entidade coletiva Universidade Estadual Paulista tem sido representada por meio da descrição de nome por extenso, nome abreviado, identificação do país de origem, localização por estado, ilustrado na figura 13.

**Figura 13** – Registro VIAF para nome de entidade coletiva UNESP

**Universidade estadual Paulista**
  
**Universidade estadual paulista São Paulo, Brésil**
  
**Universidade Estadual Paulista (Brasil)**
  
**Universidade Estadual Paulista "Julio de Mesquita Filho"**
  
 VIAF ID: 122646643 (Corporate)
   
 Permalink: <http://viaf.org/viaf/122646643>
  
 ISNI: [0000 0001 2096 9781](https://isni.org/isni/0000-0001-2096-9781)

Fonte: Virtual International Authority File (2016c, não paginado).

Por meio do *Cluster* identifica-se que as agências catalogadoras adotam diferentes formas de nomear as entidades.

A partir da figura 14, é possível identificar que a Universidade Estadual Paulista tem sido registrada a partir do formato MARC autoridade como entidade coletiva pela etiqueta 110 e como nome geográfico pela etiqueta 151, e pelo formato UNIMARC por meio da etiqueta 210 para entidade coletiva.

**Figura 14 - Cluster VIAF entrada autorizada de nome para entidade coletiva**



Fonte: Virtual International Authority File (2016c, não paginado).

As entradas autorizadas definidas em cada agência catalogadora permitem acesso ao registro MARC completo. O registro MARC elaborado pela *Wikidata*

(quadro 25) compreende o registro completo, contendo todos os campos definidos pela agência catalogadora.

**Quadro 25 – Trecho registro MARC de nome para entidade coletiva**

<a href="#">001</a>	WKP Q1817137 ( <a href="#">VIAF cluster</a> ) ( <a href="#">Authority/Source Record</a> )
<a href="#">003</a>	WKP
<a href="#">005</a>	20170123111533.0
<a href="#">008</a>	170123nneanz  abbn n and d
<a href="#">035</a>	‡a (WKP)Q1817137
<a href="#">035</a>	‡a (OCoLC)Q1817137
<a href="#">110</a>	2 ‡a Universidade Estadual Paulista ‡c Universidade pública brasileira no estado de São Paulo ‡9 pt
<a href="#">410</a>	2 ‡a Universidade Estadual Paulista Júlio de Mesquita Filho ‡9 fr
<a href="#">410</a>	2 ‡a Universidade Estadual Paulista ‡c public university in São Paulo ‡9 en
<a href="#">410</a>	2 ‡a Universidade Estadual Paulista ‡c Universität in Brasilien ‡9 de
<a href="#">410</a>	2 ‡a 聖保羅州立大學 ‡9 zh-hant
<a href="#">410</a>	2 ‡a 圣保罗州立大学 ‡9 zh-cn
<a href="#">410</a>	2 ‡a Університет штату Сан-Паулу Жуліу ді Мескіта Філью ‡9 uk
<a href="#">410</a>	2 ‡a Universitato de San-Paŭlio ‡9 eo
<a href="#">410</a>	2 ‡a 「ジュリオ・ヂ・メスキータ・フィーリョ」パウリスタ州立大学 ‡9 ja
<a href="#">410</a>	2 ‡a Universidad Estatal Paulista ‡9 es
<a href="#">670</a>	‡a wikidata authority control ‡u <a href="https://viaf.org/processed/ISNI 0000000120969781">https://viaf.org/processed/ISNI 0000000120969781</a>
<a href="#">670</a>	‡a wikidata authority control ‡u <a href="https://viaf.org/processed/LC n 84038004">https://viaf.org/processed/LC n 84038004</a>
<a href="#">670</a>	‡a wikidata authority control ‡u <a href="https://viaf.org/processed/DNB 220437-X">https://viaf.org/processed/DNB 220437-X</a>
<a href="#">670</a>	‡a wikidata authority control ‡u <a href="https://viaf.org/processed/BNF 11995524">https://viaf.org/processed/BNF 11995524</a>
<a href="#">670</a>	‡a wikidata authority control ‡u <a href="https://viaf.org/viaf/122646643">https://viaf.org/viaf/122646643</a>
<a href="#">670</a>	‡a wikidata authority control ‡u <a href="https://viaf.org/processed/SUDOC 02803967X">https://viaf.org/processed/SUDOC 02803967X</a>
<a href="#">670</a>	‡a wikidata site
	links ‡u <a href="https://fr.wikipedia.org/wiki/Universidade_Estadual_Paulista_Júlio_de_Mesquita_Filho">https://fr.wikipedia.org/wiki/Universidade_Estadual_Paulista_Júlio_de_Mesquita_Filho</a>
<a href="#">670</a>	‡a wikidata site links ‡u <a href="https://pt.wikipedia.org/wiki/Universidade_Estadual_Paulista">https://pt.wikipedia.org/wiki/Universidade_Estadual_Paulista</a>
<a href="#">670</a>	‡a wikidata site links ‡u <a href="https://de.wikipedia.org/wiki/Universidade_Estadual_Paulista">https://de.wikipedia.org/wiki/Universidade_Estadual_Paulista</a>
<a href="#">670</a>	‡a wikidata site links ‡u <a href="https://eo.wikipedia.org/wiki/Universitato_de_San-Paŭlio">https://eo.wikipedia.org/wiki/Universitato_de_San-Paŭlio</a>
<a href="#">670</a>	‡a wikidata site links ‡u <a href="https://es.wikipedia.org/wiki/Universidad_Estatal_Paulista">https://es.wikipedia.org/wiki/Universidad_Estatal_Paulista</a>
<a href="#">670</a>	‡a wikidata site links ‡u <a href="https://en.wikipedia.org/wiki/São_Paulo_State_University">https://en.wikipedia.org/wiki/São_Paulo_State_University</a>
<a href="#">670</a>	‡a wikidata site links ‡u <a href="https://it.wikipedia.org/wiki/Universidade_Estadual_Paulista">https://it.wikipedia.org/wiki/Universidade_Estadual_Paulista</a>
<a href="#">670</a>	‡a wikidata site links ‡u <a href="https://uk.wikipedia.org/wiki/Університет_штату_Сан-Паулу_Жуліу_ді_Мескіта_Філью">https://uk.wikipedia.org/wiki/Університет_штату_Сан-Паулу_Жуліу_ді_Мескіта_Філью</a>
<a href="#">998</a>	‡a Universidade Estadual Paulista ‡2 LC n 84038004 ‡3 suggested
<a href="#">998</a>	‡a Universidade estadual paulista ‡c São Paulo, Brasil ‡2 BNF 11995524 ‡3 suggested
<a href="#">998</a>	‡a Universidade Estadual Paulista ‡2 ISNI 0000000120969781 ‡3 suggested ‡3 viafid
<a href="#">998</a>	‡a Universidade estadual paulista (São Paulo, Brasil) ‡2 ISNI 0000000120969781 ‡3 suggested ‡3 viafid
<a href="#">998</a>	‡a Universidade estadual Paulista ‡2 ISNI 0000000120969781 ‡3 suggested ‡3 viafid
<a href="#">998</a>	‡a Universidade Estadual Paulista "Julio de Mesquida Filho" ‡2 ISNI 0000000120969781 ‡3 suggested ‡3 viafid
<a href="#">998</a>	‡a Universidade Estadual Paulista Júlio de Mesquita Filho ‡2 ISNI 0000000120969781 ‡3 suggested ‡3 viafid
<a href="#">998</a>	‡a Universidade estadual paulista ‡2 ISNI 0000000120969781 ‡3 suggested ‡3 viafid
<a href="#">998</a>	‡a Universidade estadual paulista (São Paulo, Brasil) ‡2 SUDOC 02803967X ‡3 suggested

Fonte: Virtual International Authority File (2016c, não paginado).

No registro, os elementos 670 e 998 são adotados para estabelecer relações de equivalência da entrada autorizada a outras entradas estabelecidas pelas demais agências catalogadoras que participam do VIAF.

A figura 15 apresenta as diferentes formas estabelecidas para descrever o nome do escritor Machado de Assis: nome artístico e data associada à pessoa, nome de batismo, nome traduzido para outros idiomas, entre outros.

**Figura 15 - Registro VIAF para pessoa Machado de Assis**



The image shows a screenshot of a VIAF record for Machado de Assis. On the left is a book cover for 'Memórias póstumas de Blás Cubas' by Joaquim Maria Machado de Assis. The main content lists various name forms with associated flags:
 

- Machado de Assis, 1839-1908 (with flags for DE, FR, NO, SE, PT, ES, IT, NL, GB, PL, CND, RU, GR, ES)
- Assis, Machado de, 1839-1908 (with flags for SE, FR, PT, NL, NO)
- Assis, Joaquim Maria Machado de, 1839-1908 (with a book icon)
- Assis, Machado de (Joaquim Maria Machado), 1839-1908 (with flag for PT)
- Joaquim Maria Machado de Assis poète brésilien, d'expression française et portugaise (with flag for FR)
- Machado de Assis, Joaquim Maria (with flags for BR, PT)
- De Assis, Machado (with flag for BR)
- ماشادو دو آسيس، جواكيم ماريا (with flag for BR)
- משדר די-אסיס, ז'ואקים מריה (with flag for IL)
- Machado de Assis, Joaquim Maria, 1839-1908 (with flag for BR)
- Assis, J. M. M. de 1839-1908 Joaquim Maria Machado de (with flag for BR)
- 1908-1839 ماشادو دو آسيس، (with flag for BR)

 At the bottom, it provides the VIAF ID: 95151633 (Personal), Permalink: <http://viaf.org/viaf/95151633>, and ISNI: 0000\_0001\_2144\_1225.

Fonte: Virtual International Authority File (2016d, não paginado).

A partir da figura, identifica-se que a forma mais adotada para descrever a entrada autorizada para Machado de Assis no mundo é “Machado de Assis, 1839-1908”, embora existam outras formas também adotadas pelas agências catalogadoras.

A figura 16 apresenta com exatidão qual a forma autorizada adotada em cada agência catalogadora.



**Quadro 26 - Trecho registro MARC para nome pessoal**

Label		00000nx aa22000373 45 0
001		NLR RU NLR AUTH 7762790 (VIAF cluster)
005		20080320160155.0
003		NLR
035		‡a (NLR)RU\NLR\AUTH\7762790
100		‡a 20080320arusy50 ca
152		‡a RCR
200	1	‡8 por ‡7 ba ‡a Assis ‡b J. M. M. de ‡f 1839-1908 ‡g Joaquim Maria Machado de
300	0	‡a Бразил. писатель.
400	1	‡8 por ‡7 ba ‡a Machado de Assis ‡b J. M. ‡f 1839-1908 ‡g Joaquim Maria ‡5 z
400	0	‡7 ba ‡a Machado de Assis ‡f 1839-1908 ‡5 z
400	0	‡8 por ‡7 ba ‡a Machado d'Assis ‡f 1839-1908 ‡5 z
400	1	‡8 por ‡7 ba ‡a Machado de Assis ‡b J. M. ‡f 1839-1908 ‡g Joaquim Maria ‡5 z
400	1	‡8 por ‡7 ba ‡a Assis ‡b M. de ‡f 1839-1908 ‡g Machado de ‡5 z
400	0	‡8 por ‡7 ba ‡a Joaquim Maria Machado de Assis ‡f 1839-1908 ‡5 z
700	1	‡3 RU\NLR\AUTH\7762787 ‡8 rus ‡7 ca ‡a Машаду де Ассиз ‡b Ж. М. ‡f 1839-1908 ‡g Жоакин Мария
801	0	‡a RU ‡b RU-19017073 ‡9 NLR ‡c 20080320
801	1	‡a RU ‡b RU-19017073 ‡9 NLR ‡c 20080320
801	2	‡a RU ‡b RU-19017073 ‡9 NLR ‡c 20150910 ‡2 rusmarc
810		‡a ГАКин РНБ: ‡b установлен. форма - Assis, Joaquim Maria Machado de; вариант. форма - Machado de Assis, Joaquim Maria, на обороте карточек ссылка на след. источники: Bul. bibliogr.bras; Index ,bibliogr. bras.
810		‡a Рассказы/ Машаду де Ассиз.- СПб, 2006: ‡b на развороте тит. л. ин. форма - Machado de Assis.
810		‡a Library of Congress Catalogs online: ‡b установлен. форма - Machado de Assis; вариант. форма - Machado de Assis, Joaquim Maria.
810		‡a Ubka : Britisher VK СОРАС (по запросу): ‡b установлен. форма - Assis, Machado de; вариант. форма - Joaquim Maria Machado de Assis.
810		‡a Ubka, Franzosischer VK, запрос: ‡b Assis, Machado de
810		‡a UBKA: каталоги Германии: HBZ VK, BVB, GBV, SWB: ‡b принятая форма - Assis, Joaquim Maria de; вариант. форма - Joaquim Maria Machado de Assis.
810		‡a UBKA:Deutsche Nationalbibliothek: ‡b Assis, Joaquim Maria de
810		‡a UBKA: KOBV Berlin-Brandenburg: ‡b Machado de Assis, Joaquim Maria.
810		‡a Краткая литературная энциклопедия. - М.,1964-1978: ‡b Machado d'Assis.
815		‡a ЭК РНБ "Издания на иностранных языках (1994- )".
996		‡2 NSK 000008601
996		‡2 ISNI 0000000069601550
996		‡2 LC no 97075947
996		‡2 PTBNP 1485062

Fonte: Virtual International Authority File (2016d, não paginado).

O registro de autoridade para o autor Machado de Assis compreende diferentes formas de transcrição de nomes que foram estabelecidas a partir das etiquetas MARC, que definem as entradas e as relações entre as formas variantes de nome.

## 5 CONSIDERAÇÕES FINAIS

O controle de autoridade surgiu da necessidade de uniformizar pontos de acesso controlados e garantir maior consistência aos catálogos, e tem sido implementado a partir de padrões, códigos, diretrizes e modelos conceituais que, ao longo dos anos, foram desenvolvidos, revistos e atualizados sob influência das tecnologias da informação e da comunicação.

O formato de intercâmbio MARC compreende um dos padrões de maior adesão internacional em bibliotecas e agências nacionais e está diretamente relacionado a sua estrutura e nível de detalhamento, que garante persistência aos registros de autoridade. Entretanto sua estrutura é estabelecida a partir dos formatos de intercâmbio e compartilhamento de dados ANSI Z39.2 e ISO 2709, que apresentam limitações e restringem a integração de registros para outros ambientes, como a Web. Isso tem impulsionado o desenvolvimento de técnicas de conversão, extração e correspondência entre padrões de metadados e possibilitado a conversão do formato de intercâmbio do padrão MARC para formatos que apresentem compatibilidade com a Web, como XML e RDF.

Dentre as instituições de maior adesão internacional do formato MARC estão as bibliotecas e agências nacionais, que, desde a década de 1970, têm sido responsáveis pelo controle bibliográfico em cada país, dando origem ao projeto Controle Bibliográfico Universal (CBU). Este, embora tenha sido interrompido em 1990, influenciou na formação de outros projetos de cooperação internacional de dados bibliográficos e de autoridade, como o VIAF, iniciado em 1998 e oficializado com um consórcio internacional em 2003.

Torna-se, portanto, imprescindível a revisão de conceitos e fundamentos estabelecidos na catalogação para o controle de autoridade e a interpretação de como a evolução tecnológica influenciou no aprimoramento das técnicas adotadas no tratamento descritivo da informação. Diante do exposto, os objetivos desta pesquisa foram estabelecidos a partir do problema de pesquisa: como o controle de autoridade no serviço VIAF contribui para a efetivação da Web Semântica de modo a proporcionar maior qualidade aos sistemas de busca e recuperação da informação?

O primeiro e o segundo objetivos – "descrever os princípios, técnicas e padrões que proporcionam o controle de autoridade na catalogação, destacando sua

evolução, características e conceitos" e "analisar o controle de autoridade de nomes para pessoas, instituições e nomes geográficos" – foram alcançados nos capítulos 1 e 2.

O terceiro objetivo – "apresentar o consórcio *Virtual International Authority File* (VIAF), sua origem, as tecnologias envolvidas em sua estrutura e sua perspectiva de contribuição para a Web Semântica" – foi apresentado nos capítulos 3 e 4.

A partir dos objetivos estabelecidos e alcançados, foi possível estabelecer algumas considerações para esta pesquisa:

- A representação de pontos de acessos estabelecidos por meio do trabalho de autoridade requer trabalho constante de atualização para que os dados de autoridade sejam consistentes. Neste ponto, considera-se que o conceito de trabalho de autoridade de Maxwell (2002), citado por Assumpção e Santos (2013), pode ser revisto e incluir a etapa de atualização dos pontos de acesso. Compreende, portanto, a definição, a manutenção, os registros e a atualização dos pontos de acesso.
- Os vocabulários estruturados por formatos interoperáveis na Internet fornecem meios de integração de informações contidas nos registros de autoridade com outros recursos informacionais; criam a possibilidade de promover maior visibilidade aos recursos mantidos e serviços realizados por bibliotecas; e agregam informações sobre o patrimônio cultural internacional a outros recursos informacionais na Web.
- A formação de registros por meio da agregação de metadados de autoridade, realizada a partir do VIAF, que vincula vocabulários e representações estabelecidas a partir do tratamento descritivo da informação internacionalmente, determina a formação de vocabulários de valor de alto nível.
- A interoperabilidade entre os vocabulários de valor formados pelos arquivos de autoridade no VIAF com vocabulários estabelecidos em outros domínios citados ao longo da pesquisa, se ocorrer de forma consistente, pode interligar pessoas, instituições e informações geográficas a recursos informacionais em sistemas distribuídos na Web.
- Os arquivos de autoridade estabelecidos por meio do VIAF poderão ser reutilizados por instituições bibliotecárias na definição de pontos de acesso consistentes, sobretudo diante dos estudos de dados interligados para

bibliotecas, definidos a partir do modelo de dados Bibframe<sup>3</sup> em desenvolvimento pela Library of Congress.

- O reconhecimento da importância do patrimônio cultural e intelectual brasileiro para a humanidade, dada sua riqueza e diversidade cultural, reforça a necessidade de o Brasil fazer parte do consórcio VIAF e se posicionar perante a comunidade internacional. Além disso, permite maior visibilidade ao conhecimento gerado por pessoas, instituições e em regiões do Brasil, assim como aos serviços de catalogação estabelecidos no país.
- Os avanços nas tecnologias e formatos adotados na infraestrutura dos catálogos indicam a necessidade de os bibliotecários adquirirem novas habilidades e conhecimentos relacionados aos formatos e modelos de representação da informação. O estudo sobre a infraestrutura do VIAF mostrou que as tecnologias da Web Semântica estão diretamente relacionadas a esse novo contexto em que os conceitos de fundamentos de catalogação têm sido atualizados. Essas tecnologias podem ser aproveitadas para o estabelecimento de serviços cooperativos entre bibliotecas e outras instituições do patrimônio cultural.

A partir das considerações apresentadas identificou-se a necessidade de novos estudos futuros que apontem sobre a real correspondência entre os elementos de metadados de autoridade e os vocabulários de ontologias aos quais as entidades têm sido integradas no VIAF. Assim como, sobre quais os conjuntos de dados os arquivos de autoridade no VIAF estão sendo vinculados no projeto *Linking Open Data*.

No Brasil, a baixa produção de estudos sobre controle de autoridade indica para a urgência de estudos de ordem teórica e aplicada.

---

<sup>3</sup> Fonte: <https://www.loc.gov/bibframe/>

## REFERÊNCIAS

- ALBÁS, L. Control de autoridades y calidad de puntos de acceso del catálogo automatizado de la Biblioteca de la Universidad de Salamanca. **Taller de Bibliotecas Universitarias Iberoamericanas**, La Habana, v. 6, 2003.
- ALVARENGA, L. Representação do conhecimento na perspectiva da Ciência da Informação em tempo e espaço digitais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, n. 15, 2003.
- ALVES, R. A. V. **Metadados como elementos do processo de catalogação**. 2010. 132 f. Tese (Doutorado em Ciência da Informação)-Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.
- ALVES, R. C. V. **Web Semântica: uma análise focada no uso de metadados**. 2005. 180 f. Dissertação (Mestrado em Ciência da Informação)-Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005.
- ALVES, R. C. V.; SANTOS, P. L. V. A. C. **Metadados no domínio bibliográfico**. Rio de Janeiro: Intertexto, 2013.
- AMERICAN LIBRARY ASSOCIATION. **RDA: Resource Description and Access**. Chicago: ALA Editions, 2013.
- AGENJO BULLON, X.; HERNANDEZ CARRASCAL, F. Influencia del ICABS en el futuro digital de las bibliotecas. **Boletín de la ANABAD**, v. 54, n. 1-2, p. 387-396, 2004. Disponível em: <<https://dialnet.unirioja.es/servlet/articulo?codigo=1198724>>. Acesso em: 14 nov. 2016.
- ANGJELI, A.; MAC EWAN, A.; BOULET, V. ISNI and VIAF – Transforming ways of trustfully consolidating identities. In: IFLA GENERAL CONFERENCE AND ASSEMBLY, 80., 2014, Lyon. **Proceedings...** Lyon: IFLA, 2014.
- ASSUMPÇÃO, F. S. **Controle de autoridade: definições, processos e componentes**. 2011. 122 f. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia)-Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2012.
- ASSUMPÇÃO, F. S.; SANTOS, P. L. V. A. C. A importância do controle de autoridade: uma abordagem baseada nos objetivos e nas funções dos catálogos. In: ENCONTRO INTERNACIONAL DE CATALOGADORES, 9.; ENCONTRO NACIONAL DE CATALOGADORES, 2., 2013, Rio de Janeiro. **Anais...** Rio de Janeiro: Fundação Biblioteca Nacional do Brasil, 2013.
- BACA, M. (Ed.). **Introduction a los metadatos: vías a la información digital**. [S.l.: s.n.], 1999.
- BAUER, F.; KALTENBOCK, M. **Linked Open Data: the essentials: a quick start guide for decision makers**. Vienna: DGS, 2011. Disponível em: <<http://www.reeep.org/LOD-the-Essentials.pdf>>. Acesso em: 21 jul. 2016.

BERNERS-LEE, T.; HENDLER, J; LASSILA, O. The Semantic Web: a new forma of Web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**, v. 284, n. 5, 2001.

BIBLIOTECA NACIONAL DA ESPANHA. Busqueda de autoridades. 2015. Disponível em: <<http://catalogo.bne.es/uhtbin/authoritybrowse.cgi>>. Acesso em: 4 dez. 2016.

BREITMAN, K. **Web Semântica**: a internet do futuro. Rio de Janeiro: LTC, 2006.

BRICKLEY, D.; MILLER, L. **FOAF Vocabulary Specification 0.99**: Namespace document. 2014. Disponível em: <<http://xmlns.com/foaf/spec/>>. Acesso em: 21 jun. 2016.

BROOKS, T. A. Watch this: LOD - linking open data. **Information Research**, v. 13, n. 4, 2008. Disponível em: <<http://www.informationr.net/ir/13-4/TB0812.html>>. Acesso em: 1 jun. 2016.

BURKE, S. K.; SHORTEN, J. Name authority work today: a comparison of types of academic libraries. **Library resources and technical services**, v. 54, n. 1, 2010.

BUTTLAR, L.; GARCHA, R. Catalogers in academic libraries: their evolving and expanding roles. **College and Research Libraries**, v. 59, n. 4, p. 311-321, 1998. Disponível em: <<https://doi.org/10.5860/crl.59.4.311>>. Acesso em: 1 jun. 2016.

CAMPELLO, B. **Introdução ao controle bibliográfico**. Brasília: Briquet de Lemos, 2006.

CAMPOS, L. F. B. Metadados digitais: revisão bibliográfica da evolução e tendências por meio de categorias funcionais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 12, n. 23, p. 16-46, 2007.

CAPURRO, R.; HJØRLAND, B. O conceito de informação. **Perspectivas em Ciência da Informação**, v. 12, n. 1, p. 148-207, 2007.

CARLOS, K. V.; ZAFALON, Z. R. Metadados e Interoperabilidade em Bibliotecas Nacionais Sul-Americanas. In: ENCONTRO INTERNACIONAL DE CATALOGADORES, 9.; ENCONTRO NACIONAL DE CATALOGADORES, 2., 2013, Rio de Janeiro. **Anais...** Rio de Janeiro: Fundação Biblioteca Nacional do Brasil, 2013.

CATARINO, M. E.; SOUZA, T. B. A representação descritiva no contexto da Web Semântica. **TransInformação**, v. 24, n. 2, p. 77-90, 2012.

CÓDIGO de catalogação anglo-americano. 2. ed. São Paulo: FEBAB, 2004.

COYLE, K. Metadata models of the World Wide Web. **Library Technology Reports**, n. 2, p. 12-19, 2010. Disponível em: <<http://dx.doi.org/10.5860/ltr.46n2>>. Acesso em: 31 maio 2016.

CYGANIAK, R.; JENTZSCH, A. **The Linking Open Data cloud diagram**. 2014. Disponível em: <<http://lod-cloud.net/>>. Acesso em: 8 dez. 2016.

DUNSIRE, G. Interoperability and Semantics in RDF representations of FRBR, FRAD and FRSAD. 2010. Disponível em: <<http://gordondunsire.com/pubs/docs/InteropSemanticsRDFFRBRPre.pdf>>. Acesso em: 20 dez. 2016.

ECO, U. **A vertigem das listas**. Rio de Janeiro: Record, 2010.

FERREIRA, J. A.; SANTOS, P. L. V. A. C. O modelo de dados Resource Description Framework (RDF) e o seu papel na descrição de recursos. **Informação & Sociedade**, v. 23, n. 2, p. 13-23, 2015.

FUSCO, E. **Aplicação dos FRBR na modelagem de catálogos bibliográficos digitais**. São Paulo: Cultura Acadêmica, 2011.

GARCÍA MARCO, F. J. Paradigmas científicos em representación y recuperación de la información. In: ENCUESTRO DE ISKO-ESPAÑA, 1., 1993, Madrid. **Anais...** Zaragoza: Universidad de Zaragoza, 1995. p. 99-112.

GARZON FARINOS, F. G. **El registros de autoridades personales tras la aparición del Web**. 2014. 525 f. Tese (Doutorado em Indústrias Culturais e da Comunicação)-Universitat Politècnica de València, Valencia, 2014.

GEONAMES. **About GeoNames**. Disponível em: <<http://www.geonames.org/about.html>>. Acesso em: 20 dez. 2016.

GORMAN, M. Authority control in the context of bibliographic control in the electronic environment. **Cataloging & Classification Quarterly**, v. 38, 2004. Disponível em: <[https://doi.org/10.1300/J104v38n03\\_03](https://doi.org/10.1300/J104v38n03_03)>. Acesso em: 20 mar. 2016.

GRÁCIO, M. C. C.; OLIVEIRA, E. F. T. Produção e comunicação da informação em CT&I - GT7 da ANCIB: análise bibliométrica no período 2003/2009. **Liinc em Revista**, v. 7, n. 1, p. 48-263, 2011. Disponível em: <<http://dx.doi.org/10.18617/liinc.v7i1.412>>. Acesso em: 15 jan. 2016.

GUEDES, V.; BORSCHIVER, S. Bibliometria: uma ferramenta estatística para a gestão da informação e do conhecimento, em sistemas de informação, de comunicação e de avaliação científica e tecnológica. In: ENCONTRO NACIONAL DE CIÊNCIA DA INFORMAÇÃO, 6., 2005, Salvador. **Anais...** Salvador: UFBA, 2006.

GUIMARÃES, J. A. C. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-americana de Ciência da Informação**, v. 1, n. 1, p. 77-99, 2008.

HERRERO PASCUAL, C. El control de autoridades. **Anales de Documentación**, n. 2, p. 121-136, 1999.

HOPPE, T. REST APIs today and tomorrow: an essay. 2015. Disponível em: <[http://vanthome.github.io/rest-api-essay-presentation/rest\\_apis.html#31](http://vanthome.github.io/rest-api-essay-presentation/rest_apis.html#31)>. Acesso em: 20 dez. 2016.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Requisitos funcionales de los datos de autoridad (FRAD)**: un modelo conceptual. Madrid: IFLA, 2009. Disponível em: <[http://www.ifla.org/files/assets/cataloguing/frad/frad\\_2009-es.pdf](http://www.ifla.org/files/assets/cataloguing/frad/frad_2009-es.pdf)>. Acesso em: 15 mai. 2015.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Declaração dos princípios internacionais de catalogação**. 2009a. Disponível em: <[http://www.ifla.org/files/assets/cataloguing/icp/icp\\_2009-pt.pdf](http://www.ifla.org/files/assets/cataloguing/icp/icp_2009-pt.pdf)>. Acesso em: 15 jul. 2015.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Directrices para registros de autoridad y referencias**. 2. ed. [S.l.]: Ministerio de Cultura, 2004.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Guidelines for authority records and references**. München: K.G. Saur, 2001.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Names of persons**: national usages for entry in catalogues. [S.l.]: UBCIM Publications, 1996.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Functional Requirements for Bibliographic Records**. 2009b. Disponível em: <[http://www.ifla.org/files/assets/cataloguing/frbr/frbr\\_2008.pdf](http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf)>. Acesso em: 21 jul. 2016.

ISAAC, A. **Correspondence between ISO 25964 and SKOS/SKOS-XL Models**. 2013. Disponível em: <[http://www.niso.org/apps/group\\_public/download.php/12351/Correspondence%20ISO25964-SKOSXL-MADS-2013-12-11.pdf](http://www.niso.org/apps/group_public/download.php/12351/Correspondence%20ISO25964-SKOSXL-MADS-2013-12-11.pdf)>. Acesso em: 20 nov. 2015.

ISAAC, A. et al. Library Linked Data Incubator Group: Datasets, Value Vocabularies, and Metadata Element Sets. **W3C Incubator Group Report**, 2011. Disponível em: <<http://www.w3.org/2005/Incubator/ld/XGR-ld-vocabdataset-20111025/>>. Acesso em: 1 ago. 2015.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 26964:2011**: Information and documentation-- Thesauri and interoperability with other vocabularies -- Part 2: Interoperability with other vocabularies. Genebra, 2011.

JACOB, E. K. Ontologies and the semantic web. **Bulletin for the American Society for Information Science and Technology**, v. 29, n. 4, p. 19-22, Apr./Mayo 2003.

JIMENEZ PELAYO, J.; GARCÍA BLANCO, R. **El Catálogo de autoridades**: Creación y gestión en unidades documentales. Gijón: TREA, 2002.

JIN, Q. **Demystifying FRAD**: Functional Requirements for Authority Data. Santa Barbara: Libraries Unlimited, 2012.

LASSILA, O.; SWICK, R. R. **Resource description framework (RDF) model and syntax specification**. [S.l]: W3C, 1999.

LEHNUS, D. J. A comparison of Panizzi's 91 rules and the AACR of 1967.

**Occasional papers**, n. 105, 1972. Disponível em:

<<https://www.ideals.illinois.edu/bitstream/handle/2142/3872/gslisoccasionalpv00000i00105.pdf?sequence=1>>. Acesso em: 22 jun. 2016.

LIBRARY OF CONGRESS. Bibliographic Framework Initiative. 2016. Disponível em: <<https://www.loc.gov/bibframe/>>. Acesso em: 02 jan. 2017.

LIBRARY OF CONGRESS. Formato MARC 21 Conciso para dados de autoridade. 2006. Disponível em: <<https://www.loc.gov/marc/authority/ecadspa.html>>. Acesso em: 19 fev. 2016.

LIBRARY OF CONGRESS. MARC 21 Authority. 2009. Disponível em: <<https://www.loc.gov/marc/authority/adinro.html>>. Acesso em: 2 out. 2016.

LIBRARY OF CONGRESS. MARC 21 Authority Format Mapping to MADS Schema version 2.0. 2011. Disponível em: <<https://www.loc.gov/standards/mads/mads-mapping.html>>. Acesso em: 19 nov. 2015.

LIBRARY OF CONGRESS. MARC 21 Format for Bibliographic Data. 1999. Disponível em: <<http://www.loc.gov/marc/bibliographic/bdintro.pdf>>. Acesso em: 05 dez. 2016.

LIBRARY OF CONGRESS. MARC 21 Format for Classification Data. 2016b. Disponível em: <<http://www.loc.gov/marc/classification/>>. Acesso em: 1 nov. 2016.

LIBRARY OF CONGRESS. MARC 21 Format for Community Information. 2016c. Disponível em: <<https://www.loc.gov/marc/community/>>. Acesso em: 01 dez. 2016.

LIBRARY OF CONGRESS. MARC 21 Format for Holdings Data. 2016a. Disponível em: <<http://www.loc.gov/marc/holdings/>>. Acesso em: 08 dez. 2016.

LIBRARY OF CONGRESS. The MARC 21 formats: Background and principles. 1996. Disponível em: <<https://www.loc.gov/marc/96principl.html#one>>. Acesso em: 02 out. 2016.

LIMA, J. C.; CARVALHO, C. L. **Ontologias**: OWL (Web Ontology Language). 2005. RT-INF 004-05. Relatório técnico. Disponível em: <[http://www.portal.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF\\_004-05.pdf](http://www.portal.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_004-05.pdf)>. Acesso em: 15 jun. 2015.

LIMA, R. A.; VELHO, L. M. L. S.; FARIA, L. I. L. Delimitação de uma área multidisciplinar para análise bibliométrica de produção científica: o caso da Bioprospecção. **TransInformação**, v. 19, n. 2, p. 153-168, 2007.

MACHADO, A. M. N. **Informação e controle bibliográfico**: um olhar sobre a cibernética. São Paulo: UNESP, 2003.

MARCONDES, C. H. "Linked data" - dados interligados - e interoperabilidade entre arquivos, bibliotecas e museus na web. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 17, n. 34, p. 171-192, 2012.

MARCONDES, C. H. Da comunicação científica ao conhecimento público: artigos científicos digitais como bases de conhecimentos. In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 6., 2005, Florianópolis. **Anais...** Florianópolis: UFSC, 2005.

MEY, E. S. A.; SILVEIRA, N. C. **Catálogo no plural**. Brasília: Briquet de Lemos, 2009.

MEY, E. S. A.; SILVEIRA, N. C. Considerações teóricas aligeiradas sobre a catalogação e sua aplicação. **InCID: Revista de Ciência da Informação e Documentação**, v. 1, n. 1, p. 125-137, 2010.

MILES, A.; BECHNOFER, S. SKOS Simple Knowledge Organization System Namespace Document - HTML Variant. 2009. Disponível em: <<https://www.w3.org/2009/08/skos-reference/skos.html#>>. Acesso em 20 dez. 2016.

MOREIRO GONZÁLEZ, J. A. **Linguagens documentárias e vocabulários semânticos para a web**: elementos conceituais. Salvador: EDUFBA, 2011.

MOREIRA, W.; RIBEIRO, T. Introdução ao uso dos protocolos SRU/SRW: ferramentas para a catalogação cooperativa. **Perspectiva em Ciência da Informação**, v. 13, n. 3, p. 167-182, 2008.

MUNNICH, M. Reuse or rule harmonization. **LRTS**, v. 44, n. 3, 2010. Disponível em: <<https://journals.ala.org/lrts/article/viewFile/5071/6141>>. Acesso em: 26 maio 2016.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **ANSI/NISO Z39.19:2005**: Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. Baltimore, 2005.

NEVES, J. L. Pesquisa qualitativa: características, usos e possibilidades. **Cadernos de pesquisas em administração**, v. 1, n. 3, 1996.

ONLINE COMPUTER LIBRARY CENTER. 2014 Annual Report to VIAF Council. 2014. Disponível em: <<http://www.oclc.org/content/dam/oclc/viaf/OCLC-2014-VIAF-Annual-Report-to-VIAF-Council.pdf>>. Acesso em: 13 mai. 2015.

ONLINE COMPUTER LIBRARY CENTER. VIAF: Virtual International Authority File. 2015. Disponível em: <<http://www.oclc.org/viaf.en.html>>. Acesso em: 13 mai. 2015.

OKUBO, Y. **Bibliometric indicators and analysis of research systems**: methods and examples. Paris: OECD, 1997.

OPEN METADATA REGISTRY. Disponível em: <<http://metadataregistry.org/>>. Acesso em: 5 dez. 2016.

PASTOR SÁNCHEZ, J. A.; MARTINEZ MÉNDEZ, F.J.M.; RODRÍGUEZ MUÑOZ. Advantages of thesaurus representation using the Simple Knowledge Organization System (SKOS) compared with proposed alternatives. **IR Information Research**, v. 14, n. 4, 2009. Disponível em: <<http://www.informationr.net/ir/14-4/paper422.html>>. Acesso em: 1 nov. 2015.

PETIYE, S. **Building a semantic web of comics: publishing linked data in HTML/RDFa**: using a comic book ontology and metadata application profiles. 2014. 189 f. Dissertation (Master of Library and Information Science and Master of Science)-College of Communication and Information, Kent State University, Kent, 2014.

PICCO, P.; ORTIZ REPISO, V. RDA, el nuevo código de catalogación: cambios y desafíos para su aplicación. **Revista Española de Documentación Científica**, v. 35, n. 1, p. 145-173, 2012.

PLASSARD, M. F. Authority control in an international environment: the UNIMARC format for authorities. In: WORKSHOP ON AUTHORITY CONTROL AMONG CHINESE, KOREAN AND JAPANESE LANGUAGES, 2., 2001, Chiyoda. **Anais...** Chiyoda: NII, 2001.

SALES, R.; CAFÉ, L. Diferenças entre tesouros e ontologias. **Perspectivas em Ciência da Informação**, v. 14, n. 1, p. 99-116, 2009.

SALGADO, D. M. **O controle de autoridade sob a norma RDA**: análise da aplicação e implicações na construção de registros de autoridade. 2015. Dissertação (Mestrado em Ciência da Informação)-Universidade de São Paulo, São Paulo, 2015.

SALGADO, D. M.; SILVA, J. F. M. AACR2 X RDA: breves reflexões acerca dos registros de autoridade. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 25., 2013, Florianópolis. **Anais...** São Paulo: FEBAB, 2013.

SANDÍN ENRIQUEZ, D. **El Fichero Internacional Virtual de Autoridades (VIAF)**: Estado de la cuestión, evaluación y propuestas de mejora. 2011. 110f. Dissertação (Mestrado em Sistemas de Informação Digital)-Facultad de Traducción y Documentación, Universidad de Salamanca, Salamanca, 2011.

SANTAREM SEGUNDO, J. E. Web semântica: introdução a recuperação de dados usando sparql. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 15., 2014, Belo Horizonte. **Anais...** Belo Horizonte: UFMG, 2014.

SANTOS P. L. V. A. C.; ALVES, R. C. V. Metadados e Web Semântica para estruturação da Web 2.0 e Web 3.0. **DataGramZero: Revista de Ciência da Informação**, v. 10, n. 6, 2009.

SANTOS, P. L. V. A. C.; PEREIRA, A. M. **Catálogo**: breve história e contemporaneidade. Niterói: Intertexto, 2014.

SAYÃO, L. F. Padrões para bibliotecas digitais abertas e interoperáveis. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, n. esp., p. 18-47, 2007.

SAYÃO, L. F.; MARCONDES, C. H. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. **TransInformação**, v. 20, n. 2, p.133-148, maio/ago. 2008.

SCHEMA.ORG. 2015. Disponível em: <<http://schemers.org/>>. Acesso em: 6 dez. 2016.

SCOPUS. Disponível em: <<http://www.elsevier.com/online-tools/scopus>>. Acesso em: 20 dez. 2015.

SEMANTIC WEB. FOAF. Disponível em: <<http://semanticweb.org/wiki/FOAF.html>>. Acesso em: 15 jul. 2016.

SILVA, M. R.; HAYASHI, C. R. M.; HAYASHI, M. C. P. I. Análise bibliométrica e cientométrica: desafios para especialistas que atuam no campo. **InCID: Revista de Ciência da Informação e Documentação**, v. 2, n. 1, p. 110-129, 2011.

SILVA, R. E. **As tecnologias da Web Semântica no Domínio Bibliográfico**. 2013. 135 f. Dissertação (Mestrado em Ciência da Informação)-Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2013.

SILVEIRA, N. C. **Análise do impacto dos Requisitos Funcionais para Registros Bibliográficos (FRBR) nos pontos de acesso de responsabilidade pessoal**. 2007. 108 f. Dissertação (Mestrado em Ciência da Informação)-Pontifícia Universidade Católica, Campinas, 2007.

SIQUEIRA, J. C. Biblioteconomia, documentação e ciência da informação: história, sociedade, tecnologia e pós-modernidade. **Perspectivas em Ciência da Informação**, v. 15, n. 3, p. 52-66, 2010.

SOUZA, R. R.; ALVARENGA, L. A web semântica e suas contribuições para a ciência da informação. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 132-141, jan./abr. 2004.

SUBIRATS, I.; ZENG, M. L. LOD-BD recommendations 2.0 report on how to select appropriate encoding strategies for producing Linked Open Data (LOD)- enabled bibliographical data. 2015. Disponível em: <[http://aims.fao.org/sites/default/files/LODE-BD-2\\_3.pdf](http://aims.fao.org/sites/default/files/LODE-BD-2_3.pdf)>. Acesso em: 20 dez. 2016.

SVENONIUS, E. The concept of authorship: past and future. In: SVENONIUS, E. **The conceptual foundations of descriptive cataloging**. San Diego: Academic Press, 1989. p. 17-40.

TALAMO, M. F. G. M.; LARA, M. L. G.; KOBASHI, N. Y. Contribuição da terminologia para a elaboração de tesouros. **Ciência da Informação**, v. 21, n. 3, p. 197-200, 1992.

TAYLOR, A. G. **Introduction to cataloging and classification**. 10th ed. Westport: Libraries Unlimited, 2006.

THE UNICODE CONSORTIUM. [2015]. Disponível em: <<http://www.unicode.org/>>. Acesso em: 31 maio 2015.

TILLET, B. B. Authority control at the level. **Association for Library Collection e Technical Services**, v. 44, n. 3, 2000.

TILLET, B. B. **Base Virtual Internacional de Autoridade**. Pretoria: IME ICC, 2007.

THOMSON REUTERS. Web of Science. 2016. Disponível em: <<http://ipscience.thomsonreuters.com/product/web-of-science/>>. Acesso em: 06 ago. 2016.

TRISTÃO, A. M. D.; FACHIN, G.R.B.; ALARCON, O. E. Sistema de classificação facetada e tesouros: instrumentos para a organização do conhecimento. **Ciência da Informação**, v. 33, n. 2, p. 161-171, 2004.

VAN HOOLAND, S.; VERBORGH, R. **Linked Data for libraries, archives and museums**: How to clean, link and publish your metadata. London: Facet, 2014.

VIRTUAL INTERNATIONAL AUTHORITY FILE. 2017. Disponível em: <<https://viaf.org/>>. Acesso em: 1 jan. 2017.

VIRTUAL INTERNATIONAL AUTHORITY FILE. **Machado de Assis, 1839-1908**. 2016d. Disponível em: <[https://viaf.org/viaf/95151633/#Machado\\_de\\_Assis,\\_1839-1908](https://viaf.org/viaf/95151633/#Machado_de_Assis,_1839-1908)>. Acesso em: 1 jan. 2017.

VIRTUAL INTERNATIONAL AUTHORITY FILE. **Brasil**. 2016a. Disponível em: <<https://viaf.org/viaf/142985214/#Brazil>>. Acesso em: 1 jan. 2017.

VIRTUAL INTERNATIONAL AUTHORITY FILE. **São Paulo**. 2016b. Disponível em: <[https://viaf.org/viaf/126223335/#São\\_Paulo\\_\(Brazil\)](https://viaf.org/viaf/126223335/#São_Paulo_(Brazil))>. Acesso em: 1 jan. 2017.

VIRTUAL INTERNATIONAL AUTHORITY FILE. **Universidade Estadual Paulista**. 2016c. Disponível em: <[https://viaf.org/viaf/122646643/#Universidade\\_estadual\\_Paulista](https://viaf.org/viaf/122646643/#Universidade_estadual_Paulista)>. Acesso em: 1 jan. 2017.

WAGNER FILHO, F.; LÓSCIO, B. F. **Web Semântica**: conceitos e tecnologias. 2009. Disponível em:

<<http://www.ufpi.br/subsiteFiles/ercemapi/arquivos/files/minicurso/mc9.pdf>>. Acesso em: 20 jun. 2015.

WAJENBERG, A. S. A cataloger's view of authorship. In: SVENONIUS, E. **The conceptual foundations of descriptive cataloging**. San Diego: Academic Press, 1989. p. 21-27.

WERTHEIM, J. A sociedade da informação e seus desafios. **Ciência da Informação**, v. 29, n. 2, p. 71-77, 2000.

WILLER, M.; DUNSIRE, G. Publishing bibliographic element sets and value vocabularies. In: WILLER, M.; DUNSIRE, G. **Bibliographic information organization in the Semantic Web**. [S.l.]: Chandos Publishing, 2013. p. 137-198.

WORLD WIDE WEB CONSORTIUM. Linked Data. 2014. Disponível em: <<http://www.w3.org/wiki/LinkedData>>. Acesso em: 16 jul. de 2015.

WORLD WIDE WEB CONSORTIUM. Linked Data. 2015. Disponível em: <<http://www.w3.org/standards/semanticweb/data#summary>>. Acesso em: 16 jul. 2015.

WORLD WIDE WEB CONSORTIUM. Use case virtual International Authority File (VIAF). 2005. Disponível em: <[http://www.w3.org/2005/Incubator/ld/wiki/Use\\_Case\\_Virtual\\_International\\_Authority\\_File\\_%28VIAF%29](http://www.w3.org/2005/Incubator/ld/wiki/Use_Case_Virtual_International_Authority_File_%28VIAF%29)>. Acesso em: 30 maio 2015.

WORLD WIDE WEB CONSORTIUM. **URIs, URLs, and URNs: Clarifications and Recommendations 1.0**. W3C/IETF URI Planning Interest Group, 2001. Disponível em: <<http://www.w3.org/TR/uri-clarification/>>. Acesso em: 31 maio 2015.

YADAGIRI, N.; RAMESH, P. Semantic Web and the libraries: an overview. **Internacional Journal of Library Science**, v. 7, n. 1, p. 80-94, 2013.

ZAFALON, Z. R. Representação e metadados em ambientes digitais. 2015. Disponível em: <<https://prezi.com/zclkbogtj2y/representacao-e-metadados-em-ambientes-digitais/>>. Acesso em: 21 mai. 2015.

ZANDHUIS, I. Towards a Genealogical Ontology for the Semantic Web. 2005. Disponível em: <<http://www.zandhuis.nl/sw/genealogy/genont.pdf>>. Acesso em: 06 jul. 2015.

ZENG, M. L.; QIN, J. **Metadata**. New York: Neal-Shuman, 2008.