

UNIVERSIDADE FEDERAL DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

Brisa Pozzi de Sousa

A NOÇÃO DE ASSUNTO E O MODELO CONCEITUAL *FUNCTIONAL*
REQUIREMENTS FOR SUBJECT AUTHORITY DATA (FRSAD)

Belo Horizonte

2019

Brisa Pozzi de Sousa

A NOÇÃO DE ASSUNTO E O MODELO CONCEITUAL *FUNCTIONAL*
REQUIREMENTS FOR SUBJECT AUTHORITY DATA (FRSAD)

Tese apresentada ao Programa de Pós-Graduação em
Ciência da Informação, da Escola de Ciência da
Informação da Universidade Federal de Minas
Gerais, como requisito para obtenção do título de
doutora em Ciência da Informação.

Linha de Pesquisa: Organização e Uso da
Informação.

Orientadora: Professora Cristina Dotta Ortega.

Coorientadora: Professora Maria da Graça Simões.

Belo Horizonte

2019

S725n	<p>Sousa, Brisa Pozzi de.</p> <p>A noção de assunto e o modelo conceitual functional requirements for subject authority data (FRSAD) [recurso eletrônico] / Brisa Pozzi de Sousa. – 2019. 1 recurso online (264 f. : il., color.) : pdf.</p> <p>Orientadora: Cristina Dotta Ortega Coorientadora: Maria da Graça Simões Tese (doutorado) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação.</p> <p>Referências: f. 179-189. Anexo: f. 190-264. Exigências do sistema: Adobe Acrobat Reader.</p> <p>1. Ciência da informação – Teses. 2. Representação do conhecimento (Teoria da informação) – Teses. 3. Cabeçalhos de assunto – Teses. 4. Catálogo de autoridade (Recuperação da informação) – Teses. 5. FRSAD (Modelo Conceitual) – Teses. I. Título. II. Ortega, Cristina Dotta. III. Simões, Maria da Graça. IV. Universidade Federal de Minas Gerais, Escola de Ciência da Informação.</p> <p style="text-align: right;">CDU: 025.4</p>
-------	--

Ficha catalográfica: Biblioteca Profª Etelvina Lima, Escola de Ciência da Informação da UFMG.



UFMG

Universidade Federal de Minas Gerais
Escola de Ciência da Informação
Programa de Pós-Graduação em Ciência da Informação

FOLHA DE APROVAÇÃO

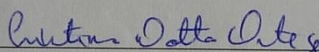
"A NOÇÃO DE ASSUNTO E O MODELO CONCEITUAL FUNCTIONAL REQUIREMENTS FOR SUBJECT AUTHORITY DATA (FRSAD)"

Brisa Pozzi de Sousa

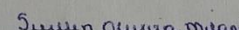
Tese submetida à Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Minas Gerais, como parte dos requisitos à obtenção do título de "**doutora em Ciência da Informação**", linha de pesquisa "**Organização e Uso da Informação**".

Tese aprovada em: 14 de maio de 2019.

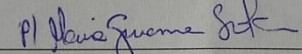
Por:



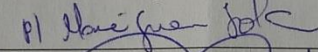
Profa. Dra. Cristina Dotta Ortega - ECI/UFMG (Orientadora)



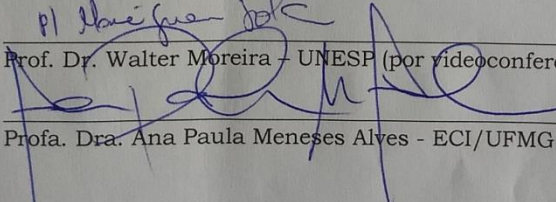
Profa. Dra. Suellen Oliveira Milani - UFF



Prof. Dr. Rodrigo de Sales - UFSC (por videoconferência)

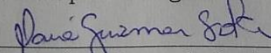


Prof. Dr. Walter Moreira - UNESP (por videoconferência)



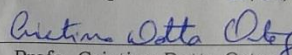
Profa. Dra. Ana Paula Menezes Alves - ECI/UFMG

Aprovada pelo Colegiado do PPGCI



Profa. Maria Guiomar da Cunha Frota
Coordenadora

Versão final aprovada em 19/07/19



Profa. Cristina Dotta Ortega
Orientadora



UFMG

Universidade Federal de Minas Gerais
Escola de Ciência da Informação
Programa de Pós-Graduação em Ciência da Informação

ATA DA DEFESA DE TESE DE **BRISA POZZI DE SOUSA**, matrícula: 2015665620

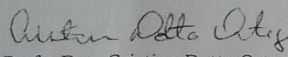
Às 14:00 horas do dia 14 de maio de 2019, reuniu-se na Faculdade de Filosofia e Ciências Humanas da UFMG a Comissão Examinadora aprovada *ad referendum* pelo Colegiado do Programa de Pós-Graduação em Ciência da Informação em 15/04/2019, para julgar, em exame final, o trabalho intitulado **A noção de assunto e o modelo conceitual Functional Requirements for Subject Authority Data (FRSAD)**, requisito final para obtenção do Grau de DOUTORA em CIÊNCIA DA INFORMAÇÃO, área de concentração: Produção, Organização e Utilização da Informação, Linha de Pesquisa: Organização e Uso da Informação. Abrindo a sessão, a Presidente da Comissão, Profa. Dra. Cristina Dotta Ortega, após dar conhecimento aos presentes do teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores com a respectiva defesa da candidata. Logo após, a Comissão se reuniu sem a presença da candidata e do público, para julgamento e expedição do resultado final. Foram atribuídas as seguintes indicações:

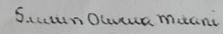
Profa. Dra. Cristina Dotta Ortega - Orientadora	APROVADA
Profa. Dra. Suellen Oliveira Milani	APROVADA
Prof. Dr. Rodrigo de Sales – por videoconferência	APROVADA
Prof. Dr. Walter Moreira – por videoconferência	APROVADA
Profa. Dra. Ana Paula Meneses Alves	APROVADA

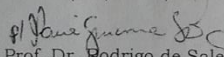
Pelas indicações, a candidata foi considerada APROVADA.

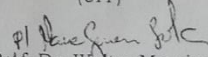
O resultado final foi comunicado publicamente à candidata pela Presidente da Comissão. Nada mais havendo a tratar, a Presidente encerrou a sessão, da qual foi lavrada a presente ATA que será assinada por todos os membros participantes da Comissão Examinadora.

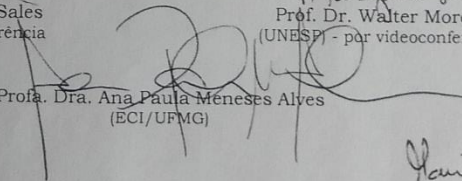
Belo Horizonte, 14 de maio de 2019.


Profa. Dra. Cristina Dotta Ortega
(ECI/UFMG)

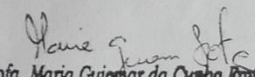

Profa. Dra. Suellen Oliveira Milani
(UFF)


Prof. Dr. Rodrigo de Sales
(UFSC) - por videoconferência


Prof. Dr. Walter Moreira
(UNESP) - por videoconferência


Profa. Dra. Ana Paula Meneses Alves
(ECI/UFMG)

Obs: Este documento não terá validade sem a assinatura e carimbo da Coordenadora.


Profa. Maria Guiomar da Cunha Frota
Coordenadora do Programa de Pós-Graduação
em Ciência da Informação

DEDICATÓRIA

A mulher incrível que me dá o título de filha;

Aos professores com quem convivi desde a infância até o doutorado, com os quais desfrutei o instigante título (que nunca me abandona): de ser aluna;

A(o)s alunos(as) da UNIRIO que me concedem o título de professora; mal sabem o quanto aprendo cotidianamente com todos e todas;

A ‘peculiar’ família Pozzi que me confere títulos incríveis e, não menos importante, ao Casagrande que tenho na minha vida – pilar de apoio e carinho.

AGRADECIMENTOS

Com o incentivo do querido amigo Fabrício J. N. da Silveira (na época 'DEPBiano') a vida de doutoranda inicia em 2015.

Do Rio a BH, da UNIRIO a UFMG,
com a companhia de amigos(as) da pós-graduação, café e pão de queijo
o percurso foi se tornando leve...
idas e vindas...

Mas, do cansaço que teimava em aparecer,
das disciplinas cursadas,
entre ser aluna e ensinar,
apoio e carinho recebi dos(as) discentes que compreenderam o meu percurso.
Idas e vindas...

Era bom estar na UFMG,
mas sentia por não aproveitá-la integralmente: cursos, palestras, eventos...
sua grandeza e potencialidade vai além dos espaços físicos.
Idas e vindas...

Aproveitei outras instituições acadêmicas no Rio:
UFF e IBICT, disciplinas cursadas com
professores instigantes.

Idas e vindas...

Em agosto de 2017 ocorreu a qualificação:
as professoras Cida Moura, Graça Simões e Marilda Lara
com proveitosos comentários ampliaram minha visão:
a pesquisa é mesmo uma atividade coletiva.

Parceria de coorientação foi realizada com a professora Graça Simões:
seus apontamentos foram de grande importância para essa pesquisa - muito obrigada!
Idas e vindas...

Colegas do DEPB incentivaram o meu percurso.

De forma colaborativa, Marcos Miranda e Tatiana de Almeida cobriram meu afastamento em
2018.

Idas e vindas...

Em 2018, uma oportunidade impar: três meses estudando na Universidade de Coimbra.
Orientações da professora Graça, companhia de amigas(os) que lá fiz, encantamento pela
Universidade, pela cultura e culinária portuguesa;
muitas contribuições foram recebidas,
não apenas em nível intelectual, mas também pessoal.

Nessas idas e vindas todo o percurso foi conduzido pela inteligência e humanidade da
professora Cristina Ortega. Seu profissionalismo é impecável! Obrigadíssima!!!

Nesse mesmo percurso coletivo de idas e vindas:
aos professores que gentilmente aceitaram examinar esse trabalho:
meu sincero agradecimento!

Recebi ajuda em diferentes perspectivas; alguns nomes não foram citados, mas guardo com
grande carinho.

Todo apoio foi fundamental para finalizar essa pesquisa
repleta de inquietações, muitas incertezas, mas de bastante aprendizado.
A todos e todas que acompanharam esse percurso: muito obrigada!!!

RESUMO

A noção de assunto em Biblioteconomia e Ciência da Informação é o tema proposto por esta pesquisa, que é delimitada por *corpus* que se caracteriza por uma forma geográfica em vertentes específicas, que marcaram a literatura da área, a partir de duas principais influências: anglo-saxã e europeia. O caráter processual e de significação da informação estabelece conexão com os atributos dos documentos, que são trabalhados por inteiro ou em suas partes, e isso inclui o entendimento de construção de sentido, em detrimento da simples reprodução pela aplicação de normas. Sendo assim, enfatiza-se a necessidade de entendimento sobre a noção de assunto e, contemporaneamente, a presença do modelo conceitual denominado *Functional Requirements for Subject Authority Data* (FRSAD), criado pela *International Federation of Library Associations and Institutions* (IFLA), reforçou a necessidade de compreensão sobre a noção em foco. Não há uma única abordagem sobre o que seja assunto, pois sua compreensão varia de acordo com a influência teórica considerada, mas tendo em vista a importância da sua manipulação, que resulta na representação temática, o problema de pesquisa centrou em investigar se o acúmulo teórico sobre a noção de assunto no processo de representação temática e seus métodos de realização, desenvolvido em diferentes vertentes da área de Biblioteconomia e Ciência da Informação, reflete-se no modelo conceitual FRSAD. Ou seja: é possível evidenciar relação entre a noção de assunto no modelo FRSAD e o desenvolvimento teórico sobre essa noção na área? O objetivo geral consistiu em investigar e analisar criticamente a base teórica e conceitual que envolve a fundamentação da noção de assunto no modelo FRSAD, no âmbito do referencial da área. Para tanto, o percurso metodológico da pesquisa, de abordagem qualitativa, subsidiou a investigação de elementos teóricos, recorrendo-se primeiramente à pesquisa bibliográfica do objeto teórico – a noção de assunto. Em seguida, se deu ênfase à análise e à interpretação do objeto empírico – o modelo FRSAD – com base, essencialmente, na análise de conteúdo de Laurence Bardin. A análise, a partir de categorias, foi moldada em duas etapas interligadas. A primeira baseou-se inteiramente em convergir perspectivas para a conceituação da noção de assunto a partir da literatura trabalhada em três principais momentos (anglo-saxão, desdobramento anglo-saxão e europeu) e, a segunda, em analisar se as categorias da primeira etapa imprimem perspectivas em face da articulação do assunto no modelo FRSAD. A exploração com foco nos momentos teóricos levou a perceber que o modelo FRSAD desconhece as diferentes vertentes da Biblioteconomia e Ciência da Informação. Ele reflete, em alguma medida, aspectos da vertente anglo-saxã. O assunto, enquanto atributo, é uma unidade linguística passível de operação, demandando, portanto, processos de desestruturação e de estruturação, na forma de uma mensagem documentária.

Palavras-chave: Noção de assunto. Representação temática. Modelo FRSAD.

ABSTRACT

The notion of subject in Library and Information Science is the theme proposed by this research, which is delimited by a corpus that is characterized by a geographic form in specific fields, that marked the literature of the area, from two main influences: Anglo-Saxon and European. The procedural and meaningful character of the information establishes connection with the attributes of the documents, which are worked in whole or in part, and this includes the understanding of construction of sense, rather than the simple reproduction by the application of standards. Thus, the need to understand the notion of subject and, presently, the presence of the conceptual model Functional Requirements for Subject Authority Data (FRSAD), created by the International Federation of Library Associations and Institutions (IFLA) is stressed, what is reinforced by the need for understanding about the notion in focus. There is not a single approach to what a subject is, because its understanding varies according to the theoretical influence considered, but given the importance of its manipulation, which results in thematic representation, the research problem consisted in investigating whether the theoretical accumulation on the notion of subject in the process of thematic representation and its methods of realization, developed in different fields of the Library and Information Science area, is reflected in the FRSAD conceptual model. That is: is it possible to evidenciate the relationship between the notion of subject in the FRSAD model and the theoretical development on this notion in the area? The general objective was to investigate and critically analyze the theoretical and conceptual basis that surrounds the foundations of the notion of subject in the FRSAD model, within the framework of the area reference. For that, the methodological path of the research, with a qualitative approach, subsidized the investigation of theoretical elements, resorting firstly to the bibliographic research of the theoretical object – the notion of subject. Thereafter, emphasis on the analysis and interpretation of the empirical object – the FRSAD model – based essentially on Laurence Bardin's content analysis was given. The analysis, from categories, was modeled in two interlinked stages. The first one was based entirely on converging perspectives for the conceptualization of the notion of subject from the literature in three main moments (Anglo-Saxon, deployment Anglo-Saxon and European), and on the second, on analyzing whether the categories of the first stage give perspective in the face of the articulation of the subject in the FRSAD model. The exploration focused on theoretical moments led to the realization that the FRSAD model does not know the different fields of the Library and Information Science. It reflects, to some extent, aspects of the Anglo-Saxon theory. The subject, as an attribute, is a linguistic unit passible of operation, demanding, therefore, processes of destructuring and structuring, in the form of a documentary message.

Keywords: Subject notion. Thematic representation. FRSAD model.

LISTA DE ABREVIATURAS E SIGLAS

FRSAD	<i>Functional Requirements for Subject Authority Data</i> ou Requisitos Funcionais para Dados de Autoridade Assunto
IFLA	<i>International Federation of Library Associations and Institutions</i>
UFMG	Universidade Federal de Minas Gerais
ECI	Escola de Ciência da Informação
FRBR	<i>Functional Requeriments for Bibliographic Records</i> ou Requisitos Funcionais para Registros Bibliográficos
FRAD	<i>Functional Requirements for Authority Data</i> ou Requisitos Funcionais para Dados de Autoridade Assunto
AACR	<i>Anglo-American Cataloguing Rules</i>
ALA	<i>American Library Association</i>
LC	<i>Library of Congress</i>
DDC	<i>Dewey Decimal Classification</i>
CDD	Classificação Decimal de Dewey
PMEST	Personalidade - <i>Personality</i> ; Matéria - <i>Matter</i> ; Energia - <i>Energy</i> ; Espaço - <i>Space</i> ; Tempo - <i>Time</i>
PRECIS	<i>PREserved Context Indexing System</i>
CRG	<i>Classification Research Group</i>
KWIC	<i>Key-word in-context</i>
CC	<i>Colon Classification</i> ou Classificação dos Dois Pontos
UNISIST	<i>United Nations International System for Information in Science and Technology</i>
UNESCO	<i>United Nations Educational, Scientific and Cultural Organization</i>
ISO	<i>International Standardization for Organizatio</i>
ABNT	Associação Brasileira de Normas Técnicas
CDU	Classificação Decimal Universal
IIB	<i>Institut International de Bibliographie</i>
RBU	<i>Répertoire Bibliographique Universel</i>
SYNTOL	<i>Syntagmatic Organization Language</i>
ECA	Escola de Comunicação e Artes
USP	Universidade de São Paulo
Unesp	Universidade Estadual Paulista
ADF	<i>Análisis Documental de la Forma</i>
ADC	<i>Análisis Documental de Contenido</i>
LD	<i>Lingüística Documental</i>
SD	<i>Semiótica Documental</i>
TDC	<i>Tratamiento Documental de Contenido</i>
LN	Linguagem Natural
LD	Linguagem Documentária
LE	Linguagens Especializadas
LRM	<i>Library Reference Model</i>
E-R	Entidade-Relacionamento
FRSAR	<i>Functional Requirements for Subject Authority Records</i> ou Requisitos Funcionais para Registros de Autoridade de Assunto
SKOS	<i>Simple Knowledge Organization System</i>
FRANAR	<i>Functional Requirements and Numbering of Authority Records</i>

LISTA DE FIGURAS

Figura 1 - As etapas do procedimento científico.....	21
Figura 2 - Capa <i>Rules for a Printed Dictionary Catalogue</i>	52
Figura 3 - Folha de rosto <i>Rules for a Printed Dictionary Catalogue</i>	52
Figura 4 - Sistematização dos três Objetivos das <i>Rules</i> de Cutter (1904)....	54
Figura 5 - Definições relacionadas às expressões do termo assunto (<i>subject</i>).....	61
Figura 6 - Elementos que compõem o livro e o documento.....	117
Figura 7 - Relações de assunto do FRSAD propostas a partir dos três grupos de entidades do FRBR.....	155
Figura 8 - Relacionamentos FRSAD.....	161
Figura 9 - Cabeçalho “ <i>Bliss Bibliographic classification</i> ”.....	163
Figura 10 - Relação <i>work-to-thema</i>	163
Figura 11 - Relação <i>thema-to-nomen</i>	164

LISTA DE QUADROS

Quadro 1 - Sistematização da pesquisa.....	16
Quadro 2 - Operadores booleanos.....	23
Quadro 3 - Termos empregados na busca bibliográfica.....	24
Quadro 4 - Relação entre os objetivos específicos e o referencial teórico da pesquisa.....	26
Quadro 5 - Esquematização com foco na análise (<i>analyse</i>).....	129
Quadro 6 - Categorização do Primeiro Momento de Construção da Noção de Assunto: o contexto anglo-saxão no século XIX.....	144
Quadro 7 - Categorização do Segundo Momento de Construção da Noção de Assunto: desdobramento do pensamento anglo-saxão no século XX.....	145
Quadro 8 - Categorização do Terceiro Momento de Construção da Noção de Assunto: a abordagem europeia no século XX.....	146
Quadro 9 - O estudo e o modelo FRASAD com as categorias resultantes dos momentos teóricos de construção da noção de assunto: convergências.....	170

SUMÁRIO

1 INTRODUÇÃO	10
2 PERCURSO METODOLÓGICO	19
3 PRECEDENTES DA CONSTRUÇÃO DA NOÇÃO DE ASSUNTO: AS INFLUÊNCIAS DOS AUTORES DO SÉCULO XIX	32
3.1 AS PRIMEIRAS PROPOSTAS NA EUROPA.....	32
3.2 A CONTRIBUIÇÃO DO FRANCÊS ADRIEN BAILLET NO SÉCULO XVII	39
4 PRIMEIRO MOMENTO DE CONSTRUÇÃO DA NOÇÃO DE ASSUNTO: O CONTEXTO ANGLO-SAXÃO NO SÉCULO XIX	45
4.1 AS INFLUÊNCIAS DE ANTHONY PANIZZI E CHARLES JEWETT	45
4.2 CHARLES CUTTER E A SISTEMATIZAÇÃO DOS CABEÇALHOS DE ASSUNTO.....	49
5 SEGUNDO MOMENTO DE CONSTRUÇÃO DA NOÇÃO DE ASSUNTO: DESDOBRAMENTO DO PENSAMENTO ANGLO-SAXÃO NO SÉCULO XX	67
5.1 JULIUS KAISER E SHIYALI RANGANATHAN: A DECOMPOSIÇÃO DO ASSUNTO.....	67
5.2 PESQUISADORES DO <i>CLASSIFICATION RESEARCH GROUP</i> E CALVIN MOOERS: A FACETA E O DESCRITOR.....	75
5.3 A INDEXAÇÃO E O ÍNDICE.....	81
5.4 CONCEITO E ANÁLISE CONCEITUAL	86
5.5 A NORMALIZAÇÃO DA INDEXAÇÃO	90
5.6 A DISPERSÃO TERMINOLÓGICA: <i>ABOUTNESS, TOPIC, CONTENT, SUBJECT</i> – ASSUNTO, TEMA, MATÉRIA – <i>CONTENIDO, MATERIA, EPÍGRAFE, TEMA</i>	95
6 TERCEIRO MOMENTO DE CONSTRUÇÃO DA NOÇÃO DE ASSUNTO: A ABORDAGEM EUROPEIA NO SÉCULO XX	111
6.1 A CONTRIBUIÇÃO DE PAUL OTLET	111
6.2 SUBSÍDIO DA ANÁLISE DOCUMENTÁRIA EM JEAN-CLAUDE GARDIN	124
6.3 A INFLUÊNCIA DA ANÁLISE DOCUMENTÁRIA: O RECORTE ESPANHOL E BRASILEIRO.....	131
7 EM FOCO A NOÇÃO DE ASSUNTO E O MODELO FR SAD: DESCRIÇÃO E ANÁLISE	142
7.1 ANÁLISE SOBRE A NOÇÃO DE ASSUNTO.....	143
7.2 O MODELO FR SAD: APRESENTAÇÃO E ESTUDOS	150
7.3 ANÁLISE SOBRE A NOÇÃO DE ASSUNTO NO FR SAD	167
8 CONSIDERAÇÕES FINAIS	174
REFERÊNCIAS	179
ANEXO – MODELO <i>FUNCTIONAL REQUIREMENTS FOR SUBJECT AUTHORITY DATA (FRSAD)</i>	190

1 INTRODUÇÃO

A representação de documentos em ambientes informacionais, como bases de dados, catálogos de bibliotecas, repositórios, dentre outros, visa à recuperação do registro que, por sua vez, remete ao documento, possibilitando seu acesso. Não é novidade afirmar que esse cenário é fundamentado na Biblioteconomia e Ciência da Informação e a promoção da organização da informação é um dos pilares da área.

Tomando como referência a organização da informação, diferentes aspectos podem ser salientados e, dentre eles, a entidade como algo que pode ser distintamente identificada. Considerando o documento uma entidade real, atributos são as propriedades que o descrevem, ou seja, a(s) característica(s) própria(s) que diferencia(m) um objeto, portanto uma entidade da outra. O documento por inteiro e/ou suas partes formam as entidades que são exploradas na organização da informação, a partir da interpretação, ação que envolve a atribuição de valor informativo com o propósito de promover o uso qualificado da informação.

Vários atributos podem ser representados em um registro, que é composto por partes, e para sua elaboração, uma estrutura deve ser seguida de acordo com a política informacional institucionalizada. A representação do assunto (ou conteúdo) em um registro resulta em um conjunto de termos que exercem a função de ponto de acesso a esse registro. Há mais de um século, foi proposta a entrada por assunto em fichas catalográficas compondo um lugar no catálogo. Com o registro on-line, a função da entrada por assunto continuou relevante, pois o contexto automatizado dos computadores não dispensa os fundamentos anteriores de organização da informação, mas reafirma a necessidade de resgate do que foi construído ao longo do tempo e, olhar para isso, reforça a compreensão sobre como a área está pautada.

Os catálogos de bibliotecas, por vezes, foram compreendidos como ferramentas para inventariar acervos, mas na Antiguidade, e em grande parte da Idade Média, a história da Catalogação mostra que não é explícita a diferença entre catálogos e inventários. Foi no final do século XV, pelo avanço das atividades bibliográficas e, no século XVIII em diante, que o catálogo passou a ser uma ferramenta de busca e recuperação. Esse contexto é permeado pelo desenvolvimento e crescimento de bibliotecas no contexto europeu.

Nesse desenvolvimento, a produção dos registros é realizada pelos aspectos do assunto abarcado pelo processo de catalogação. A Catalogação inclui a Catalogação Descritiva e a Catalogação de Assunto, as quais respondem pela construção completa do catálogo. No entanto, segundo Guimarães (2009), o aspecto temático (ou do assunto) envolve mais de uma perspectiva: além da Catalogação de Assunto, pode-se falar em Indexação e em Análise

Documentária. Vertentes constituíram-se com suas respectivas abordagens, mas elas não existem em separado, tornando necessária a reflexão sobre o que é mais consistente e funcional, haja vista que essas diferenças resultam na formação de conceitos da área. Sem esse entendimento, além de uma confusão conceitual, os fundamentos acabam restringidos a um patamar de processamento técnico, pecando contra a potencialidade de instaurar uma efetiva comunicação entre documentos e usuários(as).

Contemporaneamente, no Brasil, para a representação da informação, são adotados os termos representação descritiva e representação temática, que indicam as etapas para a construção dos registros de informação, quanto aos processos específicos em causa. Por exemplo, a primeira ocupa-se dos aspectos formais que permitem identificar o documento e também pode ser chamada de catalogação descritiva ou somente catalogação. Já a segunda opera com os processos de indexação e elaboração de resumos. Mas, em alguns currículos dos cursos de Biblioteconomia no Brasil, é possível encontrar os termos Representação Descritiva e Representação Temática de forma articulada, respectivamente, com a Catalogação e com a Indexação. No entanto, alguns instrumentos como os sistemas de classificação podem compor os currículos enquanto Sistemas de Organização do Conhecimento (SOCs), mas esse termo é empregado para abordar várias linguagens (como as de indexação e as de classificação). Nesse sentido, é preciso considerar as diferenças entre os processos e os instrumentos.

Importante ressaltar que a classificação de documentos trabalha com vários atributos, dentre eles, o assunto. A diferença entre classificação e indexação pode ser evidenciada pela constatação de que a primeira é um método para a ordenação ou arranjo de documentos, do que decorre que vários de seus atributos são considerados. Já na indexação, opera-se com a atribuição de conteúdos temáticos (assuntos) ao documento. Desse modo, os produtos também são distintos: notações classificatórias (a depender do sistema adotado) e termos de indexação, respectivamente.

Ainda relativo ao contexto brasileiro, o termo representação temática também é adotado de forma ampla para referir-se aos processos de indexação e elaboração de resumos. Apesar da complexidade teórica e prática, a construção de um registro de informação gera representações que referenciam o documento por diferentes aspectos, dentre eles o temático. Assim, na indexação, pontos de acesso por assunto remetem ao registro que, por sua vez, referenciam os documentos representados. Na mesma medida de importância, um texto objetivo e conciso do conteúdo do documento pode ser elaborado no registro, sendo chamado de resumo, com o propósito do(a) usuário(a) selecionar documento(s) que gostaria de acessar.

Eventualmente, as palavras do resumo podem compor o índice de busca, mas ele não existe em função disso.

Do ponto de vista da busca, o registro de informação pode promover vínculo(s) dos(as) usuários(as) com os documentos representados, conduzindo a um processo de comunicação em ambientes informacionais, neste caso, documentários. O resultado das atividades documentárias, caracterizadas como informativas e disponibilizadas a indivíduos em situações concretas, é obtido através da busca e recuperação via sistemas, devendo se considerar que a “[...] informação é um conceito subjetivo, mas não fundamentalmente em um sentido individual.” (CAPURRO; HJØRLAND, 2007, p. 192). A operação de busca é um processo de manipulação dos recursos em base de dados, enquanto a recuperação é a obtenção de registros de documentos de interesse.

A informação tem caráter processual e de significação. Nos processos documentários, os atributos dos documentos são trabalhados por inteiro ou em suas partes a partir da atribuição de valores. Quando reconhecido dessa forma, o processo global de construção do sistema, o que inclui a manipulação da informação documentária, é compreendido como construção de sentido, em detrimento de simples reprodução pela aplicação de normas. Com isso, enfatiza-se a preocupação de entendimento sobre a noção de assunto, pois é um atributo envolvido por diferentes perspectivas.

Nos dias de hoje, a palavra assunto também pode se relacionar a diferentes contextos, como por exemplo, ‘o assunto que está na mídia’, ‘um assunto da vida particular’, ‘ir a um buscador e recuperar um assunto’, ‘o assunto de um e-mail’, dentre uma infinidade de possibilidades. Já, assunto, na Biblioteconomia e Ciência da Informação, assume diferentes aspectos que ajudam a pensar a sua manipulação, mas deve ser entendido com a teoria que o sustenta e não isoladamente.

O assunto nos registros de informação, por exemplo, é produto documentário otimizado que, baseado em procedimentos metodológicos, permite o acesso, propiciando o uso e, assim, a ‘função social’ que caracteriza a Biblioteconomia e Ciência da Informação (CAPURRO; HJØRLAND, 2007).

A discussão sobre a noção de assunto na referida área vincula-se a diferentes vertentes que identificam e discutem o aspecto de conteúdo (ou temático) e, contemporaneamente, a presença do modelo conceitual denominado *Functional Requirements for Subject Authority Data*¹ (FRSAD), criado pela *International Federation of Library Associations and*

¹ Requisitos Funcionais para Dados de Autoridade de Assunto.

*Institutions*² (IFLA), reforçou a necessidade de compreensão sobre a natureza e finalidade da representação temática. Sendo assim, observa-se a necessidade de fundamentar a noção de assunto, objeto teórico desta investigação, conduzindo a esforços na compreensão do modelo FRSAD, constituído como objeto empírico nesta pesquisa.

Por isso, busca-se compreender se o modelo FRSAD reflete o acúmulo da área quanto à ideia de assunto no processo de representação temática e seus métodos de realização, ou, se seria uma nova proposta, com pouca aderência com a área.

Ambientes como bibliotecas e centros de documentação fazem uso de instrumentos para organização e busca da informação, situação que envolve a forma como a noção de assunto é encarada pelos profissionais que atuam nesses lugares. Nesta pesquisa, procura-se levantar um quadro teórico-metodológico que permita contribuir ao debate da noção indicada, pois diferentes momentos marcaram aquilo que se efetivou como assunto.

Amplia-se, desse modo, a percepção de que técnicas de representação do assunto devem ser subsidiadas por fundamentos da área, com abordagens que privilegiem aspectos teórico-conceituais. Tais fundamentos partem da Biblioteconomia como também da Documentação.

Assim, o foco é delimitado pelo assunto atribuído ao texto do documento. Não serão discutidos temas que digam respeito à construção de linguagens documentárias e de recuperação da informação, mas serão indicados, de algum modo, como forma de subsidiar a noção investigada.

Optou-se pelo aspecto cronológico e, em certa medida, geográfico que contemplasse o universo da pesquisa em vertentes específicas, que marcaram a literatura da Biblioteconomia e Ciência da Informação, a partir de duas principais influências: anglo-saxã e europeia (essa de origem francesa). Nesse cenário, são relevantes várias discussões, como as que emergem sobre a análise de assunto, análise conceitual e análise documentária, mas sem a pretensão de esgotá-las. A exploração mostra-se necessária, nesta pesquisa, para propiciar o entendimento da noção de assunto no modelo FRSAD, haja vista que abordagens teóricas diversas representam a construção realizada na área. Desse modo, não se trata de uma abordagem sobre a análise de assunto, mas de como o assunto, em suas diversas linhas teóricas, foi se conformando na Biblioteconomia e Ciência da Informação em torno do contexto temático.

Propostas de caráter operacional podem ser empregadas de maneira impensada, precipitadas e sem reflexão. A abordagem científica contempla discussão crítica que permite

² Federação Internacional de Associações e Instituições Bibliotecárias.

contextualizar o referencial teórico sobre a noção de assunto que, por consequência, tem reflexo no ensino e no trabalho de bibliotecários(as). Nesse contexto, deve haver fundamentação responsável por embasar a teoria e a prática profissional.

Levando em conta que, a partir da necessidade de socializar o conhecimento, os instrumentos e produtos de organização da informação se desenvolveram embasados por fundamentos, o mérito da Biblioteconomia passa de guardião dos documentos acumulados pela humanidade à prestadora de serviço de acesso e uso dos documentos. A qualidade de prestadora não induz necessariamente à efetividade do uso, mas exige a preocupação com elementos representativos do conteúdo do documento. Assim, do decorrer do século XIX à proposta do modelo FRSAD no final do século XX, observam-se manifestações que demarcam o interesse que a pesquisa persegue sobre a noção de assunto, tratados sob fundamentos que constituem, principalmente, os termos assunto, conteúdo, tema, *aboutness* e conceito. São termos do quadro referencial da área que envolvem a reflexão sobre a representação temática, constituindo tópicos essenciais tanto no ensino quanto na prática profissional. Ademais, foram selecionados outros termos a partir de três principais idiomas – português, espanhol e inglês – com o empreendimento da literatura e de teóricos que marcaram o cenário temático.

Nessa investigação, discutir a noção de assunto e, mais especificamente, compreender o *corpus* que a caracteriza, com o propósito de alcançar o modelo FRSAD, justifica-se pela dimensão teórica. Na mesma medida, uma reflexão conceitual sobre o assunto também se torna uma forma de contribuição.

Não há uma única abordagem sobre o que seja o assunto, pois sua compreensão varia de acordo com a influência teórica considerada. Tendo em vista a importância do assunto e da sua manipulação, que resulta na representação temática, pode-se indagar: o acúmulo teórico sobre a noção de assunto no processo de representação temática e seus métodos de realização, desenvolvido em diferentes vertentes da área de Biblioteconomia e Ciência da Informação, reflete-se no modelo conceitual FRSAD? Ou seja: é possível evidenciar relação entre a noção de assunto no modelo FRSAD e o desenvolvimento teórico sobre essa noção na área?

Como pressuposto, entende-se que os direcionamentos cobertos pela representação temática no modelo FRSAD devem constituir embasamento capaz de explicitar a noção de assunto, pois, caso contrário, o objetivo do modelo como proposta de âmbito internacional para o avanço no tema pode se perder. Portanto, a discussão crítica é necessária para a validação do modelo.

Nesse cenário, o objetivo geral consiste em investigar e analisar criticamente a base teórica e conceitual que envolve a fundamentação da noção de assunto no modelo FRSAD, no âmbito do referencial da Biblioteconomia e Ciência da Informação. Para tanto, os objetivos específicos buscam:

- a) Identificar e relacionar as construções teórico-conceituais pioneiras relativas à noção de assunto, no que tange ao processo, aos instrumentos e aos produtos documentários;
- b) Explorar a construção conceitual e a centralidade da noção de assunto a partir de um quadro teórico que se integre às propostas de representação temática, a qual não pode ser realizada devidamente na ausência de fundamentos com diferentes influências teóricas; e
- c) Descrever e analisar a fundamentação do modelo FRSAD por meio da investigação sobre a condição teórica da noção de assunto.

Os objetivos propostos estruturam todo o percurso do trabalho e vinculam-se aos momentos teóricos trabalhados e evidenciados na pesquisa quanto à constituição temática em Biblioteconomia e Ciência da Informação.

Ademais, investigar a noção de assunto na área envolve o percurso científico da autora iniciado no mestrado (2010), pois a dissertação concluída em 2012 centrou no processo de análise de assunto. Foi a partir da investigação deste processo que se fez presente a inquietação sobre a noção em foco. Outra motivação decorre de inquietações de estudos e leituras que subsidiassem a atividade docente, mas embora ela não constitua objeto dessa pesquisa, investigar os fundamentos teórico-conceituais sobre a noção de assunto é necessário ao resgate da literatura a ser discutida na área para a formação de futuros profissionais. A pesquisa, ao mesmo tempo que trabalha com um amplo universo teórico, também proporciona o resgate de autores substanciais da representação temática e, assim, da teoria.

Inúmeros são os referenciais que abordaram o aspecto do assunto e a escolha do que foi adotado implicou na seleção de literatura de cunho internacional, mas também nacional, situação que em alguma medida confirma a vinculação com as duas vertentes assinaladas (anglo-saxã e europeia) por uma dimensão temporal e geográfica. Notou-se que não há escassez de literatura, mas a necessidade de sistematização de diferentes abordagens tornou-se um aspecto evidente na tese.

Esse resgate, além de demonstrar como o assunto em diferentes momentos é referendado, também colabora para sua percepção histórica. A área é, antes de tudo, moldada por trajetórias do passado. A seguir, o quadro 1 possibilita uma visão geral de sistematização da pesquisa.

Quadro 1 - Sistematização da pesquisa

Problema	O acúmulo teórico sobre a noção de assunto no processo de representação temática e seus métodos de realização, desenvolvido em diferentes vertentes da área de Biblioteconomia e Ciência da Informação, reflete-se no modelo conceitual FRSAD? Ou seja: é possível evidenciar relação entre a noção de assunto no modelo FRSAD e o desenvolvimento teórico sobre essa noção na área?	Como pressuposto, entende-se que os direcionamentos cobertos pela representação temática no modelo FRSAD devem constituir embasamento capaz de explicitar a noção de assunto, pois, caso contrário, o objetivo do modelo como proposta de âmbito internacional para o avanço no tema pode se perder. Portanto, a discussão crítica é necessária para a validação do modelo.
Objetivo Geral	Investigar e analisar criticamente a base teórica e conceitual que envolve a fundamentação da noção de assunto no modelo FRSAD, no âmbito do referencial da Biblioteconomia e Ciência da Informação.	
Objetivo Específico a)	Identificar e relacionar as construções teórico-conceituais pioneiras relativas à noção de assunto, no que tange ao processo, aos instrumentos e aos produtos documentários;	Seções 3, 4, 5, 6 e respectivas subseções buscam dar conta dos dois primeiros objetivos específicos
Objetivo Específico b)	Explorar a construção conceitual e a centralidade da noção de assunto a partir de um quadro teórico que se integre às propostas de representação temática, a qual não pode ser realizada devidamente na ausência de fundamentos com diferentes influências teóricas; e	
Objetivo Específico c)	Descrever e analisar a fundamentação do modelo FRSAD por meio da investigação sobre a condição teórica da noção de assunto.	
Considerações Finais		

Fonte: elaborado pela autora.

Considerando que o percurso metodológico possibilita a investigação de elementos teóricos, a proposta da pesquisa é de abordagem qualitativa, recorrendo-se primeiramente à pesquisa bibliográfica do objeto teórico – a noção de assunto. Em seguida, se dá ênfase à análise e à interpretação do objeto empírico – o modelo FRSAD – com base, essencialmente, na análise de conteúdo de Laurence Bardin.

Cabe destacar que as traduções das citações em inglês, espanhol e francês foram indicadas como ‘tradução nossa’ e os trechos no idioma original foram inseridos em notas de

rodapé. Outra escolha que ocasionou o uso de notas foi a inclusão de referências, que em certa medida, explicitam algumas ideias, mas não constituíram fontes diretas do quadro teórico estabelecido na tese. Por sua vez, em seção própria, realizou-se a listagem das referências utilizadas.

Também cabe mencionar que no último ano de desenvolvimento da pesquisa (2018), fato que contribuiu para a coleta de material bibliográfico foi a realização do intercâmbio de três meses na Universidade de Coimbra, em Portugal, através da parceria com a coorientadora, professora Graça Simões. De outro modo, a experiência ampliou as vivências acadêmica e profissional da pesquisadora.

Em busca de cumprir os objetivos delineados, a pesquisa estruturou-se de forma que, após a ‘Introdução’, na seção dois, é apresentado o ‘Percurso Metodológico’. Para expor os ‘Precedentes da Construção da Noção de Assunto’, a seção três identifica as influências anteriores aos autores do século XIX, em que é tratada a gênese da noção de assunto no contexto europeu. A seção quatro expõe a compreensão concernente ao ‘Primeiro Momento de Construção da Noção de Assunto’, com vistas ao contexto anglo-saxão do século XIX, marcado pela interseção de teóricos como Anthony Panizzi, Charles Jewett e Charles Cutter. O último deu sustentação às listas de cabeçalhos de assunto.

Dada a complexidade de sistematização do quadro teórico, na seção cinco, o ‘Segundo Momento de Construção da Noção de Assunto’ desdobrou-se em torno do pensamento anglo-saxão no século XX. Correlacionam-se as ideias de decomposição do assunto a partir de Julius Kaiser e Shiyali Ranganathan; a influência do último no *Classification Research Group* e, após, a abordagem do descritor por Calvin Mooers, pois cabeçalhos de assunto não seriam eficazes, como se entendeu. Em seguida, apresenta-se, respectivamente, a Indexação e o Índice, devido a outros tipos de materiais e toda a produção bibliográfica, produzida pelos ingleses. Outra interação que se faz presente é o Conceito e a Análise Conceitual, para mostrar o momento em que o assunto passa a ser construído em função do termo-conceito, mas não por uma extração de palavras do texto, como foi colocado no caso da composição dos cabeçalhos de assunto. A operacionalização do percurso é exposta no movimento da Normalização da Indexação, que é procedimental, mas não é unificada, pois a última subseção resgata a Dispersão Terminológica pertinente ao quadro da representação temática.

Na seção seis, destaca-se o ‘Terceiro Momento de Construção da Noção de Assunto’, de modo a contextualizar a abordagem europeia no século XX. É relacionada a contribuição de Paul Otlet e Jean-Claude Gardin e a influência deste quadro teórico no recorte espanhol e

brasileiro. O assunto nessa vertente é aprofundado com o aporte da Linguística e da Terminologia.

A parte empírica da pesquisa é realizada na seção sete - 'Em Foco a Noção de Assunto e o Modelo FRSAD'. Inicialmente, a subseção comporta a primeira dimensão da análise que está baseada nos momentos teóricos da construção da noção de assunto; na subseção seguinte é apresentado o modelo FRSAD e alguns estudos; e, a segunda dimensão da análise, última subseção, dispõe a compreensão da noção de assunto no FRSAD. Por fim, na seção oito, as 'Considerações Finais' resultantes da investigação.

2 PERCURSO METODOLÓGICO

Para o desenvolvimento do percurso teórico-metodológico de uma pesquisa, a análise de textos pertinentes ao trabalho torna-se essencial, como modo de discutir elementos que expressem o quadro a ser alcançado. Faz-se necessário o uso de ferramentas de diálogo entre autores que balizem a discussão, na perspectiva de que nenhum percurso investigativo é neutro.

De forma ampla, a natureza do assunto na Biblioteconomia e Ciência da Informação pode ser problematizada por diferentes maneiras, pois ao longo da história não houve uma única forma de abordá-lo, ou seja, o processo de construção do assunto não foi linear. Portanto, explorar parâmetros cronológicos e geográficos auxilia na busca pelos aspectos essenciais necessários ao entendimento do que se investiga. Assim, “a problemática é a abordagem ou a perspectiva teórica que decidimos adoptar para tratarmos o problema formulado pela pergunta de partida. É uma maneira de interrogar os fenómenos estudados.” (QUIVY; CAMPENHOUDT, 2005, p. 89).

Nessa investigação, o foco envolve a noção de assunto, que tem sustentação no volume de literatura científica da Biblioteconomia e Ciência da Informação configurada por um conjunto relevante de fundamentos. A partir desse volume, optou-se em agrupar a discussão da pesquisa por momentos teóricos que subsidiaram a estrutura apresentada no sumário.

Em relação ao objeto teórico aqui estudado, o percurso metodológico pauta-se por reflexões que desencadeiem o desenvolvimento da área, pois o propósito é produzir novas indagações e articulações de ideias. “Uma investigação é, por definição, algo que se procura. É um caminhar para um melhor conhecimento e deve ser aceite como tal, com todas as hesitações, desvios e incertezas que isso implica.” (QUIVY; CAMPENHOUDT, 2005, p. 31).

No contexto da referida procura e, considerando a carga de investigação de elementos teóricos, a proposta da pesquisa é de abordagem qualitativa, recorrendo-se primeiramente a pesquisa bibliográfica do objeto teórico – a noção de assunto. Na análise do objeto empírico – o modelo FRSAD – o método de descrição analítica é empregado e se dá ênfase à análise e à interpretação, com base, essencialmente, na análise de conteúdo de Laurence Bardin. Acresce referir que quando se justifique necessário, considerar-se-á o uso da comparação, quer para convergir quer para divergir a conjugação de um quadro teórico para a análise.

Segundo apontam Martins e Theóphilo (2016, p. 141), o aspecto analítico permite “[...] a articulação entre os pressupostos teóricos do estudo e os dados da realidade” ajudando a sistematizar “a coleta de dados e favorecendo a análise destes.”

De acordo com Marconi e Lakatos (2003), na forma qualitativa, além de prevalecer interação constante com todo o processo investigativo, ancora-se a análise e a interpretação como duas atividades distintas, mas constituintes do núcleo central de investigação. A análise busca evidenciar as relações entre fenômenos estudados e fatores envolvidos e, a interpretação, dá significado ao material analisado relacionando-se aos objetivos propostos.

Para Martins e Theóphilo (2016), na pesquisa qualitativa, a análise compõe três atividades interativas e contínuas: análise do material coletado; apresentação da análise; e delineamento e busca das conclusões. A análise envolve um processo sequencial de seleção, simplificação, abstração e estruturação. Já a apresentação contempla a organização das ideias provenientes da análise, propiciando a possibilidade de tomada de decisões teóricas de forma que seja possível seguir para a próxima atividade. Sendo assim, o delineamento e busca das conclusões proporciona o desfecho contextual da investigação de forma a manter um percurso rico de interpretação e validação dos resultados.

Nesse contexto, a abordagem metodológica versa sobre a análise e a interpretação de posicionamentos teóricos sobre a noção de assunto e como a literatura vêm tratando as ideias sobre a representação temática com a influência de diferentes termos, como por exemplo, conteúdo, conceito, tema e *aboutness*. Assim, a abordagem busca nortear a análise e interpretação a respeito das variadas formas que a noção de assunto assume em Biblioteconomia e Ciência da Informação.

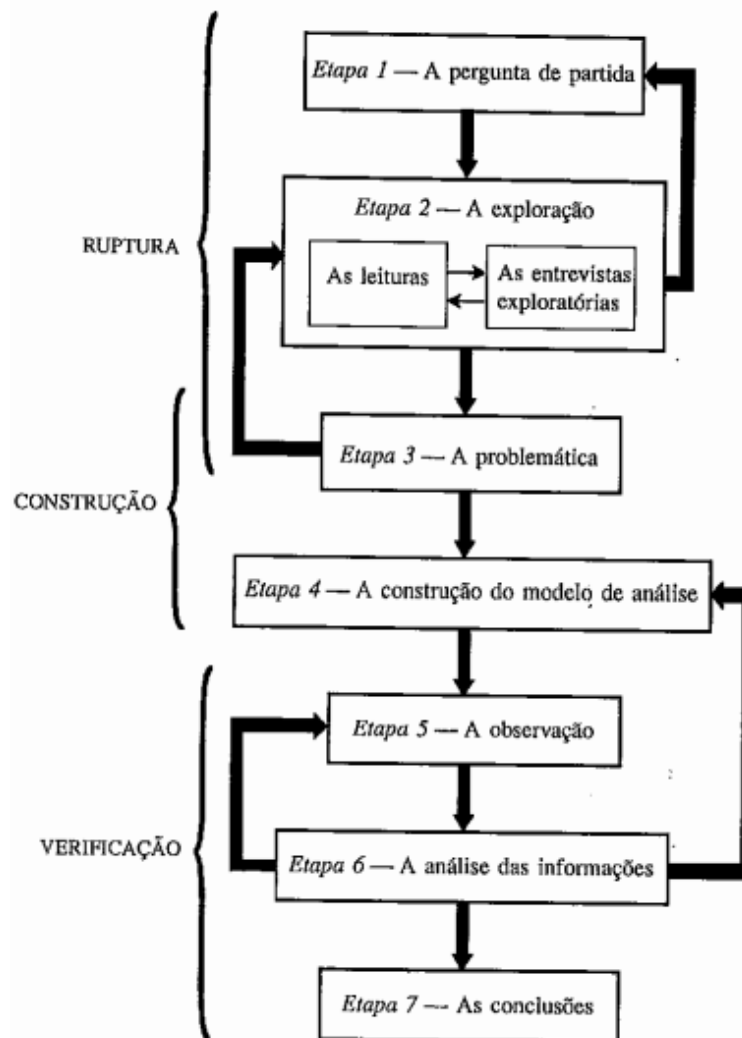
Exposta a condução, pois as dificuldades conceituais estão presentes, e retomando a proposta do percurso metodológico, Quivy e Campenhoudt (2005, p. 26) afirmam que a pesquisa pode se fundamentar na articulação de três “actos do procedimento científico” que não são independentes uns dos outros e se constituem mutuamente. Os atos são: a ruptura, a construção e a verificação fundamentados ao curso de sete etapas que formam circuitos em permanente interação.

O primeiro ato, a ruptura, consiste em romper com os preceitos e as falsas evidências de compreensão dos fatos e envolve as etapas: 1- a pergunta de partida; 2 - a exploração (que ocorre, por exemplo, a partir das leituras e entrevistas exploratórias); e 3 - a problemática. O segundo ato, a construção, efetua-se a partir de um sistema conceitual organizado que fundamenta a base do fenômeno estudado. É a construção que possibilita a experimentação e o estabelecimento de quadros teóricos. Por conta da permanente interação, relaciona-se a

etapa 3 - a problemática; e a etapa 4 - a construção do modelo de análise. O terceiro e último ato, a verificação, corresponde ao estatuto científico que a pesquisa assume envolvendo as designações pela verificação ou experimentação. Envolve as etapas 5 - a observação; 6 - a análise das informações; e 7 - as conclusões.

Os três atos – ruptura, construção e verificação – corporificam-se pelas etapas indicadas e, para essa pesquisa, ratifica o uso da abordagem qualitativa, responsável em subsidiar a análise e a discussão do problema. Para ilustrar, segue o esquema na figura 1.

Figura 1 – As etapas do procedimento científico



Fonte: Quivy e Campenhoudt (2005, p. 27).

No que concerne à retroalimentação das etapas e à dependência instaurada, destaca-se a quarta, construção do modelo de análise, momento em que o referencial teórico é corporificado pelo emprego da pesquisa bibliográfica, portanto, pela técnica procedimental.

De acordo com Cervo, Bervian e Silva (2007, p. 60), pesquisa bibliográfica é aquela que “procura explicar um problema a partir de referências teóricas publicadas em artigos, livros, dissertações e teses.” No uso das fontes secundárias indicadas, “[...] a pesquisa bibliográfica não é mera repetição do que já foi dito ou escrito sobre certo assunto, mas propicia o exame de um tema sob novo enfoque ou abordagem [...]” (MARCONI; LAKATOS, 2003, p. 183).

Com o propósito de discutir os objetivos elencados e, a fim de viabilizar a análise de referenciais teóricos relacionados à noção de assunto, foram realizadas pesquisas bibliográficas em fontes devotadas à Biblioteconomia e Ciência da Informação. Foram utilizadas tanto fontes nacionais quanto internacionais, destacando-se as seguintes:

- Repositório e-LIS (*e-prints in Library & Information Science*): <http://eprints.rclis.org/>
- Repositório Questões em Rede que abriga a coleção BENANCIB:
<http://repositorios.questoesemrede.uff.br/repositorios/>
- BRAPCI (Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação):
<http://www.brapci.inf.br/>
- SciELO (*Scientific Electronic Library Online*): <http://www.scielo.org>
- Redalyc (*Red de Revistas Científicas de América Latina y el Caribe, España y Portugal*):
<http://www.redalyc.org>
- Dialnet: <https://dialnet.unirioja.es/>
- Catálogo de teses e dissertações da Capes: <http://catalogodeteses.capes.gov.br/catalogo-teses/>
- Portal de Periódicos da Capes, com foco nas bases de dados que disponibilizam textos completos:
 - LISA (*Library & Information Science Abstract*)
 - *Library, Information Science & Technology Abstracts with Full Text*

Além dessas fontes, a busca de referências nos acervos da Biblioteca Etelvina Lima, da Escola de Ciência da Informação (ECI) da UFMG e, da Biblioteca Geral da Universidade de Coimbra e Biblioteca da Faculdade de Letras da mesma Universidade contribuíram para a recuperação de obras que não estão disponíveis on-line.

Cabe destacar que, quando possível, foram empregadas estratégias de busca por assunto a partir de operadores booleanos (*and, or, not*), nas opções de pesquisa das fontes consultadas. Uma vez que os operadores booleanos definem relações entre termos em uma consulta, há a possibilidade de expandir a busca no contexto de interesse e obter resultados com precisão. O direcionamento da busca partiu da estratégia de pesquisa por assunto,

disponível tanto no Portal de Periódicos da CAPES quanto nos catálogos on-line das bibliotecas utilizadas. No quadro 2, seguem os operadores empregados nas buscas. Mas, ainda cabe destacar que pesquisas por autores e títulos também foram amplamente realizadas, dispensando o uso dos operadores.

Quadro 2 – Operadores booleanos

<i>And</i> (e)	combina os termos da pesquisa para que cada resultado contenha todos os termos;
<i>Or</i> (ou)	combina os termos da pesquisa para que cada resultado contenha, no mínimo, um dos termos;
<i>Not</i> (não)	exclui termos para que cada resultado não contenha nenhum dos termos que o seguem.

Fonte: elaborado pela autora.

Em relação à seleção dos materiais bibliográficos nas fontes indicadas, considerou-se os termos das buscas no campo assunto em três idiomas (português, espanhol e inglês) empregados conforme seguem no quadro 3. Além disso, a estratégia de busca centrou na recuperação de materiais que subsidiassem os aspectos teóricos considerados de maior influência no ocidente, de origem anglo-saxã e europeia, compreendidos na trajetória de construção da noção de assunto. Cabe ressaltar que as buscas foram realizadas em diferentes períodos que cobriram os anos de doutoramento, mas o turno principal ocorreu após a apresentação do relatório de qualificação, em setembro de 2017, por decorrência do amadurecimento das discussões efetivadas pela banca.

Chegou-se aos termos de busca (quadro 3) a partir da vivência teórica-exploratória com a literatura. Nesse primeiro momento, foram observadas diferentes palavras que se relacionam com o termo assunto na Biblioteconomia e Ciência da Informação. Também se observou, em alguns casos, o emprego de forma sinônima para os termos análise de assunto e análise documentária, porém são distintos tanto por conta dos desdobramentos das vertentes teóricas que se desenvolveram quanto por seus fundamentos. Parte dessas diferenças podem ser constatadas pelos termos empregadas no quadro 3, pois há variação para a identificação do assunto e dos processos que por ele são envolvidos.

Quadro 3 – Termos empregados na busca bibliográfica

Idiomas	Termos de busca
Português	<p>assunto, conteúdo, análise de assunto, catalogação de assunto, indexação, análise documentária, análise documental,</p> <p>conceito, linguística documentária, linguística documental, modelo conceitual FRSAD, Requisitos Funcionais para Dados de Autoridade Assunto.</p>
Espanhol	<p><i>materia,</i> <i>contenido,</i> <i>catalogación de matéria,</i> <i>indización,</i> <i>análisis documental,</i></p> <p><i>conceptos,</i> <i>linguística documental,</i> <i>modelo conceptual FRSAD,</i> <i>Requisitos Funcionales para</i> <i>los Datos de Autoridad de</i> <i>Materia.</i></p>
Inglês	<p><i>subject,</i> <i>content,</i> <i>subject analysis,</i> <i>subject cataloguing,</i> <i>indexing,</i> <i>aboutness,</i></p> <p><i>concept,</i> <i>conceptual model FRSAD,</i> <i>Functional Requirements for</i> <i>Subject Authority Data.</i></p>

Fonte: elaborado pela autora.

As variantes de algumas palavras-chave foram empregadas, como por exemplo, na pesquisa referente à análise documentária e documental, e linguística documentária e documental, pois na tradução dos termos para o português (brasileiro) há essa inconsistência. Não houve extensão dessas palavras para o inglês, pois o foco do referencial explorado é oriundo da vertente europeia, sobretudo dos franceses que marcaram sua origem. O recorte no contexto espanhol e brasileiro resultou em privilegiar referenciais dos respectivos países e, apesar do idioma inglês estar presente universalmente, sobretudo nas publicações de eventos internacionais da área, não foram realizadas traduções dos termos citados para a busca. O foco consistiu em resgatar referenciais que marcaram o recorte proposto e, portanto, centrou na dimensão geográfica dos idiomas ao qual pertencem.

Também cabe ressaltar que o termo *aboutness* não possui uma tradução consensual nem para o português (do Brasil) e tão pouco para o espanhol, mas apesar de não constar no quadro 3 no primeiro idioma, foram eleitos de forma correspondente os termos tematicidade,

sobrecidade e atinência. Já o termo sistemas de organização do conhecimento não foi empregado na busca por tratar de diferentes tipos de linguagens, e assim, não constitui interesse direto da pesquisa.

Ademais, foram utilizados livros, artigos científicos e teses que envolvem as temáticas sobre catalogação de assunto, indexação e análise documental e que trazem representatividade ao cenário abordado a partir dos três momentos escolhidos para sustentar essa investigação. Com isso, as fontes de pesquisa citadas foram importantes para a seleção da literatura adotada, mas não constituíram a única forma de busca, pois a partir das leituras, outros referenciais foram apresentados e, assim, explorados. Nessa perspectiva, o *corpus* bibliográfico foi explorado pelos desdobramentos das buscas nas fontes, mas foi constituído principalmente no percurso de leituras. A primeira forma é um procedimento inicial e essencial a qualquer pesquisa, pois as referências precisam ser buscadas e determinadas. No entanto, à medida que as fontes iniciais são exploradas, outras referências surgem e são incorporadas em leitura e análise.

Ainda que o objetivo não fosse o de historicizar, mas discutir a noção de assunto, optou-se por trabalhar com os aspectos cronológico e geográfico, haja vista a necessidade que se colocou frente aos referenciais teóricos de diferentes vertentes. Assim, o foco da sistematização contempla diferentes teóricos que sustentam a área e se apresentou como um grande desafio diante da tamanha ramificação que a parte temática assume.

Durante três meses (setembro, outubro e novembro de 2018) a pesquisa contou com um intercâmbio na Universidade de Coimbra, em Portugal, oportunizando a consulta presencial dos acervos das bibliotecas,³ sobretudo o da Biblioteca Geral da Universidade de Coimbra e o da Faculdade de Letras da mesma Universidade, sendo que na segunda encontrase, também, parte do material bibliográfico referente à área de Biblioteconomia e Ciência da Informação. Ressalta-se o acesso a vários livros de Jean Claude-Gardin nos acervos, sendo um dos autores que representa a linha europeia, que é francesa em sua origem.

Com o propósito de estruturar o referencial compilado segundo os objetivos dessa investigação, o quadro 4 elenca autores de diferentes nacionalidades, considerando as respectivas contribuições para o entendimento da noção de assunto.

³ Para acesso ao Serviço Integrado das Bibliotecas da Universidade de Coimbra: <http://www.uc.pt/sibuc>

Quadro 4 – Relação entre os objetivos específicos e o referencial teórico da pesquisa

Objetivos específicos	Referenciais teóricos
<p>a) Identificar e relacionar as construções teórico-conceituais pioneiras relativas à noção de assunto, no que tange ao processo, aos instrumentos e aos produtos documentários;</p> <p>b) Explorar a construção conceitual e a centralidade da noção de assunto a partir de um quadro teórico que se integre às propostas de representação temática, a qual não pode ser realizada devidamente na ausência de fundamentos com diferentes influências teóricas; e</p>	<p><i>Precedentes da construção da noção de assunto:</i> Fiúza (1987), Garrido Arilla (1999), Serrai (1975), Shera e Egan (1961), Burke (2003), Strout (1956), Araújo (2015), Placer (1955), Malclès (1960), Verner (1968), Frías Montoya (1995), Buckland (2005), Ferraz (1991), López Guillamón (2011).</p> <p><i>Primeiro momento de construção da noção de assunto:</i> Garrido Arilla (1999), Fiúza (1987), López Guillamón (2011), Tillett (1989), Mann (1962), Shera e Egan (1961), Miksa (1977; 2012), Cutter (1904), Cesarino e Pinto (1978), Gomes e Marinho (1984), Tôrres ([2000?]), Guimarães (2009), Penna (1945), Dias (1967), Rovira (1966), Teixeira (1979), Taylor (1948), Sharp (1964), Litton (1971), Foskett (1973).</p> <p><i>Segundo momento de construção da noção de assunto:</i> Metcalf (1973), Foskett (1973), Tôrres ([2000?]), Cesarino e Pinto (1978), Barbosa (1972), Campos e Gomes (2003), Hjørland (1992; 2001; 2017), Dutta e Dutta (2013), Sales (2012), Fujita (1988), Guimarães (2008), Taylor (1995), Simões (2017), Lara (2001), Mooers (1951; 1985), Henderson (1996), Gil Leiva (2008), Shera e Egan (1961), Collison (1971), Wheatley (1902), Reitz (2014), Pinto (1985), Langridge (1974), Dahlberg (1978; 1993), Kaula (1982), Lancaster (2004), Princípios UNISIST (UNITED...,1985), NBR 12676 (ASSOCIAÇÃO..., 1992), Princípios Internacionais de Catalogação (STATEMENT..., 2009; 2016), Guedes (2009), Hutchins (1977; 1978), Beghtol (1986), Peixoto (1962).</p> <p><i>Terceiro momento de construção da noção de assunto:</i> Otlet (2018), Santos (2007), Tálamo (<i>et al</i>, 2002), Ortega (2008; 2017), García Gutiérrez (1990), Pinto Molina (1993), Izquierdo Alonso (2000), Gardin (1964), Kobashi (1994; 1996; 2008), Cunha (1987), Lara (1993; 2011), Chaumier (1998), Lara e Tálamo (2006; 2007), Tálamo e</p>

	Lara (2006; 2009), Cintra, Tálamo, Lara e Kobashi (1994).
c) Descrever e analisar a fundamentação do modelo FRSAD por meio da investigação sobre a condição teórica da noção de assunto.	<i>Investigação do modelo FRSAD e estudos: Modelos FRBR, FRAD e o FRSAD: International Federation of Library Associations and Institutions</i> (1998; 2009; 2010), Žumer, Salaba e Zeng (2007), Zeng e Žumer (2009a; 2009b), Žumer, Zeng e Salaba (2012; 2015), Delsey (2005), Furner (2012), Gemberling (2010; 2016), Melo e Bräscher (2014).

Fonte: elaborado pela autora.

Cada grupo de autores permitiu construir as seções e subseções que embasam essa pesquisa, não como um fim em si, mas visando elaborar uma forma de mapear teoricamente a noção que se pretende discutir.

Já em relação à forma de análise escolhida, baseia-se na análise de conteúdo, de Laurence Bardin, e fez-se uso da edição mais recente do seu livro com tradução para o português de Portugal (BARDIN, 2018).

A autora define análise de conteúdo como “um conjunto de técnicas de análise das comunicações” (BARDIN, 2018, p. 33) que busca a descrição de conteúdo das mensagens e de indicadores (quantitativos ou qualitativos) que possibilitem a inferência de conhecimentos relativos a estas mensagens. Aplicam-se procedimentos sistemáticos na descrição analítica de “conteúdos” e “continentes” extremamente diversificados. Essa “descrição analítica” envolve a provável aplicação da análise de conteúdo como um método de “categorias” pela “análise categorial”. Explica ainda pertencer “[...] ao domínio da análise de conteúdo todas as iniciativas que, a partir de um conjunto de técnicas parciais[,] mas complementares, consistam na explicitação e sistematização do conteúdo das mensagens e da expressão deste conteúdo [...]” (BARDIN, 2018, p. 44). De forma ampla, é possível dizer que a análise de conteúdo é um método com outros inseridos e que congrega algumas técnicas. Apesar de contextualização, Abbagnano (2007, p. 668) define método como “uma técnica particular de pesquisa [...] que indica um procedimento de investigação organizado, repetível e autocorrigível que garanta a obtenção de resultados válidos”. No caso da análise de conteúdo, os resultados são alcançados pela inferência do que se propõe analisar.

Segundo Bardin (2018, p. 39), “o método de categorias, [é uma] espécie de gavetas ou rubricas significativas que permitem a classificação dos elementos de significação

constitutivos da mensagem.” Para ela, uma análise de conteúdo não deixa de ser uma análise de significados, ocupando-se de uma descrição objetiva e sistemática do conteúdo extraído das comunicações e sua respectiva interpretação. Essas comunicações podem ser constituídas a partir da via escrita, como em um texto; da via oral, como uma pesquisa de satisfação falada; pela via icônica, através de uma imagem; dentre uma infinidade de possibilidades, tais como músicas, objetos e comportamentos. É uma técnica versátil que não limita o que pode ser analisado e abarca diferentes perspectivas para verificar o conteúdo do que se quer analisar.

Nessa pesquisa, emprega-se a análise categorial que consiste em construir os “núcleos de sentidos” que compõem a noção de assunto no modelo escolhido como objeto empírico - o FRSAD. Presume-se que, pela análise, será possível emergir a noção de assunto que o modelo sustenta e, nesse contexto, o referencial empregado (quadro 4).

Mas, para chegar na formação das categorias, com base em Bardin (2018), a análise de conteúdo considera a “organização da análise” a partir de três principais fases ou estágios: 1) a pré-análise; 2) a exploração do material; 3) o tratamento dos resultados, a inferência e a interpretação.

Na primeira fase, o material é organizado compondo a sistematização das ideias iniciais para o desenvolvimento do *corpus* da pesquisa. Na segunda fase, ele é explorado e codificado para a análise, a partir de critérios previamente estabelecidos. Já na terceira e última fase, o tratamento evidencia a análise diante de inferências e interpretações estabelecidas nas duas fases anteriores.

Compreende-se que todo o trabalho de busca do referencial teórico já descrito é compartilhado com as três etapas, sobretudo com a pré-análise, em que é feita a escolha dos documentos que serão submetidos à análise, caracterizada pela exploração do material, que envolve a formulação de hipóteses e objetivos fundamentais à interpretação dos resultados. A terceira fase é evidenciada na seção sete dessa pesquisa, que trata da parte empírica, ou seja, a própria análise, considerando as categorias formadas a partir do referencial teórico explorado. Ressalta-se que a construção das categorias deu-se de forma *posteriori* à realização de inferências e interpretações advindas do levantamento bibliográfico e do arranjo da tese em três principais momentos teóricos. De algum modo, esses momentos também se tornam categorias iniciais, pois não haveria como trabalhar com o referencial sem um critério de seleção e organização dos fatos a serem investigados sobre a noção de assunto.

A escolha do modelo FRSAD, emanado da Federação Internacional de Associações e Instituições de Bibliotecas, a IFLA, decorre por ele oferecer orientações sobre a representação

dos assuntos dos documentos focando nas necessidades dos usuários. Seu caráter não é normativo e procedimental, mas de referencial internacional e, com isso, o texto do FRSAD torna-se objeto fecundo de discussão.

É possível sustentar a opinião de que ele receba influência de diferentes vertentes, pois se presume que o cunho internacional da Associação demande várias perspectivas. Ademais, um modelo conceitual deve evidenciar algo, nesse caso o assunto, que pode se manifestar de diferentes maneiras. A IFLA também trabalha com modelos que exploram outros atributos e, torna-se apropriado abordar, quando necessário, o FRBR (*Functional Requirements for Bibliographic Records*) e o FRAD (*Functional Requirements for Authority Data*), à medida que se aproximam do interesse da pesquisa. Então há um ponto a observar: qual a origem do referencial do modelo? É possível haver relação entre essa origem e o modelo?

Nessa direção, o foco da análise consiste em apurar a noção de assunto no modelo FRSAD que possibilite a identificação de características que o aproxime ou o distancie do quadro teórico estabelecido a partir dos três momentos traçados na pesquisa sobre a construção da noção de assunto. Conforme esquematizado no quadro 4, vários autores foram empregados para dar conta dessa fundamentação.

Diferentes técnicas podem ser utilizadas na análise de conteúdo, como a análise categorial, análise de avaliação, análise de enunciação, análise de expressão e análise proposicional do discurso. De acordo com Bardin (2018, p. 199, destaque da autora):

No conjunto das técnicas da análise de conteúdo, é de citar em primeiro lugar a análise por categorias; cronologicamente é a mais antiga; na prática é a mais utilizada. Funciona por operações de desmembramento do texto em unidades, em categorias segundo reagrupamentos analógicos. Entre as diferentes possibilidades de categorização, a investigação dos temas, ou *análise temática*, é rápida e eficaz na condição de aplicar a discursos directos (significações manifestas) e simples.

A opção pela análise de categorias concorre mais para o emprego de interpretações e inferências sobre o referencial teórico do que a sua cronologia e quantidade de uso. A interpretação confere adequação à pesquisa, mas não se descarta a possibilidade de ocasionar fragilidade(s) ao método. Porém, antes de haver preocupação em mensurar essa intensidade, o rigor faz parte de diferentes métodos, pois é uma característica que auxilia a manutenção da homogeneidade do objeto de pesquisa, por mais heterogêneo que possa ser, como acontece com a noção de assunto. Ademais, esse movimento de interpretação, ao mesmo tempo que é contínuo, detém a necessidade de ser apreendido e analisado em um contexto histórico e socialmente condicionado. Nenhuma pesquisa se desprende disso.

Na construção das seções da pesquisa, articula-se o quadro teórico com a finalidade de dar conta dos momentos que caracterizam a noção de assunto em diferentes perspectivas. Além do estabelecido, outros quadros teóricos podem ser desenvolvidos, mas a forma de abordagem e de análise da noção pretendida não é algo dado e explícito na literatura. Cabem diferentes sistematizações e, nessa proposta, a inquietação baseia-se em não haver uma única noção de assunto. E, ainda: como entender a noção de assunto no modelo FRASAD a partir da literatura da área de Biblioteconomia e Ciência da Informação?

Com o propósito de estabelecer a análise, o uso de categorias de forma *a posteriori* permite progressivamente a definição e a categorização e “[...] é mais ou menos adaptad[a] à realidade que se nos oferece” (BARDIN, 2018, p. 146). A autora realiza uma comparação: na dúvida em ouvir a música de Bach, Ravel ou Boulez, não se utilizará o mesmo critério de uma possível escolha acerca de se ouvir violino, órgão ou piano, pois o critério de categorização difere-se, ou seja, compositor e instrumento. Mas, segundo ela, ainda é possível convergirem dois desejos para a escolha ser feita a partir de um determinado instrumento e compositor. Assim,

Classificar elementos em categorias impõe a investigação do que cada um deles tem em comum com outros. O que vai permitir o seu agrupamento é a parte comum existente entre eles. É possível, contudo, que outros critérios insistam noutros aspectos de analogia, talvez modificando consideravelmente a repartição anterior. (BARDIN, 2018, p. 146).

A categorização envolve a escolha de categorias, que por sua vez, “são rubricas ou classes, as quais reúnem um grupo de elementos (unidades de registro, no caso da análise de conteúdo)” (BARDIN, 2018, p. 145). Para a escolha de categorias adotam-se os critérios semântico (temas), sintático (verbos, adjetivos e pronomes), léxico (sentido e significado das palavras - antônimo ou sinônimo) e expressivo (variações na linguagem e na escrita). Este processo permite a junção de um número significativo de informações organizadas em duas etapas: inventário (onde se isolam os elementos comuns) e classificação (onde se divide os elementos e impõem-se organização). De acordo com Bardin (2018, p. 146-147), “a categorização tem como primeiro objetivo [...] fornecer, por condensação, uma representação simplificada dos dados em bruto.”

Para tanto, a inferência é orientada pela comunicação, que vai do emissor, ao receptor e da mensagem ao canal, porém para averiguar possíveis unificações pode haver a necessidade de se comparar enunciados e ações entre si, a partir de temas e dados. Quando os

temas encontrados são diferentes, quem investiga precisa encontrar semelhanças que possam existir entre eles.

Na mesma medida que Bardin (2018, p. 163) instiga: “sobre o que é que pode incidir este tipo de interpretação controlada que é, na análise de conteúdo, a inferência?”, também explica que a análise de conteúdo “[...] deseja distanciar-se da sua leitura ‘aderente’, para saber mais sobre esse texto.”

Na fase de interpretação, é necessário retornar ao referencial teórico procurando embasar as análises para dar sentido. Assim sendo, as interpretações pautadas em inferências buscam o que se esconde por trás dos significados das palavras para apresentarem o discurso dos enunciados. É o procedimento de inferência que permite descrever e analisar a construção de relacionamento com o quadro teórico descrito, através de um processo de análise.

Considera-se que o papel de interpretação marca a pesquisa qualitativa e fornece à análise de conteúdo o compromisso de imprimir nitidez ao quadro teórico trabalhado.

Nessa configuração, e a partir do percurso estabelecido, o referencial teórico é explorado, para em seguida, dar base à análise junto ao objeto empírico dessa investigação: o modelo FRSAD.

3 PRECEDENTES DA CONSTRUÇÃO DA NOÇÃO DE ASSUNTO: AS INFLUÊNCIAS DOS AUTORES DO SÉCULO XIX

É nítido constatar o registro da informação enquanto inscrição formulada intencionalmente com o objetivo específico de integrar ambientes informacionais, como por exemplo, os catálogos de bibliotecas. Os atributos que compõem a representação desse registro são carregados de valores e objetivos que buscam propiciar o acesso à informação.

Um catálogo enquanto um sistema de informação ou sistema documentário é um instrumento que disponibiliza índices, resumos e outros similares bibliográficos que condicionam o acesso ao documento, detentor da informação registrada. O catálogo também é o produto resultante do processo de catalogação, mas enquanto sistema articula os atributos dos documentos.

A preocupação com esses dispositivos de informação não são novas e as ideias anteriores ao século XIX permitem ampliar o olhar sobre referenciais que movimentaram este século. Portanto, parte-se do pressuposto que as questões relativas à representação documentária não são recentes e receberam diferentes influências. Nessa seção, buscar-se-á delimitação envolvendo marcas históricas do produto bibliográfico até o trabalho de Adrien Baillet (1649-1706), pois antecedeu as *Rules* de Charles Ammi Cutter (1837-1903), que marcou a vertente denominada Catalogação de Assunto.

3.1 AS PRIMEIRAS PROPOSTAS NA EUROPA

O contexto teórico que reflete a noção de assunto envolve diferentes tentativas de abordagem, incluindo variadas concepções. O que segue é uma estruturação cronológica que possibilite pensar a noção pretendida. Para dar início ao percurso, Fiúza (1987, p. 44) explica:

Nos períodos antigos da história, encontramos alguns remanescentes de listagens de livros, mas não se sabe que uso se fazia delas. Na literatura grega, há referências a coleções de livros pertencentes a indivíduos ou a bibliotecas. Calímacus, bibliotecário da Biblioteca de Alexandria, ficou famoso, na história da Biblioteconomia, porque compilou uma listagem de obras, em 250 a.C. Não se chegou, porém, a determinar, com exatidão, se esta listagem era um catálogo da biblioteca ou uma bibliografia da literatura grega.

Garrido Arilla (1999, p. 62, tradução nossa) destaca que as famosas bibliotecas da Antiguidade, como por exemplo, a de Assurbanipal, de Ebla, de Alexandria e Pérgamo possuíam catálogos na forma de “[...] listas descritivas e ordenadas cuja finalidade principal era inventariar as coleções e ordená-las segundo uma metodologia muito rudimentar.”⁴ Aponta que a biblioteca de Nínive, datada do século II a.C., possuía catálogo em tablete de argila que incluía o título, uma breve descrição de forma referente ao número de tabletas e linhas contidas, além de um registro rudimentar sinalizando o lugar do documento na biblioteca.

A autora supracitada ressalta que, nesses catálogos rudimentares, não é clara a separação entre os aspectos formais do documento, e o seu conteúdo intelectual, entendido aqui como assunto e, torna-se nítido que a preocupação principal não era a de direcionar o acesso ao documento, mas sim ordená-lo de maneira que sua integridade fosse mantida.

É possível notar uma intrínseca relação entre o tipo de escrita, o suporte e as formas rudimentares de organização dos documentos. “A escrita implica o documento, os documentos requerem não só um depósito, mas uma reunião, e esta, por sua vez, exige uma organização que facilite o encontro dos documentos toda vez que for necessário.” (SERRAI, 1975, p. 142).

Garrido Arilla (1999) explica que a busca pela organização de acervos ocorreu por listas descritivas de documentos compreendidas na antiguidade como catálogos. Serrai (1975) diz haver utilidade na organização dos documentos e, Fiúza (1987), destaca as listagens de livros com fins de catalogação ou de bibliografia para organização. Ferraz (1991, p. 97), por sua vez, aponta que:

[...] desde a Antiguidade, o arranjo por assunto é também norma, ainda que de forma menos sofisticada, pois a biblioteca do palácio de Assurbanipal tinha uma sala para cada grupo específico: uma sala para tábuas relacionadas com Histórias e Governo, outra para Lendas e Mitologia e assim por diante. Cada uma das dez salas da biblioteca de Alexandria foi designada para um departamento de ensino. Na Idade Média, os trabalhos da época, por exemplo, eram arranjados de acordo com Gramática, Lógica e Retórica e, mais tarde, Aritmética, Geometria, Música e Astronomia.

Shera e Egan (1961, p. 18), no prefácio do livro *Documentação*, de Bradford, sinalizam que no ano 731 “[...] o venerável Beda juntou à sua *Ecclesiastical History of Britain* uma relação de cerca [sic] de 40 obras, dispostas numa classificação rudimentar [...]” Outro

⁴ “[...] listas descriptivas y ordenadas cuya finalidad fundamental era inventariar las colecciones y ordenarlas según una metodología muy rudimentaria.” (GARRIDO ARILLA, 1999, p. 62).

destaque é dado pelos autores na compilação bibliográfica vinculada à igreja cristã, “[...] como o *Catalogus scriptorum ecclesiae* de John Boston [...] [que] registrou os títulos dos livros encontrados nas bibliotecas dos mosteiros e ordens eclesiásticas da Inglaterra [...]”. O destaque para a última lista sucedeu no fato de incluir compilações de 195 repositórios, elencadas de forma alfabética pelo nome do autor, com classificação parcial que seguia os padrões dos escritores segundo os livros da Bíblia (SHERA; EGAN, 1961).

Em 1550, é sinalizado que o aumento da quantidade de livros, decorrente da invenção da imprensa, alarmou grupos de estudiosos. “Mais que uma ordem dos livros, o que alguns contemporâneos percebiam era uma ‘desordem dos livros’ que precisava ser posta sob controle. Até Gesner, que cunhou a expressão *ordo libroum*, se queixava ‘dessa confusa e irritante multidão de livros’ [...]” (BURKE, 2003, p. 97).⁵

Segundo Strout (1956, p. 262, tradução nossa), “a história da catalogação no século XVI não teria muito a mostrar no caminho do progresso se tivesse sido forçada a depender unicamente dos bibliotecários e catálogos de bibliotecas.”⁶ Para a referida autora, Konrad Gesner, em 1548, ao publicar seu índice de assuntos, estabelecia um novo padrão de excelência através do *Pandectarum sive partitionum universalium*. “Ele sugeriu que as bibliotecas usassem cópias de suas bibliografias como catálogos, inserindo números de chamada ao lado das entradas que representassem seus cabeçalhos, fornecendo-se assim tanto um autor como um assunto no catálogo.”⁷ (STROUT, 1956, p. 263, tradução nossa).

O projeto bibliográfico de Gesner, *Bibliotheca Universalis*, foi elaborado no período de 1545-1555, e dividiu-se em quatro partes: 1) *Bibliotheca Universalis, sive Catalogus omnium scriptorum locupletissimus, in tribus linguis, Latina, Graeca, et Hebraica* (1545); 2) *Pandectarum sive partitionum universalium* (1548); 3) *Partitiones theologicae* (1549); e 4) *Appendix bibliothecae* (1555) (ARAÚJO, 2015).

Para o autor indicado, a primeira parte, *Bibliotheca Universalis, sive Catalogus omnium scriptorum locupletissimus, in tribus linguis, Latina, Graeca, et Hebraica* constitui um catálogo alfabético-nominal com 5031 autores de obras em três línguas eruditas (latim,

⁵ Sem objetivar exaurir as discussões em Gesner e em outros bibliógrafos citados, a presença deles é evocada de forma pontual para embasar o percurso cronológico da pesquisa.

⁶ “The history of cataloging in the sixteenth century would not have much to show in the way of progress if it had been forced to depend solely upon librarians and library catalogs.” (STROUT, 1956, p. 262).

⁷ “He even suggested that libraries use copies of his bibliographies as their catalogs by inserting call numbers beside entries which represented their holdings, thus providing themselves with both an author and a subject catalog.” (STROUT, 1956, p. 263).

grego e hebraico). Organizada de forma alfabética a partir do primeiro nome do autor, é acompanhada por uma lista a parte dos sobrenomes. Já o segundo, *Pandectarum sive partitionum universalium*, abrange a exposição classificada do conhecimento extraído dos livros (dos autores da primeira parte) com as hierarquias intelectuais da Baixa Idade Média, na seguinte ordem sistemática:

1-Gramática (e Filologia), 2- Dialética, 3-Retórica (representando o *trivium*), 4-Poética, 5-Aritmética, 6-Geometria, 7-Música e 8-Astronomia (sendo as últimas quatro classes representadas pelo *quadrivium*). Estas são seguidas pelas ciências incluindo o currículo universitário medieval: 9-Astrologia, 10-Adivinhação e Magia, 11-Geografia, 12-História, 13-Artes Mecânicas, 14-Filosofia Natural, 15-Metafísica, 16-Filosofia Moral, 17-Filosofia Econômica, 18- Política; e finalmente, 19-Direito, 20-Medicina e 21-Teologia. (ARAÚJO, 2015, p. 133-134).

Florian Trefler, monge beneditino, publicou em 1560 o tratado *Methodus exhibens per varios indices, et classes subinde, quorumlibet librorum, cuiuslibet bibliothecae, breve, facilem, imitabilem ordinationem* e relatou na introdução a dificuldade de encontrar informação nas bibliotecas, pelo motivo de os livros não serem mantidos em qualquer ordem discernível e o material não ser catalogado (STROUT, 1956).

No final do século XVII, Garrido Arilla (1999) aponta que o aparecimento da última edição do catálogo da Biblioteca Bodleiana de Oxford foi o primeiro catálogo dicionário construído, pois mesmo sendo rudimentar, continha lista em ordem alfabética de autor, título e assunto, permitindo encontrar o livro sobre diferentes aspectos.

A ordem alfabética também recebe destaque nas enciclopédias e, por sua característica de mapear o conhecimento, destacou-se no controle do conteúdo pelo sistema de referências cruzadas e remissivas. Burke (2003, p. 165) aponta que d’Alembert indica dois métodos para organização: “princípio enciclopédico” e “princípio do dicionário”. O primeiro refere-se à organização por temas, portanto temática e, o segundo, a alfabetização dos tópicos sendo também utilizado em índices. “[...] A *Biblioteca* (1545), de Gesner, arrolava os livros em ordem alfabética, e sua *História dos animais* (1551-) listava alfabeticamente os animais. O *Índex* católico dos livros proibidos seguia o mesmo critério.” (BURKE, 2003, p. 165, destaque do autor). É possível considerar que o princípio de ordem alfabética e sistemática de repertórios bibliográficos foram adotados paulatinamente e, por consequência, influenciaram as bibliotecas.

Outros exemplos são descritos por Strout (1956), sobretudo com foco no desenvolvimento de catálogos e dos códigos de catalogação, mas a autora ressalta o tempo

demandado para que certas ideias de catalogação, consideradas como simples, se fixassem, como o índice alfabético, pois “[...] lembremos que se passaram quatrocentos anos desde que foi feita pela primeira vez a sugestão de que as bibliografias impressas fossem substituídas por catálogos ou seções de catálogos e ainda parece muito ousado e radical tentar fazer tal coisa.”⁸ (STROUT, 1956, p. 263, tradução nossa).

Ainda no século XVII, segundo Garrido Arilla (1999), em bibliotecas notáveis, como a Bodleiana de Oxford e a de Mazarino na França, mudanças ocorreram por demanda dos leitores que começaram a requerer livros, mais por assuntos que por autores e, assim, demonstraram interesse por catálogos com finalidade temática.

Evidencia-se uma tímida demanda por assunto nas bibliotecas, mas esse recorte do tempo narrativo não é linear e extensivo a vários lugares, havendo diferentes maneiras para caracterizar o interesse pelo assunto de um documento a cada momento.

Essa forma de situar progressivamente os fatos resvala na relação entre a organização das bibliografias e dos catálogos, pois a quantidade de impressos conduziu a arranjos que identificassem as listagens de documentos, com indicação do autor, do título e do ano e, eventualmente, do assunto. Conforme ressaltado, Serrai (1975) aponta que a tradição da escrita subentende o documento que, por sua vez, exige organização, e a prática do registro é indissociável da tradição humana de produzir o conhecimento.

Para o público letrado, por volta de 1600, “[...] o acesso às primeiras bibliotecas modernas dependia das atitudes do bibliotecário e de sua equipe. [...] Gabriel Naudé observa que só a Bodleiana em Oxford, a Ambrosiana em Milão e a Agostiniana em Roma permitiram acesso livre aos estudiosos [...]” (BURKE, 2003, p. 160).

Ainda nesse mesmo período, grupos de estudiosos marcaram o continente europeu pela produção de documentos e uso das bibliotecas abertas ao público. Dois desses grupos diferenciaram-se em relação a outros, por serem reconhecidos como autor (*auteur*) e escritor (*écrivain*) e, sem igualdade com os membros do clero da Igreja Católica, o primeiro qualificou-se pela responsabilidade em publicar obras impressas, enquanto o segundo em escrever textos sem propor sua circulação, portanto mantendo-os na forma manuscrita. Não se pode perder de vista que na época, a censura e a perseguição religiosa cerceavam o trabalho dos estudiosos e, por isso, a autoria não era identificada pelo nome próprio. Também havia carreiras em literatura, no sentido amplo do termo, incluindo em suas atividades a elaboração

⁸ “[...] let us remember that it has been four hundred years since the suggestion was first made that printed bibliographies be substituted for catalogs or sections of catalogs and that it still seems too daring and radical a thing to try.” (STROUT, 1956, p. 263).

de dicionários e obras de história. Além desses, outro grupo pequeno de estudiosos, mas com bastante influência, que em sua composição contou com a presença de membros como Denis Diderot e Gabriel Naudé, “[...] pode ser apresentado na linguagem de nossos dias como ‘intermediários da informação’, porque punham estudiosos de diferentes lugares em contato entre si, ou como ‘administradores do conhecimento’, porque tentavam organizar o material, além de coletá-lo.” (BURKE, 2003, p. 31).

Gabriel Naudé (1600-1653), em sua *Bibliographia politica* (1633), buscou “[...] uma ideal biblioteca universal reservada à política, fornecendo uma ordem não neutral da matéria, mas selecionando, ordenando e organizando os materiais com base nos princípios que norteiam os pensamentos dos libertinos eruditos.” (CRIPPA, 2017, p. 32). O ato de organizar o registro da literatura necessária para o estudo da política contribuiu para que as “bibliografias, um tipo de obra de referência cada vez mais comum no período, [...] [fossem] descritas como ‘bibliotecas sem paredes’, que podiam viajar por toda a Europa.” (BURKE, 2003, p. 97).

Se as bibliografias detinham a finalidade de por à disposição um volume rico de informações organizadas sobre um determinado tema que concorria para o aprimoramento dos métodos de trabalho da área, por outro lado, o conjunto de técnicas seculares desses produtos contribuiu para o desenvolvimento dos catálogos das bibliotecas. Se por um lado as listagens dos acervos possuíam ligação com as bibliografias, por outro se destacavam “no sentido estrito, [como] o inventário sistematicamente organizado de material sôbre [sic] um ou vários assuntos.” (PLACER, 1955, p. 8).

Malclès (1960) destaca que é possível obter informações de documentos literários ou científicos, de forma rápida e segura, pela utilização de instrumentos intelectuais denominados bibliografias ou repertórios bibliográficos. Sendo assim, “todo repertório tem seu próprio objeto [...] [e] nele enumerados são relacionados por características comuns: mesmo lugar de origem (impressão, edição), mesma data ou período de publicação, o mesmo idioma e o mesmo assunto.”⁹ (MALCLÈS, 1960, p. 9, tradução nossa).

Cesarino e Pinto (1978, p. 274) explicam que as listas de assunto eram dispostas no final das bibliografias e “com a sistematização dessas apresentações de assunto, as listas passaram a ser organizadas alfabética ou classificadamente.” Assim, a evidência do assunto na construção de bibliografias torna-se presente, pois,

⁹ “Cada repertorio tiene objeto propio y es útil, por consiguiente, para investigaciones del mismo tipo. Dicho de otra manera, las obras allí enumeradas poseen caracteres comunes: el mismo lugar de origen (impresión, edición), la misma fecha o época de publicación, el mismo idioma o el mismo asunto.” (MALCLÈS, 1960, p. 9).

Os séculos XV e XVI representaram um grande desenvolvimento nas atividades acadêmicas, intelectuais e científicas [...]. Desde a invenção da imprensa, por Guttenberg, os acervos de bibliotecas e das livrarias precisavam ter uma organização mais cuidadosa. (FIUZA, 1987, p. 45).

Nota-se, portanto, que o aumento dos livros impressos, tanto em relação à quantidade como à variedade, desencadeou técnicas bibliográficas e, também, o crescimento do número e tamanho das bibliotecas como consequência da invenção da imprensa (BURKE, 2003).

Parece contestável que na Antiguidade e na Idade Média prevalecia o desconhecimento do aspecto temático do documento, mas as bases científicas em que, posteriormente, a Europa moderna se alicerçou foram responsáveis por metodologias bibliográficas, como a estrutura da descrição, que conduziram a produtos bibliográficos fundamentais à construção do cenário atual.

Alguns fatores como o já citado fortalecimento da impressão tipográfica, o declínio do poder da Igreja Católica e a visão humanista cooperaram para que os sistemas e serviços bibliográficos fossem postos à disposição e facilitassem o acesso do público aos documentos. Assim,

No início do século XVIII, os catálogos eram finalmente vistos como listas de descobertas e não inventários. Durante este século, eles foram por vezes classificados e por vezes alfabéticos; os índices foram considerados úteis, embora não necessários; alguns catálogos ainda eram divididos de acordo com o tamanho dos livros; os autores eram sempre inseridos pelo sobrenome e por vezes dispostos cronologicamente; a redação da página de rosto tinha assumido um grau de prestígio e passou a ser transcrita literalmente sem ser parafraseada; impressões foram incluídas; foram utilizadas notas “vinculadas”; as referências cruzadas eram bastante comuns; e algumas entradas analíticas foram usadas na maioria dos catálogos.¹⁰ (STROUT, 1956, p. 266, tradução nossa).

A despeito dessa fala de Strout (1956), é fato a influência do trabalho bibliográfico nas “listas de descobertas” adotadas em bibliotecas, com predominância do conhecimento técnico científico que resultava no produto sem confusão entre o seu processo de elaboração.

Como forma de desdobramento das metodologias expostas acima por Strout (1956), ainda por volta dos seiscentos, o bibliotecário do continente europeu – Adrien Baillet –

¹⁰ By the beginning of the eighteenth century, catalogs were at last looked upon as finding lists rather than inventories. During this century they were sometimes classified and sometimes alphabetical; indexes were considered useful, though by no means necessary; some catalogs were still divided according to the size of books; authors were now always entered under surname and were often arranged chronologically; the wording of the title page had assumed a certain degree of prestige and was now being transcribed literally and without being paraphrased; imprints were included; “bound-with” notes were used; crossreferences were quite common; and some analytical entries were used in most catalogs. (STROUT, 1956, p. 266).

compõe a outra parte do desdobramento dessa seção. O propósito é traçar um quadro para a representação de assunto anterior ao contexto anglo-saxão e suas influências.

3.2 A CONTRIBUIÇÃO DO FRANCÊS ADRIEN BAILLET NO SÉCULO XVII

Embora seja creditado a Charles Ammi Cutter o pioneirismo na sistematização de regras para organização de catálogos por assunto, Verner (1968) sinaliza Adrien Baillet como o precursor na elaboração de um conjunto de regras para um catálogo alfabético de assunto. Para esta autora, apesar de o catálogo alfabético de assunto ter se popularizado em países de língua inglesa, especialmente nos Estados Unidos, sua origem é anterior aos preceitos de Charles Coffin Jewett e Charles Ammi Cutter.

Bibliotecário francês, Adrien Baillet (1649-1706) estabeleceu regras para a compilação do catálogo da biblioteca de Chrétien-François Lamoignon (1644-1709), especializada em direito e ciências políticas. Frías Montoya (1995) indica que Baillet estruturou e expôs 17 pontos que deveriam embasar o índice da Biblioteca Lamoignana e, excetuando-se o primeiro e o último ponto, as regras constituíram normativa orgânica, que propunha uma catalogação por assunto com ordenação alfabética, mas que também oferecia orientações para a catalogação de autores.

Baillet elaborou, entre 1682 e 1685, o *In Priorem Bibliothecae Lamonianae Indicem Praefatio*, que abarcou um conjunto de regras para discutir as múltiplas entradas no catálogo, como por exemplo, de assunto, título e autoria (RIAZ, 1989), e desencadeou o volume 2 intitulado *Jugemens des Savans sur les principaux ouvrages des auteurs les plus connus* (VERNER, 1968).

Para Verner (1968), a organização de assuntos em catálogos de bibliotecas (*subject organization in library catalogs*) perpassou dois contextos: a ordenação alfabética de assuntos (*alphabetic arrangement of subjects*) no agrupamento de materiais nas prateleiras e os catálogos alfabéticos de autor (*alphabetic author catalogs*) em que obras anônimas eram representadas a partir de palavras-chave (*catchwords*) retiradas do título. Eram mais palavras significativas do texto do que propriamente palavras-chave.

Compreender o trabalho de Adrien Baillet na organização de assuntos em bibliotecas (*subject organization in libraries*), conforme Verner (1968) aponta, envolve identificar a influência que ele obteve de referenciais variados. Como exemplo, a autora indica que o catálogo do Mosteiro de Melk, na Áustria, datado de 1483, possuía três índices: de autor, de

categorias de assunto e de obras anônimas com base nas palavras mais significativas do texto. Também descreve que a partir do século XV, nos Mosteiros Rebdorf, de Eichstaett, e Aegidian, de Nuremberg, ambos na Alemanha, os catálogos eram organizados pelo nome do autor e por palavras do título, principiando assim aspectos do catálogo dicionário.

Ainda segundo Verner (1968), as bibliografias desempenharam influência na abordagem alfabética de assunto e no trabalho de Baillet, pois ele admitiu a importância delas a partir dos preceitos de Conrad Gesner, Andrew Maunsell, Florian Trefler, Johann Heinrich Hottinger, Frederic Rostgaard e Gabriel Naudé.

Ademais, Baillet se interessou pelo catálogo da *Bibliotheca Thuana*, pelo índice de assunto do catálogo de *Fabien Justinien* denominado *Index universalis alphabeticus* (1612) e pela compilação deste com Barthelemi Moser do *Trésor bibliothèque ou double bibliothèque onomastique e classique*. O catálogo da *Bodleian Library* e o do *Sion College* em Londres também influenciaram as ideias de Baillet na elaboração do catálogo da biblioteca de Lamoignon, pois “ele gostava do conceito do arranjo de documentos em prateleiras por categorias de assuntos principais, a ser espelhado na parte principal do catálogo, como era o costume na época (bem representado, por exemplo, na bibliografia de Gesner e nos primeiros catálogos de Bodleian).”¹¹ (VERNER, 1968, p. 225, tradução nossa).

Para a biblioteca de Lamoignon, Baillet buscou elaborar um catálogo diferenciado que nunca houvesse sido construído e, após observar vários métodos, construiu o índice de assuntos que infelizmente não existe mais. O fato é citado por Verner (1968) a partir de Taylor (1966).¹²

Em 1785, o catálogo da biblioteca Lamoignon foi revisado e ampliado. Estava dividido em cinco seções – teologia, lei, história, artes liberais e manuscritos – as quais possuíam subdivisões que podiam ser cronológicas ou, então, ter algum outro tipo de ordem, representando divisões de subgrupos para assuntos individuais. A abordagem de Baillet no catálogo por ordem de assunto influenciou o trabalho de Gabriel Peignot (VERNER, 1968).

O índice do catálogo deveria trazer os assuntos contidos “[...] em uma obra que poderia ser colocada somente em um ponto da prateleira e que, portanto, era representada em

¹¹ “He liked the concept of a shelf arrangement by major subject categories, to be mirrored in the principal part of the catalog, as was the custom at the time (well represented, for instance, in the Gesner bibliography and the early Bodleian catalogs).” (VERNER, 1968, p. 225).

¹² TAYLOR, Archer. **General subject-indexes since 1548**. Philadelphia: University of Pennsylvania Press, 1966.

apenas um único lugar no catálogo agrupado, espelhando o arranjo nas prateleiras.”¹³ (VERNER, 1968, p. 226, tradução nossa).

A inovação do método de Baillet, segundo Verner (1968), baseou-se na possibilidade do acesso. Para tanto, primeiro ele construiu um conciso e curto arranjo agrupado (*grouped arrangement*) denominado *Praemissa prius Bibliothecae ac per classes majores distributa synopsis, sive secundum loci positionem conspectu*,¹⁴ como suplemento para o índice de assunto que ele intitulou *dictionnaire raisonné*. Esse arranjo buscava mapear o lugar, ou seja, o ponto na prateleira que o livro ocuparia de acordo com o assunto vinculado no *dictionnaire raisonné*. Em segundo, ele considerou a necessidade de um índice de autor e catálogo próprio em sequência alfabética de sobrenome com entrada analítica através das múltiplas possibilidades (nome familiar, nome antigo, nome popular). Tais usos não eram comuns para os catálogos da época.

Segundo Baillet, todas as obras deveriam ser registradas com cabeçalho no idioma adotado na escrita, com exceção do grego e do hebraico, pois para esses idiomas deveria ser realizada a tradução para o latim. Verner (1968) aponta que, no catálogo da Bodleian, o método do cabeçalho adotado no idioma da escrita da obra era empregado, entretanto, sem o padrão da tradução para o latim.

Outras orientações empregadas por Baillet foram: o tratamento de palavras sinônimas no índice com uso de referências cruzadas (*cross-references*) pelas formas ‘não usar’ (*not used*) e ‘tal como’ (*such as*); a entrada de nomes de pessoas como assunto se houvesse relação com o contexto histórico do próprio nome; assunto com subdivisões de áreas, aspectos cronológicos e geográficos; controle do plural e singular dos assuntos; controle de nomes compostos por artigos e preposições.

Segundo Verner (1968, p. 229, tradução nossa), no contexto das regras de Baillet,

A ortografia já era considerada uma parte importante na construção de um índice alfabético, uma vez que desempenha um papel decisivo na posição da entrada. Sempre que a ortografia fosse duvidosa, Baillet indicava a entrada sob ambas as formas a fim de não causar ao leitor uma busca inútil “*ejusmodi nomina dubiae scriptionis ad utrumque titulum revocavi, ne cui de industria negotium in quaerendo facessam*”. Outra regra indicava a

¹³ “[...] in a work which could be placed only in one spot on the shelf and which was, therefore, represented in one place only in the grouped catalog mirroring the shelf arrangement.” (VERNER, 1968, p. 226).

¹⁴ Parte integrante de *In Priorem Bibliothecae Lamonianae Indicem Praefatio*.

sequência de arquivamento para *I e J, U e V* (as letras foram unidas somente na edição de 1900 do Catálogo do Museu Britânico).¹⁵

Como pode ser constatado, o legado de Baillet é significativo no que se refere ao controle alfabético do assunto (que envolve também um controle de nomes pessoais) para a busca. Ao estruturar as regras que embasaram o catálogo da Lamoignon, ele observou diferentes fontes e promoveu o movimento de acesso ao acervo da referida biblioteca.

O assunto, para Baillet, era o que constava de forma explícita na obra e deveria servir, ao mesmo tempo, para a organização nas estantes, ação que Verner (1968) denominou, conforme indicado, “organização dos assuntos em bibliotecas.”¹⁶ Nesse caso, a indicação do atributo assunto operava pela busca de regras para além das palavras-chave (*catchwords*) que formavam os títulos e cumpria com o propósito de associação do que estava explícito no documento. Baillet conjugou fundamento teórico para a reflexão sobre o assunto.

A fim de atender às especificidades da biblioteca de Lamoignon, Baillet teve a preocupação de realizar o controle de nomes dos autores, assim como, o controle dos assuntos no índice do catálogo através das regras que refletiam a disposição dos documentos nas prateleiras em ordem alfabética, dinamizando o contexto de aplicação.

No continente europeu, Verner (1968) apontou Baillet como influenciador do trabalho do alemão Martin Schrettinger (1772-1851), que em 1808, elaborou ideias sobre a ciência da biblioteca (BUCKLAND, 2005) e, no contexto anglo-saxão, motivou o trabalho de Charles Ammi Cutter (1837-1903).

Scheretinger, influenciado por Kant, evidenciou que não seria possível imitar a natureza na organização da biblioteca, pois ela reflete uma ordem da consciência humana intangível. Nesse sentido, considera a biblioteca “[...] uma coleção substancial de livros, cuja ordenação coloca cada pessoa com demanda por conhecimento em posição de usar alguma

¹⁵ Spelling was already considered an important part in the construction of an alphabetic index, since it played a decisive role in the position of the entry. Whenever the spelling was doubtful, Baillet’s entry would occur under both forms in order not to cause the reader a useless search “*ejusmodi nomina dubiae descriptionis ad utrumque titulum revocavi, ne cui de industria negotium in quaerendo facessam*”. Another rule regulated the filing sequence for *I* and *J, U* and *V* (letters interfiled as late as in the 1900 edition of the British Museum Catalog). (VERNER, 1968, p. 229).

¹⁶ “Subject organization in libraries began with the grouping of like books on the shelves.” (VERNER, 1968, p. 221).

obra conforme sua necessidade.”¹⁷ (SCHRETTINGER, 1808 *apud* BUCKLAND, 2005, p. 24).

Scherettinger estabeleceu uma proposta simples e pragmática envolvendo uma ordenação por assunto e tendo cada novo livro espaço de inclusão em uma seção que o identificasse (BUCKLAND, 2005). Com isso, o catálogo alfabético crescia pela ordem do assunto da obra na prateleira, possibilitando o acesso ao livro (item) individualmente, aproximando-se do que Baillet já havia de certa forma exposto em 1685.

Buckland (2005) considera que Schrettinger fundou a disciplina denominada Ciência da Biblioteca (*Bibliothek-Wissenschaft* em alemão; *Library Science* em inglês) e sua preocupação com a técnica, com o sistema e uso da biblioteca foi registrada no livro intitulado *Versuch eines vollständigen Lehrbuchs der Bibliothek-Wissenschaft oder Anleitung zur vollkommenen Geschäftsführung eines Bibliothekars in wissenschaftlicher Form abgefasst* (1829), cuja tradução seria: Uma tentativa de um livro didático completo de biblioteconomia ou orientação sobre o perfeito gerenciamento de um bibliotecário em forma científica.

Ferraz (1991), com base em Verona (1985),¹⁸ indica que Schrettinger foi pioneiro na produção de um catálogo alfabético para as entradas de autor e título em alemão, decorrente de um levantamento sistemático e detalhado. Todavia, Baillet e antecessores haviam trabalhado nesse esforço, mas o foco não era o idioma alemão. Ainda segundo a autora, algumas propostas de Schrettinger apresentavam inconsistências, apesar de muitas bibliotecas na Alemanha concordarem com o que foi estruturado. O trabalho de vários referenciais alemães, inclusive o de Schrettinger, culminou com a elaboração das *Instruktionen für die alphabetischen Kataloge der Preussischen Bibliotheken* (Instruções para os catálogos alfabéticos das bibliotecas prussianas), mais conhecida como Instruções Prussianas, em 1899. Foi um código de catalogação, com regras relativas às autorias corporativas e às obras anônimas, adotado amplamente na Alemanha e em outros países europeus que teve grande influência na Europa, inclusive para fins de catalogação cooperativa.

López Guillamón (2011) corrobora o trabalho de Schrettinger e o considera de forte relevância para a discussão da catalogação bibliográfica, que tem como característica a individualização de cada documento. Ainda na concepção de López Guillamón (2011, p. 360,

¹⁷ “[...] a substantial collection of books, whose arrangement sets each person who is eager to know in the position to use, without loss of time, any treatise in it according to his need.” (SCHRETTINGER, 1808 *apud* BUCKLAND, 2005, p. 24).

¹⁸ VERONA, Eva. Literacy unit versus bibliographic unit. In: CARPENTER, M.; SVENONIUS, E. **Foundations of cataloging**: a sourcebook. Littleton: Livraries Unlimited, 1985.

tradução nossa), “[...] M. Schrettinger estabelece uma localização conceitual a partir de um critério de ordem alfabética, em relação à sistematização disciplinar. Na realidade, esse pensamento é consequência de sua reflexão durante décadas, como expressa A. Serrai¹⁹ sobre esse autor.”²⁰

Buckland (2005) enfatiza no trabalho de Schrettinger o aspecto da busca por técnicas que impulsionem o acesso aos livros.

As técnicas para acesso temático aos documentos decorrem da expressão singular das demandas sociais por informação. Se não houvesse demandas de organização para o acesso, as discussões de recuperação da informação, quer seja em catálogos, bases de dados ou portais de publicações eletrônicas, não requereriam elaboração sobre o aspecto temático dos documentos. Para a presente pesquisa, o caminho dos meios e métodos que instauram as formas de pensar as técnicas são, portanto, de total interesse.

O quadro apresentado permite compreender os motivos de, em meados do século XX, Verner (1968) considerar que Schrettinger na Europa (século XVIII), e depois Cutter nos Estados Unidos (século XIX), serem os sucessores do trabalho de Baillet (século XVII). A partir desse quadro, seguem as ideias que demonstram a configuração da noção de assunto como atributo da representação da informação do século XIX aos dias de hoje.

¹⁹ SERRAI, Alfredo. **In difesa della biblioteconomia**: indagine sulla identità, le competenze e le aspirazioni di una disciplina in cerca de palingenesi. Florencia: La Nuova Italia, 1981.

²⁰ “[...] M. Schrettinger establezca una localización conceptual a partir de una razón de orden alfabética, en relación con sistematización disciplinar. En realidad, este pensamiento es consecuencia de su reflexión durante décadas, como expresa A. Serrai sobre este autor.” (LÓPEZ GUILLAMÓN, 2011, p. 360).

4 PRIMEIRO MOMENTO DE CONSTRUÇÃO DA NOÇÃO DE ASSUNTO: O CONTEXTO ANGLO-SAXÃO NO SÉCULO XIX

No decorrer dos séculos XVII ao XIX, a disseminação das bibliotecas públicas, especialmente nos Estados Unidos, ocorreu relacionada ao movimento de educação da população; neste momento também a industrialização implicou a ampliação do consumo (SHERA; EGAN, 1961). Fatores como esses, entre outros, promoveram maior demanda pelo uso de bibliotecas, por sua vez, exigindo mais recursos de acesso temático. Foi nesse contexto que o controle dos assuntos em listas de cabeçalhos se estruturou.

Alguns estudos e práticas tratados até aqui conviveram independentes, outros são correlacionados. Desde o final do século XV, avanços foram garantidos pelas atividades de bibliografia e, no século XVII, o catálogo como instrumento de busca começa a exercer o papel que conhecemos hoje. O desenvolvimento científico se fez presente na Europa principalmente depois do século XVIII e a França foi o lugar que propiciou a criação do primeiro código nacional de catalogação, em 1791, escrito por Jean-Baptiste Massieu e intitulado *Instruction pour proceder à la confection du catalogue de chacune des Bibliothèques sur lesquelles de Directoires onte dû ou doivent incessamment apposer les scellés* como decorrência da Revolução Francesa (STROUT, 1956), assim como, o uso inicial do catálogo em ficha (FIÚZA, 1987).

Mas, foi no contexto anglo-saxão no século XIX, que a vertente denominada Catalogação de Assunto (*Subject Cataloguing*) se desenvolveu com ênfase na elaboração de catálogos como produto, com forte influência dos princípios alfabéticos e das listas de cabeçalhos de assunto (GUIMARÃES, 2009). Assim, esse século marcou a Biblioteconomia pelas contribuições que recebeu de teóricos da Catalogação, como por exemplo, de Anthony Panizzi (1797-1879) e Charles Jewett (1816-1868), que se envolveram com a formulação de princípios para as práticas realizadas. Da mesma forma, Charles Cutter (1837-1903) se destacou por investir esforços na abordagem dos assuntos e na tradição do catálogo dicionário.

4.1 AS INFLUÊNCIAS DE ANTHONY PANIZZI E CHARLES JEWETT

A produção dos registros documentários é denominada catalogação, termo frequentemente empregado para referir-se à produção de catálogos de bibliotecas. A forma

‘catalogação’ é aqui compreendida como processo documentário envolvido por normativas que subsidiam a prática profissional; já ‘Catalogação’ é adotada para designar fundamentos teóricos, conceituais e metodológicos que congregam o processo. Essa, por sua vez, cobre aspectos da Catalogação Descritiva e da Catalogação de Assunto. Em certa medida, estudiosos da primeira fomentaram o desenvolvimento da segunda.

De forma ampla, a predominância da Catalogação tem reconhecimento nos estudos de Anthony Panizzi (1797-1879), considerado pioneiro na consolidação de princípios de Catalogação. Seu trabalho no *British Museum* subsidiou um compêndio de regras para a elaboração do catálogo da biblioteca e, publicado em 1841, foi denominado *Catalogue of Printed Books in the British Museum*, amplamente conhecido como as “91 Regras de Catalogação de Panizzi” (GARRIDO ARILLA, 1999).

Seu trabalho de confecção do catálogo cobriu todos os livros, mapas e partituras existentes no *British Museum*, iniciando em 1837, quando foi nomeado bibliotecário chefe. Nos quatro anos subsequentes permaneceu devotado ao trabalho e às discussões perante uma comissão do Parlamento Britânico encarregada de aprovar o novo catálogo e, pelas inúmeras defesas obteve êxito, demonstrando a importância das suas Regras (FIÚZA, 1987). Segundo López Guillamón (2011), a equipe de Panizzi era composta por Thomas Watt, J. Winter Jones, John Humffreys Parry e Edward Edward.

Para o registro do documento, até a primeira metade do século XIX, prevalecia o uso de catálogos na forma de livro e, da segunda metade em diante, intensificou-se o uso das fichas. As entradas eram arranjadas de maneira que melhor pudessem ser recuperadas sendo, então, datilografadas ou escritas à mão. Essa era a tecnologia disponível à época que influenciou o trabalho de Panizzi (TILLET, 1989).

Em suas Regras, Panizzi chamou atenção às instruções para três principais aspectos: a descrição completa de um registro, de forma única, por uma entrada; especificação do uso de referências cruzadas em três tipos – nome para nome, nome para obra e obra para obra (TILLET, 1989); e a criação de catálogos alfabéticos, principalmente de autores pessoais ou corporativos, para propiciar a localização rápida e fácil de uma obra ou o seu agrupamento por um mesmo autor com suas distintas edições e traduções (GARRIDO ARILLA, 1999).

Segundo Tillett (1989), a descrição, feita uma única vez por meio de uma entrada, foi consequência das restrições econômicas observadas à época em relação ao tamanho e à disposição dos catálogos na forma de livros e, após, fichas. Por sua vez, a orientação para os três tipos de referências cruzadas propiciaria o acesso de forma completa às entradas, pois os catálogos deveriam ser mais do que listas de localização ou de inventários. Assim, essas

referências se constituíram como auxílio a quem pesquisa e, em função da primeira, nome para nome, se estabeleceu para a variação de um nome a forma escolhida para a entrada; na segunda, nome para obra, os títulos das obras foram considerados importantes para a entrada principal, incluindo para a entrada da obra nomes pessoais, corporativos, de conferência, geográficos, bem como o próprio nome da obra, ou seja, seu título; já a terceira, obra para obra, deu-se na direção de uma obra a outra, quer seja por inteiro ou em suas partes. Ainda de acordo com a mesma autora, as referências cruzadas, a partir das Regras de Panizzi, foram a base para as orientações normativas do *Anglo-American Cataloguing Rules* (AACR, ou, Código de Catalogação Anglo-Americano).

Por sua vez, a criação de catálogos alfabéticos, principalmente de autores pessoais ou corporativos, decorreu do fato de Panizzi ter se preocupado com as entradas cuja autoria era desconhecida, como das obras anônimas, as que eram concebidas por instituições, as concebidas por vários autores, assim como as formas distintas das obras em suas edições e variações, como as traduções, adaptações, com comentários, entre outras.

Fiúza (1987, p. 46, destaque da autora) explica que Panizzi sistematizou a noção de obra, iniciada na Biblioteca Bodleyana em 1732 e, para ele, “o livro procurado por uma pessoa não é realmente, na maioria das vezes, o objeto de seu interesse, mas a **obra** nele contida; esta **obra** pode ser encontrada em outras edições, traduções e versões, publicada sob diferentes nomes do autor e diferentes títulos [...]”. A noção de obra, proposta por Panizzi, como parte dos princípios para a produção de catálogos em bibliotecas é retomada contemporaneamente pelo *Functional Requirements for Bibliographic Records* (FRBR, Requisitos Funcionais para Registros Bibliográficos), que será abordado à frente junto ao modelo FRSAD. No FRBR, as entidades “manifestação” e “expressão” de uma mesma obra são agrupadas de modo a permitir ao público, durante a busca no catálogo, selecionar a que melhor corresponda às suas necessidades informacionais.

De modo geral, as Regras de Panizzi chamaram atenção para as entidades que deveriam ser observadas, como o título e os nomes das obras e os responsáveis por elas. Também sinalizou que os termos referentes aos nomes para entrada da ficha, como pessoas, entidades, nomes geográficos, título e, de algum modo, até o conteúdo eram formas que compunham as possibilidades de recuperação nos catálogos (PANIZZI, 1841).

Desde seu ingresso na instituição que remonta a 1831, os anos de trabalho de Panizzi no *British Museum* resultaram na organização dos catálogos históricos, datados de 1787, e na elaboração de outros, mas principalmente no que Garrido Arilla (1999) explica ser o primeiro código para uma grande biblioteca nacional. Ainda segundo a autora, as Regras de Panizzi

influenciaram os códigos de catalogação posteriores, tanto os europeus como o anglo-americano.

Essa influência se ramificou nos Estados Unidos, no final do século XIX e início do seguinte, pois Charles Coffin Jewett (1816-1868) realizou a revisão das 91 Regras de Panizzi e publicou um código de catalogação para a biblioteca da *Smithsonian Institution* (FIÚZA, 1987).

De acordo com Garrido Arilla (1999), o *Smithsonian Institution*, criado pelo governo estadunidense em 1846, foi uma agência de promoção do desenvolvimento científico. Jewett foi convidado para dirigir um projeto audacioso no *Institution* sobre a produção de um catálogo nacional centralizado com abrangência de todos os acervos das bibliotecas públicas dos Estados Unidos. Para isso, recorreu a Panizzi e, a partir das suas Regras, criou a normativa *Rules for preparing catalogues*, dividida em duas partes. De acordo com López Guillamón (2011, p. 411, tradução nossa):

Sua estrutura é igualmente significativa: aspectos descritivos sobre a transcrição dos títulos (I-XII); cabeçalhos (XIII-XXIX); relações internas do catálogo (XXX-XXXI); ordenação das entradas bibliográficas em fichas (XXXII-XXXVII); e indicações para a catalogação de materiais especiais (XXXVIII-XXXIX); e uma série importante de exemplos representativos da mesma normativa e um índice de materiais que incluem os exemplos citados.²¹

Entretanto, apesar do apoio inicial obtido e tendo como parte do resultado as *Rules*, segundo Garrido Arilla (1999), Jewett saiu do projeto, pois o *Smithsonian Institution* na pessoa de Joseph Henry, que era seu chefe, não foi favorável à utilização de boa parte dos recursos financeiros no projeto de produção de um catálogo nacional centralizado, optando por investir nas pesquisas desenvolvidas pelo Instituto. Todavia, a proposta de Jewett foi aceita e implementada anos depois na *Library of Congress* (LC).

Outra inovação que coube a Jewett, implementada somente mais tarde pela LC, foi a gravação de cada ficha catalográfica em placa estereotipada, a qual permitiria a reprodução das fichas, possibilitando flexibilidade ao catálogo e a cooperação entre várias bibliotecas (FIÚZA, 1987).

²¹ Su estructura es igualmente significativa: aspectos descriptivos sobre la transcripción de los títulos (I-XII); encabezamientos (XIII-XXIX); relaciones internas del catálogo (XXX-XXXI); ordenamiento de los asientos bibliográficos en fichas (XXXII-XXXVII); e indicaciones para la catalogación de materiales especiales (XXXVIII-XXXIX); más una serie importante de ejemplos representativos de la normativa misma y un índice de materias incluidas en los ejemplos citados. (LÓPEZ GUILLAMÓN, 2011, p. 411).

Ademais, em 1855, Jewett iniciou suas atividades como bibliotecário na *Public Library of Boston* ocupando o cargo de diretor por 10 anos, de 1858 a 1868, e na época era a terceira biblioteca mais importante do país. Ele foi o responsável em publicar índices do catálogo, no triênio de 1858 a 1861, e as entradas foram organizadas da seguinte forma: por ordem alfabética de autor, pela primeira palavra significativa do título e por uma ou mais palavras que expressassem o conteúdo abordado no documento, mas que eram retiradas do título (LÓPEZ GUILLAMÓN, 2011). Por isso, de acordo com o mesmo autor, o trabalho de Jewett foi inicial para a institucionalização do catálogo dicionário de Cutter.

4.2 CHARLES CUTTER E A SISTEMATIZAÇÃO DOS CABEÇALHOS DE ASSUNTO

Charles Ami Cutter (1837-1903) é considerado o grande teórico da Catalogação de Assunto (*Subject Cataloguing*) na literatura de origem anglo-saxã. A organização de catálogos por assunto nos Estados Unidos foi tratado por Cutter, criador de regras que culminaram com a difusão do catálogo dicionário (TAYLOR, 1948). Buonocore (1976, p. 118, tradução nossa) define esse catálogo como a fusão de três diferentes tipos: de autor, de assunto e de título. Para este autor, a dificuldade a ser enfrentada em relação ao catálogo dicionário, é a “[...] escolha apropriada da palavra-chave [...] que sirva para caracterizar o assunto tratado na obra, escolha que o catalogador faz com o auxílio de listas [...] para tais fins.”²²

Mann (1962, p. 151) explica que no catálogo dicionário “[...] tôdas [sic] as palavras de entrada (autor, título, assunto) são arrumadas numa só lista alfabética. [...] Esse catálogo é assim chamado por sua arrumação semelhante a um dicionário.” Ainda segundo a mesma autora, palavra de entrada é o mesmo que palavra-chave, responsável em ocupar a parte da entrada sob a qual a ficha é arquivada no catálogo. Logo, a ficha, que é o suporte do registro de um documento, leva em consideração diferentes atributos para sua entrada.

Cutter trabalhou com Jewett na *Public Library of Boston*, objetivando a construção de um catálogo para uma coleção especial (FIÚZA, 1987). Todavia, quando Cutter ainda era aluno da *Harvard Divinity School*, tornou-se um auxiliar de biblioteca, entre 1857-1859, e o cargo foi determinante para sua carreira na criação da sua obra mais conhecida, as *Rules for a*

²² “[...] acertada elección de la palabra clave [...] que sirve para caracterizar el asunto de que trata la obra, elección que hace el catalogador con el auxilio de listas [...] para tales fines.” (BUONOCORE, 1976, p. 118).

Printed Dictionary Catalogue. O então bibliotecário e catalogador chefe, Ezra Abbot, notou seu desempenho e, em 1860, convidou-o para a equipe do *Harvard College*. Já no final do ano de 1868, em decorrência do trabalho desenvolvido com Abbot, Cutter recebeu outro convite: ser bibliotecário do *Boston Athenaeum* que, junto com *Harvard*, eram os pontos principais da influente comunidade intelectual de Boston-Cambridge (MIKSA, 1977; WORLD ENCYCLOPEDIA OF LIBRARY AND INFORMATION SERVICES, 1993). Ao assumir o cargo no *Boston Athenaeum*, ele destacou como primordial “[...] a compilação de um catálogo impresso. Usou, então, a sua experiência profissional influenciada diretamente por Jewett e, indiretamente por Panizzi, [...] para estabelecer suas <<Regras para um Catálogo Dicionário Impresso>>.” (FIÚZA, 1987, p. 48).

Garrido Arilla (1999, p. 80, tradução nossa) aponta que as *Rules* constituíram um código que permitia “[...] elaborar um registro bibliográfico por três entradas: autor, título e assunto.”²³ Anteriormente, no entanto, Panizzi estruturou nas 91 Regras orientações para a criação de catálogos alfabéticos, principalmente de autores pessoais e corporativos, que permitissem ao público a localização rápida e fácil de uma obra, além do agrupamento das obras de um mesmo autor com suas distintas edições e traduções.

Mas, Cutter, ao desenvolver as *Rules for a Printed Dictionary Catalogue* (1876), que ficou conhecida como *Rules for a Dictionary Catalog*,²⁴ preocupou-se em construir regras (369 ao total) que norteassem a decisão das entradas no catálogo, incluindo a de assunto (GARRIDO ARILLA, 1999). Sem tais orientações, o catalogador adotaria diferentes cabeçalhos para os mesmos assuntos, gerando dúvidas no processo de busca pelo público. Dentre as regras de Cutter, algumas serão citadas a frente para dar conta do que ele predispôs ser a elaboração do assunto, sendo que muitas delas têm relação direta com a entrada no catálogo.

Segundo Mann (1962, p. 176) “por entrada de assunto queremos dizer o termo ou termos usados num catálogo [...] para exprimir o assunto ou assuntos dos livros.” De acordo com essa autora, as *Rules* tornou-se a principal referência, em inglês, para o estabelecimento de regras na elaboração de cabeçalhos de assunto, que formaram as listas. Além de termos, palavras, frases, nomes ou expressões foram empregadas para reunir as listas de cabeçalhos de

²³ “[...] redactar un registro bibliográfico por tres entradas: autor, título y materia.” (GARRIDO ARILLA, 1999, p. 80).

²⁴ As regras para catálogos do tipo dicionário de Cutter receberam o título *Rules for a Printed Dictionary Catalogue* nas primeiras edições: 1ª (1876), 2ª (1889) e 3ª (1889). Já a 4ª e última edição, de 1904, foi intitulada *Rules for a Dictionary Catalog* (WORLD ENCYCLOPEDIA OF LIBRARY AND INFORMATION SERVICES, 1993).

assunto, as quais serão retomadas a frente. Ademais, Cutter (1904, p. 21) também definiu cabeçalho (*heading*) como a palavra por ordem alfabética que fornece o lugar da entrada no catálogo e pode ser determinado pelo nome do autor, do assunto ou palavra do título.

Em referência ao cenário retratado até aqui, nota-se que elaborar orientações para as entradas dos catálogos não foi uma preocupação apenas de Cutter, mas ele criou orientações normativas para dar forma à representação alfabética no catálogo e, assim, por mais de meio século as *Rules* permaneceu como a única sistematização disponível para pensar e estruturar os assuntos pelos cabeçalhos.

Com o acesso às bibliotecas pela população,²⁵ constata-se a continuidade de alteração da percepção do catálogo, de simples ferramenta de registro e inventário do acervo, para instrumento propiciador do livre acesso às consultas das bibliotecas, inclusive as públicas. Conforme demonstrado, tais mudanças decorreram ao longo dos séculos sob variadas influências até alcançar orientações consistentes para a organização dos catálogos. No entanto, o trabalho é tomado como útil quando indivíduos incorporam essa organização para alterar o seu estado de conhecimento. É nesse ponto que a demanda por informação organizada, necessitou de parâmetros para pensar os assuntos dos documentos e se, anteriormente, os próprios bibliotecários atuavam como catálogos vivos, para auxiliar o acesso aos acervos, tal situação se tornou inconcebível (SHERA; EGAN, 1969).

Cutter recebeu apoio do *Bureau of Education*, órgão do governo dos Estados Unidos, que por sua vez, solicitou um relatório sobre a história, condição e administração das necessidades das bibliotecas públicas. Com isso, as *Rules* foram concluídas em 1875, e publicadas em primeira edição como a parte dois (*Part II*) do relatório *Public Libraries in the United States of America: their history, condition and management*, conforme é possível observar nas figuras 2 e 3 (MIKSA, 1977; WORLD ENCYCLOPEDIA OF LIBRARY AND INFORMATION SERVICES, 1993).

Essa divulgação sinalizou, mesmo que indiretamente, a importância de um instrumento sistematizador, aprovado por um órgão do governo, que desse conta de indicar os

²⁵ A ‘população’ que fazia uso das bibliotecas era constituída por um público restrito de pessoas, ou seja, um público singular assíduo a esse ambiente com uma maneira própria de buscar informação. Cutter (1904) usou o termo ‘*the users*’. Ao utilizar o artigo *the* antes do substantivo, o autor dá indícios de considerar a existência de outros grupos de usuários, mas presume-se a homogeneidade de um público à época: homem branco, heterossexual e com certo nível de escolaridade. Para Olson (2001, p. 641, tradução nossa), a lógica utilizada por Cutter “[...] introduz uma aparente imposição democrática de que o catálogo deveria ser construído para a conveniência do público ao qual ele serve”.

OLSON, H. A. The power to name: representation in library catalogs. *Journal of Women in Culture and Society*, Chicago, v. 26, n. 3, p. 639-668. 2001.

atributos dos documentos e a forma de organização na ficha. Os acervos gerais das bibliotecas públicas demandaram uma organização para o seu uso.

Figura 2 – Capa *Rules for a Printed Dictionary Catalogue*

PUBLIC LIBRARIES
 IN THE
 UNITED STATES OF AMERICA
 THEIR
 HISTORY, CONDITION, AND MANAGEMENT
 SPECIAL REPORT
 DEPARTMENT OF THE INTERIOR, BUREAU OF EDUCATION

Fonte: Cutter (1876).

Figura 3 – Folha de rosto *Rules for a Printed Dictionary Catalogue*

—————
 PART II
 —————
 WASHINGTON
 GOVERNMENT PRINTING OFFICE
 1876

Fonte: Cutter (1876).

Também se constata que o referido documento forneceu visibilidade para a Biblioteconomia, além de provar certa preocupação do governo estadunidense com as bibliotecas. Além das *Rules*, Cutter contribuiu para a Biblioteconomia com o seu sistema de classificação denominado *Expansive Classification* (Classificação Expansiva) e a Tabela de Autores. Suas contribuições se deram, ainda, como membro fundador da *American Library*

Association (ALA) e editor de seu periódico *Library Journal* (WORLD ENCYCLOPEDIA OF LIBRARY AND INFORMATION SERVICES, 1993).

Fiúza (1980, p. 142-143) explica que “conforme predissera Cutter, na introdução da 4ª edição de suas regras, as preocupações teóricas foram subordinadas às considerações práticas da adoção das fichas da *Library of Congress*.” De fato, devido à forte influência da ALA e da LC na crescente difusão das fichas catalográficas impressas nos primeiros anos do século XX, persistiu a prática de elaboração de produtos em detrimento da discussão teórica realizada anteriormente, conforme atesta Guimarães (2009). A venda das fichas centrou-se no desenvolvimento de listas de cabeçalhos de assunto que pudessem ser rapidamente aplicadas, já que as bibliotecas recebiam as fichas prontas da LC.

Mas, em que medida o assunto foi uma preocupação nas *Rules*, visto ter sido um dos documentos mais conhecido em língua inglesa e responsável por indicar essa sistematização? Para essa compreensão, primeiro é necessário observar à abordagem da *Rules*, pois sua divisão preponderou em duas grandes partes, sendo a primeira (ou A) dedicada à Entrada (*A. Entry – Where to enter*, ou, Onde entrar) e, a segunda (ou B), dedicada ao Estilo (*B. Style – How to enter*, ou, Como entrar). Antecedendo as duas, encontram-se os Comentários Gerais (*General Remarks*), que se subdivide em Objetivos (*Objects*), Meios (*Means*) e Razões para escolha (*Reasons for choice*).

Formas para a ‘Entrada’ no catálogo já haviam sido previstas por Panizzi, mas o ‘Estilo’ toma significância por propor critérios pragmáticos para uma política do catálogo. Toma-se proximidade à Entrada e a busca no catálogo. Cutter propõe os Objetivos do catálogo, como segue (CUTTER, 1904, p. 12, tradução nossa):

1. Permitir a uma pessoa encontrar um livro do qual

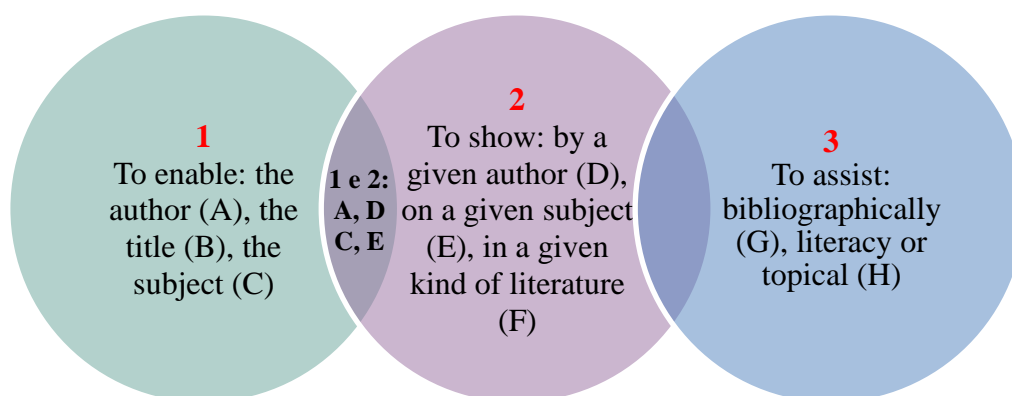
(A) o autor	}	é conhecido
(B) o título		
(C) o assunto		
2. Mostrar o que a biblioteca possui
 - (D) por um autor determinado
 - (E) com um assunto determinado
 - (F) em um tipo de literatura determinado
3. Ajudar na escolha de um livro
 - (G) quanto a sua edição (bibliograficamente)
 - (H) quanto ao seu caráter (literário ou tópico).²⁶

²⁶ 1. To enable a person to find a book of which either
 (A) the author
 (B) the title } is known
 (C) the subject
 2. To show what the library has

O primeiro item propõe permitir o encontro de um livro através de ao menos um dos três diferentes elementos: autor-A, título-B e assunto-C. O segundo item busca mostrar o que a biblioteca possui, por ao menos um dos três elementos: autor-D, assunto-E e tipo de literatura-F, sendo que autor-D e assunto-E, também estão inseridos na possibilidade anterior (A e C). Já o terceiro item consiste em auxiliar a escolha de um livro por dois elementos: bibliograficamente-G e literário ou tópico-H. O esquema apresenta os elementos necessários para buscas no catálogo e isentam relações menos habituais à época.

De outro modo, no primeiro objetivo, pela identificação de um livro no catálogo é possível alcançar o segundo objetivo, que mostra o que a biblioteca tem, possibilitando o agrupamento de atributos dos livros. Já o terceiro objetivo ajuda a escolher um livro por sua edição, ou caráter literário ou tópico, pois o usuário ao encontrar livros similares necessitará realizar uma seleção, por exemplo, entre as várias edições disponíveis. É possível refletir, ainda, que no século XIX não havia abundância de muitas edições dos livros, mas foi ponto merecedor de atenção. Na figura 4 é sistematizada a relação entre os três Objetivos.

Figura 4 – Sistematização dos três Objetivos das *Rules* de Cutter (1904)



Fonte: elaborado pela autora.

-
- (D) by a given author
 (E) on a given subject
 (F) in a given kind of literature
 3. To assist in the choice of a book
 (G) as to its edition (bibliographically)
 (H) as to its character (literacy or topical). (CUTTER, 1904, p. 12).

A interseção apresentada na figura 4 é explícita em A e D, C e E, mas não da mesma maneira em F e H. Fiúza (1987, p. 49), explica que para alcançar os Objetivos, que ela denominou de “Objetos”, o catálogo conta com seis Meios: “1. Entradas de autor, com as referências necessárias (A e D); 2. Entradas de título ou referências ao título (B); 3. Entradas de assunto, referências cruzadas (C e E); 4. Entrada de forma e língua (F); 5. Transcrição da edição e imprensa, com notas, quando necessário (G); 6. Notas especiais (H).” Em alguma medida, os Objetivos aproximam-se dos Meios, e os aspectos principais são estabelecidos pelos atributos autor, título e assunto que moldam a forma da entrada do catálogo dicionário. O assunto propiciaria a referência cruzada (*cross reference*), que remete um assunto para outro, e essa recomendação de Cutter consistiu na criação de uma rede de remissivas, pelo Princípio Sindético, para os catálogos. Mas, a preocupação central não foi com o assunto e sim com a sua forma de entrada.

Por isso, Cesarino e Pinto (1978) elucidam que a abordagem alfabética de Cutter consistiu em uma rede bem construída de referências cruzadas com o intuito de unir os assuntos relacionados. Mas,

Cutter, entretanto, na prática, limitava os vínculos a remissivas descendentes, isto é, de assuntos mais genéricos para assuntos mais específicos, sugerindo que devíamos ignorar, por razões econômicas, tanto os vínculos ascendentes (de assuntos mais específicos para assuntos mais genéricos) quanto os vínculos colaterais de um termo para outro de igual categoria. Assim, Cutter teria remissivas como:

Literatura *see also* drama

Mas não

Drama *see also* Literature

ou

Drama *see also* Poetry. (FOSKETT, 1973, p. 48).

Se a rede era limitada por questões econômicas, a preocupação com a fundamentação do assunto ficou em segundo plano e não houve a difusão de algo diferente do que já era realizado. Conforme já citado com base em Strout (1956), desde o início do século XVIII as referências cruzadas já eram bastante comuns. A complexidade das técnicas de organização cresceu “à medida que aumentavam os impressos tanto em número como em variedade [...]” (SHERA; EGAN, 1969, p. 11).

Buonocore (1976, p. 367) explica que o termo referência (*referencia*), no sentido amplo, está relacionado com o processo de catalogação nas chamadas referências cruzadas (*referencias cruzadas*) “ver” (*véase*) e “ver também” (*véase también*). Portanto, há que se considerar nesse contexto, remissiva e referência como sendo palavras correlatas.

Ainda em relação aos elementos F, G e H dos Objetivos, um tipo de literatura determinado, tanto à edição (bibliograficamente) quanto o caráter (literário ou tópico) do livro poderiam compor uma outra forma de reunião, pois os livros que detém várias edições poderiam ser representados não como uma entidade separada, mas sim de forma agrupada, situação que comungaria com o Princípio de Uso.

Pautando-se em Garrido Arilla (1999), segue a descrição dos três princípios de Cutter:

- 1) Princípio Específico: o livro deve ser incluído no catálogo com o cabeçalho de assunto mais específico possível, fator esse determinante para a economia e uniformidade do instrumento;
- 2) Princípio de Uso: os cabeçalhos de assunto devem utilizar termos que os usuários do catálogo esperem encontrar;
- 3) Princípio Sindético: os cabeçalhos de assunto empregados nos catálogos devem ter conexão com referências que conduzam o usuário de uma forma que não é autorizada ou utilizada, para a que seja autorizada e utilizada pelos usuários. Verifica-se na *Rules* que os Princípios em sua ordem são guiados, respectivamente, pelos números §161, §169 e §187 (CUTTER, 1904).

Ao tratar das Razões para escolha (*Reasons for choice*), a fim de alcançar os Objetivos propostos, Cutter (1904) explica que pode haver o emprego de vários métodos. Entretanto, ele relaciona três pontos e os vincula à forma de entrada no catálogo dicionário. Primeiro, indica a necessidade de observar o público de pessoas que usam a biblioteca; segundo, a consistência e cobertura das entradas; e, por último, se for uma entrada ampla para um grande público, que ele chamou de entradas de massa (*mass entries*), deve-se perceber as possíveis dificuldades em empregar essa entrada e adotar a que melhor auxilie a busca. Os três tópicos são bastante abrangentes, não especificam nenhum método, não aludem aos elementos dos Objetivos, mas se aproximam dos três Princípios.

Na *Rules* ainda é possível encontrar uma lista alfabética de termos com respectivas definições e destaca-se: “*assunto*, o tema ou temas do livro, esteja ou não colocado em seu título.”²⁷ (CUTTER, 1904, p. 23, tradução nossa, destaque do autor). Mann (1962, p. 150) apesar de não citar Cutter, concorda com o autor, pois define “assunto – o tema ou os temas principais de um livro declarados ou não no título.”

Essa definição breve acentua a associação de assunto com outra palavra – tema – sem, contudo, os autores aprofundá-las. Também é notado que o tema pode ser atribuído a partir do título, fato que remete à indicação das palavras-chave (*catchwords*). Esse aspecto, conforme

²⁷ “*Subject*, the theme or themes of the book, whether stated in the title or not.” (CUTTER, 1904, p. 23, destaque do autor).

apontado por Verner (1968), foi anterior a Cutter, como citado no caso do catálogo do Mosteiro de Melk (Áustria), datado de 1483, que possuía três formas de entradas para os índices, sendo uma delas a de obras anônimas com base nas palavras-chave (*catchwords*) do título, que constituíam mais palavras significativas do texto do que propriamente *catchwords*. Buonocore (1976, p. 328, tradução nossa, itálico do autor), define palavra-chave, da seguinte forma: “é a palavra com característica representativa ou mais significativa do título do livro. Em inglês: *catchword*. O catalogador a seleciona para formar as listas de *epígrafes*.”²⁸

Ainda nas definições, as palavras catálogo, entrada e cabeçalho demonstram o papel do assunto nas *Rules*:

Catálogo de assunto, um catálogo de assuntos, organizado em classes ou alfabeticamente por nomes de assuntos.
 Entrada de assunto, registro de um livro sob seus assuntos. (Ver §161-188)
 Cabeçalho de assunto, o nome de um assunto usado como cabeçalho sob o qual livros relacionados a esse assunto são inseridos. (CUTTER, 1904, p. 23).²⁹

A despeito dessas e, conforme já descrito, entrada também se refere a uma parte da ficha incorporada no catálogo, que através do cabeçalho ganha um nome controlado para o assunto atribuído ao livro. A coesão entre catálogo, entrada e cabeçalho levaria às indicações de assuntos escolhidos de forma uniforme em fichas catalográficas e, neste sentido, constituíram palavras recorrentes e estáveis. Seguindo essa proposta, é sabido que a *Rules* foi pensada para o catálogo dicionário (que é alfabético) e, de acordo com Gomes e Marinho (1984), a principal característica para esse tipo de catálogo é a entrada direta pelas palavras. Com isso, a primeira regra para a entrada de assunto, a 161,³⁰ impunha a condição: “Insira uma obra sob o cabeçalho de assunto, não sob o cabeçalho da classe com inclusão desse assunto.” (CUTTER, 1904, p. 66, tradução nossa).

Gomes e Marinho (1984, on-line), apontam que no catálogo dicionário “[...] a entrada de assunto é direta, sem mediação de um termo mais amplo, sendo o cabeçalho tirado do uso comum.” Ainda segundo os autores, “[...] foi para atender ao cidadão comum que se

²⁸ “Es la palabra típica característica o más significativa del título del libro. En inglés: *catchword*. El catalogador las selecciona para formar las listas de *epígrafes*.” (BUONOCORE, 1976, p. 328, itálico do autor).

²⁹ *Subject catalog*, a catalog of subjects, whether arranged in classes or alphabeted by names of subjects.
Subject entry, registry of a book under its subjects. (See §161-188)
Subject heading, the name of a subject used as a heading under which books relating to that subject are entered. (CUTTER, 1904, p. 23, destaque do autor).

³⁰ “161. Enter a work under its subject-heading, not under the heading of a class which includes that subject.” (CUTTER, 1904, p. 66).

desenvolveram técnicas e serviços biblioteconômicos tais como o catálogo dicionário [...].” A despeito disso, a entrada do assunto de forma direta e a prioridade de quem utiliza o catálogo fazem parte, respectivamente, dos Princípios Específico e de Uso de Cutter.

A entrada direta sem subdivisão e a forma usual das palavras envolveram a linguagem natural, mais próxima possível do usuário como base para o estabelecimento dos cabeçalhos a partir das listas. Se “[...] na Europa, o catálogo sistemático e o alfabético por classe, [eram tidos como os] mais apropriados a bibliotecas usadas por um público erudito [...]” (GOMES; MARINHO, 1984, on-line), no contexto anglo-saxão, a partir do trabalho de Cutter, é possível pressupor influência do uso de um vocabulário natural. Com isso:

A filosofia do cabeçalho de assunto tem sido, desde Cutter a de preferir a forma usada pelo usuário. *“Usage is the supreme arbiter the usage, in the present case, not of the cataloger but of the public in speaking of subjects”*. A gramática e a sintaxe do cabeçalho de assunto foram estabelecidas, portanto, a partir da gramática (da língua inglesa) e o vocabulário ditado pelo uso do público. (GOMES; MARINHO, 1984, on-line).

Mann (1962, p. 176), com base em Cutter, afirma que o vocabulário das listas de cabeçalhos deve preferir a forma usada pelo usuário, mas “[...] no que diz respeito à arrumação das entradas de assunto sua relação com as outras partes do catálogo-dicionário” precisa ser considerada. Reconhece-se a proximidade dessa afirmação com os Objetivos propostos por Cutter, como também com seus Princípios, sobretudo o sindético. Por outro lado, Mann (1962, p. 182) chamou atenção na escolha de termos para os cabeçalhos: “deve ficar estabelecido logo de início que não há regras firmes e definitivas para a escolha de termos [sic] que se apliquem em todos os casos. Os livros não são todos iguais e os assuntos mudam à medida que se alargam os conhecimentos.” Porém, segundo a mesma autora, quem cataloga, ao usar os termos que representam os assuntos precisa considerar que “a palavra ou frase deve ser apropriada para cada livro e ao mesmo tempo não deve estar limitada àquele livro, mas ser aplicável a um grupo de livros.” (MANN, 1962, p. 183).

Mann (1962) apontou a necessidade de flexibilidade nas regras. Além disso, considerou as particularidades dos assuntos dos livros, mas, em seguida, opôs-se ao indicar um limite que não individualiza, mas agrupa. Centra preocupação na arrumação das fichas no catálogo e na forma de unir os assuntos pelas palavras ou frases que constam nas listas de cabeçalhos de assunto. Portanto, sua preocupação foi expressa por uma forma de agrupamento por encaixe e não pela transferência da informação. Porém, a autora não explicita se essas escolhas decorrem por medida econômica de elaboração das fichas.

A flexibilidade não caracterizou a *Rules* nem em função do Princípio de Uso, pois ela levou em conta o Princípio anterior, ou seja, o Princípio Específico. Neste último, conforme exposto, um livro deve ser representado no catálogo sob o cabeçalho mais específico (entrada específica), sem intermediação de termos mais amplos, como por exemplo, uma classe que inclua o assunto.

De acordo com Cutter (1904, p. 67, tradução nossa), a “[...] ‘entrada específica’ é a principal distinção entre o catálogo dicionário e o catálogo alfabético classificado.”³¹ Porém, Tôres ([2000?]) aborda o fato de Cutter não ter definido o que é essa forma de entrada e aponta na *Rules*, em §161, haver somente uma contraposição dessa com a entrada por classe. Assim, Tôres ([2000?], on-line) explica:

É curioso observar que, ao rejeitar o catálogo alfabético classificado, Cutter tentou abolir a entrada por classe. Entretanto, todo cabeçalho de assunto é nome de classe ou representa classes de referentes (coisas, materiais, processos, propriedades, agentes, lugares, etc.), com exceção dos cabeçalhos individuais (ex.: nomes próprios personativos, locativos, etc.).

Nesse contexto, a efetividade das palavras pelos cabeçalhos, quer seja por uma classe ou não, se dá por nomes. Para Cutter (1904, p. 67, tradução nossa), “alguns assuntos não tem nome; eles são falados apenas por meio de uma frase ou de várias frases ainda não definitivas o suficiente para serem usadas como cabeçalho.”³²

Com isso, considera-se o assunto enquanto uma palavra, ou, uma frase, nomeado a partir das listas de cabeçalhos de assunto. Esse nome gera uma entrada na ficha catalográfica para ser organizada alfabeticamente no catálogo dicionário. É antes um mecanismo de entradas por agrupamento alfabético de palavras do que um agrupamento das ideias dos assuntos (TÔRES, [2000?]). Mais grave é não ter sido estabelecida na *Rules* uma base de sustentação entre nomes e assuntos. Esses nomes de assuntos existiam enquanto houvesse aceitação geral e fossem de uso do público.

Se, em Cutter (1904), os documentos são nomeados por palavras que agrupam os assuntos alfabeticamente, esse contexto é generalizado pela utilização das listas de cabeçalhos. Por sua vez, as listas “[...] como instrumento de representação e recuperação da

³¹ “[...] ‘specific entry’ is the main distinction between the dictionary-catalog and the alfabético-classed.” (CUTTER, 1904, p. 67).

³² “Some subjects have no nome; they are spoken of only by a phrase or by several phrases not definite enough to be used as a heading.” (CUTTER, 1904, p. 67).

informação registrada em documentos tem apresentado, no entanto, uma série de inconsistências relacionadas à sua construção e ao seu uso.” (TÔRRES, [2000?], on-line).

Segundo a mesma autora, os cabeçalhos compostos são formados por mais de uma palavra, enquanto os simples por apenas uma. É pela sintaxe que a palavra “[...] evidentemente só se manifesta no universo dos cabeçalhos compostos, de que são exemplos típicos os cabeçalhos-frase ou as expressões compostas.” (TÔRRES, [2000?], on-line). Assim,

Como não há consistência lógica na construção dos cabeçalhos compostos, as inversões, por exemplo, são tentativas frustradas de solução para o problema de elaboração da sintaxe. Cabeçalhos invertidos com vírgula ou hífen coexistem com cabeçalhos na ordem direta ou com parênteses, o que evidencia a falta de critérios consistentes para a determinação ao do ponto-de-acesso e a combinação dos componentes. (TÔRRES, [2000?], on-line).

Cabe destacar que os cabeçalhos compostos, da mesma forma que os simples, são lineares e foram pensados para dar conta de uma situação que os assuntos nomeados por cabeçalhos simples não fossem suficientes. Fato é que ambos estão no nível alfabético.

Em relação às outras regras na *Rules* relativas à palavra assunto, encontra-se: referência de assunto (*subject reference*), entrada do assunto pela palavra (*subject-word-entry*) e referência do assunto pela palavra (*subject-word-reference*). A primeira é usada para palavras sinônimas, cabeçalhos relacionados com alguma divisão necessária, vinculando-se aos números §187-188. Em relação à segunda, vale quando é feita sob uma palavra do título que indica o assunto do livro (§151). Quanto à terceira, trata-se de referências cruzadas (*cross-reference*) de um assunto para outro, no caso de assuntos correlatos. Todas as expressões subsidiam a ramificação das palavras em assunto, por mais que Cutter tenha indicado a entrada direta dele como sua principal forma de entrada. Sobre as três formas de uniformização das fichas – catálogo, entrada e cabeçalho de assunto (*subject catalog*, *subject entry* e *subject heading*) – conforme citado acima, demonstram o papel do assunto nas *Rules*, mas também pressupõe a relação entre a organização alfabética como forma de não dispersar os livros e os assuntos estabelecidos a partir das entradas. Porém, conforme já exposto, a forma alfabética une palavras, mas não aproxima ideias. Na proposta do levantamento realizado foi possível encontrar, ao total, seis expressões relacionadas ao termo assunto, conforme apresentação na figura 5.

Figura 5 – Definições relacionadas às expressões do termo assunto (*subject*)



Fonte: elaborado pela autora, com base no *Index das Rules* (CUTTER, 1904, p. 172).

A entrada direta pelo o que Cutter chamou de assunto e a sua disponibilização no catálogo dicionário envolveu uma forma múltipla de ramificação das palavras para uniformização das fichas. Por outro lado, influenciou o instrumento de controle dessas entradas por assunto, no caso, as listas de cabeçalhos de assunto. Nota-se uma forte preocupação com a sintaxe das palavras empregadas para representar os assuntos.

Na proposta de Cutter, os assuntos eram tratados de forma direta através das palavras que os constituem, tanto no catálogo quanto nas listas de cabeçalhos. Por isso, observa-se que, em fins de década de 1890, “[...] catalogadores podiam escolher cabeçalhos de assuntos e referências cruzadas sem passar pelo esforço intelectual de criá-los [...]”³³ (MIKSA, 2012, p. 10, tradução nossa), pois várias listas de cabeçalhos foram surgindo.

Nesse contexto, Guimarães (2009, p. 106), explica que a abordagem teórica da Catalogação de Assunto,

³³ “[...] cataloguers could choose subject headings “and cross-references without going through the intellectual effort of creating them [...]” (MIKSA, 2012, p. 10).

[...] na medida em que remonta à segunda metade do século XIX, [...], em muito norteadas pelos princípios de catalogação alfabética de Cutter e da tradição de cabeçalhos de assunto da *Library of Congress*, cuja ênfase reside no catálogo enquanto produto do tratamento da informação em bibliotecas [...].

Por isso, apesar de Cutter ter trabalhado diretamente com um conjunto de regras, que versaram sobre a manipulação do assunto, sua influência é grandemente percebida na elaboração de listas de cabeçalhos de assunto. Seguem exemplos de listas que marcaram o cenário dessa vertente.

Considerada a primeira, a *List of Subject Headings for Use in Dictionary Catalogs*, foi publicada em 1895. Produzida pela ALA e com três edições (1895, 1898, 1911),³⁴ foi substituída pela lista da LC. A popular *Library of Congress Subject Headings* (LCSH) recebeu esse nome a partir da oitava edição e, na sua concepção em 1898, foi intitulada *Subject Headings Used in the Dictionary Catalogues of the Library of Congress* (MIKSA, 2012).

A demanda das bibliotecas de pequeno porte levou à primeira edição da *List of subject headings for samll libraries* ou *Sears list of subject headings*, compilada por Minnie Earl Sears,³⁵ em 1923, a partir da terceira e última edição da lista da ALA (PENNA, 1945). Teve o nome alterado a partir da sexta edição para *List of subject headings* (DIAS, 1967). No Brasil, Wanda Ferraz (1977) realizou a primeira tradução da *Sears*, em 1944, com grande aceitação e uso nas bibliotecas.

Elva S. Smith³⁶ compilou no ano de 1933, em inglês, uma lista de cabeçalhos de assunto especial para acervo infantil. Em língua espanhola, Javier Lasso de la Vega (1939)³⁷ produz as *Reglas para la formación y redacción de los catálogos diccionarios en las bibliotecas, seguida de una lista de encabezamientos de materia y numerosos apéndices*, mas que se diferencia por ter abordado a elaboração de catálogos e as formas de cabeçalhos de

³⁴ A terceira edição foi a última publicada: AMERICAN LIBRARY ASSOCIATION. **List of Subject Headings for Use in Dictionary Catalogs**. 3th revised by Mary Josephine Briggs. Chicago: ALA, 1911. 398 p.

³⁵ SEARS, Minnie Earl. **List of subject headings for samll libraries**. New York: The H. W. Wilson Company, 1923.

³⁶ SMITH, Elva S. **Subject headings for children's books in public libraries and in libraries in sementary and junior high schools, with and introduction on the cataloging of children's books**. Chicago: American Library Association, 1933. 255 p.

³⁷ LASSO DE LA VEGA, Javier. **Reglas para la formación y redacción de los catálogos diccionarios en las bibliotecas, seguidas de una lista de encabezamientos de materia y numerosos apéndices**. Santander: [s. n.], 1939.

assunto (*encabezamientos de materia*) e subcabeçalhos para nomes de pessoas, cidades, estados e contexto histórico contemplando o idioma espanhol (PENNA, 1945).

Já Rovira (1966), explica que Javier Lasso de la Veja,³⁸ em parceria com Cesáreo Goicoechea, traduziu e adaptou de forma pioneira, a lista *Sears*. Ainda segundo a autora, na Espanha, a lista foi considerada notável, pois envolveu a primeira formulação de princípios para o catálogo dicionário, pouco difundido no país.

Segundo Teixeira (1979, p. 17), houve a partir de 1930, nos Estados Unidos, “[...] a não aceitação tácita das práticas da Biblioteca do Congresso, levando muitos bibliotecários a reclamarem a divulgação das regras em que a Biblioteca se baseava para determinar seus cabeçalhos de assunto.” Ainda segundo o autor, somente em 1951, a obra de Haykin³⁹ trouxe explicações de “[...] muitos pontos da prática seguida pela Biblioteca do Congresso.”

Porém, em 1946, e antes de Haykin, Julia Pettee fundamentou a discussão sobre os cabeçalhos de assunto e sua obra *Subject headings: the history and theory of the alphabetical subject approach to books*⁴⁰ é considerada a primeira de importância fundamental sobre o tema (PENNA, 1945). Entretanto, Tôrres ([2000?]) diz ter havido um corolário pouco aprofundado em Haykin e Pettee sobre a fundamentação da construção dos cabeçalhos de assunto. Pressupõem-se os autores não terem apresentado soluções mais aprofundadas por ausência de fundamentos que permitissem o amadurecimento teórico da área. Seria a ausência de uma fundamentação anterior refletida na teoria.

As ideias em prol da adoção das listas de cabeçalhos e do catálogo alfabético de assunto avançaram e, de acordo com Teixeira (1979, p. 17-18, destaque do autor), o instrumento recebeu popularidade em vários países. Assim, tem-se que:

[...] na França, o “Biblio” passou a adotá-lo e decidiu adaptar os cabeçalhos da Biblioteca do Congresso ao idioma francês, publicando posteriormente uma lista de cabeçalhos; na Itália, o catálogo alfabético de assunto teve também seus adeptos, apesar das “Normas” da Vaticana; em 1956, foi publicada a lista de cabeçalhos da Biblioteca Nacional de Florença; na Espanha, a lista de LASSO DE LA VEGA e o catálogo alfabético de assunto da Biblioteca Central de Barcelona. Foi na Inglaterra, porém, que os estudos teóricos do catálogo alfabético de assunto e do sistemático atingiram estágio

³⁸ LASSO DE LA VEGA, Javier; GOICOECHEA, Cesáreo. **Reglas para la formación y redacción de los catálogos-diccionarios en las bibliotecas, seguidas de una lista de encabezamientos de materia y numerosos apêndices.** [S. l.]: Vitória, 1939.

³⁹ HAYKIN, David J. **Subject heading: a practical guide.** Washington, D. C.: Govt. Print Off., 1951. 140 p.

⁴⁰ PETTEE, Julia. **Subject headings: the history and theory of the alphabetical subject approach to books.** New York: H. W. Wilson, 1946.

mais alto, embora não se note neles grande interesse pela composição das listas de cabeçalhos. As obras do inglês COATS [sic], do australiano METCALF [sic] e do sul-africano COETZEE, entre outros, [...] [discordam], quase constantemente, das práticas norte-americanas [...]

Essas listas de cabeçalhos de assunto foram tentativas de contemplar a realidade nacional dos países em detrimento das práticas da LC. Na Biblio,⁴¹ lista de cabeçalhos em língua francesa, expôs-se os princípios empregados na seleção dos cabeçalhos, mas segundo Dias (1967), não deixou de ser uma tentativa de adaptação do método de Cutter. Quanto à lista italiana, conhecida como *Soggettario*,⁴² ela é bastante ampla e só perdeu em quantidade de termos para a lista da LC, contudo seu destaque foi seguir características próprias (DIAS, 1967). No contexto das bibliotecas de língua espanhola, Rovira (1966) ao levantar as listas nesse idioma, declarou que a de Juana Manrique de Lara⁴³ foi a primeira a ser publicada, no México em 1934, e houve uma segunda edição em 1953, com forte influência das regras de Cutter.

Rovira (1966) e Dias (1967) indicam outras listas, tanto de cunho geral quanto de temáticas específicas, ilustrando que houve um movimento que priorizou a prática e a construção do instrumento, mas sem o alinhamento de discussões teóricas e de fundamentação sobre a noção de assunto.

Parece válido inferir concomitância, a partir da década de 1940, entre a elaboração e difusão das listas de cabeçalhos e as obras com foco na Catalogação envolvendo o processo de catalogar o assunto (PENNA, 1945; TAYLOR, 1948; MANN, 1962; SHARP, 1964; SHERA; EGAN, 1969; DIAS, 1967; LITTON, 1971). As abordagens dos autores se aproximam nas indicações de como fazer a catalogação descritiva e de assunto envolvendo desde regras para as entradas dos registros no catálogo (TAYLOR, 1948; SHARP, 1964), passando às funções e formas dos catálogos e o sistema de fichas principal e secundárias (PENNA, 1945; DIAS, 1967) à vinculação da catalogação e a sistematização da classificação (PENNA, 1945; MANN, 1962; SHERA; EGAN, 1969; LITTON, 1971). Entretanto, nenhum desses referenciais abandonou a explicitação de métodos e atributos dos tipos de catálogos e o modelo persistente da LC na crescente difusão das fichas catalográficas impressas.

⁴¹ BIBLIO. **Liste des vedettes-matière de Biblio**. Paris: Hachette, 1954.

⁴² BIBLIOTECA NAZIONALE CENTRALE (Firenze). **Soggettario per i cataloghi delle biblioteche italiane**. Firenze: Il Cenalo, 1956.

⁴³ MANRIQUE DE LARA, Juana. **Guía de encabezamientos de materia para los catálogos diccionarios**. México: Secretaría de Educación Pública, 1934.

As obras citadas não constituíram manuais práticos para a execução do processo de catalogação, mas incorporam a sistematização de regras e consideraram características do referido processo. Essa situação é constatada em Penna (1945), Taylor (1948) e Dias (1967). Assim, mesmo sem ocorrer indicação direta, mas tomando por base o referencial é possível evidenciar essas obras como manuais práticos para execução da catalogação de assunto. Ademais, o que fica explícito é o fato de indicarem Cutter como figura unânime do processo e da vertente denominada Catalogação de Assunto.

Como produto do processo de catalogação, os catálogos em fichas se dividiram de acordo com seus objetivos, como por exemplo, catálogo de autor (ou autoridade, ou onomástico), de título (ou didascálico), de assunto (ou ideográfico) e receberam uma entrada principal (que é a primeira informação registrada) responsável em subsidiar o que a literatura denominou sistema de ficha única. Esse sistema compõe um conjunto de fichas principal e secundárias que deve manter a uniformidade na redação (PENNA, 1945; DIAS, 1967).

No que concerne ao assunto, seu foco incidiu na forma secundária na ficha que terá “[...] na sua pista, a indicação dos desdobramentos a serem feitos [...]” (DIAS, 1967, p. 24). Assim sendo, “[...] as fichas secundárias de assunto tem por objetivo auxiliar a localização do livro, no catálogo dicionário ou no sistemático, no grupo de assuntos que seja correspondente. No caso de usar um catálogo sistemático [...] são colocados os números ou símbolos de classificação [...]”⁴⁴ (PENNA, 1945, p. 26, tradução nossa).

Houve a difusão de organização das fichas no catálogo por uma perspectiva classificatória hierárquica de assuntos refletidos nos números, através das notações classificatórias, partindo gradativamente de um arranjo sistemático geral para os tópicos específicos. Nos Estados Unidos o surgimento da Classificação Decimal de Dewey (CDD, ou, *Dewey Decimal Classification* - DDC), em 1876, que ocorreu no mesmo ano da primeira edição das *Rules* de Cutter, propiciou a organização sistemática das fichas no catálogo (FOSKETT, 1973).

Posto o cenário das *Rules* de Cutter e da geração de produtos como catálogos e listas de cabeçalhos de assunto, sobretudo a partir da influência da LC, o assunto foi guiado por regras que buscavam padronizar as decisões no processo de catalogação, o qual envolvia desde o momento da elaboração à escolha dos cabeçalhos. A LC se beneficiou da uniformidade procedimental difundida pela *Rules*, tornando-se em certa medida, referência

⁴⁴ “[...] as fichas secundarias de materia tienen por objeto ubicar el libro, ya sea en el catálogo dicionario o en el sistemático, dentro del grupo de materias a que corresponda. En el caso de utilizarse un catálogo sistemático [...] se colocan los números o símbolos clasificadores [...]” (PENNA, 1945, p. 26).

para o processo de catalogação de assunto. A lista de cabeçalho de assunto por ela produzida foi base de estudos críticos, conforme é possível observar em Gomes e Martinho (1984) e Tôrres ([2000?]).

Todavia, a partir do século XX, a forma de representar o assunto passa a compreender aspectos da análise, com implicações no modo de realizar o processo, assim como, nos resultados. Se na catalogação, os assuntos são elaborados na forma de cabeçalhos, a partir de outras vertentes, a tendência é o detalhamento e a análise do assunto para posterior representação.

É fato que os aspectos do assunto discutidos nessa seção, com forte vínculo nos eventos do século XIX, estão atrelados à construção de cabeçalhos. A abordagem se baseou em processos centrados nas regras e no resultado de padronização dos assuntos nas fichas.

5 SEGUNDO MOMENTO DE CONSTRUÇÃO DA NOÇÃO DE ASSUNTO: DESDOBRAMENTO DO PENSAMENTO ANGLO-SAXÃO NO SÉCULO XX

O levantamento dessa seção centra-se em um quadro teórico estabelecido por forte influência anglo-saxônica, permitindo relações que contribuíram para a construção da noção de assunto. É o momento em que ele, como sistematizado por Cutter, na forma de um cabeçalho, passa a ser realizado pela decomposição. O fato principal é que essas relações, assim como outras, não aconteceram de forma linear, ou seja, ocorreram com o envolvimento de influências variadas e, em determinados momentos, com fragmentos.

Pela revisão, foi possível mapear um quadro teórico que não almeja limitar, mas demonstrar as orientações que fazem parte do universo da representação temática, que coexistiram em diferentes linhas de pensamento.

5.1 JULIUS KAISER E SHIYALI RANGANATHAN: A DECOMPOSIÇÃO DO ASSUNTO

De acordo com Metcalfe (1973), em 1911, foi interessante como Julius Otto Kaiser (1868-1927), bibliotecário alemão que atuou em instituições empresariais do Reino Unido e dos Estados Unidos, desenvolveu uma forma sistemática de indexação. A partir da sua obra *Systematic Indexing*, tomou assuntos, como coisas em geral, reais ou imaginárias, buscando designá-los como “concreto” e como “processo”. Portanto, Kaiser (*apud* METCALFE, 1973, p. 307, tradução nossa) percebeu que “os assuntos de nossa observação e raciocínio são coisas em geral, reais ou imaginárias, e as condições a eles associadas. Vamos chamá-los de concretos e processos, respectivamente [...]”⁴⁵

A preocupação de Kaiser foi com a aplicabilidade de um sistema eficiente para o cotidiano da organização de documentos, relativos ao comércio e à indústria e, por esse motivo, segundo Foskett (1973), suas aplicações são limitadas. Mas, naquele período voltado para os catálogos alfabéticos de assuntos, como o catálogo dicionário, a fundamentação da indexação sistemática de Kaiser apresentou algumas nuances que se diferenciaram das *Rules* de Cutter. Uma importante característica que diferenciou o primeiro do segundo, conforme

⁴⁵ “The subjects of our observation and reasoning are things in general, real or imaginary, and the conditions attaching to them. We shall call them concretes and processes respectively [...]” (KAISER *apud* METCALFE, 1973, p. 307).

Sales (2012), foi a proposta da decomposição de partes do assunto a partir de duas principais categorias.

Kaiser buscou resolver os problemas da inconsistência presente na catalogação de assunto e, por isso, conforme Metcalfe (1973, p. 310, tradução nossa):

[...] enquanto Kaiser distinguia formal e consistentemente Concreto e Processo, ou em termos mais gerais, objeto e aspecto, como cabeçalho e subcabeçalho, Cutter pensava quase exclusivamente em palavras compostas como *bookbinding* [encadernação], ou *bookkeeping* [escrituração contábil], ou agricultura, em que um nome de objeto era o primeiro elemento, ou de frases com o mesmo efeito, como fertilização de flores, história antiga e história medieval, ornitologia na Nova Inglaterra. A única exceção de Cutter foi seu agrupamento por tópico ou divisão de numerosos títulos sob cabeçalho de assunto ou palavra de entrada.⁴⁶

Para Cutter, a forma específica para a entrada incluía palavras que poderiam formar cabeçalhos simples (constituídos por apenas uma palavra) ou cabeçalhos compostos (constituídos por mais de uma palavra), como mencionado anteriormente.

Na abordagem que Kaiser propôs, haveria uma decomposição e o ponto de partida seria sempre o concreto, que envolveria assuntos modificáveis, atualizados conforme a expansão do conhecimento através de novas descobertas. Na combinação concreto-processo, de acordo com Foskett (1973, p. 50), “[...] o mais importante é o concreto, o qual deve ser citado em primeiro lugar.” Entretanto, Kaiser não definiu o termo concreto e forneceu maior atenção ao processo.

Mas é Kaiser quem propõe pela primeira vez a possibilidade do “[...] agrupamento de assuntos semelhantes, desde que a informação fosse representada segundo expressões construídas artificialmente, com base em uma fórmula pré-estabelecida.” (TÔRRES, [2000?], on-line). Partindo desse ponto, o estabelecimento de qualquer assunto combinaria, por um lado, a dimensão estagnada (concreto) e, por outro, o movimento (processo).

A este respeito, deve se considerar o concreto em relação à forma de subdivisão dos assuntos. Segundo Foskett (1973, p. 50), Kaiser propôs “[...] uma completa rede de remissivas a fim de superordenar e coordenar, bem como subordinar os cabeçalhos.” Outra dimensão foi estabelecida, referente a lugar, como complemento para denotação de localidades como

⁴⁶ [...] whereas Kaiser formally and consistently distinguished Concrete and Process, or in more general terms, object and aspect, as heading and subheading, Cutter thought almost exclusively of compound words such as *bookbinding*, or *bookkeeping*, or agriculture, in which an object name was the first element, or of phrases to the same effect, such as Flower fertilization, Ancient history and Medieval history, New England ornithology. Cutter's only exception was his topical grouping or division of numerous titles under subject heading or entry word. (METCALFE, 1973, p. 310).

países, por exemplo, com o propósito de particularizar os concretos dos assuntos (TÔRRES, [2000?]). Se o concreto abrangia uma dimensão estagnada do assunto, um lugar reforçaria a posição acerca da localização dele em um processo.

Tôrres ([2000?], on-line) afirma que:

A teoria semântica subjacente à linguagem criada por Kaiser, segundo Svenonius [1978],⁴⁷ é a teoria nominalista do significado, que foi introduzida por Platão em seu “Cratylus”, na qual as palavras são vistas como referindo-se a[s] coisas e, conseqüentemente, como nomes ou rótulos de coisas. Não obstante a época em que foi formulada e embora esquecida ou ignorada pelos catalogadores de assunto da LC, a teoria de Kaiser trata de modo inovador a questão da sintaxe dos cabeçalhos compostos e da escolha de seus respectivos pontos-de-acesso: toda informação deve ser ordenada segundo o concreto de que ela trata e deve ser expressa numa seqüência [sic] de nomes ou termos, cuja ordem canônica é concreto-processo.

As palavras, ao se referirem às coisas, consistiram na ideia de rótulos a objetos. Nesse sentido, como propor uma relação entre um significante e um objeto? A preocupação de Kaiser foi a de rotular os assuntos por palavras, mas suas ideias avançaram em relação às de Cutter, porque ele propôs o agrupamento dos assuntos por combinação, segundo uma fórmula, em detrimento da ordem alfabética do catálogo dicionário.

Se nas *Rules*, Cutter produziu normativas, pois sem regras para manter certa consistência nas entradas, o público não encontraria através do cabeçalho o nome do assunto de interesse, com Kaiser, foi proposta a ordem de citação nos cabeçalhos e a apresentação sistemática dos subcabeçalhos (CESARINO; PINTO, 1978).

A partir da contribuição de Kaiser, entre outros, pode-se falar em uma vertente denominada Indexação, caracterizada pelo trabalho com assunto em um nível especializado e por uma literatura sobre fundamentos conceituais e sobre o processo da análise.

De algum modo, o processo de indexação foi concebido tendo por fim a criação de índice em arquivo classificado, ou seja, em arranjo sistemático, mas que é acompanhado de um cabeçalho. De acordo com Needham (1971, p. 96, tradução nossa), houve ampliação do entendimento, porque os processos passaram a cobrir “[...] a análise do conteúdo do assunto de documentos, classificação e criação de entradas de catálogo.”⁴⁸

⁴⁷ SVENONIUS, E. Facet definition: a case study. **International Classification**, v. 5, n. 3, p. 134-141. 1978.

⁴⁸ “[...] the analysis of the subject content of documents, classifying, and the creation of catalogue entries.” (NEEDHAM, 1971, p. 96).

A partir da década de 1930, Shiyali Ramamrita Ranganathan (1892-1972) fundamentou a combinação de aspectos de um assunto, ou de conceitos simples, para a formação de conceitos complexos (BARBOSA, 1972; METCALFE, 1973; CAMPOS; GOMES, 2003), influenciando assim a categorização dos assuntos de forma ampla.

Ranganathan, indiano e matemático de formação, em 1925, teve a oportunidade de estudar com Berwick Sayers na *School of Librarianship*, do *University College* de Londres. Refletiu um movimento que considerou o conteúdo dos documentos por uma divisão de assunto, em múltiplos aspectos, simples ou complexos, que ele denominou de facetas (BARBOSA, 1972) e “[...] nos *Prolegomena*⁴⁹ [fundamentou] [...] o processo de formação de conceitos e sua relação com o que denomina universo das idéias [sic] ou do conhecimento e sua influência no trabalho da classificação.” (CAMPOS; GOMES, 2003, p. 152).

O desenvolvimento de uma proposta para definição de assunto em Ranganathan, segundo Hjørland (2017), pode ser observado no seu sistema de classificação *Colon Classification* (CC, em português, Classificação de Dois Pontos), que é um esquema analítico-sintético baseado na combinação de elementos únicos de facetas para designação de assuntos. De forma a caracterizar o esquema CC, Barbosa (1972, p. 72) explica que a coordenação de conceitos é feita a partir da análise em facetas, “[...] significando que um assunto, por mais complexo que seja, pode ser representado pela síntese de mais de uma faceta, cada uma indicando conceitos diferentes. Eis a razão porque esse tipo de classificação é chamado de analítico-sintético.”

Ao explorar definições que Ranganathan teria fornecido para o assunto, Hjørland (2017, p. 57, tradução nossa) apresenta três possíveis abordagens:

“Assunto: termo assumido” (Ranganathan 1963, 27).

“Assunto: conteúdo do pensamento de um documento” (Ranganathan 1964, 109).

“Assunto: um corpo organizado de ideias, cuja extensão e intensão tendem a cair coerentemente dentro do campo de interesses e confortavelmente dentro da competência intelectual do campo da inevitável especialização de uma pessoa normal” (RANGANATHAN, 1967, 82).⁵⁰

⁴⁹ RANGANATHAN, S. R. **Prolegomena to library classification**. Bombay: Ásia Publishing House, 1961.

⁵⁰ “Subject: assumed term” (Ranganathan 1963, 27).

“Subject: thought-content of a document” (Ranganathan 1964, 109).

“Subject: an organized body of ideas, whose extension and intention are likely to fall coherently within the field of interests and comfortably within the intellectual competence and the field of inevitable specialization of a normal person” (Ranganathan 1967, 82). (HJØRLAND, 2017, p. 57).

A primeira definição considera o assunto como algo evidente, e foi desenvolvida em *Documentation and its Facets*.⁵¹ A segunda apresenta o ponto de vista do conteúdo do pensamento de um documento, tratado no artigo *Subject Heading and Facet Analysis*⁵² e, a terceira, foi apresentada na obra *Prolegomena to Library Classification*⁵³ (HJØRLAND, 2017).

É a partir da terceira abordagem indicada por Hjørland (2017) que Ranganathan trata cada assunto como uma síntese de vários conceitos interligados, segundo uma espiral de desenvolvimento. O universo do conhecimento, também na configuração de espiral, portanto como algo infinito, é regido de forma contínua, abarcando várias entidades, com muitos atributos que indicam as características de divisão. Por isso,

[...] se o movimento da espiral do conhecimento propicia o ato de perceber os fatos que ocorrem no mundo fenomenal, com a espiral do desenvolvimento de Assuntos é possível verificar a relação entre este perceber e a produção de conhecimento que, no nosso caso, é conhecimento registrado. (CAMPOS; GOMES, 2003, p. 155).

O assunto presente em um documento remete a uma análise intelectualmente necessária que pode ou não ser codificada em uma linguagem controlada, que é estruturada e sistematizada em palavras, descritores ou símbolos, como os numéricos dos sistemas de classificação, por exemplo. Com Ranganathan, houve a proposta de um mapeamento de atributos para fundamentação de conceitos, levando em consideração categorias fundamentais conhecidas pelo acrônimo PMEST (Personalidade - *Personality*; Matéria - *Matter*; Energia - *Energy*; Espaço - *Space*; Tempo - *Time*).

Embora seja possível encontrar outras sistematizações, como o mapeamento do conhecimento apresentado pela Classificação Decimal de Dewey (CDD), que apresenta assuntos a partir de disciplinas, na *Colon Classification*, segundo Hjørland (2017), os conceitos, transcendem o mapeamento, pois possibilitam o relacionamento entre as facetas dos assuntos de forma ilimitada. Com isso, a compreensão de termos, como por exemplo, “isolado” e “conceito básico” revelam que:

⁵¹ RANGANATHAN, S. R. **Documentation and its facets**. London: Asian Publishing House, 1963.

⁵² RANGANATHAN, S. R. Subject heading and facet analysis. **Journal of Documentation**, n. 20, p. 109-19. 1964.

⁵³ RANGANATHAN, S. R. **Prolegomena to library classification**. 3. rd. London: Asia Publishing House, 1967.

Os conceitos de Ranganathan são altamente idiossincráticos, por exemplo, a afirmação que o ouro não pode ser um assunto (mas é alternativamente denominado “um isolado”). O conceito “disciplina” é substituído por “assunto básico” definido [...] como assunto, que não tem ideias isoladas como um componente (exemplo: matemática). (HJØRLAND, 2017, p. 57, tradução nossa).⁵⁴

Metcalfe (1973), embasado em Ranganathan, afirma que assunto é o conteúdo do pensamento do documento a partir dos isolados. Por sua vez, os conceitos resultam da associação de duas ou mais percepções, que se associam na memória.

Assim sendo, na memória se aloja no que Ranganathan denominou perceptos puros, que são impressões produzidas através de um sentido primário simples, enquanto a impressão pela associação de dois ou mais perceptos puros resultam em perceptos compostos (CAMPOS; GOMES, 2003).

Com Ranganathan, o assunto torna-se a representação sintetizada dos conceitos tratados nos documentos, oriundos da percepção humana, podendo abranger diferentes facetas. Para tanto, o autor leva em consideração a definição de faceta “[...] como o componente do assunto [...] é um termo genérico usado para denotar qualquer componente - seja um assunto básico ou um isolado - de um assunto composto, e também suas respectivas formas classificadas, termos e números.”⁵⁵ (DUTTA; DUTTA, 2013, p. 82, tradução nossa).

A palavra *prolegonema*, parte do título da obra de Ranganathan, é o plural do termo latino *prolegomenum*, que remete a princípios básicos para o estudo de um assunto. No dicionário Aulete (on-line), *prolegomenos* refere-se a “1. Princípios gerais de qualquer ciência ou arte expostos preliminarmente. 2. Introdução expositiva de algum tratado científico ou artístico.” Embora a obra *Prolegomena to Library Classification* articule princípios, postulados e leis para a teoria da *Colon Classification*, sua principal característica foi envolver aspectos e conceitos empregados no sistema de classificação proposto e, não a própria teoria, a qual se encontra presente em outra obra, intitulada *Colon Classification*.

Em relação a esses fundamentos, em que Ranganathan aborda que isolados e facetas podem identificar tanto o assunto básico como o composto, Metcalfe (1973), afirmou que a dúvida persistia. O autor explica que, em 1953, na *10th All India Library Conference*, a partir

⁵⁴ Ranganathan’s concepts are highly idiosyncratic, for example, the claim that gold cannot be a subject (but is alternatively termed “an isolate”). The concept “discipline” is substituted with “basic subject” defined [...] as a subject, that does not have isolate ideas as a component (example: mathematics). (HJØRLAND, 2017, p. 57).

⁵⁵ “[...] as the component of subject [...] is a generic term used to denote any component - be it a basic subject or an isolate - of a compound subject, and also its respective ranked forms, terms and numbers”. (DUTTA; DUTTA, 2013, p. 82).

de documentos discutidos e apresentados pelo próprio Ranganathan, Vickery compilou um glossário corrente com termos técnicos sobre classificação. Como resultado, “cerca de 148 termos de 320 foram atribuídos a Ranganathan. Assunto por si só não foi listado [...]”⁵⁶ (METCALFE, 1973, p. 315, tradução nossa). Para o mesmo autor, em outro glossário publicado em 1959, no *Annals of Library Science*, Ranganathan aponta que assunto (*Subject - Assumed term*) foi o termo assumido como referência para conhecimento (*Knowledge item - Subject*). Já em 1964, no *Indian Standard Glossary of Classification Terms*, Ranganathan declarou, na parte intitulada Universo do Conhecimento, a situação do assunto como um termo alternativo assumido para massa de conhecimento (*Knowledge mass*) e o isolado parece não ter uma definição básica em relação ao assunto, mas foi conduzido como um termo assumido.

Dutta e Dutta (2013, p. 82, tradução nossa) consideraram que “a definição de Ranganathan de ‘assunto’ é baseada no sistema de Classificação de Dois Pontos [...] um esquema analítico-sintético que é baseado na combinação de elementos únicos a partir de facetas que designam o assunto.”⁵⁷

Se a CC propõe que entre dois pontos existem possibilidades infinitas de estruturar os assuntos para serem combinados quando necessário, Hjørland (2017, p. 57, tradução nossa) argumenta que a despeito de Ranganathan definir “[...] o conceito de ‘assunto’ como algo que favorece seu próprio sistema [...] [esse] conceito [...] deveria possibilitar a comparação das diferentes formas de estabelecer acesso à informação.”⁵⁸ O autor ainda afirma que de forma independente do tipo de sistema (que é o tipo de linguagem), se pré ou pós-coordenado, ou facetado, o assunto pode ser definido considerando o exame do seu grau de organização em casos específicos.

Hjørland (2017) compreende que, na CC, Ranganathan determinou o assunto por meio de aspectos abrangentes, visando contemplar necessidades gerais, sem aprofundamento de especialidades, que seriam os casos específicos.

⁵⁶ “About 148 terms in 320 were attributed to Ranganathan. Subject by itself was not listed [...]” (METCALFE, 1973, p. 315).

⁵⁷ “Ranganathan’s definition of “subject” is based on Colon Classification system [...] an analytico-synthetic scheme, which is based on the combination of single elements from facets to subject designation.” (DUTTA; DUTTA, 2013, p. 82).

⁵⁸ “[...] the concept of “subject” in a way that favors his own system [...] concept [...] should make it possible to compare different ways of establishing access to information.” (HJØRLAND, 2017, p. 57).

Nesse aspecto, observa-se que a discussão sobre assunto não deve se misturar com demandas específicas de uso, pois, de um lado, há que considerar o que sua noção representa e, de outro, sua descrição a fim de atender à demanda de busca e recuperação. São particularidades que se complementam, mas cada qual com seus critérios.

Se a predileção de Ranganathan foi o fornecimento de métodos para a reunião de conceitos em categorias, abrangendo, por um lado, o universo classificatório e o critério das facetas dos assuntos, por outro, a fundamentação da noção de assunto não se tornou central, como afirma Metcalfe (1973). Mas, as ideias de Ranganathan avançaram quanto às de Cutter e de Kaiser na fundamentação da noção de assunto, na abordagem do isolado, da faceta, dentre outros aspectos.

Ranganathan é apontado como um dos precursores da visão sistêmica da representação de documentos, sendo que, de acordo com Dahlberg (1978), a classificação facetada pode ser entendida como um sistema de conceitos, já que as facetas e seus elementos são refletidos nas categorias. A combinação da análise de facetas demonstra a abordagem fecunda de Ranganathan por conduzirem à organização de campos conceituais pelos assuntos. O teórico propôs a ideia de universo do conhecimento no formato de espiral que possibilita perceber e verificar os fatos e a produção do conhecimento por outra espiral de desenvolvimento, de assuntos (CAMPOS; GOMES, 2003).

Ainda que seja recorrente considerar como precursora a ideia de faceta de Ranganathan, Cesarino e Pinto (1978), Dousa (2010) e Sales (2012) atribuem a Kaiser a formulação teórica que influenciou Ranganathan na abordagem das categorias. Guimarães (2008, p. 80) corrobora explicando que Kaiser “[...] trouxe importantes elementos metodológicos que influenciaram, inclusive, os trabalhos de Ranganathan, uma vez que as facetas Personalidade e Energia parecem decorrer diretamente das facetas Concreto e Processo [...]”.

Sem exaurir o ponto apresentado, mas compreendendo que o amadurecimento teórico é decorrente das discussões que foram desenvolvidas, nessa subseção o movimento apresentado centrou na decomposição do assunto. Outras pesquisas convergem para a compreensão do assunto em uma abordagem conceitual, como se observará à frente. Houve diferença no modo de abordar o conceito, mas isso não significou sua ausência.

E, apesar de nas listas de cabeçalho os assuntos não serem construídos com base nos conceitos, o objetivo final era o mesmo, e a diferença centrou no modo de elaboração dos assuntos.

5.2 PESQUISADORES DO *CLASSIFICATION RESEARCH GROUP* E CALVIN MOOERS: A FACETA E O DESCRITOR

Foskett (1973) cita Eric J. Coates, o sistema PRECIS (*PREserved Context Indexing System*), de Derek Austin e Jason Ferradane como referenciais que envolveram a abordagem temática da informação.

Coates, em 1960, publicou o livro *Subject Catalogues Headings and Structure*, que abordou a ordem dos cabeçalhos compostos e buscou compreender a razão que levou Kaiser a optar pelo concreto ao invés do processo como ponto de entrada dos cabeçalhos. Seu desafio consistiu em buscar a aplicação da sua teoria no *British Technology Index*, pois foi o redator responsável do projeto. Coates estabeleceu a ordem de importância das entradas pela fórmula Coisa-Parte-Material-Ação.

O PRECIS foi idealizado por Derek Austin, a partir de 1968, e consistiu na mecanização da produção de índices em inglês, como também em outras línguas, caracterizando-se por ser um sistema de indexação pré-coordenado, feito para produzir automaticamente os índices de assunto da *British National Bibliography* (FUJITA, 1988). Nesse sentido, o avanço “[...] que o PRECIS [...] [representou foi a] elaboração de um enunciado básico que pode ser processado pelo computador, a fim de proporcionar um conjunto completo de cabeçalhos.” (FOSKETT, 1973, p. 62). O sistema foi elogiado por propiciar ao usuário visualizar de forma completa cada cabeçalho.

Já Farradane, segundo Foskett (1973), elaborou um sistema de operadores relacionais embasado em conceitos sobre o processo de aprendizagem. Com base na psicologia infantil adotou nove operadores relacionais que vinculava termos através de aspectos delimitados no tempo e no espaço.

O movimento que congrega esses autores é o de institucionalização dos estudos de Ranganathan, na Inglaterra, a partir de 1952, realizado pelo *Classification Research Group* (CRG). O grupo contou com a participação de inúmeros pesquisadores, dentre eles Derek Austin (Inglaterra), Eric Coates (Estados Unidos), Jason Farradane (Canadá), Ingetraut Dahlberg (Alemanha), dentre outros nomes (FUJITA, 1988).

Para Metcalfe (1973), Farradane teria introduzido o isolado de Ranganathan como um termo técnico, pois em 1950, ele apresentou o que chamou de teoria científica da classificação e suas aplicações práticas. Já Coates, no *British Technology Index*, ainda segundo Metcalfe (1973), considerou o artigo de Ranganathan⁵⁹ sobre cabeçalho de assunto e análise de facetas

⁵⁹ RANGANATHAN, S. R. *Prolegomena to library classification*. Bombay: Ásia Publishing House, 1961.

(discutido por Hjørland (2017), como indicado na subseção anterior), mas não incluiu qualquer terminologia de Ranganathan em seu livro (*Subject Catalogues Headings and Structure*). Todavia, Coates abordou Melvil Dewey (criador da CDD), Cutter e Kaiser como autores que foram anteriores a Ranganathan, no sentido de ter havido uma continuidade das discussões entre eles, apesar dos diferentes aspectos que os distanciam e as particularidades teóricas que os sustentam. Mas, foi em Kaiser que Coates viu a solução “[...] do problema da importância relativa dos componentes de um assunto composto no ponto em que Cutter tinha deixado.”⁶⁰ (METCALFE, 1973, p. 324, tradução nossa).

Em sentido amplo, a partir de Guimarães (2008, p. 80), é possível inferir que o quadro teórico apresentado indicou

[...] as experiências investigativas de Kaiser (EUA, 1911) relativamente à composição do binômio concreto/processo para a estruturação de enunciados de assunto e o trabalho teórico de Ranganathan (Índia, 1933) quanto a uma análise em facetas – Personalidade, Matéria, Energia, Espaço e Tempo – do conteúdo temático dos documentos.

Igualmente há de se registrar as experiências investigativas do *Classification Research Group* (Inglaterra, 1952), que ampliou para doze as cinco categorias fundamentais de Ranganathan, e de Derek Austin (Inglaterra, 1968) e do grupo de Bangalore (Índia, 1969) que, por meio dos sistemas PRECIS e POPSI, lançaram as bases para uma aplicação da análise facetada em sistemas automatizados.⁶¹

Embora haja outros referenciais que tenham contribuído com diferentes discussões no contexto temático, o ponto em comum foi a transição de algo explícito no documento com fins de organização dos acervos. A decomposição e os possíveis relacionamentos entre os assuntos foram percebidos à medida que o contexto linear de sintaxe dos cabeçalhos não deram conta de conferir sentido à padronização das formas de entrada.

Por outro lado, os cabeçalhos de assunto não se tornavam eficazes, pois sua elaboração foi marcada pelo aspecto morfológico da palavra. Taylor (1995, p. 484, tradução nossa) exemplifica que “[...] o conceito de busca por assunto permaneceu com má reputação.”⁶² A autora relata que na conferência da *American Library Association* (ALA), em 1994, houve um

⁶⁰ “[...] the problem of the relative importance of the components of a compound subject at the point at which Cutter had left it.” (METCALFE, 1973, p. 324).

⁶¹ POPSI (*POstulated-based Permuted Subject Indexing Language*) originou-se de estudos realizados, a partir de 1969, por A. Neelamegham e colaboradores no *Documentation Research and Training Center* de Bangalore, Índia. Neelamegham fez parte do *Classification Research Group* (FUJITA, 1988).

⁶² “[...] the concept of subject searching remained a disreputable one.” (TAYLOR, 1995, p. 484).

movimento que contestou a ênfase dada ao uso da *Library of Congress Subject Headings* (LCSH).

Taylor (1995) ainda menciona que Ray R. Larson apontou em artigo na mesma conferência da ALA a obviedade sobre o declínio nas porcentagens de uso do índice de assuntos da LCSH e a compensação das buscas realizadas no índice de palavras-chave por título. A autora continua explanando que, desde 1985, a LC busca reestruturar regras para a LCSH, como por exemplo, de caminhos hierárquicos para os termos dos cabeçalhos. Em 1991, a LC realizou uma conferência com o propósito de cooptar sugestões sobre como poderiam melhorar o sistema existente de subdivisões da LCSH e a ação resultou em mudanças na referida lista.

A discussão que Taylor (1995) propõe envolve um olhar permanente para o registro de assunto no catálogo, visando à busca por conteúdo temático dos documentos pelos usuários. Não se falou na possibilidade de se mudar o instrumento, mas isso não era usual, pois seria uma demanda de alto custo.

Ademais, presume-se que o contexto propiciou demanda por linguagens pós-coordenadas. Foskett (1973) refere-se que o tipo de linguagem de indexação reflete o resultado do processo na saída da informação para o usuário, mas também na análise para a entrada do assunto. As listas de cabeçalhos de assunto são linguagens pré-coordenadas, pois os termos de busca empregados no sistema de recuperação prescrevem a combinação. Já uma linguagem pós-coordenada permite a combinação de termos no sistema de recuperação, como acontece com o uso dos tesouros. Nota-se que o assunto no sistema de recuperação evidencia-se pelo controle do vocabulário na forma apresentada pela linguagem.

No mínimo dois caminhos podem ser apontados em relação ao aspecto da análise de assunto: um que visa coletar termos para construir um vocabulário controlado, portanto uma linguagem, e outro que envolve a sistematização de aspectos metodológicos da indexação de assuntos.

Porém, ainda com a estrutura das listas de cabeçalhos de assunto, sobretudo da LCSH, sustentando entre os cabeçalhos apenas relações de sinonímia (*see – ver*) e de associações (*see also – ver também*), predominou a necessidade de mecanismos que dessem conta de aperfeiçoar o sistema de recuperação com vínculo semântico.

Simões (2017, p. 166) afirma que “as remissivas foram-se alterando ao longo do tempo, quer no que respeita à nomenclatura, ao princípio e à função.” A autora alerta que a nomenclatura alterou-se para relações semânticas.

Mas, as relações para um vocabulário envolvem um controle, que

[...] não significa apenas padronizar as formas de entrada (significantes) ou registrar a sinonímia: significa operar sobre a significação (significante/significado) dos termos, o que depende inteiramente de delimitações de sentido conferidas pela rede de relações lógico-semânticas amparadas por definições. (LARA, 2001, p. 6).

Com a difusão do uso do computador e das fichas perfuradas, Calvin Mooers, na década de 1950, introduz o uso do termo descritor (*descriptor*). Fundou a *Zator Company* e desenvolveu um sistema próprio de recuperação da informação, o *Zatocoding*, no contexto dos estudos de *information retrieval* (IR).

Segundo Mooers (1951, p. 21, tradução nossa), seu sistema foi de “[...] especial importância para o documentalista porque lhe dá o poder de uma nova ferramenta de especificação de ideias para a recuperação de informações armazenadas.”⁶³ Seu propósito foi o de agilizar a recuperação por assunto, através da sobreposição de padrões aleatórios de códigos no registro. Esses padrões de código formavam a representação de descritores por ideias, ou seja, o conhecimento sobre o assunto do documento a ser registrado em informação.

Em linhas gerais, o sistema operava através da

[...] memória de filme com o campo de codificação de cada quadro possuindo 216 posições que podem ser marcadas por opacidades. Assumindo uma coleção de 5.000.000 documentos - comparáveis à Biblioteca do Congresso - um padrão do *Zatocoding* de 8 pontos por assunto pode ser usado. O campo de codificação pode conter até 18 padrões de assunto, e seleções podem ser feitas em qualquer combinação desses padrões. O tamanho do vocabulário descritivo do *Zatocoding* é ilimitado. (MOOERS, 1951, p. 22, tradução nossa).⁶⁴

Nesse panorama, o tratamento empregado pelo sistema *Zatocoding*, através das fichas perfuradas, usava as suas margens e máquinas seletoras, que a partir de uma lista de codificação registrava os assuntos. O padrão da codificação proporcionava a recuperação pelo código numérico de cada descritor.

O emprego do termo descritor por Mooers ocorreu para uma aplicação específica do seu sistema, com funcionamento independente, mas ele entendia que os “[...] descritores [...]

⁶³ “[...] special importance to the documentalist because it gives him powerful new tools of idea specification for retrieving information from storage.” (MOOERS, 1951, p. 21).

⁶⁴ “[...] film memory in which the coding field for each frame has some 216 positions that can be marked by opacities. Assuming a collection of 5,000,000 documents--comparable to the Library of Congress--a *Zatocoding* pattern of 8 marks per subject idea can be used. The coding field can hold up to 18 such subject patterns, and selections can be made upon any combination of these patterns. The size of the *Zatocoding* descriptive vocabulary is unlimited. (MOOERS, 1951, p. 22).

constituem as melhores unidades conceituais possíveis (científica ou tecnologicamente) para o trabalho específico de recuperação.”⁶⁵ (MOOERS, 1985, p. 253, tradução nossa). De acordo com Henderson (1996, p. 660, tradução nossa),

Mais tarde Mooers observou que o neologismo, “descriptor”, sofreu um triste destino: o termo foi introduzido para “sintetizar um conceito cuidadosamente elaborado no uso de unidades de significado para recuperação”, mas veio a ser usado como sinônimo de entradas de índice de qualquer tipo. Por exemplo, desde a publicação dos anais da Conferência Internacional sobre Informação Científica, realizada em Washington, D.C. em 1958, o termo “descriptor” foi equiparado a “termos de índice” e “palavras-chave” no índice dos anais. Na visão de Mooers, isso era “a maior corrupção semântica.”⁶⁶

No entanto, conforme exposto por Henderson (1996), o termo descriptor, empregado por Mooers, inicialmente na concepção do seu sistema de recuperação, baseado em um processo mecânico, perdeu-se em meio ao emprego de outros produtos documentários. De fato, Mooers demonstrou interesse sobre métodos mecânicos de recuperação por assunto em coleções de documentos, como demonstram suas investigações sobre recuperação da informação.

À medida que sua investigação deflagrou interesse, conforme trabalho apresentado no ano de 1962, na *University of Minnesota*, mas reimpresso em 1985, Mooers abordou a linguagem de indexação em sistemas de recuperação da informação. Para ele, a linguagem de indexação é o meio responsável por mediar as informações armazenadas nos documentos, tornando-se “[...] a ponte, a conexão entre os usuários e as informações.”⁶⁷ (MOOERS, 1985, p. 249, tradução nossa). Ele demonstrou interesse por essas linguagens, mas centrou-se em entender como as máquinas poderiam manipulá-las.

Para esse autor, uma lista de cabeçalho de assunto, por exemplo, determinaria um arquivo científico de palavras com sintaxe específica, conforme “*Lubrication - high temperature - non-hydrocarbon*”. Explica que esse cabeçalho possui diferentes conceitos unitários, como “*lubrication*”, “*high temperature*”, “*non-hydrocarbon*”, que permitem outras

⁶⁵ “[...] the descriptors are the best possible conceptual units (scientifically or technologically) for the particular job of retrieval.” (MOOERS, 1985, p. 253).

⁶⁶ Mooers later noted that that neologism, “descriptor”, suffered a sad fate: The term was introduced to “epitomize a carefully-drawn concept in the use of units of meaning for retrieval” purposes, but it came to be used as synonymous with index entries of any sort. For example, as early as the publication of the proceedings of the International Conference on Scientific Information, held in Washington, D.C. in 1958, the term “descriptor” was equated to “index terms” and “keywords” in the proceedings index. In Mooers' view this was “the uttermost semantic corruption.” (HENDERSON, 1996, p. 660).

⁶⁷ “It is the bridge, the connection between the users and the information.” (MOOERS, 1985, p. 249).

combinações, mas que ele chamou de descritor. Propõe o descritor como composto por duas partes: um rótulo ou etiqueta (*label*) e a definição (*definition*), que pode ser um verbo ou uma palavra. A partir do rótulo e da definição, considerou alguns princípios básicos úteis à recuperação da informação.

Segundo Mooers (1985), a linguagem de indexação auxiliaria o isolamento dos conceitos através de um rótulo (*label*), exemplificando com o termo “*lubrication*” que pode ser um *descriptor label* (rótulo do descritor), mas que também precisaria, segundo ele, de um significado. Por sua vez, esse significado relacionado com o sistema de recuperação não estaria baseado em definições de dicionários, pois “[...] toda organização ou laboratório com um sistema de recuperação de informações tem seu ponto de vista especial sobre cada tópico ou conceito de importância para ele. Portanto, este ponto de vista especial deve ser escrito na definição do rótulo.”⁶⁸ (MOOERS, 1985, p. 253, tradução nossa).

Seria possível, a todo descritor, a inclusão de sua definição anexada ao rótulo como uma ideia pronta, o que o autor denominou ideia empacotada (*packaged idea*) “[...] como um conceito que deve ser usado principalmente para recuperação em algum sistema de recuperação específico.”⁶⁹ (MOOERS, 1985, p. 253, tradução nossa). Com isso, haveria a constituição de uma lista, com rótulos e definições aclopadas a um conjunto de seis características definidas.

A primeira característica refere-se ao vocabulário com os rótulos associados a um conjunto pequeno de descritores (250 a 350), sendo que, em casos especiais, esse número poderia ser maior, chegando a mil. A segunda é a do uso de um descritor ou mais para formar um conjunto a ser adotado na caracterização dos documentos. A terceira é a de que cada descritor seria definido por uma nota de escopo (*scope note*). A quarta é a necessidade de sutileza, segundo denominação de Mooers, pois o conjunto de descritores, como também cada um deles individualmente, deveriam ser adaptados para um trabalho específico de recuperação. A quinta característica é a de que seria desejável organizar o conjunto de rótulos dos descritores em uma folha de papel, com o propósito de programar e exibir, agrupar ou classificar; a classificação serviria apenas para individualizar os descritores, não implicando uma hierarquia, que na visão de Mooers seria incômoda e deveria ser evitada. A sexta e

⁶⁸ “Every organization or laboratory with an information retrieval system has its own special viewpoint about each topic or concept of importance to it. Therefore this special viewpoint should be written into the definition of the label.” (MOOERS, 1985, p. 253).

⁶⁹ “[...] as a concept which is to be used primarily for retrieval in some particular retrieval system.” (MOOERS, 1985, p. 253).

última característica é a escolha dos descritores para a busca. Ao escolher um, dois ou três descritores, seria possível usá-los de forma concomitante, como um “conjunto prescrito” (*prescribing set*), com a finalidade de conduzir a busca pela informação desejada. A pesquisa considerando a prescrição de descritores aprimoraria a busca dos documentos, tornando-a mais específica.

De acordo com Mooers (1985), seu sistema não tinha equivalência com o *Uniterm* (Unitermo), criado por Mortimer Taube, em 1952, o qual adotou o método de indexação a extração de palavras diretamente do texto. Para o autor, a maior dificuldade do Unitermo seria o controle de sinônimos, pois o sistema baseou-se em palavras ao invés de conceitos. Da mesma forma, as duas propostas não se aproximavam do tesouro, em função da ausência de relacionamentos entre os termos, controle terminológico, definições, dentre outros aspectos a serem considerados.

De forma geral, o trabalho de Mooers relacionou a linguagem de indexação à recuperação da informação, pois considerou os descritores a partir da coleção. O descritor foi o fundamento para distinguir o que ele propunha e que não era idêntico ao contexto de outros sistemas, como aqueles constituídos por cabeçalhos de assunto, palavras-chave, dentro outros elementos de controle. Mas, seu interesse primeiro foi o de subsidiar um sistema que tivesse a capacidade de melhorar a recuperação da informação.

5.3 A INDEXAÇÃO E O ÍNDICE

O desenvolvimento da elaboração de índices englobou tanto os livros quanto os catálogos de bibliotecas, no entanto, a intensificação da atividade de indexação decorreu do crescimento das publicações periódicas, fato que gerou a necessidade de mecanismos que abordassem os assuntos de forma mais específica.

Para Gil Leiva (2008), o surgimento de bibliotecas especializadas e centros de documentação no início do século XX, sobretudo junto com as grandes bases de dados especializadas, demandou que o conteúdo dos documentos fosse detalhado, a partir de campos das bases de dados que permitissem identificar assuntos, pelo processo de indexação e, assim, facilitar a recuperação.

De maneira geral, os catálogos podem ser caracterizados pela forma alfabética e sistemática, variando de acordo com as entradas arranjadas nas fichas. Dentre os mais

conhecidos, é possível citar os catálogos caracterizados pelo tipo de entrada por autor, assunto, título e geográfico, no caso dos catálogos alfabéticos, sendo que o catálogo dicionário congrega todas essas entradas no mesmo conjunto. Já os catálogos sistemáticos podem ser de assunto, topográfico e cronológico. Assim, conforme a forma assumida nas entradas, o índice do catálogo pode ser gerado para indicar a posição de cada registro e facilitar o acesso (SHERA; EGAN, 1969).

A elaboração de índices é o resultado do processo de indexação com seus aspectos teórico-metodológicos. O arranjo em ordem alfabética ou sistemática dos catálogos acompanhou os índices. Cada um dos elementos definidos como critério para a entrada na ficha receberia uma ou mais palavras, em correspondência ao assunto retirado da lista de cabeçalho, no caso do arranjo alfabético e, um número, reproduzido de acordo com a configuração no sistema de classificação, mas que também é acompanhado de um arranjo alfabético como modo de auxiliar o acesso.

Em relação à elaboração de índices de livros, o processo consiste na ação de identificar e descrever o seu conteúdo com termos representativos dos seus assuntos, produzindo-se uma lista destes assuntos, cada um seguido do número da página em que está citado no livro.

Já na elaboração de índices de periódicos, assim como acontece com os livros, descreve-se o conteúdo com a indicação das páginas. Mas, há diferença por conta da periodicidade e devido ao fato de a paginação recomeçar em cada novo número. Por isso, é importante que na indicação de referência à página, conste também os elementos do fascículo do periódico para a identificação dos assuntos elencados. Os índices dos periódicos impulsionaram a elaboração de índices de bases de dados de itens bibliográficos, sobretudo as bases de dados de artigos.

Os assuntos para a formação dos índices são registrados na forma de entradas e, independente se feitos para livros, periódicos ou bases de dados, sua finalidade é ser um recurso para a recuperação da informação. Essas entradas caracterizam o tipo de arranjo do índice, podendo ser alfabético, sistemático, cronológico, geral (combinação de diferentes tipos de entradas), dentre outros.

Quanto à prática de construção dos índices, Shera e Egan (1961, p. 21) descrevem que:

Em meados do século XIX, o periódico tinha uma história de quase dois séculos, mas em 1850 sua importância para os estudiosos tornara-se tão grande que os bibliotecários se convenceram de que era necessário elaborar uma técnica para a organização, por assunto, do conteúdo de tantas publicações periódicas. Em 1848, publicou-se *An Alphabetical Index to Subjects Treated in the Reviews and other Periodicals*, sob os auspícios da Universidade de Yale. Seu compilador, William Frederick Poole [...].

Foi Poole quem preparou o caminho para os índices dos periódicos. Sua iniciativa consistiu na expansão da estrutura já materializada pelo processo, pois “[...] ao invés de um índice para cada periódico referente a um único ano, Poole introduziu a idéia [sic] de um índice para muitos periódicos abrangendo um considerável número de anos.” (COLLISON, 1971, p. 10).

Em 1901, foi lançado, por H. W. Wilson, o *Readers' Guide to Periodical Literature*⁷⁰ com a indexação de artigos por autor e assunto, com numerosas remissivas, vinculando assuntos entre si, de modo uniforme. O aumento da produção científica, quer através dos periódicos quanto dos livros, foi fator propulsor para a qualidade na elaboração dos índices destes tipos de documentos e de outros, como é o caso da *Encyclopedia Britannica*. A produção desta Enciclopédia no século XIX é considerada admirável (COLLISON, 1971).

A obra de Wheatley (1902) tornou-se uma fonte de consulta para profissionais da indexação, sobretudo na *Index Society*. O autor sistematizou de forma minuciosa o contexto histórico dos índices e elucidou que:

O uso cuidadoso de referências cruzadas é importante para a seleção de cabeçalhos apropriados. Grande julgamento, no entanto, é necessário, já que os consultores ficam naturalmente irritados ao serem encaminhados para trás e para frente, particularmente em um grande índice. Ao mesmo tempo, se criteriosamente inseridas, essas referências são de grande ajuda. O Sr. Poole diz, em um artigo sobre seu próprio índice no *Library Journal*: “Se todo assunto possuir referências cruzadas como aliadas, a obra será principalmente um livro de referências cruzadas em vez de um índice de assuntos.”⁷¹ (WHEATLEY, 1902, p. 71, tradução nossa).

Embora o autor não realize aproximação com as referências cruzadas de Cutter, ele explica que no referido processo, esse elemento assume caráter metodológico, todavia devem-se observar os excessos, conforme consta na citação.

Cutter (1904) recomendava ocasionalmente o uso de “[...] remissivas do específico para o geral, para superar a carência de especificidade [...]; todavia, não preconizava o

⁷⁰ O *Readers' Guide to Periodical Literature* ainda hoje fornece indexação de 300 periódicos populares de relevância publicados nos Estados Unidos e no Canadá. Os artigos indexados abrangem tópicos de interesse atual e histórico abordando temas variados como Astronomia, Informática, Educação, Moda, Cinema, História etc. É comercializado pela empresa de mesmo nome do seu idealizador. Ver: <https://www.hwwilsoninprint.com/periodicals.php>

⁷¹ The careful use of cross references is next in importance to the selection of appropriate headings. Great judgment, however, is required, as the consulters are naturally irritated by being referred backwards and forwards, particularly in a large index. At the same time, if judiciously inserted, such references are a great help. Mr. Poole says, in an article on his own index in the *Library Journal*: “If every subject shall have cross references to its allies, the work will be mainly a book of cross references rather than an index of subjects.” (WHEATLEY, 1902, p. 71).

procedimento inverso, isto é: estabelecer uma remissiva do geral para o específico.” (SIMÕES, 2017, p. 164). A autora destaca que esse princípio de remissivas também foi uma forma de antever as relações hierárquicas em instrumentos documentários utilizados no processo de indexação.

De forma geral, é possível definir o índice como um guia sistematizado por entradas acompanhadas de indicativos que sinalizam o lugar (página, volume, ano, etc) do assunto no documento. Para Reitz (2014, on-line), índice (*index*) é uma lista organizada, geralmente em ordem alfabética, de nomes pessoais, lugares e assuntos que são abordados em um trabalho escrito, com os respectivos números das páginas para encaminhar o leitor ao ponto no texto no qual as informações pertencentes às entradas podem ser localizadas. Wheatley (1902) afirma que a palavra índice a partir do latim é relativamente recente, mas não precisa a data. Segundo Buonocore (1976, p. 256), ela é originária do latim *index*, *indicis*, *de indiciere* e seu significado tem relação com livro e biblioteca. Para ele, índice é sinônimo de catálogo, pois designa a ordem desse instrumento de busca.

Para Houaiss (2001, p. 652), a etimologia da palavra índice origina-se do latim *index*, que se refere a “[...] o que anuncia, mostra, denuncia, indicador, revelador, índice, inscrição, registro, catálogo, lista, inscrição.”

O instrumento mais importante em um sistema de recuperação de informações, segundo Pinto (1985, p. 170), é o índice, considerado “[...] uma ponte ou elo entre o conteúdo de um arquivo de informações e seus usuários.” Para ele existir, deve haver o processo de indexação, que “[...] consiste na indicação do conteúdo informativo de um documento através da determinação de um ou mais termos que representarão esse conteúdo.”

Gil Leiva (2008, p. 64, tradução nossa) explica que a função de um índice é ligar uma ‘coisa’ com a sua localização e assim estabelecer uma conexão exata entre dois pontos. Para o autor, “quando falamos de documentos, elaborar um índice é tirar palavras ou frases de seu contexto, arranjá-las segundo um objetivo e indicar seu lugar no documento onde se localiza cada palavra ou frase extraída.”⁷²

O processo de elaborar índices, a indexação, segundo Langridge (1974, p. 78), envolve distinção entre a forma de indexação por palavra e por conceito. Para o autor, na primeira “é possível indexar documentos empregando simplesmente as palavras encontradas neles próprios.” Explica ainda que essa é a forma mais rudimentar de indexar e, cita como exemplo,

⁷² “Cuando hablamos de documentos, confeccionar un índice es sacar palabras o frases de su contexto, ordenarlas según convenga y señalar el lugar del documento en donde se localiza cada palabra o frase entresacada.” (GIL LEIVA, 2008, p. 64).

o *Key-word in-context* (KWIC). Nesse tipo de índice, palavras-chave dos títulos dos artigos dos periódicos eram extraídas para que cada uma fosse usada alfabeticamente como entrada. O processo era rápido e chegava com velocidade até os usuários, mas pecou por ser superficial, uma vez que não fornece consistência aos assuntos. Já pela indexação por conceito, se “[...] reconhece os assuntos, sejam quais forem as palavras em que estejam expressos [...]” demandando mais tempo. O autor ainda ressalta que esse tipo de indexação “[...] requer um vocabulário controlado [...]” (LANGRIDGE, 1974, p. 79).

A indicação do assunto pelo processo de indexação exige a atividade de análise. Abbagnano (2007, p. 51, destaque do autor) define de forma ampla a análise como a maneira de descrever ou interpretar uma situação ou um objeto “[...] nos termos dos elementos mais simples pertencentes à situação ou ao objeto em questão. A finalidade desse processo é *resolver* a situação ou o objeto nos seus *elementos*, de modo que um processo analítico é considerado bem-sucedido quando tal resolução é realizada.”

Valendo-se da definição de Abbagnano (2007) para o contexto proposto, o assunto é resultante da análise no processo de representação temática do documento na forma de um registro na base de dados. Entretanto, por um longo período na história, a representação temática de documentos esteve atrelada às operações empíricas profissionais do “[...] ‘bom senso’, pelas quais se atribuía, a um texto/documento qualquer, uma ou várias palavras-chave destinadas a facilitar a recuperação, no momento da pesquisa, sob um dado tema.” (CUNHA, 1990, p. 59).

É pela análise que o processo de indexação é iniciado, podendo receber denominações baseadas na linguagem utilizada para representar o assunto. Nesse sentido, a conversão da análise pode ser realizada através dos sistemas classificatórios, denominando-se classificação ou, por tesouros, designando-se indexação. Em sentido amplo, a análise de assunto para ambos os processos não se diferencia. Mas, cabe salientar que os sistemas classificatórios não se referem exclusivamente aos aspectos temáticos, como demonstra a CDD, a partir da qual é possível considerar a tipologia dos documentos (como nas obras literárias, dicionários, periódicos, etc), pontos de vista particulares, como em uma abordagem histórica, teórica, dentre outros elementos.

5.4 CONCEITO E ANÁLISE CONCEITUAL

A palavra conceito é amplamente explorada em Biblioteconomia e Ciência da Informação no âmbito do tema assunto, haja vista que a elaboração de conceitos é etapa necessária à organização do conhecimento. Por sua vez, essa organização envolve uma via dupla, tanto em relação à construção de instrumentos e linguagens quanto de assuntos em registros de informação.

Dahlberg (1993) relaciona a Teoria da Classificação Facetada de Ranganathan como uma relevante contribuição para o universo do conceito. Kaula (1982, on-line) aborda que dentre os três Planos de Ranganathan (das ideias, verbal e notacional),

O trabalho no plano das idéias [sic] é considerado superior. Ele provém do trabalho da mente, que é o lugar onde se originam as idéias [sic]. O criador de idéias [sic] precisa de auto-comunicação no interior da mente a fim de criar mais idéias [sic]. O trabalho no plano das idéias [sic] pode ser tomado como análise do conceito. Uma idéia [sic] é um conceito que ao tomar forma concreta pode levar a alguma informação. A análise conceitual é uma tarefa difícil que tem que ser esgotada na concepção do esquema de classificação. Um conceito pode ser um isolado, um quase isolado ou um assunto e é a identificação de conceitos, sua posição no universo de assuntos, seu arranjo sistemático entre outros conceitos, etc., que faz do trabalho uma tarefa árdua.

De acordo com Campos e Gomes (2003), para Ranganathan, o conhecimento é contínuo, resultando em um universo infinito de assuntos. Segundo Kaula (1982) a análise conceitual deve ser “esgotada na concepção de um esquema de classificação”, mas ela tem início no plano das ideias. Segundo o mesmo autor, identificar um conceito na “posição do universo dos assuntos” é tarefa árdua. Ele ainda explica:

Creio que o termo ‘Classificação para Bibliotecas’ deveria ser substituído por Classificação de Conceitos[,] pois a classificação de assunto cria homonímia. Já a Dra. I. Dahlberg trabalhou na análise conceitual endossando a teoria básica da classificação como foi realizada pelo Dr. Ranganathan. Ela é de opinião[o] que a estrutura do sistema deveria ser baseada na categoria de descritores de objetos e campos de aspectos, o que é um esclarecimento avançado para a categoria fundamental e para os conceitos de relação de fase usados na Classificação dos Dois Pontos. (KAULA, 1982, on-line).

A Classificação dos Dois Pontos (ou, *Colon Classification*, CC) se diferenciou por envolver uma abordagem analítico-sintética rompendo com a estrutura dos métodos tradicionais de classificação. “Em vez de enumerar classes [...], a CC lista somente objetos e idéias relativamente simples bem como propriedades e características gerais, ou ‘facetas’, cujas marcas de classe podem ser combinadas para expressar exatamente o assunto de um

documento.”⁷³ (WORLD ENCYCLOPEDIA OF LIBRARY AND INFORMATION SERVICES, 1993 p. 210, tradução nossa).

Em relação à Dahlberg (1978), ela desenvolveu a Teoria do Conceito com base na análise das relações lógicas e na estruturação em hierarquias abordando um conjunto de subordinação e superordenação para formação de relações a partir de abstrações, com o objetivo de representar o real. Para a autora, é através da linguagem que os enunciados dos objetos são elaborados e, o objeto pensado como único, em relação ao espaço e tempo, torna-se individual. Já os objetos gerais, que estão fora do tempo e espaço, segundo Dahlberg (1978, p. 102), tem correspondência com os chamados “conceitos gerais”.

Dahlberg (1978, p. 102) afirma que pelos enunciados é possível elaborar os conceitos dos objetos, mas “cada enunciado verdadeiro representa um elemento do conceito”. De outro modo, os conceitos são formados por objetos que os enuncie fidedignamente. Ademais, o conceito constitui unidade básica de pensamento, nomeado posteriormente pela autora como unidade de conhecimento.

Os enunciados “[...] identificam-se com as chamadas características *dos conceitos*. Traduzem os atributos das coisas designadas [...]” como resultado do processo de análise conceitual (DAHLBERG, 1978, p. 102, destaque da autora). Ao atribuir as características, que são propriedades dos objetos, no nível do conceito, essas passam a ser características do conceito. Também é possível tratar de vários atributos que corresponderão a muitas características e, dessa forma, “[...] o predicado de um enunciado pode tornar-se sujeito de novo enunciado e assim sucessivamente até [...] uma característica tão geral que possa ser considerada uma *categoria*. (Entende-se aqui por categoria o conceito na sua mais ampla extensão).” (DAHLBERG, 1978, p. 102, destaque da autora).

Por sua vez, essas características quando se diferenciam estabelecem relações lógicas e diferentes tipos de relacionamentos semânticos, por exemplo, hierárquico, partitivo, de oposição e funcional, fato que auxilia as comparações entre os conceitos, contribuindo para a organização de sistemas de classificação e tesauros.

De forma ampla, sistemas de classificação e tesauros são sistemas conceituais, estruturados a partir de conceitos e suas características. Mesmo em alguns sistemas de classificação estruturados por disciplinas, presume-se que implicitamente há presença

⁷³ “Instead of enumerating classes [...], the CC lists only relatively simple objects and ideas as well as general properties and characteristics, or “facets”, whose class marks can be combined to express exactly the subject of a document.” (WORLD ENCYCLOPEDIA OF LIBRARY AND INFORMATION SERVICES, 1993, p. 210).

conceitual para tal caracterização. Segundo Dahlberg (1978), a análise conceitual consiste na identificação das características desses conceitos e das relações que podem ser estabelecidas.

De outro modo, Lancaster (2004, p. 9) explica que a indexação de assuntos envolve duas etapas principais: análise conceitual e tradução. Para o autor, “a análise conceitual, [...] implica decidir do que trata um documento – isto é, qual o seu assunto.” O propósito, segundo o autor, deve alcançar as necessidades de quem utiliza e, para tanto, a decisão deve consistir em não somente analisar o que é tratado no documento, mas em como ele poderá interessar a grupos de usuários. Ainda segundo o mesmo autor,

Em outras palavras, não há um conceito ‘correto’ de termos de indexação para documento algum. A mesma publicação será indexada de modo bastante diferente em diferentes centros de informação, e deve ser indexada de modo diferente, se os grupos de usuários estiverem interessados no documento por diferentes razões. (LANCASTER, 2004, p. 9).

Assim, para Lancaster (2004), a análise conceitual constitui a primeira etapa do processo de indexação, como uma ação intelectual que contribui com a etapa seguinte, que é a tradução dos assuntos identificados pela análise em termos descritores de uma linguagem. A palavra conceito, citada pelo autor, tem reflexo na ação mental desencadeada na primeira etapa e, como consequência, impacta na escolha dos termos retirados da linguagem de indexação. Entretanto, de sua parte não houve aprofundamento teórico sobre o uso desse termo e, o autor define: “‘Conceito’ é outra palavra sobre a qual alguns autores gostam de filosofar (ver, por exemplo, Dahlberg [1979]). [...]. ‘Análise conceitual’, portanto, significa nada mais do que a identificação dos assuntos estudados ou representados num documento.” (LANCASTER, 2004, p. 15). O autor enfatiza que a análise conceitual é a própria identificação do assunto.

Pinto (1985, p. 171) também divide o processo de indexação em dois estágios:

- a) estabelecimento dos conceitos tratados e seleção dos que serão indexados, tendo em vista os objetivos do sistema;
- b) tradução dos conceitos selecionados em termos indexadores de acordo com algum padrão consistente, ou seja, de acordo com a linguagem de indexação usada no sistema.

A autora não faz uso do vocábulo ‘análise conceitual’, porém acentua que a primeira etapa consiste no estabelecimento de conceitos com foco nos objetivos do sistema, que devem estar de acordo com a demanda de quem faz uso, portanto, conforme Lancaster (2004) declarou, dos grupos de usuários. Em seguida, a autora coloca que os conceitos devem ser

traduzidos para termos indexadores e, não para assuntos, diferindo-se de Lancaster (2004). Porém, ela afirma que “qualquer conjunto de termos e/ou símbolos usados para denominar classes de assuntos em sistemas de recuperação da informação constitui uma linguagem de indexação.” (PINTO, 1985, p. 171). Não especifica o que são classes de assuntos, mas deixa entender que estão relacionados aos conceitos selecionados no segundo estágio, depois transformados em termos indexadores.

Langridge (1974) explica que o processo de indexação é melhor executado ao ser trazido ao nível da consciência e, de alguma forma, envolvendo o instrumento adotado. Nesse sentido, aborda o *Roget's Thesaurus*, que é diferente do arranjo alfabético de um dicionário por agrupar termos segundo seu significado. Apesar de possuir um índice alfabético com o fim de permitir localizar palavras relacionadas, o instrumento foi devotado às relações semânticas.

Outros instrumentos foram sucessivamente desenvolvidos com foco em objetivos particulares e, considerando os sistemas de classificação bibliográfica e os tesouros, Lara (2001, p. 6), a partir de Grolier (1962),⁷⁴ explica:

De um modo geral, a tendência da maioria das propostas está relacionada à decomposição dos assuntos complexos em “fatores” simples, ao abandono das classificações fortemente hierarquizadas e a idéia [sic] de que é necessário não apenas relacionar os termos elementares que servem à análise dos documentos, mas principalmente as relações que os conformam.

Se a estrutura semântica do instrumento ganhou força a partir do *Roget's Thesaurus*, ao contrário do movimento hierarquizador dos sistemas classificatórios, o tesouro volta-se para os conceitos. Para Lara (2001, p. 8, destaque da autora), o tesouro tornou-se o instrumento documentário pioneiro de representação ao trabalhar “[...] com a noção de controle do vocabulário. [É] resultado do reconhecimento de que operamos no universo da linguagem, o que significa dizer que a *transferência da informação* está atrelada a sistemas de *significação*.”

Nesse contexto, o tesouro passou a ser compreendido como instrumento construído em função do termo-conceito, não como uma consequência da ideia de um cabeçalho de assunto, mas de uma decomposição do assunto pautada por parâmetros semânticos.

⁷⁴ GROLIER, E. de. **Étude sur les catégories générales applicables aux classifications et codifications documentaires**. Paris: UNESCO, 1962.

A compreensão sobre o assunto foi se desvencilhando da palavra presente no texto, de onde seria apenas retirado, extraído, para ser percebido como algo construído a partir do texto. O instrumento reflete a escolha do termo para representar o assunto. Para alcançá-lo, uma sequência de etapas foi desenvolvida – análise, síntese e representação – porque o assunto não existe *a priori*, mas é construído por quem indexa. Esses avanços tornaram-se mais nítidos na vertente europeia, que será aprofundada à frente.

5.5 A NORMALIZAÇÃO DA INDEXAÇÃO

A representação dos assuntos dos documentos em um conjunto de códigos representativos, tópicos ou conceitos demandaram critérios de padronização metodológica. Conforme aponta Lancaster (2004), devem ser indexadas as ideias do autor e não as palavras do texto. Para tanto, instruções de caráter internacional e nacional foram desenvolvidas com recomendações para execução de todo processo, visto que ele sofre interferência de diferentes elementos, como cognitivos, institucionais, sociais, dentre outros.

De acordo com Gil Leiva (2008), desde a metade dos anos 1960, critérios foram pensados como regras de indexação para representação dos conceitos presentes no documento. Ainda segundo o mesmo autor, os esforços normalizadores iniciais aconteceram na década de 1970, com a publicação na França da NF 47-102 1978, intitulada *Principles généraux pour l'indexation des documents* (Princípios gerais para a indexação de documentos). Houveram esforços parecidos em outros países.

Mas, a despeito das normas e, considerando a necessidade de formalização metodológica de princípios para o embasamento do processo de indexação, compreende-se que essa formalização ocorreu com os Princípios de Indexação do UNISIST. O programa *World Scientific Information Programme*, nomeado com a sigla UNISIST (*United Nations International System for Information in Science and Technology*), foi responsável pela criação dos Princípios de Indexação. Desenvolvido em nível internacional, em 1975, e vinculado a *United Nations Educational, Scientific and Cultural Organization* (UNESCO),⁷⁵ culminou em 1985, com a criação da primeira *International Standardization for Organization* (ISO), responsável em fornecer diretrizes ao processo de indexação. A ISO 5963/1985 foi denominada *Documentation – methods for examining documents, determining their subjects*,

⁷⁵ Organização das Nações Unidas para a Educação, a Ciência e a Cultura.

and selecting indexing terms. No Brasil, em 1992, a Associação Brasileira de Normas Técnicas (ABNT) publicou a tradução da ISO 5963, intitulada Métodos para análise de documentos – determinação de seus assuntos e seleção de termos de indexação – com a numeração NBR 12676/1992.

Considerando que os Princípios UNISIST constituíram a primeira iniciativa de âmbito global a providenciar recomendações sobre a análise no processo de indexação, observa-se que:

A indexação é vista como a ação de descrever e identificar um documento de acordo com seu assunto. [...]
Durante a indexação, os conceitos são extraídos do documento através de um processo de análise, e então traduzidos para os termos de instrumentos de indexação (tais como tesauros, listas de cabeçalhos de assunto, esquemas de classificação, etc.). (UNITED NATIONS INTERNATIONAL SYSTEM FOR INFORMATION IN SCIENCE AND TECHNOLOGY, 1981, p. 84).

Observando a análise como meio para alcançar o assunto, pode-se atribuir a ela o resultado da síntese. Os dois momentos, análise e síntese, são responsáveis por desencadear a construção metodológica do assunto, desconstruindo a ideia de ele estar pronto no documento. Por consequência, o percurso propiciará a tradução, pois através dos instrumentos será feita a escolha de termos que representem tematicamente o documento. Esse panorama geral subsidia o ato de indexar sob dois pontos principais: um enquanto processo embasado na análise de conceitos que determinarão o assunto do documento e outro quanto à finalidade de propiciar a sua recuperação e acesso.

Os Princípios UNISIST explicitam de forma ampla que o processo de indexação consiste em dois estágios principais: “-Estabelecimento dos conceitos tratados num documento, isto é, o assunto; -Tradução dos conceitos nos termos da linguagem de indexação.” (UNITED NATIONS INTERNATIONAL...,1981, p. 85).

Sobre o primeiro, depreende-se o estabelecimento do nível conceitual que não é o mesmo das palavras que designam os assuntos. Por meio do processo mental, um conjunto de características são estabelecidas para elaboração conceitual que possa representar o assunto do documento. O estabelecimento do assunto de um documento pode ser dividido em três estágios: 1-Compreensão do conteúdo do documento como um todo; 2-Identificação dos conceitos que representam este conteúdo; e, 3-Seleção dos conceitos válidos para a recuperação (UNITED NATIONS INTERNATIONAL...,1981).

Os três estágios são abordados pelos Princípios de forma interligada e o primeiro deve considerar a compreensão do documento pela forma, apontando que podem ser gráficos

(como por exemplo, livros, monografias, relatórios, artigos, teses) e não gráficos (como por exemplo, audiovisuais, visuais, sonoros). Na identificação de conceitos, deve-se reconhecer as categorias que tenham relação com a cobertura do documento, como por exemplo, o fenômeno, o processo, a propriedade, dentre outras categorias envolvidas. Por último, na seleção de conceitos, ocorrerá a escolha desses para sua tradução em termos descritores que possam ser empregados com o objetivo do processo, que é a recuperação.

Essa situação implica a decisão em torno de dois parâmetros principais a serem considerados no processo de indexação: exaustividade e especificidade. A exaustividade refere-se à identificação de um amplo número de termos de indexação com valor potencial de recuperação. A especificidade implica a adoção de termos de indexação identificados de forma mais específica quanto a cada documento analisado.

Apesar de conceito e assunto não serem elementos idênticos, há uma relação entre eles e se subentende pelos Princípios UNISIST que o primeiro proporciona a construção do segundo. Essa consonância entre a compreensão, a identificação e a seleção de conceitos para a determinação do assunto assegura o próximo estágio, que é o da tradução, ou seja, a representação de conceitos por termos de uma linguagem de indexação. Segundo os Princípios, as linguagens são instrumentos de tipo verbal (como tesouros e listas de cabeçalhos), e de tipo simbólico (conceitos representados por símbolos de classificação), devendo quem indexa estar familiarizado com os respectivos instrumentos, além dos seus procedimentos e regras. A ressalva feita pelos Princípios é a de que alguns desses instrumentos, como as listas de cabeçalhos de assunto, podem impossibilitar a representação exata do conceito. Outros, como o tesouro, permitem o desdobramento de relações através dos termos descritores. No texto dos Princípios, explica-se ainda que, dependendo do instrumento, caso ele não possibilite a tradução do conceito para um descritor específico, um genérico deverá ser adotado e um novo descritor será candidato para a atualização do instrumento.

Os Princípios UNISIST são recomendações gerais que propiciam o entendimento da indexação de forma ampla e, mesmo sem referendar outros aspectos importantes, como por exemplo, as políticas de indexação, e detalhar formas de avaliação do processo, respaldam movimento em prol da abordagem sistemática da representação de assuntos.

Na indexação, ocorre a identificação de aspectos que deem conta de uma representação mais especializada, conforme foi se estabelecendo na *Index Society* e, após, no contexto contemporâneo, no texto dos Princípios UNISIST e nas normas de indexação que dele resultaram.

Na *Rules* de Cutter (1904), um livro deve receber a entrada pelo assunto específico de forma direta. Por outro lado, na ficha catalográfica, um ou dois assuntos ao serem selecionados para a entrada secundária consistem em palavras-chave contidas na lista de cabeçalho de assunto e, a rigor, são delimitadas ao espaço disponível da própria ficha. Já nos índices das bases de dados não há quantidade de entradas dos termos, o que não necessariamente possibilita maior detalhamento dos assuntos tratados, mas agilidade para fins de recuperação.

Em relação à NBR 12676, ela também trabalha com assunto e conceito, e os define como:

Conceito

Qualquer unidade de pensamento. O conceito pode ter o seu conteúdo semântico reexpresso pela combinação de outros conceitos, que podem variar de uma língua ou de uma cultura para outra.

Assunto

Tema representado num documento por um conceito ou combinação de conceitos. (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992, p. 1).

A mesma Norma aponta como sendo três os estágios compreendidos no processo de indexação: exame do documento e estabelecimento do assunto de seu conteúdo, identificação dos conceitos presentes no assunto e tradução desses conceitos nos termos de uma linguagem de indexação (ASSOCIAÇÃO BRASILEIRA..., 1992).

No primeiro estágio, que se dá pela análise do documento, assim como nos Princípios UNISIST, a NBR 12676 afirma que o documento ao ser analisado dependerá da sua forma física, podendo ele ser impresso ou não, isso por conta das partes que serão analisadas. Assim, não recomenda realizar a indexação por qualquer uma das partes do documento isoladamente, sendo necessária uma leitura mais apurada.

Na identificação dos conceitos, próximo estágio da norma, há instrução para o indexador adotar uma abordagem sistemática de identificação dos conceitos essenciais na descrição do assunto. A Norma sugere perguntas que devem ser elaboradas pelo indexador, as quais servirão para identificar os aspectos importantes na área coberta pelo índice. Os exemplos sugeridos são: a) o assunto principal do documento; b) como definir o assunto em termos de teorias, hipóteses; c) a ação, a operação ou processo do assunto (se possuir); d) a definição do agente da ação, operação, processo, etc.; e) a constatação de métodos, técnicas e instrumentos especiais; f) a consideração dos aspectos no contexto de um local ou ambiente especial; g) a identificação de variáveis dependentes ou independentes; h) a consideração do

assunto em relação a um ponto de vista interdisciplinar (ex.: um estudo sociológico da religião).

Se antes, Kaiser e os autores do *Classification Research Group* imputaram cientificidade em seus estudos, os exemplos sugeridos pela NBR 12676 (que é um documento traduzido), não enfocam a mesma proporção de discussão. É possível presumir ausência de relação com a carga teórica desenvolvida sobre a indexação em momentos anteriores.

Ainda em relação à identificação dos conceitos, ao fazer a escolha, quem indexa deve levar em consideração as consultas que podem ser realizadas pela comunidade usuária. Também conforme apontado nos Princípios UNISIST, a NBR 12676 explica que nem todos os conceitos identificados serão aplicados, dependendo do grau de exaustividade e especificidade do sistema de informação. A Norma explica que a abrangência atribuída aos assuntos deve considerar os diferentes grupos de usuários.

Em continuidade, ainda na identificação dos conceitos, a Norma aponta que não se deve estabelecer limite ao número de termos ou descritores que possam ser atribuídos ao documento (exaustividade). Nesse caso, se um serviço de indexação necessita limitar a quantidade de termos, a seleção dos conceitos deve ser orientada pela política de indexação e não pelo julgamento de quem indexa, de acordo com o papel que cada conceito desempenha na expressão do assunto global do documento. Em relação ao peso dado ao conceito, se o indexador considerar que uma ideia não está completamente desenvolvida, ou se houver ineditismo e/ou necessidade de consolidação, também poderá realizar uma indexação em nível geral.

O último estágio de indexação proposto pela Norma é a seleção de termos de indexação, que é explicada ao indexador por meio das seguintes práticas: a) usar descritores convenientes com a linguagem de indexação; e b) para os termos que representem novos conceitos, deve-se verificar sua aceitação em outros instrumentos como: dicionários e enciclopédias reconhecidas, tesouros e tabelas de classificação.

A partir do exposto, pode-se perceber que os autores da primeira parte do século foram mais longe quando propuseram categorias, com base em fórmulas, que já apresentavam algum tipo de 'base linguística' e na observação da constituição de um enunciado. Mas, nada disso foi considerado pelos Princípios e pela Norma. Ambas usam o termo conceito, mas não tratam do caráter conceitual do processo de indexação.

Paralelo a isso, outro aspecto a se considerar é o assunto enquanto ponto de acesso em bases de dados. O marco da sua padronização decorreu da Conferência de Paris, em 1961, também conhecida como Conferência Internacional sobre Princípios de Catalogação

(*International Conference on Cataloguing Principles*). O processo de revisão dos Princípios de Paris se fez por meio de várias conferências no decorrer dos anos de 2003 e 2007 e, por consequência, resultou na Declaração dos Princípios Internacionais de Catalogação de 2009 (*Statement of International Cataloguing Principles*), com versão atualizada no ano de 2016. (STATEMENT OF INTERNATIONAL CATALOGUING PRINCIPLES, 2009; 2016).

Na Declaração dos Princípios, tanto de 2009 quanto de 2016, é nítida a composição sobre a importância do ponto de acesso, fato que encontra recepção nos modelos conceituais da IFLA (FRBR, FRAD e FRSAD), que serão abordados à frente.

Em todo conjunto das perspectivas apontadas, prevaleceu, de um lado, a secundarização de fundamentos sobre aspectos que envolveram o assunto e sua representação, pois o destaque ocorreu nas técnicas operacionais do processo de indexação atrelado ao uso da linguagem. Esse contexto marcou o cenário temático da Biblioteconomia, a despeito dos fundamentos subjacentes às sistematizações realizadas sobre o processo e às regras propostas para realizá-lo.

5.6 A DISPERSÃO TERMINOLÓGICA: *ABOUTNESS*, *TOPIC*, *CONTENT*, *SUBJECT* – ASSUNTO, TEMA, MATÉRIA – *CONTENIDO*, *MATERIA*, *EPÍGRAFE*, *TEMA*

Apesar da diferença no modo de abordar o processo, segundo a literatura que adota a denominação catalogação de assunto e a que adota a denominação indexação, que se faz, como se vê, por nomes próprios, percebe-se coincidência no objetivo de representar o conteúdo para propiciar a recuperação.

Seja em ambiente analógico, como nos catálogos em fichas, ou em ambiente digital, como nos catálogos on-line e bases de dados, os índices exercem o papel de facilitar o acesso aos documentos.

Quanto a estes diferentes termos para tratar do processo, há propostas diversas também no que tange à construção da noção de assunto, conforme desenvolvido nas seções anteriores.

Na década de 1970, discussões relacionadas à análise de documentos para transformá-los em representações seguiram o que alguns pesquisadores chamaram de *aboutness*. Guedes

(2009), a partir de Guimarães e Moraes (2006),⁷⁶ explica que inicialmente foi Hutchins (1977; 1978) o responsável em dar destaque ao termo. Por sua vez, Hutchins (1977) indica responsabilidade a Fairthorne (1969).⁷⁷

A preocupação inicial de Hutchins (1977, p. 17) partiu de como compreender o “tópico de um documento” (*topic of a document*) no contexto de um sistema de informação. Conforme o mesmo autor, uma resposta superficial seria entender esse tópico como assunto de uma entrada em um índice, com correspondência ao documento. Mas, a descrição do assunto simplesmente não contempla uma fórmula ou equação, tornando-se instigante alcançar “o que o documento trata” (*what the document is about*).

Segundo o mesmo autor, a linguagem documentária (*documentary language*) adotada pode ser um facilitador da “formulação de descrições de assunto” (*formulation of subject descriptions*), pois o propósito seria expressar “todo o tópico” (*whole topic*) do documento. De outro modo, o tópico de um documento (*the topic of a document*) envolve a sumarização do seu conteúdo para a recuperação em um sistema, de forma independente da linguagem documentária empregada.

Ao considerar o documento num contexto temático específico, sua possibilidade de recuperação será abarcada por esse contexto, porém antes das ferramentas e recursos disponibilizados, há uma descrição particular do conteúdo. Nesse entendimento, Hutchins (1977, p. 18, tradução nossa) interroga-se: “O que queremos dizer com o conteúdo de um documento? Para responder a isso, precisamos ser claros sobre a distinção entre o ‘sentido’ de uma expressão linguística e a ‘referência’ dessa expressão.”⁷⁸

Ao vincular o contexto linguístico, Hutchins (1977, p. 18, tradução nossa) explica que “o sentido de uma palavra determina o alcance de seus potenciais referentes.” De forma isolada, “uma palavra tem sentido, mas não tem um referente real; pode ter um referente apenas em um contexto textual específico.”⁷⁹

⁷⁶ MORAES, João Batista; GUIMARÃES, José Augusto Chaves. Análise documental de conteúdo de textos literários: em busca do diálogo entre as concepções de *aboutness/meaning* e de percurso temático/percurso figurativo. *Scire*, Zaragoza, v. 12, n. 1, en./jun. 2006.

⁷⁷ FAIRTHORNE, R.A. Content analysis, specification and control. *Annual Review of Information Science and Technology*, v. 4, p. 73-109. 1969.

⁷⁸ “What do we mean by the content of a document? To answer this we need to be clear about the distinction between the ‘sense’ of a linguistic expression and the ‘reference’ of that expression.” (HUTCHINS, 1977, p. 18).

⁷⁹ “[...] the sense of a word determines the range of its potential referents.”
 “[...] a word has a sense, but it has no actual referent; it can have a referent only in a particular textual context.” (HUTCHINS, 1977, p. 18).

O autor traça uma distinção entre o sentido de um texto e a sua referência. Explica que o entendimento do texto não implica as referências ou o que o autor teve em mente ao escrevê-lo e, se a preocupação consistir na busca da verdade ou falsidade do que foi escrito, o conhecimento dos sentidos das expressões não se torna suficiente. Cita, como exemplo, Bertrand Russell e a condição de ‘rei da França é careca’. Hutchins (1977) explica não ser possível dizer se a afirmação é verdadeira ou falsa fora de uma situação referencial específica (*outside a particular referential situation*). O referente seria a França do século XVIII e, com isso, haveria a possibilidade de testar a verdade ou a falsidade da declaração. “Assim, enquanto o sentido permanece constante, a referência varia - e assim também, em consequência, o valor de verdade.” (HUTCHINS, 1977, p. 18, tradução nossa).⁸⁰ O autor completa, com base em Fairthorne (1961), que os indexadores centram atenção no sentido dos textos que “significa o conteúdo dos documentos” (*this is what is meant by the content of documents*) e, para ele, assim deve ser. Afirma que o sentido do texto independe do autor, como se alcançasse uma existência autônoma, “como parte do que Karl Popper chamou de ‘Mundo 3’” (*as part of what Karl Popper has called 'World 3'*) (HUTCHINS, 1977, p. 18).

Com isso, o autor afirma que o problema linguístico de sumarização necessita de compreensão da estrutura do texto. Isso envolve uma progressão temática e semântica no que Fairthorne (1969), segundo Hutchins (1977), estabeleceu haver uma distinção entre o caráter extensional de um texto e sua intensionalidade. O primeiro é definido pelos tópicos de partes componentes de um texto, os tópicos de seus parágrafos, seções, capítulos, etc. O último é o tópico do texto como um todo, representando algo a mais que os tópicos de suas partes.

De acordo com Beghtol (1986), para Fairthorne, o aboutness extensional (*extensional aboutness*) “é o assunto inerente ao documento” (*is the inherent subject of the document*); já o aboutness intensional (*intensional aboutness*) é a razão ou propósito para o qual o documento foi adquirido de acordo com o interesse dos usuários. Explica que qualquer documento possui mais de um assunto, dependendo de quem o solicita, por que e para quê. Nesse contexto de forte aproximação com o assunto, ainda segundo Beghtol (1986), com base em Boyce (1982),⁸¹ os termos ‘*topicality*’ (*aboutness*) e ‘*informativeness*’ (*meaning*), concatenam-se ao significado do documento, que não é algo pronto e fixo, pois variará de acordo com diferentes fatores, dentre eles, sociais, culturais, institucionais, etc. Esse recorte demonstra que no

⁸⁰ “Thus, while the sense remains constant, the reference varies – and so too, in consequence, does the truth value.” (HUTCHINS, 1977, p. 18).

⁸¹ BOYCE, B. Beyond topicality: a two-stage view of relevance and the retrieval process. **Information Processing and Management**, 18, p. 105-109. 1982.

aboutness extensional e no ‘*topicality*’ permanece a ideia de algo pronto, estável, enquanto no aboutness intensional e ‘*informativeness*’ há presença de uma ideia mutável e não fixa. Cabe destacar a dificuldade de tradução dos termos para língua portuguesa.

Ademais, o objetivo de Beghtol (1986) foi aproximar o *aboutness* dos sistemas de classificação bibliográfica, pois segundo a autora os classificacionistas projetam os sistemas, realizam todo um trabalho conceitual e notacional, mas não consideram os significados particulares intrínsecos, pois o documento muda quando um significado particular é atribuído a ele ou quando são usados de modo particular por usuários.

Segundo Lancaster (2004, p. 13), o *aboutness*, traduzido no seu livro por ‘atinência’, busca subsidiar a expressão ‘de que trata um documento’ que, para ele, é sinônimo de ‘tem por assunto’. De acordo com o autor, “estas expressões talvez não sejam muito precisas e não é fácil definir ‘trata de’ [*about*] e ‘tem por assunto’ [*aboutness*]. Apesar disso, são expressões que soam aceitáveis para a maioria das pessoas, sendo por elas compreendidas.” Lancaster (2004, p. 13) ainda afirma ser difícil explicar a situação no que tange à indexação de assuntos, e autores como “Beghtol (1986) e Hutchins (1978) recorreram à linguística do texto ao examinar essa questão”, porém houveram autores que recorreram ao recorte probabilístico e até outros que afirmaram tratar de um aspecto “‘intratável’, visto ser tão difícil decidir do que trata um documento.”

O assunto, considerado como um dos atributos do documento, pois é uma característica que permite a estrutura do registro de informação compondo sua natureza temática, é a forma de expressar os potenciais epistemológicos do documento, porém, conforme a ótica de quem o descreve. Nessa concepção, Hjørland (1992) considera o conceito de assunto (*subject* ou *subject matter*) em dois aspectos: idealístico subjetivo e idealístico objetivo.

O primeiro aspecto envolve percepções ou visões de um ou mais indivíduos, pois assuntos são subjetivamente compreendidos e, assim, o seu conceito reside no estudo das mentes de algumas pessoas, como por exemplo, os autores ou usuários de documentos. “Do ponto de vista da epistemologia, o idealismo subjetivo é caracterizado por tornar a percepção e o pensamento independente de uma maneira subjetivista. O positivismo é o representante mais comum do idealismo subjetivo.”⁸² (HJØRLAND, 1992, p. 174, tradução nossa). Esse ponto de vista enfatiza a interpretação do documento, acarretando no que Hjørland (1992)

⁸² “From the point of view of epistemology, subjective idealism is characterised by making perception and thinking independent in a subjectivistic manner. Positivism is the most common representative of subjective idealism.” (HJØRLAND, 1992, p. 174).

nomeia ser uma concepção agnóstica do assunto (*agnostic conception of subject*), a qual leva à impossibilidade de saber como ele deve ser determinado.

O autor faz crítica ao idealismo subjetivo, pois aponta ser impossível determinar assuntos examinando a mente de autores, usuários ou qualquer outro grupo específico de pessoas. No documento há

[...] as intenções do autor, seu ponto de vista do assunto, e outros novos aspectos a ele se relaciona. Isso deu origem ao conceito de “*aboutness*” na literatura da biblioteconomia e ciência da informação, interesse que, a meu ver, representa um beco sem saída, uma tentativa de escapar das dificuldades do conceito de assunto [...]”⁸³ (HJØRLAND, 1992, p. 174, tradução nossa).

O que chama atenção no extrato acima é a dificuldade sobre o conceito de assunto. O termo *aboutness*, que não possui tradução consensual para o português do Brasil, não deve ser considerado sinônimo de assunto. Situação parecida acontece no espanhol, pois se encontra dificuldade em estabelecer definição para *contenido* e *materia*, quando se refere à palavra assunto em inglês: *subject*.

Quanto ao aspecto idealístico objetivo, o assunto é enfatizado a partir da análise teórica, de um ponto de vista absoluto, independentemente da consciência humana. Nessa concepção, os documentos compartilham ideias que estão expressas em um assunto e “essas ideias existem fora da consciência humana (ou dentro dela como percepções *a priori*) e também são anteriores aos conceitos individuais expressos nos documentos individuais.”⁸⁴ (HJØRLAND, 1992, p. 177, tradução nossa).

Ainda segundo Hjørland (1992), o aspecto idealístico objetivo expressa que o assunto pode ser identificado de maneira independentemente do contexto que será empregado, mesmo a partir de um sistema de classificação.

A fim de compreender o que seja o assunto dos documentos, torna-se necessária algum tipo de análise, pois o objetivo é realizar a representação temática: “[...] a informação é

⁸³ “[...] the intentions of the author, his view of its subject, and what new things he has to relate. This has given rise to the concept of 'aboutness' in library and information science literature, an interest which in my view represents a blind alley, an attempt to escape from the difficulties in the concept of subject [...]” (HJØRLAND, 1992, p. 174).

⁸⁴ “These ideas exist outside the human consciousness (or within it as *a priori* perceptions) and are also prior to the individual concepts expressed in the individual documents.” (HJØRLAND, 1992, p. 177).

procurada em bibliotecas ou bases de dados em que os documentos (transportadores de informação) são registrados por assunto.”⁸⁵ (HJØRLAND, 1992, p. 179, tradução nossa).

Acontece que a construção do conceito de assunto não elimina ou descarta seu caráter intangível e, com isso, a dificuldade em defini-lo. Duas razões, segundo Hjørland (1992), constituem esse quadro: os documentos possuem infinitas propriedades e essas não podem ser estabelecidas de uma única vez. Sendo assim, apesar do registro de informação possibilitar a recuperação temática, o assunto depende da análise de fatores contextuais.

O termo assunto, no contexto das listas de cabeçalhos, foi bastante adotado de uma forma mais prática que teórica, como produto do processo de catalogação de assunto.

Em relação a esse termo, no Dicionário da Língua Portuguesa (2004, p. 166), assunto é definido como “matéria de que se trata; objecto; tema; motivo (Do lat. *assumptu*-, <<assumido>>).” No Dicionário de Latim-Português (FERREIRA, 1987, p. 139) a etimologia *assumptus* refere-se ao *assumo* - “1. Tomar, tomar para si, tomar consigo, associar a. 2. Adaptar, atribuir, arrogar, aplicar.”

No dicionário Aulete (on-line, destaque do dicionário), sua origem também é identificada a partir do latim “*assumptus*, a, um”, mas no sentido em “que se elevou; ELEVADO”. Outra definição: “aquilo sobre o que se conversa, fala ou escreve; TEMA; MATÉRIA” (destaque do dicionário). Indica-se primeiro sua origem latina, vinculando-se as regras da língua portuguesa, pois pelo radical é possível incluir, por exemplo, afixos⁸⁶ para formação de outras palavras. Observam-se também outros conceitos, como “tema” e “matéria”.

Faria e Pericão (2008, p. 111), no Dicionário do Livro, definem assunto como “aquilo que, numa obra literária, constitui o conteúdo de pensamento sobre o qual se exerceu o talento criador do autor. Aquilo sobre o que se aplicou a reflexão do escritor numa obra científica ou didáctica. O tema, matéria ou argumento de uma obra, mencionado ou não no título.” Essa definição de um dicionário especializado sobre o livro, seja ele apresentado em suporte tradicional ou eletrônico, aborda o assunto relacionado ao documento desde a escrita ao conteúdo intelectual, entendido como “tema, matéria”.

⁸⁵ “[...] the information is searched for in libraries or databases in which documents (carriers/conveyors of information) are registered by *subject*.” (HJØRLAND, 1992, p. 179, destaque do autor).

⁸⁶ “Os afixos são elementos que se agregam ao radical para modificar-lhe o significado. Os afixos que se antepõem ao radical chamam-se prefixos; os que a ele se pospõem, sufixos.” (CUNHA, on-line).

Já no dicionário de Biblioteconomia, em idioma espanhol, a definição para assunto, que corresponde à palavra *materia* é: “assunto, tema ou argumento que uma obra literária é composta, etc.”⁸⁷ (BUONOCORE, 1976, p. 298, tradução nossa).

Com exceção dos dois primeiros dicionários, que não são especializados, ou seja, são dicionários de língua, as respectivas definições trazem em comum os termos tema e matéria.

Para Hjørland (1992, p. 172, tradução nossa), “uma ligação muito estreita existe entre o que os assuntos são, e como estamos a conhecê-los.”⁸⁸ De acordo com a abordagem até aqui adotada, é possível observar nitidamente as variações terminológicas para abordá-lo.

Como apontado, a relação entre a ideia de um assunto e os documentos foi sistematizada no trabalho de Cutter pelo catálogo dicionário e, segundo Miksa (2012, p. 15, tradução nossa), constitui-se como “[...] referente primário [...] [e] seu status formal em relação ao domínio classificatório natural do conhecimento humano. Em contraste, a ideia de um assunto como um atributo de um livro ou de qualquer documento tendia a ser de menor importância.”⁸⁹ Portanto, a ideia de assunto, naquele momento, não foi uma proposição teórico-metodológica de elementos de manipulação do texto.

Levando em consideração a palavra assunto em inglês, *subject*, no dicionário de Biblioteconomia e Ciência da Informação, Reitz (2014, on-line, tradução nossa), define:

Qualquer um dos tópicos ou temas de uma obra, explicitamente declarado no texto ou título ou implícito em sua mensagem. [No processo de] catalogação em bibliotecas, é atribuído a um livro ou outro item um ou mais cabeçalhos de assunto como pontos de acesso, para ajudar os usuários a localizar seu conteúdo por assunto. Nos serviços de indexação e resumos, os cabeçalhos designados para representar o conteúdo de um documento são chamados de descritores. Abreviação subj. Veja também: *aboutness* e análise de assunto.

Em um sentido mais geral, qualquer tópico de estudo ou discussão, tema expresso por escrito, ou objeto ou cena retratado na pintura, desenho, fotografia, etc. Na Classificação Decimal de Dewey (CDD), os assuntos são organizados por disciplina. Porque um assunto pode ser estudado em mais de uma disciplina (exemplo: casamento em lei, psicologia, religião, sociologia, etc.), as escolhas feitas na classificação são governadas por regras.⁹⁰

⁸⁷ “Materia. Asunto, tema o argumento de que se compone una obra literaria, científica, etc.” (BUONOCORE, 1976, p. 298).

⁸⁸ “A very close connection exists between what subjects are, and how we are to know them.” (HJØRLAND, 1992, p. 172).

⁸⁹ “[...] primary referent [...] formal status in relationship to the natural classificatory realm of human knowledge. In contrast, the idea of a subject as an attribute of a book or of any document tended to be of much less significance.” (MIKSA, 2012, p. 15).

⁹⁰ Any one of the topics or themes of a work, stated explicitly in the text or title or implicit in its message. In library cataloging, a book or other item is assigned one or more subject headings as access points, to assist

Nesse extrato, tópico e tema são igualmente identificados como assunto. O autor resgata a tradição, segundo Cutter, de como os assuntos dos livros podem ser representados com uso de palavras sistematizadas em cabeçalhos. Apresenta também os descritores adotados nos serviços de indexação e resumos (em geral, sistemas especializados) e faz referência aos sistemas de classificação bibliográfica, como a CDD, referindo-se aos assuntos, agrupados por disciplinas, mas sem especificar o processo que este instrumento contempla. Quanto aos termos relacionados, cita *aboutness* e análise de assunto. A definição é ampla e abarca o que já foi tratado neste trabalho. Portanto, aproxima as vertentes Catalogação de Assunto e Indexação. Se, “para Cutter, a estabilidade dos assuntos depende do processo social em que seu significado é estabilizado em um nome ou designação”⁹¹ (HJØRLAND, 2017, p. 56, tradução nossa), o cenário muda de configuração ao observar conceitos como *aboutness* e análise de assunto.

Hjørland (2001) explica não ser fácil entender o conceito de assunto (*subject*), pois seu contexto histórico imprime abordagem no processamento de livros em catálogos de bibliotecas. O autor pondera ao especificar que o assunto de um documento é o objetivo de um serviço de informação, definindo-o⁹² “[...] como os *potenciais* epistemológicos ou informativos dos documentos. [Afirma ainda que] A melhor análise de assunto é aquela que faz o melhor prognóstico do uso futuro do documento.”⁹³ (HJØRLAND, 2001, p. 776, itálico do autor, tradução nossa). O autor destaca a palavra potencial (*potentials*), que significa possibilidade, algo que pode vir a ser, portanto não real (AULETE, on-line), pois deverá corresponder ao uso do documento.

Desse modo, Hjørland (2001) relaciona outros termos que por vezes são considerados sinônimos, como *aboutness*, tópico (*topic*), tema (*theme*) e conteúdo (*content*). Com exceção do primeiro, os demais corroboram as orientações que seguem. Ao abordar o termo tópico, o

users in locating its content by subject. In abstracting and indexing services, the headings assigned to represent the content of a document are called descriptors. Abbreviated *subj*. See also: *aboutness* and subject analysis. In a more general sense, any topic of study or discussion, theme expressed in writing, or object or scene depicted in painting, drawing, photography, etc. In Dewey Decimal Classification (DDC), subjects are arranged by discipline. Because a subject can be studied in more than one discipline (*example*: marriage in law, psychology, religion, sociology, etc.), the choices made in classification are governed by rules. (REITZ, 2014, on-line).

⁹¹ “For Cutter, the stability of subjects depends on a social process in which their meaning is stabilized in a name or a designation.” (HJØRLAND, 2017, p. 56).

⁹² O autor usa como base artigo que publicou em 1992, também utilizado nessa pesquisa. Consultar seção de referências.

⁹³ “[...] as the epistemological or informative *potentials* of documents. The best subject analysis is the one that makes the best prognosis of the future use of the document.” (HJØRLAND, 2001, p. 776, itálico do autor).

autor usa como explicação ele ser uma parte específica em relação ao que um cientista trabalha em um campo mais amplo e, assim, remete à concepção do específico em relação a uma cobertura ampla. Para explicar o termo tema, faz aproximação com o que é individualizante e caracterizador. Cita-o como um elemento que direciona uma conferência ou um número especial de periódico e, portanto, fornece o aspecto central sobre o que está sendo abordado. Já em relação ao conteúdo, o autor faz aproximação com um texto, indicando que ele tem semelhança com um comentário que pormenoriza um determinado assunto. Sendo assim, aproxima à análise de conteúdo a determinação de assuntos dos documentos. Essas definições básicas podem caracterizar e relacionar os termos, porém são limitadas frente às demandas que poderiam ser colocadas para discussão da noção de assunto.

A conjugação do assunto como parte do registro de informação em catálogos e bases de dados deve reconhecer aspectos conceituais presentes no documento, os quais variam com a experiência de quem realiza a análise, o contexto, o público que fará uso, dentre outros fatores. Desse modo, a defesa dos assuntos serem potenciais epistemológicos ou informativos dos documentos (HJØRLAND, 2001) concorre para uma forma que seja procedimental e significativa, como por exemplo, o assunto gerado a partir da análise.

Embora Hjørland (1992, p. 187, tradução nossa) explore diferentes linhas epistemológicas para estruturar o conceito do assunto como um elemento informativo, pensar a sua descrição envolve “[...] avaliação e atribuição de prioridades para as propriedades de um documento com relação a uma categorização e uma descrição do assunto desse documento.” Se em termos gerais, essa descrição é precedida por alguma forma de análise, ela é “[...] contínua e decisiva para a ‘visibilidade’ do documento em bibliotecas e bases de dados, e assim para seu papel potencial no desenvolvimento futuro do conhecimento.”⁹⁴

Se o assunto abarca, de algum modo, adequação para indicar o que se quer representar e recuperar nos documentos, a análise de assunto, também denominada, conforme Dias e Naves (2013), análise conceitual, análise temática, análise documentária e análise de conteúdo deve almejar implicações significativas para a constituição de uma representação contextualizada. Entretanto, é importante destacar que essa variação terminológica tem reflexo a partir de diferentes vertentes teóricas, mas também se relaciona à imprecisão que

⁹⁴ “[...] evaluation of and assignment of priorities to the properties of a document with regard to a categorisation and a subject description of that document.”

“[...] go on is decisive for the ‘visibility’ of the document in libraries and databases, and thereby for its potential role in future development of knowledge.” (HJØRLAND, 1992, p. 187).

assumem as palavras conceito, tema, assunto, conteúdo e matéria. Além disso, há distinções de significado relacionado aos termos análise de assunto e análise documentária.

Com base na citação acima e a partir do livro de Dias e Naves (2013, p. ix), encontra-se na ‘Apresentação’ um alerta ao leitor para “os problemas de variação terminológica” que segundo os autores é percebida, inclusive, na literatura de língua estrangeira. Mas, na literatura nacional o cenário não é tão diferente. A despeito de os autores fazerem o alerta, no decorrer do livro, percebe-se ausência de uma discussão sobre essas diferenças. Da mesma forma que no inglês, é difícil estabelecer uma única definição para *subject* e *aboutness*, fato similar acontece no espanhol, em relação às palavras *contenido* e *materia*.

Na Catalogação, o termo *epígrafe*, em espanhol, é usado em linhas gerais para definir palavra ou palavras que são empregadas nos cabeçalhos das fichas de assunto dos catálogos dicionários, expressando o que a obra trata (ROVIRA, 1966; LITTON, 1971; BUONOCORE, 1976).

Penna (1945) preferiu usar *encabezamientos de materia* em detrimento de *epígrafes de materia*, da mesma forma que Litton (1971) realizou essa escolha. Buonocore (1976, p. 189) destaca que os *encabezamientos de materia* ou *asuntos* recebem o nome de *epígrafe*, traduzido para o português brasileiro como ‘rubrica’.

O principal responsável em propagar a definição de *epígrafe*, segundo Rovira (1966), foi Jorge Aguayo,⁹⁵ objetivando que a Biblioteconomia formasse seu próprio vocabulário em espanhol. Na visão de Rovira (1966), pareceu aceitável usar *epígrafe* em detrimento de *encabezamiento de materia*, entretanto, a incorporação do termo evidenciou-se na literatura de Cuba, país que Aguayo desenvolveu longo trabalho na *Universidad de La Habana*.

Buonocore (1976, p. 198, tradução nossa), ao definir o termo *epígrafe*, demonstrou reflexão sobre a prática profissional iniciada no século XIX, e explica que,

Embora os termos título e *epígrafe* sejam sinônimos em linguagem natural, a terminologia profissional os diferencia. O título é a palavra ou frase com a qual o autor denomina oficialmente sua obra e como decide estampar, como se sabe, na capa da mesma. A *epígrafe*, por outro lado, é o título atribuído pelo catalogador da obra e para isso ele escolhe a palavra mais específica e expressiva do conteúdo da mesma, usando uma lista previamente preparada - a de Lasso de la Vega, a Sears, por exemplo, - ou, caso contrário, mantendo

⁹⁵ AGUAYO, Jorge. **Manual práctico de clasificación y catalogación de bibliotecas**. La Habana: Jesús Montero, 1943.

um registro dos cabeçalhos usados no catálogo da biblioteca para padronizar critérios e estabelecer relações entre as diferentes *epígrafes*.⁹⁶

Neste panorama e, mesmo com o exemplo de Aguayo que buscou consolidar o termo citado em língua espanhola, o fato não alterou a influência do arcabouço anglo-saxão na Catalogação. Considerando a obra de Cutter antiquada em muitos aspectos, Rovira (1966, p. 7, tradução nossa) expressou que os autores de textos da área “[...] se limitavam a aconselhar a adoção de listas impressas, a ensinar seu uso e a dar algumas indicações práticas e princípios gerais deduzidos das mesmas.”⁹⁷

Em Portugal, constata-se, a partir de Peixoto (1962, p. 26), referência à *epígrafe* como sinônimo de encabeçamento ou cabeçalho de assunto, compondo “[...] a palavra ou palavras que, postas na parte superior das fichas de matérias, nos dão o tema ou assunto concreto [...]”. Ainda para o autor, “a expressão terminológica do assunto” (p. 26, destaque do autor) é a *epígrafe* e, “a expressão numérica” (p. 27, destaque do autor) é a classificação, disposta a partir da notação (também numérica) de um sistema classificatório, como por exemplo, a Classificação Decimal Universal (CDU). Desse modo, o autor trata das modalidades de *epígrafes*, conforme definições a seguir:

- 1) Epígrafes temáticas que designam a principal expressão terminológica do contexto da espécie bibliográfica a ser tratada catalograficamente. Ex.: Educação.
- 2) Epígrafes geográficas, que designam a expressão terminológica do espaço geográfico do contexto. Ex.: Portugal – Educação.
- 3) Epígrafes cronológicas ou históricas, que designam a expressão terminológica da época histórica ou cronológica. Ex.: Renascimento; Século XIII.
- 4) Epígrafes onomásticas ou biográficas, que designam uma pessoa individual ou colectiva tratada no seu aspecto biográfico ou histórico, etc., usando-se sempre o mesmo nome, quer se trate do catálogo de autores ou de assuntos. Ex.: Tomás de Aquino (Santo).
- 5) Epígrafes de título ou didascálicas, que se usam para designar como tema ou assunto uma obra apenas conhecida pelo título. Ex.: Diário de Notícias; Funchal e as lendas marítimas. (PEIXOTO, 1962, p. 27).

⁹⁶ Aunque los vocablos título y epígrafe son sinónimos en el lenguaje vulgar, la terminología profesional los diferencia. El título es la palabra o frase con que el autor denomina oficialmente su obra y va estampado, como se sabe, en la portada de la misma. El epígrafe, en cambio, es el título que le asigna el catalogador de la obra y para ello elige la palabra más específica y expresiva del contenido de la misma, utilizando una lista confeccionada de antemano – la de Lasso de la Vega, la de Sears, por ejemplo, – o, de lo contrario, llevando un registro de los encabezamientos usados en el catálogo de la biblioteca para uniformar criterio y establecer relaciones entre los distintos epígrafes. (BUONOCORE, 1976, p. 198).

⁹⁷ “[...] se limitaban a aconsejar la adopción de las listas impresas, a enseñar su uso y a dar unas cuantas indicaciones prácticas y principios generales deducidos de las mismas.” (ROVIRA, 1966, p. 7).

Para Peixoto (1962), a disposição do catálogo alfabético de matérias (ou seja, de assuntos) é oriunda da tradição anglo-saxã do catálogo dicionário, que é a fusão do catálogo de autores, títulos e de matérias. Sobre o catálogo dicionário, Peixoto (1962, p. 26) explica que bibliotecários europeus mostraram-se “[...] relutantes em o aceitar. [...] nota-se uma certa tendência para o abandonar, [...] em virtude de ele crescer desmedidamente. [...] outros afirmam que o catálogo-dicionário, ao juntar o de autores e o de matérias, confunde duas fases da investigação, que são bem distintas.” De fato, a tradição europeia separou a elaboração do catálogo de assunto, por meio da construção do catálogo sistemático.

Litton (1971) explica que pelo exame do livro é possível indicar seu assunto (*materia*) e, assim, elegê-lo em uma lista de cabeçalho, o que compatibilizaria, segundo ele, a natureza da publicação. Contudo, também explicitou que pelo cabeçalho é possível descrever “[...] o conteúdo dos livros registrados no catálogo da biblioteca apenas em grau relativo. De todo modo, são mais uniformes e corretos para a interpretação do conteúdo das obras quando a instituição adota e adere estritamente a uma lista autorizada.”⁹⁸ (LITTON, 1971, p. 81, tradução nossa). Nessa condição, o assunto do documento seria pré-concebido a partir da lista de cabeçalho mantendo vinculação com um plano explícito de palavras do documento. Essa relação, permeada por “[...] variações praticamente ilimitadas dos temas abordados nos livros, [e] a análise destes através dos cabeçalhos e das tabelas de classificação não é tão suscetível à formulação de um corpo de regras [...]”⁹⁹ (LITTON, 1971, p. 77, tradução nossa).

No âmbito das fichas, a representação temática dos documentos comportou a identificação do que se denominou *contents* (conteúdo). Taylor (1948) abordou que essa compreensão envolveu aspectos de uma obra possuir vários volumes e, assim, a identificação requerer uma nota de conteúdo (*contents note*). Outra situação para o reconhecimento desse aspecto, segundo a mesma autora, seria o de uma obra apresentar capítulos escritos por diferentes autores. Ainda segundo Taylor (1948, p. 91, tradução nossa, destaque da autora), na ficha catalográfica:

A nota de conteúdo é iniciada meio centímetro abaixo da última linha da entrada. A palavra *conteúdo* começa na margem interna, é indicada em itálico [...] e seguida por dois pontos. Depois vêm os detalhes do conteúdo,

⁹⁸ “[...] el contenido de los libros registrados en el catálogo de la biblioteca sólo en un grado relativo. De todos modos, son más uniformes y correctos para la interpretación del contenido de las obras cuando la institución adopta y se adhiere estrictamente a una lista autorizada.” (LITTON, 1971, p. 81).

⁹⁹ “[...] variaciones prácticamente ilimitadas en los temas de que tratan los libros, el análisis de éstos a través de los encabezamientos y la tabla de clasificación no es tan susceptible de la formulación de un cuerpo de reglas [...]” (LITTON, 1971, p. 77).

sendo estes levados de volta para a margem externa e cada item separado por um longo traço. Um centímetro para cima a partir da parte inferior da ficha deve ser deixado em branco, exceto para as palavras (continua na próxima ficha), pois a parte inferior é difícil de ser vista quando se está em pé em uma gaveta do catálogo. Ao continuar numa segunda ficha, numere esta (2), repita o cabeçalho e o título brevemente, seguidos por um espaço de dois centímetros e *Conteúdo (continuação)*. Deixe meio centímetro e comece a listar o restante do conteúdo sem qualquer recuo de parágrafo na margem interna.¹⁰⁰

A identificação do *contents* (conteúdo) na ficha catalográfica alcançou relativa estabilidade através da nota de conteúdo, e não propriamente no tange à representação temática. Por sua vez, Buonocore (1976, p. 143, tradução nossa), com base em Jorge Aguayo, afirma que a descrição do *contenido* “[...] é o índice colocado nos livros seguindo a mesma ordem dos capítulos.”¹⁰¹ Neste caso, ele está se referindo ao recurso que chamamos mais frequentemente de sumário.

O conteúdo, através da nota concebida logicamente por palavras extraídas do documento, apresentou uma proposta de padronização envolvendo a pontuação e o lugar de apresentação na ficha, porém não houve embasamento metodológico para essa construção.

Ainda no que tange às fichas, a entrada secundária, explica Litton (1971), rege o desdobramento, ou seja, a elaboração de fichas secundárias, a partir da ficha principal ou única (*tarjeta única*). Nessa tradição, ao identificar o assunto (*materia*), quem catalogou deve encontrá-lo na lista de cabeçalho e, assim, indicá-lo no lugar da entrada secundária na ficha principal. Após essa escolha e, seguindo o padrão da ficha secundária de assunto,

O cabeçalho selecionado é transcrito no espaço acima da linha do autor, a partir da segunda margem. Para destacá-lo, uma dessas duas regras é geralmente seguida: a) as letras são todas escritas em maiúsculas e em preto; e b) todas as letras são transcritas em maiúsculas e em vermelho.

Quando um cabeçalho não se encaixa na linha a ele destinada, terá que elevá-lo mais para a borda da ficha, de modo que a segunda parte possa ocupar a 2ª linha a partir da terceira margem.

[...]

¹⁰⁰ The contents note is begun half a centimetre below the last line of the entry. The word *contents* starts at the inner margin, is put in italics [...] and is followed by a colon. Then come the details of content, these being taken back to the outer margin and each item separated by a long dash. Quite one centimetre upwards from the bottom of the card should be left blank, except for the words (continued on next card), as the lower part is difficult to see when standing upright in a catalogue drawer. When continuing on a second card, number this (2), repeat heading and brief title, followed by a space of two centimetres and *Contents (continued)*. Leave the half a centimetre and then begin to list the remainder of contents without any paragraph indention on the inner margin. (TAYLOR, 1948, p. 91, itálico da autora).

¹⁰¹ “[...] es el índice puesto en los libros siguiendo el mismo orden de los capítulos.” (BUONOCORE, 1976, p. 143).

Quando um cabeçalho é composto por várias partes, elas podem ser separadas umas das outras por meio de: a) um ponto e dois espaços; e b) uma linha com dois traços (nunca um).¹⁰² (LITTON, 1971, p. 138, tradução nossa).

De modo análogo, conforme Penna (1945, p. 85, tradução nossa) expressa, “[...] as fichas de assunto são as que representam os livros de acordo com seu conteúdo.”¹⁰³ Para o autor, essa aproximação parte das *fichas de materia* (fichas de assunto) de um catálogo dicionário se relacionarem com o modo de ser do catálogo sistemático através dos seus símbolos classificadores (*símbolos clasificadores*). Ainda segundo o mesmo autor, antes de decidir esse símbolo e o cabeçalho, quem cataloga deve buscar o *tema* (tema) tratado na obra através da leitura do seu prefácio, sumário, dentre outros elementos e, se necessário, recorrer à consulta de terceiros, ou a fontes adequadas como bibliografias e outros catálogos. Após determinar o tema da obra, o autor explica que as tarefas de indicação do símbolo classificador e do cabeçalho não são idênticas, pois no primeiro caso “[...] será necessário ter um raciocínio que nos leve exatamente ao lugar do sistema que, pela disposição lógica das ciências, corresponda ao livro que se busque localizar [...]”¹⁰⁴ Já para o segundo, “[...] devemos procurar a palavra que expresse de forma mais determinante o conteúdo do livro [...]” e dessa forma ambos os processos “[...] se diferem de forma tão fundamental.”¹⁰⁵ (PENNA, 1945, p. 85, tradução nossa).

A disposição das fichas nesse tipo de catálogo segue a ordem alfabética agrupando autores, títulos e assuntos, mas, os cabeçalhos de assunto (*encabezamientos de materia*) assumem relevância “[...] pois são os que agrupam em determinadas partes do catálogo todas

¹⁰² El encabezamiento escogido se transcribe en el espacio que va encima de la línea de autor, partiendo del segundo margen. Para hacerlo resaltar, se suele seguir una de estas dos normas: a) las letras se escriben todas en mayúsculas y en negro; y b) todas las letras se transcriben en mayúsculas y en rojo.

Cuando un encabezamiento no cabe en la línea que tiene destinada, habrá que subirlo más hacia el borde de la tarjeta, para que la segunda parte pueda ocupar la 2ª línea desde el tercer margen.

[...]

Cuando un encabezamiento consta de varias partes, éstas se pueden separar una de otra por medio de: a) un punto y dos espacios; y b) una raya e dos guiones (nunca uno solo). (LITTON, 1971, p. 138).

¹⁰³ “[...] las fichas de materia son las que representan los libros de acuerdo con su contenido.” (PENNA, 1945, p. 85).

¹⁰⁴ “[...] será necesario hacer un razonamiento que nos conduzca exactamente al lugar del sistema que, por la lógica disposición de las ciencias, corresponda al libro que se quiere ubicar [...]” (PENNA, 1945, p. 85).

¹⁰⁵ “[...] debemos buscar la palabra que exprese de manera más determinante el contenido del libro [...]” “[...] difieren en forma tan fundamental.” (PENNA, 1945, p. 85).

as fichas de obras que tratam do mesmo tema.”¹⁰⁶ (PENNA, 1945, p. 85, tradução nossa). Ainda segundo esse autor, o conteúdo seria a substância do texto formando as partes do livro e, através do seu exame, se determinaria o tema. Portanto, a partir do tema, o assunto seria indicado pelo cabeçalho e, representado na ficha, congregaria o primeiro vinculando-o com a descrição controlada do instrumento. Com esse aspecto, depreende-se que o assunto seria o resultado do tema, pois o primeiro é esquematizado no cabeçalho e, já o segundo, resultaria da análise.

A despeito da relação apontada por Penna (1945), a atribuição de um símbolo classificador ao documento forneceria uma localização no acervo, de acordo com o mapeamento das ciências dado no sistema bibliográfico e, com o cabeçalho, o uso de uma palavra específica resultaria no assunto. Em ambos os casos, a representação continuou privilegiando aspectos genéricos dos instrumentos e se baseou no documento como um todo.

Sharp (1964), ao abordar as tarefas de classificação e catalogação de assunto, explica que a entrada secundária de assunto nas fichas fornece um caminho que amplia o processo de análise por conta das referências cruzadas que podem ser construídas. Critica apontando que houve frequente negligenciamento na relação dessas construções e que “muito já foi dito sobre os méritos de disposição dos assuntos nas respectivas formas dos catálogos dicionário e classificado [...]”¹⁰⁷ (SHARP, 1964, p. 334, tradução nossa). Mas, o autor sob o viés anglo-saxão, também se restringiu metodologicamente em torno da prática tão difundida e igualmente já apontada nos livros de Catalogação (PENNA, 1945; TAYLOR, 1948; MANN, 1962; LITTON, 1971).

A sistematização para o assunto fornecida por Cutter, na forma de um cabeçalho, foi influente na construção da vertente da Indexação, mas tomou outra direção a partir da ideia de índice e das metodologias de indexação.

Nessa diferença entre vertentes, o processo de indexação é comum, mas a ideia de uma entrada secundária, não é. Nos catálogos em fichas ou on-line, há entrada secundária de assunto, mas nas bases de dados científicas elas não existem como tal, pois se trabalha com pontos de acesso de assunto. Com isso, o ponto de acesso tido como principal empregado por um nome, termo ou signo são utilizados para encontrar uma entidade ou um conjunto de entidades em uma ficha catalográfica e seus desdobramentos são possibilitados pelas entradas

¹⁰⁶ “[...] pues son los que agrupan en determinadas partes del catálogo todas las papeletas de obras que tratan del mismo tema.” (PENNA, 1945, p. 85).

¹⁰⁷ “a good deal has been said about the collocation of subjects and the respective merits of the dictionary and the classified forms of catalogue [...]” (SHARP, 1964, p. 334).

secundárias. Porém, com as bases de dados vários pontos podem ser disponibilizados para a recuperação e a palavra principal perdeu a função original.

A exposição efetuada permite verificar que o desdobramento da vertente anglo-saxã é constituída por fragmentos, mas com aspectos comuns, como o da entrada por assunto. Ademais, o assunto é um atributo relevante, mas prevalece a dispersão terminológica em torno da sua noção. Do ponto de vista da organização, instrumentos condicionam o assunto, como a lista de cabeçalhos, que centram na morfologia e na sintaxe, enquanto o tesauro na semântica. Do ponto de vista da representação, o assunto é construído e não extraído do documento.

6 TERCEIRO MOMENTO DE CONSTRUÇÃO DA NOÇÃO DE ASSUNTO: A ABORDAGEM EUROPEIA NO SÉCULO XX

De modo a tornar o conhecimento registrado em documentos disponível e acessível a quem viesse necessitar deles, na Bélgica, esforços foram desenvolvidos para a elaboração de metodologias que efetivassem esse trabalho.

A articulação de ideias que permitiram a formação de um novo cenário ocasionou, por exemplo, a constituição de cooperação internacional e a elaboração de um instrumento documentário que desencadearam o fomento da Documentação europeia. Dessa forma, e diferente do contexto anglo-saxão, houve outro modo de dispor as informações, em especial, para uso da documentação científica.

Nesse sentido, as discussões ocasionadas na vertente europeia favoreceram outros cenários de trabalho documentário que se opunha à dispersão da produção intelectual. Para tanto, de um lado, as ideias propostas por Paul Otlet (1868-1944) e, de outro, o de Jean-Claude Gardin (1925-2013) alcançou, sobretudo, a Espanha e o Brasil pela necessidade de se desenvolver meios e linguagens para o acesso de documentos de forma integrada.

6.1 A CONTRIBUIÇÃO DE PAUL OTLET

O desenvolvimento pela perspectiva da Documentação teve início do final do século XIX para o XX, e, decorreu especialmente, da concepção de Paul Otlet no âmbito europeu. Otlet, em seu *Traité de Documentation: le livre sur le livre: théorie et pratique*, publicado em 1934, propôs um conjunto de fundamentos, dentre os quais, o que relaciona o livro/documento e os modos de abordagem e tratamento dos seus atributos, como por exemplo, sua análise e descrição. O *Traité* é resultado da sistematização das reflexões e do trabalho de Otlet, em parceria com Henri La Fontaine, realizado no *Institut International de Bibliographie* (IIB). Mas, antes do *Traité*, Otlet publicou outros trabalhos relevantes para a Documentação, como por exemplo, os artigos: *Les sciences bibliographiques et la documentation* (1903) e *L'organisation rationnelle de l'information et de la documentation en matière économique* (1905). Para essa investigação fez-se uso da tradução brasileira do Tratado de Documentação (2018).

Advogado e com 23 anos, Otlet ingressou na *Société des Études Sociales et Politiques*, em Bruxelas (Bélgica), e lá conheceu La Fontaine, também advogado e diretor do programa

bibliográfico da *Société*. O segundo acumulou outra função, de carreira política, exercendo o cargo de senador belga e, junto do primeiro, desenvolveram profícuos projetos. Com suas respectivas experiências bibliográficas, estabeleceram metodologias e instrumentos para a área da Documentação, como por exemplo, a elaboração da Classificação Decimal Universal (CDU).

Foram muitas as preocupações teóricas de Otlet, mas, as que mais interessam a essa investigação centram-se no encontro de alternativas para organizar os documentos produzidos. O aumento significativo da produção desses documentos envolveu cientistas, pesquisadores, bibliotecários e bibliógrafos e, conhecido como Movimento Bibliográfico, marcou a história da Documentação por subsidiar seu caráter científico, através de metodologias e técnicas, com o aporte de outras áreas, como a Sociologia, Psicologia, Lógica, Linguística e Estatística. A visão ampla de Otlet revolucionou a forma de trabalhar com a informação naquele tempo resultando em fundamentos que permanecem presentes (SANTOS, 2007).

Em relação ao IIB, seu objetivo consistiu em aperfeiçoar métodos documentários para a distribuição das informações coletadas através do *Répertoire Bibliographique Universel* (RBU). Quase 18 milhões de registros foram elaborados para o RBU entre 1895 e o final da década de 1930 e, o desenvolvimento desse trabalho exigiu a elaboração de padrões, tanto em termos de registros bibliográficos quanto da classificação (SANTOS, 2007).

Otlet (2018, p. 630-631) fundamentou que:

[...] O Repertório Bibliográfico Universal (RBU) ou *Bibliographia Universalis* foi projetado como um catálogo no qual deve estar registrada integralmente e ordenada toda a produção intelectual, [*]em forma de inventário, e que a torna largamente acessível a todos e para todos os propósitos[*].¹⁰⁸ Ele deve abarcar todas as obras, incluir os escritos de todos os tempos, produzidos em todos os países, em todas as formas, em quaisquer materiais. Dele farão parte livros, folhetos, folhas avulsas, revistas, jornais, publicações oficiais e obras didáticas. Estarão excluídas as partituras musicais, as estampas, a numismática, a epigrafia e os materiais de arquivos. Não nos limitaremos a mencionar as obras como unidades. Quando formarem coleções, ali estarão representadas com todos os seus elementos componentes, como acontece com as obras poligráficas. É o caso[,] sobretudo dos periódicos. O repertório conterá o resultado de sua análise, isto é, os registros dos artigos que compõem cada um de seus fascículos.

¹⁰⁸ Confrontando o trecho destacado a partir do asterico [*], com a versão original em francês, observa-se que para a tradução caberia: “sob uma forma que constitui o inventário e que o torna amplamente acessível a todos e para todos os propósitos.”

A bibliografia universal consistiu em repertoriar as formas documentais de produção científica, como livros, revistas, publicações oficiais, entre outros, e foi a maneira concreta de organizar a produção intelectual em um único lugar. Esse projeto de cunho universalista decorreu para o uso com independência da finalidade dada, se científica ou artística, por exemplo.

Otlet considerou como unidade elementos do texto, o conteúdo e o suporte físico do documento, de forma integrada, não isolada, mas interagindo com a coleção. Os documentos, a partir da diversidade de objetos de texto, formariam múltiplos. Assim, “no que concerne à documentação, a unidade será o livro, seus múltiplos serão os conjuntos formados pelo livro, como as coleções (bibliotecas), e seus submúltiplos serão as divisões, como suas partes (capítulos, etc).” (OTLET, 2018, p. 63). Nesse sentido, compreende-se a natureza dinâmica de um documento e a possibilidade de interação de maneira independente da sua forma, se literária, científica ou artística.

A análise seria a maneira de alcançar a representação de um documento e, nesse sentido, o reconhecimento do seu aspecto intelectual e físico compõe o que Otlet denominou de Princípio da Monografia, ou, Princípio Monográfico. Além desse, outros dois princípios foram elaborados por ele, conforme segue no extrato a seguir:

- a) *Princípio da monografia*. Cada elemento intelectual de um livro é (após ter sido seccionado do conjunto do texto) incorporado em um elemento material correspondente.
- b) *Princípio da continuidade e da pluralidade de elaboração*. Lembrando que um livro é elaborado intelectualmente por um ou vários colaboradores e, uma uma [sic] vez terminada a redação, será encerrado, as fichas permitem que com ele trabalhe um número ilimitado de pessoas. Jamais se deve considerá-lo como acabado.
- c) *Princípio da multiplicação de dados*. As fichas são multiplicadas, a fim de que os diferentes dados figurem em diferentes critérios de ordenação (por exemplo, critérios ideológicos, geográficos, cronológicos, etc.). (OTLET, 2018, p. 602, itálico do autor).

No Princípio da Monografia, fica explícito o elemento do conteúdo do documento, que Santos (2007) denomina elemento intelectual e, com o elemento material, ou seja, o suporte físico, integram-se. A forma de alcançar o que fosse informativo no documento pelo elemento intelectual e físico, mesmo que retirando dele palavras significativas, não elimina a necessidade da análise, pois ela é o meio de conduzir à representação no registro do repertório. De acordo com Santos (2007), os documentos abordam diferentes aspectos e, para economizar o tempo de quem fará uso deles, o Princípio da Continuidade e da Pluralidade

envolvem a redação de fichas analíticas com o resultado da análise de elementos do livro (documento), como autor, título, dentre outros, que compõem o registro. Já em relação ao último Princípio, o da Multiplicação dos Dados, a duplicação das fichas com o uso de um instrumento para organização, como a CDU, visava o acesso à informação por vários pontos.

De acordo com Tálamo (*et al*, 2002), o Princípio da Monografia opera com a decomposição do texto em unidades intelectuais que podem ser representadas em fichas e, assim, confluir com as unidades físicas, que são a parte material e manipulável do documento.

A aplicação dos Princípios não se faz de maneira isolada, já que, para a estruturação dos três há “[...] uma estrutura lógica de organização da informação representada pela CDU, instrumento que serviria tanto para armazenar, quanto para recuperar informação” (SANTOS, 2007, p. 56), que permite a manipulação de todo repertório.

Em relação à organização do RBU, compunha-se de fichas em formato universal, ou seja, fichas padronizadas, divididas em duas partes: uma por autor e outra por assunto (OTLET, 2018, p. 632). Estes aspectos exigiam a adoção de mecanismos que orientassem as formas de organização para acesso e, assim,

[...] toda a produção bibliográfica internacional deveria ser descrita, de acordo com o projeto, em fichas de 7,5 x 12,5.¹⁰⁹ [...] Rapidamente Otlet se dá conta da importância da adoção de padrões que deveriam nortear o trabalho descritivo, uma vez que desde o início ele imaginou o RBU como o resultado de um intenso intercâmbio internacional, pelo qual cada Biblioteca Nacional elaboraria as fichas descritivas da produção bibliográfica nacional e cada Sociedade Científica organizaria a descrição dos artigos de periódicos especializados de sua área. Estas fichas deveriam ser enviadas ao IIB. Era o início da catalogação cooperativa, [os esforços de catalogação cooperativa entre bibliotecas na Europa são anteriores] naquela época ainda não denominada desta maneira, mas sustentada pela idéia [sic] do intercâmbio de informações. (TÁLAMO *et al*, 2002, p. 2).

Para o objetivo de organização do conhecimento acumulado pela humanidade, fez-se necessário a padronização dos registros dos documentos pela ficha catalográfica. Nesse sentido, atributos foram escolhidos para pautar a sistematização do RBU. Por sua vez, esses padrões seguiriam critérios revestidos de preceitos metodológicos pelo estudo do livro/documento e, não a forma contrária, ou seja, do repertório para o documento, como por exemplo, a adaptação de um tipo de catálogo, mas não foi esse o caso. Mesmo que o resultado fosse o de alcançar o máximo de registros, compreendendo um conjunto de elementos, levou-se em consideração o que constituiu a representação, ou seja, a evidência do livro

¹⁰⁹ De acordo com as autoras, a ficha constituía o padrão da *Library of Congress* e foi adotada para o RBU.

(documento). Na ficha, foram apresentados atributos de identificação do documento que facilitassem a composição e organização das informações, o que possibilitou o acesso e a circulação dessas informações. Nesse contexto, o objeto reflete o processo que o gerou. A ideia de análise e síntese é colocada como fundamental para chegar a uma representação do documento e uma configuração entre unidade física e intelectual, operadas pela representação nas fichas, é proporcionada. Ademais, a identificação da unidade intelectual possibilitaria a construção do conteúdo do documento, portanto do aspecto temático, do assunto.

Dessa forma, o propósito documentário principal consistiu em servir de apoio ao trabalho científico no tocante à seleção de “cada elemento intelectual de um livro” e sua manipulação “em um elemento material correspondente” (Princípio da Monografia) (OTLET, 2018, p. 602).

Santos (2007) identifica o Princípio da Monografia como uma importante contribuição de Otlet para a Documentação, que influenciou o que depois passou a denominar-se Ciência da Informação. A autora ressalta que a manipulação física e intelectual acarreta a construção de novos documentos que contêm informação nova. Com isso, um duplo movimento é gerado: desmaterialização e virtualização para a transformação dessa nova informação, entendida como o resultado potencial do documento de passar pelo processo de análise por quem o manipula.

No *Traité*, Otlet sistematizou o estudo do livro e do documento, dentre outras formas de transmissão da informação, como o telégrafo, telefone, rádio, cinema, dentre outros, os quais apontou como sendo substitutos do livro. Mas, para estudá-lo, recorreu à Bibliologia “[...] como ponto de partida de suas reflexões, aproveitando seu objeto de estudo e metodologia.” (SANTOS, 2007, p. 57). Otlet usa livro e documento enquanto expressão e vai construindo o conceito de documento, falando em livro, depois mencionando documento. Nesse sentido, ele não tratou de duas tipologias documentais distintas ao usar essa expressão.

Apesar de não ter sido o responsável em cunhar o termo bibliologia, Otlet compreendeu o livro em toda sua complexidade como “[...] a arte de escrever, de publicar e difundir os dados da ciência.” Ademais, explica que “[...] os conhecimentos relativos ao livro, à informação e à documentação permaneceram durante muito tempo no estado onde se encontrava a biologia há um século.” (OTLET, 2018, p. 11). Ele enfatiza a necessidade de esforços para a constituição de uma “[...] ciência autônoma [de] todos os conhecimentos teóricos e práticos relacionados ao livro: a bibliologia. Esta ciência destina-se a retirar do empirismo as aplicações e as realizações.” (OTLET, 2018, p. 11).

Para Santos (2007, p. 57), a dificuldade terminológica de compreensão do cenário pode ter relação “[...] com a problemática [da] definição dos objetos de estudo: o livro, o documento e os substitutos do livro.” Para essa autora, os substitutos do livro, com base em Otlet, são diferentes “[...] meios que servem para informar e comunicar algo e que não tenham a escrita como principal meio de expressão. Dessa forma, entram nessa categoria os objetos de museus, maquetes, o telégrafo, o telefone, o rádio, a televisão, o cinema, os discos entre outros.”

No *Diccionario de Bibliología*, Martínez de Sousa (1989, p. 238, tradução nossa), a partir de autores como Briet, Amat i Noguera, Lasso de La Vega, López Yepes, entre outros, define: “documentação. Ação e efeito de documentar ou documentar-se. Conjunto de documentos que servem para este fim.”¹¹⁰ E, a partir de Peignot, Finó, Estivals, dentre outros, o mesmo autor define bibliologia como “ciência que se ocupa do livro em seus aspectos internos e externos, materiais e imateriais, históricos, terminológicos e técnicos.”¹¹¹ (MARTÍNEZ DE SOUSA, 1989, p. 73, tradução nossa).

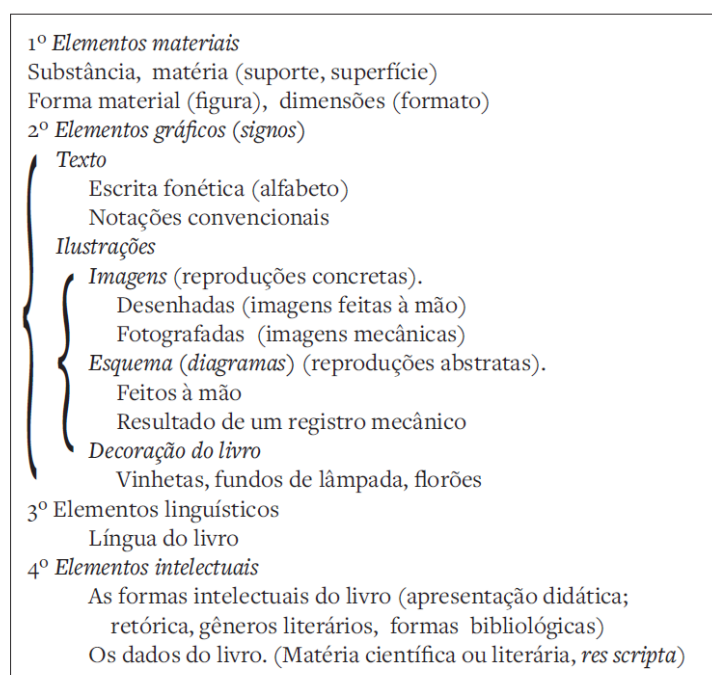
Em seu *Traité*, Otlet (2018, p. 11) usa “Bibliologia ou Documentologia” como “ciências do livro e da documentação”. A primeira, como a descrição dos documentos, mas que se relaciona com a ciência do próprio documento, ou seja, da documentação. Para ele, “a documentação no sentido lato do termo abrange o livro, isto é, meios que servem para representar ou reproduzir determinado pensamento, independentemente da forma como se apresente.” Assim, livro e documento resultam do trabalho intelectual capaz de produzir efeitos em quem utiliza; o primeiro termo é recuperado por Otlet, segundo Santos (2007), para revitalizar a “ciência do livro”. Esse aproveitamento como ponto de partida ampliou-se em seu objeto de estudo e metodologia.

A fundamentação do conhecimento sobre o livro e o documento, de acordo com Otlet, demonstra a sua organicidade. Ele considerou quatro elementos principais na composição do livro – materiais, gráficos, linguísticos e intelectuais – que conforme esquematização na figura 6, existem de forma integrada. Embora os elementos sejam apresentados nesta ordem, não foi atribuída prioridade entre eles, mas há o relacionamento, sobretudo por esses elementos vincularem-se às unidades física e de conteúdo que constituem o livro (e o documento).

¹¹⁰ “documentación. Acción y efecto de documentar o documentarse. Conjunto de documentos que sirven para este fin.” (MARTÍNEZ DE SOUSA, 1989, p. 238).

¹¹¹ “Ciencia que se ocupa del libro en sus aspectos internos y externos, materiales e inmateriales, históricos, terminológicos y técnicos.” (MARTÍNEZ DE SOUSA, 1989, p. 73).

Figura 6 – Elementos que compõem o livro e o documento



Fonte: Otlet (2018, p. 64).

A partir desses elementos, segundo Otlet, decorre a materialização do pensamento armazenado em signos e vinculado à matéria (suporte), como também ao aspecto linguístico e intelectual. De forma geral, ele explica que um livro é “[...] um suporte, feito de determinado material e com determinada dimensão, eventualmente resultado de determinada dobragem ou de montagem em rolo, onde se colocam os signos representativos de certos dados intelectuais.” (OTLET, 2018, p. 59).

O autor ainda especifica de maneira aprofundada no *Traité* cada um dos elementos. Orientando-se pela figura 6, e pela explicação acima sobre a composição do livro, nota-se que o cunho material e intelectual configuram relações, que se manifestam no Princípio da Monografia. Pode-se dizer que, a partir da unidade física, observa-se a unidade intelectual e, assim, alguns elementos desempenham influência de forma concomitante em ambas as unidades, como por exemplo, os elementos gráficos, que acompanham a unidade física, pois o estabelecimento de ilustrações vincula-se à análise da matéria, mas também a forma intelectual, ou seja, o conteúdo representativo das imagens, como nas Artes, Arquitetura, Geografia, Geologia, Design, Fotografia, dentre outras áreas. Já o elemento linguístico acompanha a unidade textual. Pode-se considerar que houve permissão de análise dos elementos do documento, segundo Otlet. Ao contrário, a partir desses elementos, um modelo poderia ter sido proposto ou analisado.

Nesse sentido, segundo Otlet, prevalece uma relação integrada entre o aspecto material e o aspecto intelectual do livro em dois níveis - a forma como os conteúdos se organizam no texto e os conteúdos propriamente ditos - e o que ele considerou ser essencial como forma de organizar a informação, a partir do que, busca-se, nesta pesquisa, pela compreensão do assunto.

Retomando os elementos que compõem o livro, Otlet (2018, p. 139) explica que os elementos intelectuais são definidos como “[...] as ideias (concepções, sentimentos, atividades, imaginações), são as formas como se exprimem as ideias (exposições científicas e didáticas, de um lado, e, de outro, exposições literárias e artísticas).” O autor relaciona as ideias (pensamento) e como essas são expressas na criação de obras “ricas de substância” e “impecáveis na forma” que não se perdem, pois estão registrados nos documentos. Afinal, “[...] o pensamento é tão sutil e tão fugaz que é preciso saber fixá-lo.” (OTLET, 2018, p. 139).

A expressão das ideias da obra coloca questões intelectuais complexas, pois o livro, nesse contexto, “[...] torna-se o meio de elaboração do pensamento humano, a concretização desse pensamento em seus graus mais elevados.” (OTLET, 2018, p. 140). Dito de outro modo, os livros são resultado de uma realidade do autor que transforma suas ideias em texto. Dessa forma, os documentos são consequência da construção pessoal de cada autor, de acordo com seu percurso acadêmico, científico, artístico, literário, dentre outros fatores.

Para analisar a estrutura de um documento, Ortega (2008) trabalha com três categorias: unidade física ou objeto físico informacional (parte material), unidades informacionais ou textuais (parte textual) e conteúdos (informação propriamente registrada nas unidades). Essas podem ser aproximadas dos elementos que compõem o livro, mas segundo a autora, a categoria conteúdo envolve uma noção “[...] extremamente frágil e imprecisa, já que pode não ser determinado de forma apriorística e descontextualizada.” (ORTEGA, 2008, p. 10).

Com base nas três categorias, a mesma autora concordando com Tálamo (*et al*, 2002), identifica que Otlet, a partir do Princípio da Monografia, “[...] implica a decomposição do texto em ‘unidades intelectuais’ [...]” (ORTEGA, 2008, p. 10) que transcritas nas fichas propuseram “[...] a coincidência entre as ‘unidades físicas’ e as ‘unidades intelectuais’ do documento.” (ORTEGA, 2008, p. 10).

Compreende-se que no contexto da natureza documentária, entre a unidade física e intelectual, há manifestação do conteúdo, pois o livro (e o documento) “[...] são o objeto da maior divisão do trabalho e cooperação que existam [...] [tanto no que] tange a sua produção e ao trabalho posterior que eles desencadeiam.” (OTLET, 2018, p. 385). Para a presente investigação, a segunda perspectiva é de interesse, pois ela “[...] concentra-se em identificar

uma série de operações (atos, atividades) que, agrupadas segundo seu objeto, constituem de alguma forma as ‘funções’ da documentação.” (OTLET, 2018, p. 385).

Por natureza documentária, entende-se os meios materiais e conceituais que o documento potencializa, a partir da análise e representação, que permite uma estruturação na forma de registros para posterior acesso, portanto de uma nova informação. As operações são o reconhecimento das funções da documentação, que asseguradas pela formalização, projetam a produção científica, artística, literária, dentre outras formas intelectuais na sociedade. É pela organização que se tem acesso à informação. Não obstante, “[...] diferentemente das ciências naturais, o objeto de estudo da documentação é [...] um produto da vontade humana; a especulação aqui se traduziria na imaginação e na invenção [...]” (SANTOS, 2007, p. 60), que tem por consequência a fundamentação de princípios que sustentam teórica e metodologicamente a documentação.

Em relação ao aspecto do trabalho suscitado pela produção do livro, na perspectiva de Otlet (2018, p. 397), envolveria a “Escolha do tema a ser tratado. Assunto” (no original, *Choix du sujet traite. Matière*). Prossegue explicando que o tema pode ser livre ou encomendado e que os escritores possuem uma infinidade de assuntos possíveis de serem abordados. Considerando as ciências, o autor aborda que os “temas a serem tratados” condizem com a forma com que os problemas de cada área apresentam-se aos pesquisadores. Nesse sentido, o assunto constitui o aspecto intelectual registrado na escrita.

A palavra assunto ainda é operada por Otlet para qualificar o tipo de bibliografia e uma das formas de repertoriar o catálogo. No primeiro caso, ele cita alguns exemplos, em “bibliografia especializada por assuntos” que vincula “[...] temáticas de todas as ciências, de quase todos os assuntos especializados” (OTLET, 2018, p. 458), e que, portanto, são específicas, como por exemplo, o *Manuel de bibliographie historique* de Langlois, o *Manuel bibliographique de la littérature française moderne* (1500–1900). Em seguida, aborda serem quatro os objetivos de um catálogo: primeiro o de constituir “um inventário do conteúdo de uma coleção”; segundo, compor a “descrição das obras, que delas destaca certas características essenciais”; terceiro, representar de forma concisa as obras para facilitar “qualquer tipo de intervenção, que assim pode ser praticada ou planejada não sobre as ‘coisas significadas’, mas sobre ‘signos’ ”; e, quarto, ser “um dispositivo especialmente indicado para recuperar facilmente os livros procurados, mesmo quando só se conhece uma de suas características, como, por exemplo, o nome do autor ou o assunto tratado.” (OTLET, 2018, p. 470).

Chama atenção, na exposição dos objetivos do catálogo, o aspecto da descrição da obra (segundo objetivo), e nelas facilitar “as coisas significadas” (terceiro objetivo). Sobre a forma de repertoriar o catálogo, uma das possibilidades que Otlet indica é a sistematização por assunto, ação que “[...] tem por objetivo dar a conhecer as obras existentes ou possuídas pela biblioteca sobre uma matéria determinada: ciência, técnica, questões sociais, personagens, época histórica, etc.” (OTLET, 2018, p. 471). A sistematização por assunto atrela-se à classificação, ou seja, ao arranjo, à forma de estruturação do catálogo sistemático. Cabe ressaltar que no *Traité*, portanto, no texto original em francês, Otlet faz uso do termo *classification* para identificar o esquema de classificação e, *classement*, para a atividade de classificação. Destaca-se que na tradução brasileira, essa diferenciação não ficou nítida, mas entende-se que a atividade de classificação envolve tanto os aspectos de ordenação da coleção quanto do sistema de recuperação, como o catálogo.

Cabe sinalizar que *classement* é o arranjo ou ordenação. *Classification* é um método para o arranjo, que pode ser alfabético ou numérico, ou, misto. Coloquialmente é comum o emprego da classificação como se fosse o instrumento e não o processo, o que não é o ideal. Essa terminologia compõe o conhecimento do campo em língua francesa até os dias atuais.¹¹²

Ainda em relação às tarefas documentárias para descrição do livro, Otlet (2018, p. 466) indica os resumos como “exposição da documentação” a partir da análise, que “[...] de forma sucinta, [...] [demonstra o] que está contido nos documentos (concisão, condensação, extração da ‘essência medular’).” Assim, nos “resumos e análises”, o primeiro é produto resultante da segunda.

O procedimento de elaboração de resumo decorre da avaliação sobre sua utilidade, e assim, a técnica de resumir foi-se constituindo de recomendações e caráter metodológico. Segundo Otlet (2018, p. 466), “o resumo tem um valor em si. A concisão de sua exposição obriga a ressaltar os princípios fundamentais e adotar um plano que leve a realçar ao mesmo tempo o estado de um assunto e sua dinâmica interior.” O resumo, enquanto resultado de uma operação, primeiro responde pela função de seleção e, de condensação, enquanto processo, para após ser produto da informação e facilitador do acesso. Tem “um fato e tem um papel a cumprir”, pois informam, assim “[...] como uma revista científica que se lê para saber das

¹¹² Dissertação que se apropria dessa discussão:

SILVA, Camila Mariana Aparecida da. **Para uma abordagem contemporânea sobre ordenação de documentos:** propostas do século XIX e início do XX. 2016. 230 f. Dissertação (Mestrado em Ciência da Informação)-Programa de Pós-Graduação em Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2016.

‘novidades’, para ficar a par, *grosso modo*, do que está acontecendo, para não ignorar por completo algum fato essencial.” (OTLET, 2018, p. 466, itálico do autor).

Mas, ainda na perspectiva das tarefas documentárias para a descrição do livro, Otlet indicou para o repertório ou catálogo de assunto a forma sistemática para o arranjo, pela Classificação Decimal, e, assim, desenvolveu o que pode ter sido sua principal percepção sobre a questão do assunto, conforme é possível observar:

Quando uma obra trata de vários assuntos, ela é representada no *catálogo sistemático* por várias fichas ordenadas pelos números respectivos desses assuntos. Contudo o volume é colocado em ordem nas estantes somente pelo número correspondente ao *assunto principal* da obra, ou, se os diversos assuntos forem de igual importância, pelo número que aparece primeiro na ordem da Classificação Decimal. E este número ou, se for o caso, o do assunto principal, que será usado como indicador de localização [...]. Para evitar confusão, é necessário, em casos como este, reproduzir esse indicador de localização em cada uma das fichas desdobradas da ficha principal tanto do repertório alfabético quanto do repertório sistemático. (OTLET, 2018, p. 471, itálico do autor).

Por essa exposição, é possível compreender que o avanço teórico e prático dos três Princípios (da Monografia, da Continuidade e da Multiplicação dos Dados) e a questão do assunto posto por Otlet, ao mesmo tempo que desenha o seu sistema de informação de forma normatizada, também considera o percurso do mapeamento do conhecimento produzido pela contextualização em um sistema classificatório. Em função disso, o assunto é fixado em uma estrutura sistemática, efetivando também a capacidade de localização dos documentos. A estrutura classificatória trouxe contribuições, como segue:

Por meio da objetivação e contextualização da informação, segundo Otlet, seriam criados novos discursos, do ponto de vista do próprio sistema. Em contrapartida, o sistema poderia oferecer várias possibilidades de acesso às informações, já que o discurso, fragmentado e tratado pelas tabelas da CDU, poderia ser recontextualizado pelo pesquisador. Otlet estava, na verdade, tentando escrever um novo tipo de livro, aquele que seria lido por toda a humanidade [...] (SANTOS, 2007, p. 56).

Os fundamentos envidados por Otlet demonstram seu interesse em propor metodologias para a organização da informação. Sua abordagem teórica recupera a bibliologia e amplia o seu objeto de estudo, que vai do livro (documento), aos princípios que sustentam a Documentação.

Para a organização da informação, a análise do documento sob a égide de processos que são operados, ao longo do tempo, pode ser explorada, segundo Ortega (2008, p. 10), da seguinte forma:

- apresentação (elaboração da aparência da unidade física documental);
- representação (organização das unidades informacionais relativas à forma e ao conteúdo);
- indicação dos pontos de acesso (por meio da seleção dos itens de representação citados acima).

Os três processos foram evoluindo em decorrência das contingências históricas de cada época e, com a Documentação, não foi diferente. Como foi possível constatar nas seções anteriores dessa pesquisa, no decorrer do tempo a intenção de organizar a informação envolveu diferentes estudiosos, entretanto, foi a manipulação para a recuperação que movimentou relações que deflagraram a organização adotada para um sistema documentário. Isso não foi diferente com Otlet, pois, quer em comunidades de bibliotecas ou em serviços especializados de informação científica, os processos articulam-se com elementos a serem representados. Desse modo, os processos pela apresentação da unidade física, a representação da forma e do conteúdo e a indicação de pontos de acesso pela unidade física e intelectual relacionam o que pode ser evidenciado na produção do registro de um documento em um sistema.

Outro ponto que Otlet (2018, p. 140) aborda é: “na apresentação do assunto aqui, é preciso distinguir as seguintes questões: a) as regras da composição literária em geral, a palavra ‘literária’ estendendo-se aqui a tudo que é letra ou escrito, portanto, a ciência e a técnica, e não menos a literatura.” Nesse sentido, o conjunto de livros científicos ou literários mantém a mesma relação de formato intelectual pelo texto, mas essa forma varia segundo a abordagem, “[...] conforme se trate de uma obra literária ou de uma obra científica. Fantasia e imaginação em um caso; rigor científico e racionalização em outro.” Entretanto, mesmo havendo o contexto de cada texto, “[...] as formas de exposição, sejam elas literárias ou científicas, possuem em comum inúmeros elementos que convém examinar em conjunto.”

Esses textos são produzidos segundo tipologias documentais, constituídas por partes. Segundo Otlet (2018, p. 140), essa relação é estruturada pela aplicação da Classificação Decimal, especificamente por meio das tabelas de subdivisões comuns de formas. O autor explica que a “forma do livro não é arbitrária”, mas “amplamente determinada por necessidades, ou seja, por objetivos a serem alcançados”. Por sua vez, a forma do livro (documento) define o registro, a ser recuperado no âmbito do sistema.

Sob essa perspectiva, compreende-se que a forma é seccionada a partir da identificação de elementos pertinentes ao livro (e ao documento), recurso que subsidia seu tratamento. Segundo Otlet (2018, p. 62),

É preciso considerar as características do livro da mesma maneira pela qual o naturalista considera as espécies animais, vegetais e minerais. A concepção de um tipo geral e abstrato, o livro, se destaca da mesma maneira como a zoologia, a botânica e a mineralogia imaginam o animal paralelamente aos animais, a planta paralelamente as plantas, o mineral paralelamente aos minerais. Deve-se examinar sucessivamente:

1° os elementos constitutivos do livro ou documento;

2° suas diversas partes e sua estrutura;

3° as espécies ou famílias de obras.

O exame desses dados tem sua razão de ser em si mesmo e para qualquer finalidade. Também serve de base para as operações de colação, de bibliografia, de catálogo e de classificação e lhes dá um fundamento científico e racional.

A determinação das características de um livro é indispensável para reconhecê-lo e identificá-lo. Essa determinação individual somente se conseguirá fazer em função das características gerais.

Da mesma forma como o naturalista examina os fatores que compõem seu problema para elucidá-lo, o modo proposto por Otlet permite compreender o livro (e o documento) de forma integrada, ao organizar as suas informações, observando um triplo movimento: os elementos, as partes e estruturas, e, as famílias. O relacionamento desses no tratamento da informação depende do sistema documentário em questão e dos objetivos a serem alcançados. Ademais, as reflexões de Otlet permitem observar com quais áreas é possível dialogar, formulando condições teóricas e metodológicas que possam ser operadas. O exame se relaciona com uma forma de análise metódica que pode aproximar outros documentos, reunindo semelhanças em função das características, portanto atributos do documento.

Compreende-se que todo o conjunto teórico de Otlet contribuiu de forma pragmática para a representação dos elementos dos documentos, priorizando a forma sistemática, pois para o funcionamento do RBU, métodos haviam de ser escolhidos para ser possível a composição da rede de acesso à informação.

A seguir, será exposto o trabalho da vertente conhecida como *analyse documentaire* que, em linhas gerais, também considerou o movimento de acesso à informação científica.

6.2 SUBSÍDIO DA ANÁLISE DOCUMENTÁRIA EM JEAN-CLAUDE GARDIN

A vertente nomeada Análise Documentária (português) e *Análisis Documental* (espanhol) sobressai no contexto de pesquisas que envolveram metodologias e instrumentos de recuperação da informação com foco na documentação científica e com fundamento linguístico-documental, a partir do que na França denominou-se *Analyse Documentaire*. Ressaltam-se os nomes de Jean-Claude Gardin, Marcel Van Dijk, Maurice Coyaud, Jacques Chaumier, entre outros. Como exemplo dos experimentos franceses, destaca-se a rede semântica *Syntagmatic Organization Language* (SYNTOL), construída na década de 1960 por Gardin e sua equipe, visando expressar o conteúdo dos documentos. Ocupante do cargo de direção no *Centre de Recherches Archéologiques* (CNRS), Gardin buscou possibilidades de imprimir rigor à sistematização de métodos para análise e representação de documentos científicos. A tônica de abordagem dessa corrente foi caracterizada, sobretudo na Espanha, através de inúmeros pesquisadores, como os de interesse para essa investigação: Sagredo Fernández e Izquierdo Arroyo (1983), García Gutiérrez (1990), Pinto Molina (1993) e Izquierdo Alonso (2000), além de outros nomes de destaque.

A Análise Documentária constitui-se no Brasil com as pesquisas do Grupo Temma, fundado em 1986, na Escola de Comunicação e Artes da Universidade de São Paulo (ECA/USP), a partir de iniciativa da professora Johanna Smit. A publicação do livro *Análise Documentária: a análise da síntese*, em 1987, por ela organizado, é considerado marco oficial do grupo (LARA, 2011). Ademais, salienta-se o fato de Johanna Smit ter se doutorado na França, sob orientação de Gardin, e sua parceria com demais professores e pesquisadores da USP resultou na expansão e fortalecimento da corrente que caminhou com intensidade para a Universidade Estadual Paulista (Unesp), câmpus de Marília (GUIMARÃES, 2011).

A *analyse documentaire* é por vezes no Brasil denominada análise documental, mas faz-se uma observação em relação à essa terminologia, conforme trata Ortega (2017). Guimarães (2008), com base em Guimarães, Nascimento e Moraes (2005), adere ao termo análise documental, considerando a derivação de adjetivos na língua portuguesa, a partir dos substantivos finalizados em -nto ser feita em -al. Porém, Ortega (2017 p. 12) explica que “o [...] sufixo -ário é adotado para a ideia de ‘relativo a’, além de baseado em vocábulos de origem latina, entre outras situações”.

Ainda segundo a autora, o referido sufixo comporta o raciocínio pelas regras da língua portuguesa para envolver o tratamento de processos, instrumentos e produtos “relativos à” significação de objetos. No entanto, “[...] originalmente em francês, o adjetivo documentário

contempla tanto o documento – objeto que é produto da atividade documentária – quanto aquilo que é relativo à documentação, ou seja, os processos que resultam em documentos.” (ORTEGA, 2017, p. 12). Com isso, é reafirmada a escolha pelo sufixo -ário, ao invés do -al.

Segundo Gardin (1964, p. 12, tradução nossa, itálico do autor),

Por <<análise documentária>> – ou por abreviação, <<análise>> – entendemos qualquer operação ou sequência de operações para *representar* um determinado documento de uma forma diferente da forma original, seja para traduzir, para resumir, indexar, etc., a fim de facilitar sua consulta ou identificação pelos especialistas interessados.¹¹³

A posição do autor tem em vista que a análise subjaz um conjunto de operações que apoiam a representação de um documento. Deve-se considerar que o foco da análise para Gardin é o processo, pois seu trabalho centrou em sistemas automáticos para a organização da documentação científica, que pode ser explicado por seu envolvimento na época enquanto diretor na *Section d'automatique documentaire* do *Centre National de la Documentation* e no *Centre Français de Documentation Scientifique*. Ademais, na década de 1960, o uso dos computadores intensificaram atividades que eram feitas manualmente, como por exemplo, a busca e recuperação da informação.

Para Gardin (1964, p. 14), a análise é a etapa necessária à produção do resumo e à indexação. Na mesma medida, a produção do resumo e a indexação constituem, de forma implícita e geral, o próprio processo de análise, pois resultam em produtos substitutos dos documentos, portanto, de forma diferente do original, mas que o representam com o propósito de facilitar o acesso à informação científica.

Na “perspectiva da automatização”, o autor anuncia que os resumos são envolvidos por debates intensos, por conta do seu custo de elaboração ser alto. Porém, a partir de três pontos de vista, ele defende os resumos científicos pela utilidade (*quant à l'utilité des résumés*), necessidade (*quand à la nécessité des résumés*) e oportunidade (*quant à l'opportunité*). Justifica, respectivamente, que a utilidade científica envolve os aspectos para a “[...] revisão bibliográfica [...] [pois, pelo resumo a] informação [é] mais facilmente acessível e, por isso, a mais merecedora de atenção.”¹¹⁴ (GARDIN, 1964, p. 14, tradução nossa). A

¹¹³ Par <<analyse documentaire>> – ou par abrégé, <<analyse>> – nous entendrons toute opération ou suite d'opérations visant à *représenter* un document donné sous une forme différente de la forme originale, qu'il s'agisse de le traduire, de le résumer, de l'indexer, etc., pour en faciliter la consultation ou le repérage par les spécialistes intéressés. (GARDIN, 1964, p. 12, itálico do autor).

¹¹⁴ “[...] et la revue bibliographique [...] l'instrument d'information le plus facilement accessible, et pour cette raison le plus digne d'attention.” (GARDIN, 1964, p. 14).

essência dos resumos é puramente informativa, pois fornece a quem necessita o conteúdo informacional básico do texto científico. O processo de automatização dos resumos não isenta problemas de natureza teórica (como critérios de relevância, duração e custo de uma leitura automática), e a forma manual de condensação é necessária, precisamente por causa dos limites da máquina para ‘compreender’ o conteúdo e gerar o produto documentário. O modo de elaboração dos resumos, a partir da condensação do texto original, contempla um percurso metodológico e intelectual que propicia a difusão e o progresso científico. Relacionado ao ponto de vista econômico, é um produto documentário de custo elevado e, por isso, segundo Gardin (1964), justificam-se as investigações sobre o resumo automático, a partir de cálculo estatístico, conforme trabalho desenvolvido por H.P. Luhn.¹¹⁵ Entretanto, os avanços desses estudos necessitam do “[...] desenvolvimento de uma <<entrada>> automática de textos *completos* (fitas perfuradas, máquinas de leitura, etc.), que está em fase experimental.”¹¹⁶ (GARDIN, 1964, p. 15, destaque do autor, tradução nossa).

Atualmente, os estudos sobre automatização de resumos continuam em desenvolvimento e, em algumas áreas, como por exemplo, as exatas, há mais chance da máquina ter sucesso na elaboração em detrimento de outras áreas, como nas ciências sociais.

Em relação à indexação, para Gardin (1964, p. 15, tradução nossa), consiste em:

[...] toda caracterização de um documento particular por meio de palavras ou grupos de palavras extraídas de um vocabulário limitado, para descrever a fonte (autor, data, referências, etc.) ou o seu conteúdo (assunto tratado). A quantidade de informação expressa varia de acordo com a necessidade. Se for uma questão de atribuir um código de localização [cota, em português de Portugal] a um documento, tendo por fim seu arranjo em uma biblioteca ou de sua indicação em um arquivo, a indexação é geralmente sumarizada. Ao contrário, se os dados extraídos constituírem uma representação real de cada documento, considerado do ponto de vista de seu conteúdo, a indexação torna-se imediatamente mais complexa.¹¹⁷

¹¹⁵ A técnica de Luhn relaciona-se com a frequência de ocorrência das palavras no texto. Para mais detalhes consultar:

LUHN, H. P. The automatic creation of literature abstracts. **IBM Journal of Research and Development**, v. 2, n. 2, p. 159-165, abril. 1958.

¹¹⁶ “[...] mise au point d'une <<entrée>> automatique des textes *intégraux* (bandes perforées, machines à lire, etc.), qui en est elle-même au stade de l'expérimentation.” (GARDIN, 1964, p. 15, destaque do autor).

¹¹⁷ “[...] toute caractérisation d'un document donné au moyen de mots ou groupes de mots empruntés à un vocabulaire généralement limité, pour en décrire la source (auteur, date, références, etc.) ou le contenu (sujet traité). La quantité d'informations ainsi exprimée varie selon les besoins. S'il s'agit d'attribuer une <<cote>> à un document, en vue de son rangement dans une bibliothèque ou de sa signalisation dans un fichier, l'indexation demeure généralement sommaire. Au contraire, si les données retenues doivent constituer une représentation véritable de chaque document, envisagé du point de vue de son contenu, l'indexation devient immédiatement plus complexe. (GARDIN, 1964, p. 15).

A análise para fins de representação necessita do procedimento para dar conta do modo de realização. A indexação é o procedimento que implica o isolamento das características do documento e sua complexidade consiste na forma de representação do conteúdo. Segundo o autor, o aspecto lógico-linguístico do texto é elemento de complexidade, situação que envolve, por exemplo, redundância e ambiguidade, aspectos que interferem na busca por informações.

Gardin (1964) discorre sobre a dúvida de alguns autores quanto à suficiência da indexação para a representação dos aspectos lógico-linguísticos do texto científico. A dúvida deles, segundo Gardin (1964, p. 16, *itálico do autor, tradução nossa*), assemelha-se ao que foi citado sobre os resumos, pois seria possível adquirir benefícios em “[...] gravar e explorar por meios automáticos *todas as* informações encontradas em textos científicos, sem prévia condensação ou indexação, etc.”¹¹⁸ Todavia, ressaltando que essa posição não é sua, Gardin (1964) apresenta duas objeções, como segue.

A primeira objeção é a de que a indexação seria um processo amplo para organização da literatura científica. Essa organização envolveria a exploração automática como auxílio à classificação dos documentos, ou, alternativamente, a constituição dos termos da indexação na compilação de índices e tabelas com conteúdos analíticos de periódicos, dentre outros. A segunda objeção é a de que, mesmo com os avanços da tecnologia e formas de armazenamento do texto completo dos documentos científicos, para Gardin (1964, p. 16-17, *itálico do autor, tradução nossa*), o problema que envolve a indexação não desapareceria, pois:

De fato, o objetivo da indexação não é essencialmente a concisão, mas sim a normalização de expressões registradas, para descrever o conteúdo de textos científicos (ver a expressão <<normalização de conceitos>> frequentemente usada para a indexação). Se esses textos fossem armazenados em sua forma literal bruta, sem normalização - ou seja, sem indexação -, seria fácil mostrar que a exploração da informação se depararia com problemas linguísticos equivalentes, por causa da diversidade permitida na formulação do conteúdo documental. Em outras palavras, por termos renunciado a algumas <<reduções>> na fraseologia dos textos científicos, seríamos forçados a compensar com <<amplificações>> na expressão das *questões*, de modo a encontrar todas as correspondências solicitadas entre o *conteúdo* das questões e o conteúdo dos documentos, quaisquer que sejam as variantes observadas na forma literal desses últimos.¹¹⁹

¹¹⁸ “[...] enregistrer et explorer par des moyens automatiques la *totalité* des informations trouvées dans les textes scientifiques, sans condensation ou indexation préalables, etc.” (GARDIN, 1964, p. 16, *itálico do autor*).

¹¹⁹ En effet, l'objectif de l'indexation n'est pas essentiellement la concision, mais plutôt la normalisation des expressions enregistrées, pour décrire le contenu des textes scientifiques (cf. l'expression <<normalisation des concepts>> souvent employée à propos de l'indexation). Si ces textes devaient être mis en mémoire sous leur forme littérale brute, sans normalisation - i.e. sans indexation - il est facile de montrer que l'exploration des

Se para Gardin (1964), a análise documentária envolve procedimentos de manipulação do conteúdo dos textos científicos, através da elaboração de resumos e da indexação, o relacionamento com a normalização de conceitos assume duplo papel: auxiliar a representação e a busca. Termos normalizados isolam o conteúdo dos textos e os reduzem para sua representação. Por sua vez, essa redução dialoga com as questões de busca, sobretudo pelo relacionamento que deve haver entre o conteúdo procurado e o conteúdo dos documentos.

De acordo com Gardin (1964), o texto científico, ao ser indexado, passa por um processo de tratamento semântico e, para atender às questões de busca, ele indicou a necessidade de normalização das expressões do conteúdo. Para o autor, esse processo semântico, que envolve a manipulação e a representação dos textos científicos, necessita de um léxico ou de uma linguagem artificial especializada para o tratamento de informações de um domínio específico. A linguagem especializada é restritiva por implicar maior nível de especificidade e, com isso, caracteriza bem um domínio científico. Aponta ainda que a análise documentária no tratamento mecânico dos textos científicos, através da indexação, gera produtos que formam um conjunto de representações codificadas para uma massa de documentos, como por exemplo, índices, bibliografias, dentre outros.

No *Centre Français pour la Documentation Scientifique*, Gardin (1964) explicou existirem meios (*moyens*) para a organização dos procedimentos que envolvem métodos (*méthodes*), equipamentos (*équipements*), técnicos (*techniciens*) e organização (*organisation*) para a coleta (*collecte*), análise (*analyse*), tratamento (*traitement*) e difusão (*diffusion*) da informação científica. Focando na análise (*analyse*), ele considerou a esquematização presente no quadro 5.

informations se heurterait à son tour à des problèmes linguistiques équivalents, du fait même de la diversité admise dans la formulation des contenus documentaires. En d'autres termes, pour avoir renoncé à certaines <<reductions>> dans la phraséologie des textes scientifiques, on se trouverait contraint à des <<amplifications>> compensatrices, dans l'expression des *questions*, de manière à retrouver toutes les correspondances cherchées entre *contenu* des questions et contenu des documents, quelles que soient les variantes observées dans la forme littérale de ces derniers. (GARDIN, 1964, p. 16-17, itálico do autor).

Quadro 5 – Esquematização com foco na análise (*analyse*)

	<i>Moyens</i>				
	<i>Fonctions</i>	<i>Méthodes</i>	<i>Equipements</i>	<i>Techniciens</i>	<i>Organisation</i>
A n a l y s e	<i>Resumé</i>			<i>chercheurs</i>	<i>réseau d'extraction</i>
	<i>Traduction</i>			<i>chercheurs</i>	<i>réseau d'traduction</i>
	<i>Indexation</i>	<i>systèmes linguistiques d'indexation</i>	<i>indexation automatique</i>	<i>analystes linguistes</i>	<i>Organisation interne du Centre</i>
	<i>Codification</i>		<i>codification (et lecture) automatiques</i>		

Fonte: Gardin (1964, p. 28).

É importante destacar os elementos que Gardin (1964) arrolou para a análise, pois nos demais - coleta, tratamento e difusão - não houve o mesmo nível de desdobramento. Ademais, foi na indexação que ele explorou os meios linguísticos, envolvendo preocupação com o léxico e a síntese das palavras que compõem o texto. Para o léxico, destacou a importância da normalização de uma terminologia científica e organizada. Já em relação à sintaxe, fundamentou:

Até os últimos anos, às vezes se assumia que o único problema de indexação era escolher, em um léxico do tipo precedente, as <<palavras-chave>> que designavam os principais elementos do conteúdo, para um determinado documento. A única combinação dessas palavras-chave deveria então fornecer uma representação suficiente do conteúdo, sem precisar especificar, como na linguagem natural, a natureza das *relações* lógicas entre os termos adotados. (GARDIN, 1964, p. 30, itálico do autor, tradução nossa).¹²⁰

De acordo com o autor, o quadro empírico da indexação necessitou de alterações que transpusessem a combinação de palavras-chave para a representação do conteúdo de documentos especializados, e para isso houve a exigência de formalizar procedimentos abarcados pela análise. Se por um lado, a automação pôde propiciar equipamentos que

¹²⁰ Jusqu'à ces dernières années, on admettait parfois que le seul problème de l'indexation consistait à choisir, dans un lexique du type précédent, les <<mots-clé>> désignant les principaux éléments du contenu, pour un document donné. La seule combinaison de ces mots-clé était alors censée fournir une représentation suffisante du contenu, sans que l'on ait à spécifier, comme dans le langage naturel, la nature des *relations* logiques entre les termes retenus. (GARDIN, 1964, p. 30, itálico do autor).

geraram rapidez na recuperação, por outro, o desenvolvendo das técnicas de análises linguísticas (*analystes linguistes*) envolveu formas de transpor o texto científico de uma linguagem natural para a linguagem documentária, que é uma linguagem construída artificialmente. Assim, a terminologia científica de um domínio, o vocabulário normalizado e organizado, a sintaxe documentária e a linguagem modulada de forma simbólica com vistas às questões de busca auxiliaram o desenvolvimento do SYNTOL e da parceria dos estudos sobre *thesaurus* em *Cambridge*, intitulado *Language Research Unit*, com J. Needdham (GARDIN, 1964, p. 40).

O trabalho de Gardin sobre a análise linguística e os aspectos de mecanização da documentação científica, iniciado na década de 1960 (GARDIN, 1967; 1968), com subsequente desenvolvimento em 1970 até 1990, gerou inúmeros trabalhos. Também embasou outros, como o de Belly (*et al*, 1970), que estruturou procedimentos de análise semântica para a mecanização da indexação.

O pensamento de Gardin com a proposição da análise documentária culmina com o rigor semântico que não estava presente nos instrumentos tradicionais de representação da informação. De acordo com Kobashi (1996), a ele credita-se a denominação do campo conhecido como Análise Documentária. Seu interesse relacionou os processos de passagem do texto original para sua representação considerando as linguagens documentárias, que se constituem “[...] por um conjunto de termos (**o léxico**), por relações entre as unidades lexicais, determinadas *a priori* (**o eixo paradigmático**) e por uma sintaxe que articula os encadeamentos entre os termos da linguagem (**o eixo sintagmático**).” (CROSS; GARDIN; LÉVY, 1964 *apud* KOBASHI, 1996, p. 7, destaque da autora).¹²¹

Observa-se que os parâmetros linguísticos foram reconhecidos no trabalho de Gardin para construção de um léxico documentário, que estruturou a elaboração de linguagem documentária, com o propósito de servir à indexação. Kobashi (1996) afirma que as reflexões mais significativas entre linguagem documentária e linguística, inaugurada em fins dos anos de 1960, encontram-se na produção de García Gutiérrez, pesquisador espanhol que forneceu subsídios para os trabalhos do grupo Temma no Brasil.

¹²¹ CROSS, R. C.; GARDIN, J. C.; LÉVY, F. **L'automatization des recherches documentaires: um modele général, le SYNTOL**. Paris: Gauthiers-Villars, 1964.

6.3 A INFLUÊNCIA DA ANÁLISE DOCUMENTÁRIA: O RECORTE ESPANHOL E BRASILEIRO

Do ponto de vista metodológico, a Análise Documentária (AD) define-se, fundamentalmente, como uma operação com textos (KOBASHI, 1994). De acordo com Cunha (1987, p. 40) é “[...] um conjunto de procedimentos efetuados com o fim de expressar o conteúdo de documentos, sob formas destinadas a facilitar a recuperação da informação.”

Pinto Molina (1993) considera que, em documentos textuais, o *contenido* é a *materia* intrínseca do texto, indicando que o primeiro é permanente no documento, entretanto, envolve um número variado de significados. Pelo fato de os documentos serem usados por diferentes pessoas, a autora evidencia que os potenciais significados não devem ser tomados com excessiva rigidez. Com isso, no contexto da *análise documental* (AD), Pinto Molina (1993, p. 155-156, tradução nossa) propôs serem “[...] cinco as variáveis que definem o marco de referência conceitual da AD de conteúdo: texto, contexto, conhecimento base do analista, objetivos documentais do momento e procedimento de validação de resultados.”¹²²

Embora a *análise documental* tenha estruturado, por um lado, a representação temática dos textos em vários aspectos, conforme os indicados por Pinto Molina (1993), ela afirma a existência de “extrema anarquia conceitual” referente às operações de AD e aponta que os problemas de ambiguidade e falta de precisão ocorrem por se tratar de um domínio que carece de definições precisas. Entretanto, a autora também elucida que o *corpus* teórico da AD abrange terminologia própria composta tanto por termos aceitos de outras disciplinas afins quanto termos oriundos da linguagem cotidiana. Ademais, apresenta que a AD detém níveis de análise que estão relacionados com os elementos dicotômicos do documento - conteúdo (mensagem) e forma (suporte) - direcionando, assim, a *Análisis Documental de la Forma* (ADF) como responsável pela abrangência da catalogação e a *Análisis Documental de Contenido* (ADC) relativa ao aspecto do assunto pela indexação, classificação e pelo processo de resumir. Conforme já exposto, vale lembrar que o atributo assunto constitui um, dentre outros, relacionados ao processo de classificação.

Para Izquierdo Alonso (2000), a *Análisis Documental de Contenido* (ADC) refere-se à *Lingüística Documental* (LD)/*Semiótica Documental* (SD) responsável por abranger o *Tratamiento Documental de Contenido* (TDC), que por sua vez, utiliza como marco

¹²² “[...] cinco las variables que definen el marco de referencia conceptual de AD de contenido: texto, contexto, conocimiento base del analista, objetivos documentales del momento y procedimiento de validación de resultados.” (PINTO MOLINA, 1993, p. 155-156).

referencial modelos linguísticos do estruturalismo ao gerativismo, além de correntes pragmáticas, de análise de discurso, sociolinguística e psicolinguística. Consta-se que esse cenário teórico recebeu grande influência do professor José María Izquierdo Arroyo, sobretudo em relação à Semiótica.

É fato que a atividade documentária sofre influência da linguagem, pois evidencia aspectos simbólicos que marcam as diferentes práticas da sociedade. É percebida no tratamento da informação, pois visa ser insumo do processo social de geração de sentido textual do processo. Para tanto,

Os investimentos metodológicos realizados pela Análise Documentária partiram do princípio de que a formalização dos procedimentos nessas operações poderia alterar o quadro empírico de representação no âmbito dos sistemas documentários. As operações de substituição de textos em Linguagem Natural por representações supõem necessariamente a existência de uma metalinguagem, que é constituída por unidades lexicais (símbolos designando noções ou conceitos) e convenções sintáticas (para expressão das relações lógicas presentes nos textos em Linguagem Natural). (LARA, 2011, p. 98).

Nessa perspectiva, a influência gardiniana é assimilada por envolver a passagem de um texto original para uma representação por operação semântica que substitui o que está em linguagem natural (LN), por uma metalinguagem, ou seja, uma linguagem controlada, ou, documentária (LD). Nesse sentido o termo metalinguagem não substitui o termo LD, mas enfatiza uma das suas características. Em suma, textos analisados são representados por um sistema simbólico, ou seja, uma metalinguagem convencionalizada.

Antes dos fundamentos estabelecidos pela vertente europeia e, por um bom tempo, segundo Cunha (1987, p. 40), os métodos de análise de textos/documentos foram considerados operações empíricas de “[...] bom senso, pelas quais se atribuíam [...] uma ou várias palavras-chave destinadas a facilitar a recuperação do documento, no momento da pesquisa, sobre um dado tema.” Outro preceito, de acordo com a mesma autora, considerava a neutralidade na atribuição de palavras-chave, como algo desprovido de interferência humana.

Ainda segundo Cunha (1987), uma operação semântica envolve a análise na passagem de um texto original para sua representação, por palavras-chave, por exemplo. Para quem analisa, mesmo que de forma intuitiva e sem seguir nenhuma metodologia, ocorrerá variação na seleção das palavras representativas, pois variáveis precisam ser consideradas: “[...] as do discurso (com pretensões científicas ou não) realizadas através de uma língua dada e que compreende a sua organização interna e a linguagem/ideologia.” Considerando as variáveis,

elas estão presentes em duplo contexto e “[...] surgem não só na fase de análise, como [também] na de obtenção do produto documentário.” (CUNHA, 1987, p. 40-41).

De acordo com o processamento e a produção mencionada, para além de funcionar como codificação, a linguagem opera com estruturas de informação dos textos propiciando sua organização para posterior uso. Quer na forma manual ou automatizada, o produto gerado pela análise vai além da atribuição de algumas palavras-chave. Cunha (1987, p. 41) explica que apesar da automação ter gerado muitas vantagens em isolar por ocorrência e frequência os conceitos dos textos, ela ainda deixa muito a desejar, sobretudo em áreas de Ciências Humanas, “[...] mesmo quando esses textos se dizem científicos, o que pressuporia uma preocupação lógica/metodológica maior e uma maior precisão nos conceitos e termos utilizados.”

Sendo assim, o desenvolvimento de léxicos intermedia a representação e a busca por informação ocupando “[...] o papel de sistemas de conversão entre conceitos apresentados de forma independente nas diversas linguagens, e conceitos de leitura ‘universal’ definidos pela própria Análise Documentária.” (CUNHA, 1987, p. 41). A autora considera que no momento da análise existem três linguagens com fins documentários e não linguísticos: a linguagem natural (LN), as linguagens especializadas (LE) e as linguagens documentárias (LD). Como pretensão de resultado, pela análise documentária se almeja que analistas diferentes ao processar um texto gerem representações semelhantes, para assim haver um cenário propício para validação da busca e dos instrumentos de controle do vocabulário, como os tesouros. Os três tipos de linguagens são diferentes em vários aspectos, inclusive quando relacionadas a diferentes áreas, como nas Ciências Exatas e Humanas, pois entre ambas prevalece variação dos sistemas simbólicos. Se na linguagem especializada da primeira área há grande chance de se transformar em documentária, na segunda área, a linguagem natural confunde-se com a documentária. Dessa forma, a especialização do vocabulário apresenta, por exemplo, diferentes escolas e vertentes vinculadas à própria linguagem natural (CUNHA, 1987).

Face às especificidades apresentadas e relacionadas às atividades documentárias, que são informacionais, é possível compreender que a construção de linguagens torna-se primordial para organização, busca e recuperação de informação. Conforme citado no início dessa subseção, a AD pode estar vinculada à análise da forma e à análise de conteúdo e, por vezes, ela é confundida com a indexação, embora todos esses processos tenham surgido antes da constituição da própria AD.

O processo de indexação, no âmbito da AD, comporta a geração de produtos como resumos e a elaboração de índices, e envolve a análise, a síntese e a representação pela

manipulação do texto. Considerando o índice de forma abrangente e responsável por elencar pontos de acesso em bases de dados, no que concerne o processo de indexação, ele envolve metodologias de organização sobre o que é abordado tematicamente no documento, definido como operação de representação documentária, com o propósito de tornar disponíveis as informações tratadas, possibilitando seu acesso. Nesse sentido, o registro da informação consiste na informação documentária, que “é, contudo, uma representação particular porque, sob as suas diversas formas de manifestação, deve procurar responder às exigências impostas pela recuperação da informação, mantendo, com o texto original, uma relação de similaridade.” (KOBASHI, 1996, p. 12).

Em relação ao texto do documento, ao ser indexado, ele será desestruturado, isto é, analisado e sintetizado para posteriormente ser representado e transformado em informação documentária. Primeiro, estrutura-se as informações selecionadas e, após, submetidas a um processo de combinação, são transformadas. Assim, um novo texto é gerado, no caso do resumo, e símbolos são produzidos, a partir de uma linguagem documentária, no caso da indexação (KOBASHI, 1994).

É na busca, em meio a um universo variado de registros, que a seleção responde à necessidade de informação e se mostra a importância do produto documentário. O documento passa pelo processamento documentário para ser transformado em informação, tornando-se disponível para recuperação, e entre ambos “[...] existe uma relação de pressuposição não recíproca [...]” (KOBASHI, 1994, p. 22), pois a recuperação da informação não é exequível sem antes quem indexa efetuar a análise, a síntese e a representação do documento.

Segundo Chaumier (1998), a parte mais importante da análise documentária é a indexação, responsável por valorar um sistema documentário. Explica que a indexação opera em nível conceitual a partir do conteúdo do documento e sua transcrição ocorre pelo uso de uma linguagem documentária. Ainda para esse autor, o processo de indexação envolve a demanda de quem faz uso buscando alternativas metodológicas que diminuam “[...] ‘RUÍDOS’ (isto não corresponde ao que eu procurava), [...] ‘SILÊNCIOS’ (o documento existe, mas está perdido).” (CHAUMIER, 1988, p. 74, destaque do autor).

Lara (1993, p. 41) explica que a análise do texto do documento é o processo que fomenta procedimentos de desestruturação, estruturação e representação documentária. Como a representação transforma a mensagem do texto em uma forma diferente do estado original, entende-se “[...] que não se recupera qualquer tipo de informação [...] mas apenas informação documentária. Esta se apresenta como produto específico obtido através de procedimentos

documentários a partir da informação original - presente no texto - operada por instrumentos documentários.”

Para que haja correspondência entre informações textuais e as representações documentárias, sejam resumos ou índices, segundo Kobashi (1994, p. 45), o processo global de análise documentária defronta-se com ao menos dois objetivos: “a) o Texto, como a sua unidade de análise; b) a Linguagem Documentária como instrumento comutador.”

Ao colocar-se a frente desse cenário, a AD assume ser, de acordo com Kobashi (1994, p. 15), citando Gardin,

[...] uma disciplina de natureza teórica e prática que integra o domínio conhecido por Biblioteconomia e Documentação, mais modernamente denominado Ciência da Informação. Enquanto parte constitutiva de uma área de conhecimentos, a Análise Documentária assimila, necessariamente, os objetivos globais do tronco do qual é membro, qual seja o de determinar e criar princípios e mecanismos que promovam a circulação de informações e de documentos. No interior desse quadro mais geral, a Análise Documentária responde pelo ‘conjunto de procedimentos utilizados para exprimir o conteúdo dos documentos científicos sob formas destinadas a facilitar a sua localização ou consulta’. (GARDIN, 1987, p. 48-49).

Se, por um lado, a Análise Documentária abarca diferentes aspectos, quer seja enquanto área, quer seja do ponto de vista metodológico, ela também abrange métodos e processos referentes à Linguística Documentária (*Linguística Documental*, em espanhol). Lara (2011, p. 100), explica que “atualmente, a Análise Documentária faz parte de um vasto conjunto dos estudos que podem ser incluídos no subcampo da Linguística Documentária que tende a integrar, por sua vez, o conjunto dos estudos de Organização e Representação do Conhecimento.”

É creditado a García Gutiérrez (1990) a denominação *Linguística Documental*, a partir do seu livro *Estructura Lingüística de la Documentación: teoría y método*. Anterior a este autor, Coypaud (1972), publicou *Linguistique et documentation: les articulations logiques du discours*,¹²³ mas, sob a ótica do seu trabalho, outras investigações também desenvolveram-se anteriormente, como por exemplo, *Introduction a l'étude des langages documentaires* (1966),¹²⁴ associando-se à proximidade e reflexão induzida diretamente a partir de Gardin.

¹²³ COYAUD, Maurice. **Estructura Lingüística de la Documentación: teoría y método**. Paris: Librairie Larouse, 1972.

¹²⁴ COYAUD, Maurice. **Introduction a l'étude des langages documentaires**. Paris: C. Klincksieck, 1966.

Assim, é no tocante à influência gardiniana pelo recorte espanhol que a compreensão da Linguística Documentária é indicada de forma bastante objetiva a partir de García Gutiérrez (1990). O autor considerou o método linguístico como fonte de conhecimento para o processo documentário que envolve a organização de conteúdos para o processo comunicativo. Segundo ele, o termo

[...] ‘Linguística documentária’ surge no momento em que os elementos que intervêm no processo comunicativo-documentário excedem sua tradicional natureza lógica e mecanicista e são descritos, se transformam e se organizam, ao menos essa é a nossa proposta, como elementos de comportamentos fundamentalmente linguísticos mantendo e reforçando seus objetivos habituais. Quanto à análise documentária deixa de ser uma representação formal e superficial do documento, e a linguagem documentária abandona seus modos inventariais e metalinguísticos, o universo da Teoria da Documentação busca auxílio em outras disciplinas. (GARCÍA GUTIÉRREZ, 1990, p. 11, tradução nossa).¹²⁵

A comunicação da informação torna-se possível a partir de uma linguagem que tenha sentido tanto para quem processa como para quem faz uso dos produtos documentários. Desse modo, García Gutiérrez (1990) sugere uma Teoria da Documentação com base na Linguística. Em linhas gerais, sua teoria consistiu em estudar os problemas que caracterizam a Linguagem Documentária como uma forma específica de linguagem inscrita no universo da linguagem geral. A essa dimensão inseriu-se o funcionamento da Linguagem Documentária, tanto no que interessa à sua construção quanto na medida em que envolve os propósitos metodológicos para a confecção de produtos documentários.

O autor considera que a Documentação opera com estruturas de informação e explica que a *análisis documental* trabalha com o aspecto da *análisis de contenido* com fins documentais. Por um lado, produz uma colisão e, por outro, a síntese das estruturas cognitivas de quem analisa e de quem faz uso da informação processada, pois nesse percurso deve haver coincidência de interesse entre o primeiro e o segundo sujeito, independente do último ser qualificado com necessidades informacionais gerais ou especializadas. A partir da análise, é realizado o isolamento de significados para a elaboração de mensagens e é nesse ponto que García Gutiérrez (1990) demonstra interesse em captar e transformar essas mensagens.

¹²⁵ [...] “Linguística documental” surge en el momento en que los elementos que intervienen en el proceso comunicativo-documental exceden su tradicional naturaleza lógica y mecanicista y son descritos, se transforman y organizan, al menos esa es nuestra propuesta, como elementos de comportamiento fundamentalmente lingüístico manteniendo y potenciando sus objetivos habituales. Cuando el análisis documental deja de ser una representación formal y superficial del documento, y el lenguaje documental abandona sus modos inventariales y metalingüísticos, el universo de la Teoría de la Documentación halla auxilio em otras disciplinas. (GARCÍA GUTIÉRREZ, 1990, p. 11).

Para o autor, a “linguagem documentária (ld), é o referencial dessa mensagem e, portanto, a norma em termos linguísticos frente à faculdade de uso que é a análise.”¹²⁶ (GARCÍA GUTIÉRREZ, 1990, p. 25, tradução nossa).

García Gutiérrez (1990, p. 31) criou fundamentos para a construção de linguagem documentária via modelos de Análise Documentária. Segundo o autor, a Linguística Documentária é compreendida como a linguagem do sistema, pois transmite mensagens documentárias através de elementos significantes; é considerada em nível científico e informativo através dos processos estabelecidos por meio dos documentos; considera o interesse social do uso da informação em detrimento do uso individual; e, trabalha com vocabulários, normas e sistemas controlados em detrimento da linguagem natural. A efetivação desse contexto decorreu dos problemas relacionados ao caráter simbólico da informação, das particularidades da mediação documentária e das possibilidades de prover a circulação social da informação.

A linguagem assume, no contexto informacional, diferentes perspectivas com relação à expressão significação, pois pela análise do texto será realizada a extração de signos que pode diferir das palavras usadas por quem o escreveu. Já quem busca informação, não necessariamente usará a mesma forma (de palavras) de quem escreveu e de quem analisou. Nesse sentido, a linguagem documentária, pela apropriação da linguagem, aproxima-se da Linguística.

Em relação ao exposto acima, e anteriormente à centralidade da mensagem no processo comunicativo que García Gutiérrez (1990) coloca, destaca-se a aproximação que ele propõe a partir do Estruturalismo. Vale-se dessa vertente desenvolvida por Ferdinand de Saussure para propor o estudo descritivo da língua no sentido metodológico e não ideológico. Assim,

A estrutura pressupõe estratificação e hierarquia conceitual, algo que deve ser analisado em detalhes, e seus elementos são definidos pelas relações que eles mantêm. Texto, enunciados e termos são níveis da língua e ocupam um lugar respectivo. A estrutura é gerada quando se observa a materialização das relações que os envolvem no mesmo processo discursivo. O método estruturalista pode ser usado em LD, conseqüentemente, criando relações artificiais entre os elementos constituintes de um vocabulário, a partir [de relações] naturais, a fim de criar uma estrutura de representação e síntese de uma determinada área do conhecimento. (GARCÍA GUTIÉRREZ, 1990, p. 35, tradução nossa).¹²⁷

¹²⁶ “Lenguaje documental (ld), es el marco referencial de ese mensaje y, por tanto, la norma en términos lingüísticos frente a la facultad de uso que es el análisis.” (GARCÍA GUTIÉRREZ, 1990, p. 25).

¹²⁷ La estructura supone estratificación y jerarquización conceptual, algo que deberá ser analizado al detalle, y sus elementos se definen por las relaciones que mantienen. Texto, enunciados y términos son niveles de la

Pelo mecanismo de funcionamento da linguagem natural, as possibilidades de comunicação, interpretação e representação são consideradas pela Linguística Documentária com vinculação ao funcionamento da língua. Entretanto, nesta linha de estudos, não se reforça o uso desse tipo de linguagem e a referência é apenas uma maneira de modelização inicial, pois a linguagem natural também tem estrutura, pois esta é inerente à linguagem. Sendo assim, a relevância está totalmente centrada na estrutura.

A língua como sistema estrutural de relações permite a operação de estruturas de informação. Também pode se relacionar com outras funções sógnicas que introduzidas em uma cadeia (com outros elementos) podem desempenhar outras funções também sógnicas, como por exemplo, acontece na Semiótica (GARCÍA GUTIÉRREZ, 1990 *apud* ECO, 1975).

A proposta de García Gutiérrez (1990, p. 53, tradução nossa) envolveu a Análise Documentária como parte da Linguística Documentária no que se refere ao significado do texto, “[...] que pertence por coerência a um universo discursivo-documentário (apoiado materialmente ou documentalizado em sentido otletiano) [...].”¹²⁸ Para o autor, a AD com ênfase nos processos e seus respectivos resultados supõe a abordagem discursiva de um documento que ele chamou de *linguo-documental*. O quadro elabora-se através da organização de conteúdos codificáveis e decodificáveis, dividindo-se entre a estrutura cognitiva do produtor e a estrutura interpretativa de quem usa. Portanto, a influência do estruturalismo em paralelo ao objeto de investigação (signo) propõe o vínculo da documentação e da língua (*linguo-documental*).

As ideias encontram na linguagem documentária seu núcleo composto por conceitos, enquanto instrumento que fundamenta a organização dos conteúdos. No que tange ao cenário brasileiro, a Linguística Documentária teve inserção a partir do grupo Temma, que conforme já citado, foi o grupo de pesquisa que envolveu investigadores da ECA/USP.

Lara (2011, p. 114) explica que “a Linguística Documentária constitui um subcampo da Ciência da Informação que tem como objetivo estudar os problemas que caracterizam a Linguagem Documentária como uma forma específica de linguagem inscrita no universo da

lengua y ocupan un lugar respectivo. La estructura se genera cuando se observa la materialización de las relaciones que los involucran en el mismo proceso discursivo. El método estructuralista puede utilizarse en LD, en consecuencia, creando relaciones artificiales entre los elementos constituyentes de un vocabulario, a partir de las naturales, con el fin de crear una estructura de representación y síntesis de un área determinada del conocimiento. (GARCÍA GUTIÉRREZ, 1990, p. 35).

¹²⁸ “[...] que pertenece por coherencia a un universo discursivo-documental (apoyado materialmente o documentalizado en sentido otletiano) [...]” (GARCÍA GUTIÉRREZ, 1990, p. 53).

linguagem geral.” Se a Língua Documentária foi alicerçada na Linguística Estruturalista, a autora afirma que também nessa direção o grupo Temma procurou o

[...] entendimento do funcionamento da linguagem para o tratamento da informação, quer no aspecto metodológico de construção do instrumento mediador, quer na elaboração de produtos documentários que funcionem na comunicação documentária, ou comunicação em ambientes documentário-informacionais.

A proposição da Linguística Documentária como campo de estudos no Brasil conheceu seus desenvolvimentos específicos. Mantendo-se os objetivos de substituir a atividade empírica de representação de conteúdos por outra, via Análise Documentária e uso da Linguagem Documentária, esse desenvolvimento tem focalizado, com maior ênfase, a construção do instrumento de mediação, mobilizando para isso as referências, linguístico-semióticas e estabelecendo um diálogo mais frequente com a Terminologia. (LARA, 2011, p. 115).

A apropriação do referido grupo pela Terminologia resultou em várias pesquisas, orientações e publicações, como por exemplo, Lara e Tálamo (2006; 2007) e Tálamo e Lara (2006; 2009). As autoras consideram que a proposta de García Gutiérrez (1990) avança com a Terminologia devido à possibilidade de robustez teórico-metodológica que ela apresenta. Afirmam que a importância da Linguística Documentária consiste em integrar Linguística e Análise Documentária para um ambiente de cunho interpretativo que não deve ser constituído pelo “bom senso”, mas por políticas institucionais que pautem os espaços de “trânsito da informação”.

Conforme exposto, os esforços do ponto de vista metodológico oriundos da AD respondem pelo tratamento do conteúdo do texto. Para a representação desse conteúdo um instrumento deve ser adotado e, com isso, houve a difusão do conceito de tesouro documentário, pois com ele a linguagem de tratamento da informação passa a ser considerada efetivamente sob seu aspecto sistêmico.

Cintra, Tálamo, Lara e Kobashi (1994)¹²⁹ evidenciam o texto e o discurso no que refere à construção de descritores no âmbito da atividade terminológica. Para tanto, consideram que o texto corresponde a uma unidade de sentido, coerente e coesa, sendo um objeto acabado, portador de uma estrutura e de um léxico. Já discurso consiste em uma atividade comunicativa que se manifesta através do texto. Não está acabado, pois é um objeto em construção, cuja realização se dá na interlocução. Nesse sentido, para a construção de

¹²⁹ Nesse caso, optou-se em não usar a expressão ‘*et al*’ pela representatividade das autoras junto ao grupo Temma.

tesauros, texto e discurso devem ser considerados, pois os textos relacionam os elementos integrantes do discurso, mas a operacionalização torna-se um desafio.

Ao largo da interação entre as autoras, é importante perceber a diferença entre linguagem de especialidade e terminologia. Em relação à primeira, ainda existe o diferencial entre ela e a linguagem especializada, que é restritiva; a de especialidade pode ser caracterizada por discursos de modo geral, inclusive os não científicos. Os termos dos textos de especialidade podem facilitar a interface com os usuários. Já as terminologias de área, por sua vez, constituem o produto da pesquisa científica, representando, de modo sistematizado, uma linguagem especializada. Por isso, as autoras recomendam as terminologias de áreas para uso na construção de tesauros e, em especial, no que se refere à seleção e organização de descritores.

Para tanto, as autoras relacionam a dinâmica do sentido no isolamento das unidades linguísticas e semânticas, que são formas de expressão e formas de conteúdo, para a construção das linguagens documentárias. Portanto,

[...] a função das Linguagens Documentárias para o tratamento de textos se dá no âmbito de categorias de assunto: não há como operar, documentariamente, a especificidade das apropriações individuais consubstanciadas no texto, através de linguagens de intermediação. (CINTRA; TÁLAMO; LARA; KOBASHI, 1994, p. 78).

De acordo com as autoras, não é possível recuperar “a dinâmica do sentido de textos individuais”. Sendo assim, a representação pelas linguagens documentárias decorre por assuntos, por classes de assuntos e não por textos individuais. Portanto, o assunto é visto como agrupador de sentido para a interseção entre as terminologias e as linguagens documentárias.

Ainda segundo as autoras, o texto ao ser submetido ao tratamento de elaboração de resumo é desestruturado, para após ser organizado sob a forma de um novo texto, com a reconstrução potencial do sentido original. Na indexação, o tratamento envolve outros procedimentos, como o de análise e coleta de dados. Assim,

Ao se procurar caracterizar o assunto global do texto ou, em certos casos, destacar dados factuais ou conceitos nele presentes, utiliza-se um código comutador, a Linguagem Documentária. A condensação, neste caso, é expressa por elementos de um código exterior ao texto submetido à conversão.

Esse código comutador, normalmente composto por um conjunto limitado de termos denominados “descritores”, prescreve as formas de entrada e de

busca a serem utilizadas pelo indexador ou pelo usuário, num sistema documentário. Por ser assim constituído, tal instrumento integra elementos resultantes de escolhas feitas em um dado universo lexical. Suas unidades têm origem tanto em Linguagens de Especialidade, como na linguagem de uso corrente e nas Terminologias de área. (CINTRA; TÁLAMO; LARA; KOBASHI, 1994, p. 77).

O assunto construído a partir do texto é incorporado e colabora para análise de conceitos. Por outro lado, um tesouro documentário também é embasado por conceitos, mas torna-se importante destacar que as terminologias não podem ser confundidas com as linguagens documentárias, pois são referência para estas. Ademais, na indexação, pelo uso da linguagem documentária, o resultado é a compactação do texto e Kobashi (2008, p. 50), afirma que “do processo de análise resultam palavras ou sintagmas indicadores do conteúdo.” Dessa forma, assunto e conteúdo podem ser compreendidos como sinônimos.

Ao considerar a indexação pela análise do texto, permite-se a construção da informação documentária, que gera uma relação de similaridade. Kobashi (2008, p. 52) explica que esse tipo de informação é ocasionado, “[...] de um lado, pela neutralização do poder expressivo do texto e, de outro, moldada de acordo com formas previamente determinadas, [...] apesar de ser formalmente diferente do original, portanto ‘representação’, é equivalente a ele, do ponto de vista do conteúdo informacional.” Nesse sentido, conteúdo informacional é o resultado do processo gerado, de elaboração da informação documentária.

Diante do cenário apresentado, é nítido observar que a vertente europeia e seus desdobramentos detêm fundamentos que a particulariza e a diferencia da anglo-saxã. De forma cada vez mais acentuada essa particularidade é percebida, conforme pode-se constatar, em grande parte dos trabalhos publicados nos *Anais do 11º Colloque ISKO France 2017*¹³⁰ (EL HADI, 2018). De outro modo, Hudon (2018), em artigo recente no periódico *Knowledge Organization* resgata o referido evento e alguns nomes de referenciais aqui tratados, como Otlet e Gardin, e também o grupo Temma.

¹³⁰ Evento realizado pelo capítulo francês da *International Society for Knowledge Organization*: www.isko-france.asso.fr/colloque2017/en/homepage/

7 EM FOCO A NOÇÃO DE ASSUNTO E O MODELO FRSAD: DESCRIÇÃO E ANÁLISE

Nessa parte, inicia-se o foco empírico da pesquisa. Conforme delineado na segunda seção, percurso metodológico, recorre-se à abordagem qualitativa que se pautou na descrição dos três momentos de construção da noção de assunto. Considera-se que uma proposta de representação do assunto na Biblioteconomia e Ciência da Informação não pode ser realizada na ausência de fundamentos com diferentes influências.

Ao pensar nos conceitos de uma determinada área através da literatura, que especificamente nesse caso envolve a noção de assunto abstraída, a reflexão deve propiciar a interpretação para compreender características, estruturas e/ou modelos que estão por trás das mensagens. Assim, “fragmentos de mensagens” são levados em consideração e, através da intencionalidade, permitem a criação de categorias (BARDIN, 2018).

Construir categorias envolve operar sobre características comuns aos objetos considerados. “A categorização é um processo de tipo estruturalista e comporta duas etapas: • o inventário: isolar os elementos; • a classificação: [...] procurar ou impôr uma certa organização às mensagens.” (BARDIN, 2018, p. 146). No sentido de conduzir a uma compreensão do que é investigado, a classificação, ao mesmo tempo que reparte, realiza a organização de mensagens, proporcionando um entendimento a cada categoria.

A categorização é vista como uma possibilidade de interpretação sobre o objeto investigado. Sendo assim, contribui para analisar a noção de assunto incorporada no modelo FRSAD.

De todo modo, em análise de conteúdo, “[...] a mensagem pode ser submetida a uma ou várias *dimensões* de análise.” (BARDIN, 2018, p. 146, itálico da autora). A seguir são realizadas duas dimensões: a primeira está baseada nos momentos teóricos de construção da noção de assunto e a segunda faz uso dos resultados da primeira para analisar o modelo FRSAD.

Para tanto, a subseção seguinte trata da primeira dimensão da análise; a próxima, sobre o modelo FRSAD, engloba sua apresentação e estudo; e, posteriormente, uma segunda dimensão de análise é realizada, com resgate a inicial, pois busca a condição teórica da noção de assunto no modelo.

7.1 ANÁLISE SOBRE A NOÇÃO DE ASSUNTO

Antes de partir para a fundamentação do modelo FRSAD e, sua posterior análise, faz-se necessário identificar modos diferentes de realizar os processos relativos à mesma ‘coisa’, ou seja, a noção que está em foco, que é a noção de assunto. Como já explorado, o assunto não está pronto no documento, ou seja, ele é construído, mas as mudanças de enfoques teóricos deflagram diferentes modos de tratar esta questão. Em linhas gerais, essa construção é feita a partir da análise do assunto, sendo ele compreendido e sintetizado para gerar a representação (temática), mantendo vínculo de semelhança com o conteúdo do documento original.

Nessa subseção, citações são trazidas para auxiliar a identificação das categorias de análise, cuja sistematização foi registrada em quadros específicos. Conforme exposto acima, constituem os “fragmentos de mensagens”.

Sendo assim, o quadro 6 considera o referencial explorado que caracteriza essa construção a partir do primeiro momento – o contexto anglo-saxão no século XIX; o quadro 7 considera o segundo momento – desdobramento do pensamento anglo-saxão no século XX; e, o quadro 8, o terceiro momento – a abordagem europeia, que é de origem francesa, no século XX.

Para dar conta da análise, fez-se a construção de três categorias: estruturação do assunto, o controle do assunto para a entrada e o instrumento adotado, para cada momento.

No decorrer do tempo, propostas diversas de aprimoramento foram feitas e as intenções mudavam a cada vez, mas explícita ou implicitamente, as categorias imprimem perspectivas em face da articulação do assunto. Nessas categorias convergem elementos para a referida noção, mas em cada momento a abordagem se apoia na relação com o contexto e as circunstâncias de trabalho com o assunto na época respectiva.

Quadro 6 – Categorização do Primeiro Momento de Construção da Noção de Assunto: o contexto anglo-saxão no século XIX

Categorias	Referencial teórico da tese
- Estruturação do assunto	<p>De acordo com Cutter (1904, p. 67, tradução nossa), a “[...] ‘entrada específica’ é a principal distinção entre o catálogo dicionário e o catálogo alfabético classificado.” Porém, Tôrres ([2000?]) aborda o fato de Cutter não ter definido o que é essa forma de entrada e aponta na <i>Rules</i>, em §161, haver somente uma contraposição dessa com a entrada por classe. Assim, Tôrres ([2000?], on-line) explica:</p> <p style="text-align: center;">É curioso observar que, ao rejeitar o catálogo alfabético classificado, Cutter tentou abolir a entrada por classe. Entretanto, todo cabeçalho de assunto é nome de classe ou representa classes de referentes (coisas, materiais, processos, propriedades, agentes, lugares, etc.), com exceção dos cabeçalhos individuais (ex.: nomes próprios personativos, locativos, etc.).</p> <p>Nesse contexto, a efetividade das palavras pelos cabeçalhos, quer seja por uma classe ou não, se dá por nomes. Para Cutter (1904, p. 67, tradução nossa), “alguns assuntos não tem nome; eles são falados apenas por meio de uma frase ou de várias frases ainda não definitivas o suficiente para serem usadas como cabeçalho.”</p>
- O controle do assunto para a entrada	<p>Cutter preocupou-se em construir regras (369 ao total) que norteassem a decisão das entradas no catálogo, incluindo a de assunto (GARRIDO ARILLA, 1999).</p> <p>Por isso, Cesarino e Pinto (1978) elucidam que a abordagem alfabética de Cutter consistiu em uma rede bem construída de referências cruzadas com o intuito de unir os assuntos relacionados.</p> <p>A coesão entre catálogo, entrada e cabeçalho levaria às indicações de assuntos escolhidos de forma uniforme em fichas catalográficas e, neste sentido, constituíram palavras recorrentes e estáveis. Seguindo essa proposta, é sabido que a <i>Rules</i> foi pensada para o catálogo dicionário (que é alfabético) e, de acordo com Gomes e Marinho (1984), a principal característica para esse tipo de catálogo é a entrada direta pelas palavras.</p>
- Instrumento adotado	<p>Com isso, considera-se o assunto enquanto uma palavra, ou, uma frase, nomeado a partir das listas de cabeçalhos de assunto. Esse nome gera uma entrada na ficha catalográfica para ser organizada alfabeticamente no catálogo dicionário. É antes um mecanismo de entradas por agrupamento alfabético de palavras do que um agrupamento das ideias dos assuntos (TÔRRES, [2000?]). Mais grave é não ter sido estabelecida na <i>Rules</i> uma base de sustentação entre nomes e assuntos. Esses nomes de assuntos existiam enquanto houvesse aceitação geral e fossem de uso do público.</p>

Fonte: elaborado pela autora.

Quadro 7 – Categorização do Segundo Momento de Construção da Noção de Assunto: desdobramento do pensamento anglo-saxão no século XX

Categorias	Referencial teórico da tese
- Estruturação do assunto	<p>Kaiser apresentou algumas nuances que se diferenciaram das <i>Rules</i> de Cutter. Uma importante característica que diferenciou o primeiro do segundo, conforme Sales (2012), foi a proposta da decomposição de partes do assunto a partir de duas principais categorias: processo e concreto.</p> <p>Ranganathan refletiu um movimento que considerou o conteúdo dos documentos por uma divisão de assunto em múltiplos aspectos, simples ou complexos, que ele denominou de facetas (BARBOSA, 1972) e “[...] nos <i>Prolegomena</i> [fundamentou] [...] o processo de formação de conceitos e sua relação com o que denomina universo das idéias [sic] ou do conhecimento e sua influência no trabalho da classificação.” (CAMPOS; GOMES, 2003, p. 152).</p> <p>Estudos de Ranganathan, na Inglaterra, a partir de 1952, realizado pelo <i>Classification Research Group</i> (CRG) - a decomposição e os possíveis relacionamentos entre os assuntos foram percebidos à medida que o contexto linear de sintaxe dos cabeçalhos não deram conta de conferir sentido à padronização das formas de entrada.</p>
- O controle do assunto para a entrada	<p>Para Gil Leiva (2008), o surgimento de bibliotecas especializadas e centros de documentação no início do século XX, sobretudo junto com as grandes bases de dados especializadas, demandou que o conteúdo dos documentos fosse detalhado, a partir de campos das bases de dados que permitissem identificar assuntos, pelo processo de indexação e, assim, facilitar a recuperação.</p> <p>De acordo com Campos e Gomes (2003), para Ranganathan, o conhecimento é contínuo, resultando em um universo infinito de assuntos. Segundo Kaula (1982), a análise conceitual deve ser “esgotada na concepção de um esquema de classificação”, mas ela tem início no plano das ideias. Segundo o mesmo autor, identificar um conceito na “posição do universo dos assuntos” é tarefa árdua.</p> <p>Na indexação, ocorre a identificação de aspectos que deem conta de uma representação mais especializada, conforme foi se estabelecendo na <i>Index Society</i> e, após, no contexto contemporâneo, no texto dos Princípios UNISIST e nas normas de indexação que dele resultaram.</p>
- Instrumento adotado	<p>De forma ampla, sistemas de classificação e tesauros são sistemas conceituais, estruturados a partir de conceitos e suas características. Para Dahlberg (1978), a análise conceitual consiste na identificação das características desses conceitos e das relações que podem ser estabelecidas.</p>

Fonte: elaborado pela autora.

Quadro 8 – Categorização do Terceiro Momento de Construção da Noção de Assunto: a abordagem europeia no século XX

Categorias	Referencial teórico da tese
- Estruturação do assunto	<p>Otlet (2018) considerou como unidade elementos do texto, o conteúdo e o suporte físico do documento, de forma integrada, não isolada, mas interagindo com a coleção.</p> <p>Sobre a forma de repertoriar o catálogo, uma das possibilidades que Otlet indicou foi a sistematização por assunto, ação que “[...] tem por objetivo dar a conhecer as obras existentes ou possuídas pela biblioteca sobre uma matéria determinada: ciência, técnica, questões sociais, personagens, época histórica, etc.” (OTLET, 2018, p. 471). A sistematização por assunto atrela-se à classificação.</p> <p>Para Gardin (1964, p. 14), a análise é a etapa necessária à produção do resumo e à indexação, pois resultam em produtos substitutos dos documentos, portanto, de forma diferente do original, mas que o representam com o propósito de facilitar o acesso à informação científica.</p> <p>Se, por um lado, a Análise Documentária abarca diferentes aspectos, quer seja enquanto área, quer seja do ponto de vista metodológico, ela também abrange métodos e processos referentes à Linguística Documentária (<i>Linguística Documental</i>, em espanhol). Lara (2011, p. 100), explica que “atualmente, a Análise Documentária faz parte de um vasto conjunto dos estudos que podem ser incluídos no subcampo da Linguística Documentária que tende a integrar, por sua vez, o conjunto dos estudos de Organização e Representação do Conhecimento.”</p> <p>García Gutiérrez (1990) aproxima o que ele propõe a partir do Estruturalismo. Vale-se dessa vertente desenvolvida por Ferdinand de Saussure para propor o estudo descritivo da língua no sentido metodológico e não ideológico.</p>
- O controle do assunto para a entrada	<p>Em relação à organização do RBU, compunha-se de fichas em formato universal, ou seja, fichas padronizadas, divididas em duas partes: uma por autor e outra por assunto (OTLET, 2018, p. 632).</p> <p>Na ficha, foram apresentados atributos de identificação do documento que facilitassem a composição e organização das informações, o que possibilitou o acesso e a circulação dessas informações. Nesse contexto, o objeto reflete o processo que o gerou. A ideia de análise e síntese é colocada como fundamental para chegar a uma representação do documento e uma configuração entre unidade física e intelectual, operadas pela representação nas fichas, é proporcionada. Ademais, a identificação da unidade intelectual possibilitaria a construção do conteúdo do documento, portanto do aspecto temático, do assunto.</p> <p>Em relação ao texto do documento, ao ser indexado, ele será desestruturado, isto é, analisado e sintetizado para posteriormente ser representado e transformado em informação documentária. Primeiro, estrutura-se as informações selecionadas e, após, submetidas a um processo de combinação, são transformadas. Assim, um novo texto é gerado, no caso do resumo, e símbolos são produzidos, a partir de uma linguagem documentária, no caso da indexação (KOBASHI, 1994).</p>

- Instrumento adotado	<p>Já em relação ao Princípio da Multiplicação dos Dados (OTLET, 2018), a duplicação das fichas com o uso de um instrumento para organização, como a CDU, visava o acesso à informação por vários pontos.</p> <p>De acordo com o Gardin (1964), o quadro empírico da indexação necessitou de alterações que transpusessem a combinação de palavras-chave para a representação do conteúdo de documentos especializados, e para isso houve a exigência de formalizar procedimentos abarcados pela análise. Se por um lado, a automação pôde propiciar equipamentos que geraram rapidez na recuperação, por outro, o desenvolvendo das técnicas de análises linguísticas (<i>analyses linguistes</i>) envolveu formas de transpor o texto científico de uma linguagem natural para a linguagem documentária, que é uma linguagem construída artificialmente. Assim, a terminologia científica de um domínio, o vocabulário normalizado e organizado, a sintaxe documentária e a linguagem modulada de forma simbólica com vistas às questões de busca auxiliaram o desenvolvimento do SYNTOL e da parceria dos estudos sobre <i>thesaurus</i> em <i>Cambridge</i>, intitulado <i>Language Research Unit</i>, com J. Needdham (GARDIN, 1964, p. 40).</p> <p>Cintra, Tálamo, Lara e Kobashi (1994) evidenciam o texto e o discurso no que se refere à construção de descritores no âmbito da atividade terminológica. Para tanto, consideram que o texto corresponde a uma unidade de sentido, coerente e coesa, sendo um objeto acabado, portador de uma estrutura e de um léxico. Já discurso consiste em uma atividade comunicativa que se manifesta através do texto. Não está acabado, pois é um objeto em construção, cuja realização se dá na interlocução. Nesse sentido, para a construção de tesouros, texto e discurso devem ser considerados, pois os textos relacionam os elementos integrantes do discurso, mas a operacionalização torna-se um desafio.</p>
-----------------------	--

Fonte: elaborado pela autora.

Pelo o que foi exposto nas seções e, conforme a sistematização do quadro 6, relativo ao primeiro momento em que se discutiu a vertente anglo-saxã, pode-se inferir que em Cutter, através da influência advinda de Panizzi e Jewett, a construção do assunto decorre da entrada no catálogo dicionário através de um nome. Esse assunto é nomeado a partir das listas de cabeçalhos de assunto por meio da atividade de palavras extraídas do próprio texto. Sendo assim, o assunto nomeado a partir das listas consiste em cabeçalho autorizado no catálogo. A preocupação com o nome voltou-se para a morfologia da palavra, articulando o ato de nomear e seu significado em substantivos, adjetivos, verbos, dentre outros. Esse aspecto pressupõe um universo linear, ordenado e estático para os assuntos.

Para alcançar um padrão que dê conta da equivalência do conteúdo (do texto) do documento, com o que o usuário necessita, Cutter, através da *Rules*, construiu regras (369 ao total), sendo que algumas referiam-se a uma rede de remissivas, mas com limitações de ordem econômica e semântica. A entrada do assunto de forma direta no catálogo dicionário marcou o que se entendeu na época como forma de acesso por demanda de uso, pois de acordo com Cutter, o assunto deveria ter aceitação do público para ser nomeado.

Essa forma de entrada direta do assunto (estruturação do assunto) tem a ver com as palavras escolhidas e sua ordem, o que leva a sua disponibilização no catálogo dicionário envolverem aspectos morfológicos e de sintaxe das palavras para a uniformização das fichas (o controle do assunto para a entrada). Esses aspectos influenciaram o instrumento adotado para as entradas por assunto, no caso, as listas de cabeçalhos de assunto.

O quadro 7, correspondente ao segundo momento teórico da pesquisa, envolve o desdobramento do primeiro momento sobre a vertente anglo-saxã. Kaiser, Ranganathan e os pesquisadores do *Classification Research Group* avançaram em relação ao que foi proposto por Cutter. A decomposição dos assuntos decorreu de uma fórmula pré-estabelecida em Kaiser e em Ranganathan. Para Kaiser, o concreto e o processo, e também, o lugar, são a base da fórmula; para Ranganathan, a noção de faceta é essencial e a fórmula constitui-se pelo acrônimo PMEST (em português, Personalidade - Matéria - Energia - Espaço - Tempo). Paralelo a isso, estudou-se a entrada dos assuntos nos índices e o processo de indexação tendo em vista uma representação mais especializada. Os autores da primeira parte do século foram mais longe quando propuseram categorias que já apresentavam algum tipo de base linguística. Eles estavam observando como se constitui um enunciado. O assunto passa a estruturar-se na forma de enunciados, a partir dos quais, uma rede semântica pode ser elaborada. No entanto, o enunciado não foi privilegiado no processo posterior de normalização da análise de assunto, como é possível inferir na abordagem dos Princípios de Indexação do UNISIST e na Norma

Brasileira de Análise de Assunto. Na análise sobre esse momento, as categorias evidenciam: estruturação do assunto considerando as fórmulas, o controle do assunto para a entrada no índice e, o instrumento adotado, como o tesauro, embasado por relações semânticas, diferentemente das listas de cabeçalhos de assuntos, que são caracterizadas pela ordenação alfabética.

No quadro 8, foi explorado o terceiro momento, relativo à segunda vertente – a europeia – que possui muitas particularidades em relação à anglo-saxã, evidenciando-se métodos que buscam ser mais rigorosos de modo a fundamentar a etapa da análise, segundo aspectos que caracterizam um texto. O avanço teórico e prático dos três Princípios de Otlet (da Monografia, da Continuidade e da Multiplicação dos Dados) e a questão do assunto posto por ele, ao mesmo tempo que desenha o seu sistema de informação, também considera o mapeamento do conhecimento produzido no contexto de um sistema classificatório. Algumas décadas depois, Gardin trata do texto científico, que ao ser indexado, passa por um processo de tratamento semântico e, para atender às questões de busca, necessita de formalização do processo através das expressões do conteúdo dos textos. O objetivo consistiu em isolar o sentido do texto, pois essa operação proporcionaria a representação do conteúdo dos textos, ou, a significação desses textos.

Para o autor, esse processo semântico, que envolve a representação dos textos científicos, necessita de um léxico ou de uma linguagem especializada construída artificialmente para o tratamento de informações de um domínio específico. O pensamento de Gardin, com a proposição da análise documentária, culmina com a passagem do texto original à representação e, para isso, considerou necessária uma linguagem com autonomia. Nesse sentido, seus estudos foram inovadores e se diferenciaram, sobretudo dos instrumentos tradicionais de representação da informação, como as listas de cabeçalhos de assunto. Ademais, em relação a todo o processo de representação, as etapas foram estruturadas buscando-se contemplar o modo que devem ser realizadas na perspectiva de um processo linguístico.

Mais tarde, García Gutiérrez sugere uma Teoria da Documentação com base na Linguística. Com isso, os estudos voltam-se à fundamentação da linguagem documentária, com destaque para o tesauro, e embasados pelos estudos da linguagem. No entanto, ao pensar a comunicação da informação, deve-se partir de uma linguagem que tenha sentido tanto em relação aos documentos quanto em relação ao público. Desse modo, os estudos passaram a ser amparados também pela Terminologia, para além do controle que uma lista de palavras proporciona. Para tanto, a ideia de uma representação, que começou a ser buscada com Kaiser

e Ranganathan, agrega às fórmulas propostas, a necessidade de análise e síntese do texto, para a estruturação do assunto. Paralelo a isso, o controle do assunto para a entrada parte de uma desestruturação do texto, cujo processo seguinte é o de estruturação por meio da combinação das unidades identificadas. O instrumento adotado é responsável por modular de forma simbólica a representação temática com vistas às questões de busca, e remete ao tesauro, que também foi foco do momento anterior. O tesauro é o responsável por proporcionar relações que permitam a representação e não a reprodução do assunto.

Pode-se sintetizar os três quadros do seguinte modo: no primeiro momento (quadro 6) – trabalhou-se na estruturação do assunto em relação ao catálogo; no segundo momento (quadro 7) – essa estruturação tem atenção voltada aos índices; e, no terceiro momento (quadro 8) – observou-se a estruturação do assunto com o aporte da Linguística, sob a consideração de que esse aporte permite um nível maior de elaboração.

É preciso considerar que a investigação não privilegia nenhum momento teórico analisado. No entanto, a vertente europeia, que é a mais recente, evidencia o desenvolvimento de metodologias rigorosas, centrando-se no processo como modo de obter produtos consistentes.

Pensando a construção do assunto de maneira conceitual, foi possível demonstrar que ele, incluindo os contrapontos entre os autores, é trabalhado segundo uma certa forma, compõe uma entrada de um índice, é um item de representação de um documento, e se organiza, em conjunto, em um instrumento que se aproxima mais ou menos de uma linguagem, a depender da vertente em questão.

Dada essa etapa inicial de análise, segue-se com o objetivo de abordar o modelo FRASAD e, para tanto, ele e alguns estudos são trabalhados.

7.2 O MODELO FRASAD: APRESENTAÇÃO E ESTUDOS

Como a noção de assunto compõe o processo global de confecção de catálogos, a *International Federation of Library Associations and Institutions* (IFLA) inseriu a modelagem conceitual nesse contexto, visando apresentar uma forma de fundamentar a representação da informação documentária, por meio dos *Functional Requirements for Bibliographic Records* (FRBR, em português, Requisitos Funcionais para Registros Bibliográficos). No ano de 1998, a IFLA publicou o relatório sobre o modelo FRBR, que abrange conceitualmente o universo bibliográfico, ressaltando o usuário das bases de dados

bibliográficas. O FRBR foi apresentado como um esquema para aperfeiçoar a organização da informação documentária em catálogos e bases de dados, pois direciona e potencializa as estruturas e relações bibliográficas, refletindo assim, no modo de representação dos documentos.

De acordo com o relatório da IFLA sobre o FRBR, a proposta do modelo decorre, em um primeiro nível, no fornecimento de um quadro estruturado com o intuito de relacionar dados especificados em registros bibliográficos às necessidades informacionais de usuários. Também recomenda, como segunda abordagem, a estrutura básica de funcionalidade para a criação de registros que sejam criados por entidades bibliográficas nacionais.

No contexto da proposição dos modelos conceituais pela IFLA, iniciada pelo modelo FRBR, fala-se em família FRBR, a qual é constituída pelo próprio FRBR, o FRAD (*Functional Requirements for Authority Data*, em português, Requisitos Funcionais para Dados de Autoridade) e o FRSAD. Este último – foco da pesquisa – tem a finalidade de operar em nível conceitual os aspectos referentes às entidades que representam os assuntos dos documentos.

Os estudos iniciais sobre o FRSAD, antes denominado *Functional Requirements for Subject Authority Records* (FRSAR, em português, Requisitos Funcionais para Registros de Autoridade de Assunto), foram constituídos em abril de 2005, pelo grupo que estava envolvido com o desenho do modelo FRAD.¹³¹ Isto porque, notou-se a necessidade de investigar as possíveis vinculações propostas conceitualmente entre o FRBR e o FRAD com as relações bibliográficas dos dados de autoridade de assunto.

Percebe-se que da primeira nomenclatura, *Functional Requirements for Subject Authority Records* (FRSAR), para a atual, *Functional Requirements for Subject Authority Data* (FRSAD), houve uma mudança na palavra ‘registros’ (*records*) para ‘dados’ (*data*). ‘Registros de autoridade’ refletem um conjunto estruturado de atributos enquanto unidades informacionais na área de Biblioteconomia e Ciência da Informação, diferente de ‘dados de autoridade’. É presumido que ‘dados’ seja preterido pelo contexto digital dos catálogos e bases de dados, mesmo que os termos que caracterizam a área tenham se desenvolvido no contexto analógico.

O documento que estrutura o modelo FRSAD, aprovado em junho de 2010, é um relatório oriundo da atual Seção de Análise de Assunto e Acesso (antiga Seção de

¹³¹ Passou por várias denominações, sendo primeiramente nomeado *Functional Requirements and Numbering of Authority Records* (FRANAR), depois *Functional Requirements Authority Records* (FRAR) até ser nomeado *Functional Requirements for Authority Data* (FRAD).

Classificação e Indexação) da IFLA, editado por Marcia Lei Zeng, Maja Žumer e Athena Salaba. A alteração do nome da seção ocorreu em setembro de 2016 ¹³² (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2016b, on-line; ŽUMER; ROE, 2016).

No início de 2016, outro movimento de mudança é constatado, pois em fevereiro, a Seção de Catalogação anunciou a integração dos três modelos (FRBR, FRAD e FRSAD) em um único, denominado *FRBR-Library Reference Model* (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2016a, on-line). Atualmente, o nome do modelo integrado é *IFLA Library Reference Model (LRM)*, conforme consta na última publicação de dezembro de 2017.¹³³

O objetivo do FRSAD de tratar a noção de assunto presume a necessidade de teorização, para assim indicar o papel que essa noção pode exercer na área. Essa compreensão resvala na integração do FRSAD no LRM, pois se presume que uma alteração de tamanha amplitude deve ser embasada por discussão e amadurecimento teórico. Isso porque, o FRSAD, juntamente com o FRBR e o FRAD, foram unidos no LRM antes mesmo da validação que permitisse a difícil tarefa de atribuir sentido teórico e prático a estes modelos enquanto saberes científicos da Biblioteconomia e Ciência da Informação.

Diante dessas mudanças, compreender quais direcionamentos sustentam a noção de assunto no modelo FRSAD continuou necessário. Coube continuar a pesquisa com o modelo FRSAD, pois ele compõe o desenho da área, tendo como base, por exemplo, a concepção proposta na Declaração dos Princípios Internacionais de Catalogação de 2009 (*Statement of International Cataloguing Principles*), com versão atualizada no ano de 2016 (STATEMENT..., 2009; 2016).

Levando em consideração que os catálogos de bibliotecas são sistemas de informação que comportam o contexto da representação de documentos, a IFLA apropriou-se do modelo conceitual proveniente da Ciência da Computação para os registros de bases de dados e, propôs a investigação do modelo Entidade-Relacionamento (E-R) a fim de priorizar as

¹³² We propose “Subject Analysis and Access” as the new name of our section. It is short and a good compromise among the different proposals identified so far. In addition to the end-user focus (“Access”) we also include the role of the professional. “Analysis” refers to the work and thought that goes into determining “aboutness” (or “determining what a document is about”) and then creating subject access – be it assigning classification or headings, developing thesauri, or devising tools and other automatic ways to achieve subject access. (ŽUMER; ROE, 2016, p. 21).

¹³³ INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **IFLA Library Reference Model (LRM)**. December 2017. Disponível em: <https://www.ifla.org/publications/node/11412>. Acesso em: 28 dez. 2018.

necessidades dos usuários dos catálogos (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 1998).

O modelo E-R, desenvolvido por Peter Chen, embasa estruturalmente a família FRBR e é utilizado para construir e projetar bases de dados relacionais. É conceitual, pois implica a modelagem de ‘coisas’, processos ou abstrações objetivando sintetizar e sistematizar sistemas, teorias ou fenômenos com intuito de ser aplicável. Identifica-se no modelo E-R as entidades, os atributos e os relacionamentos. Considera-se entidade como algo, ou alguma coisa que pode ser distintamente identificada, relacionamento como a associação entre as entidades, e atributos como características próprias de uma entidade (CHEN, 1990).

Dessa forma, um registro de informação com base na família FRBR é composto por entidades que se relacionam, e as entidades são compostas por atributos. O modelo FRSAD representa a análise das entidades que constituem os assuntos de uma obra, baseado em determinar um marco que proporcione o entendimento dos registros bibliográficos e os respectivos nomes de autoridades de assunto com o objetivo de atender as buscas dos usuários (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2010). O modelo faz perceber que as entidades e os relacionamentos já estão dados no universo bibliográfico, ou seja, não desenvolve as ideias que subsidiaram a proposição desses elementos.

O modelo destaca os dados de autoridade de assunto que constituem informação sobre os arquivos de assuntos autorizados (*information about subjects from authority files*). Esses arquivos são constituídos pelo trabalho de controle de nomes de autoridades que asseguram consistência na representação de um valor - um nome de uma pessoa, um nome de lugar ou um termo ou código, representando um assunto - nos elementos usados como pontos de acesso na recuperação de informações.

No modelo, há, ainda, a indicação da catalogação ou indexação (*cataloguing or indexing*) como processos subsidiários de criação do ponto de acesso e a menção aos instrumentos desses processos que respondem pelos assuntos. Mas, apesar do foco ser o mesmo, há distinções, seja entre os processos, seja nos instrumentos, quer por não serem linguagens de fato (caso das listas de cabeçalhos de assunto), ou por, além disso, discorrerem por outros atributos que não somente assuntos (caso dos sistemas de classificação bibliográfica).

Após o escopo, são apresentados no modelo os termos *aboutness* e *ofness*. Para o primeiro, há uma breve explicação a partir de Hjørland (1992), na qual, sob o ponto de vista filosófico nominalista ou idealista e realista, encontra-se uma possível justificativa para a “[...]”

análise apropriada do *aboutness* – ou seja, a relação entre uma obra e a matéria do assunto.”¹³⁴ (INTERNATIONAL..., 2010, p. 10, tradução nossa). No ponto de vista nominalista ou idealista, o *aboutness* deve ser concebido concretamente, como um nome. Mas, pelo realismo é o contrário, pois o assunto seria independente das expressões linguísticas utilizadas para nomeá-lo. Quanto ao segundo termo, o modelo:

[...] reconhece que qualquer declaração sobre “*ofness*” de uma obra pode ser uma **declaração de assunto** e é provável que seja o alvo da pesquisa de um usuário no catálogo. O modelo FRSAD é, portanto, aplicável a situações em que as declarações *ofness* (representação) estão incluídas no acesso por assunto. (INTERNATIONAL..., 2010, p. 11, negrito do modelo, tradução nossa).¹³⁵

É importante salientar que a definição apresentada para *aboutness* e para *ofness* no FRSAD remetem, ambas, à ideia de representação que, por algum tipo de controle, indica o significado dos assuntos. Entretanto, o termo *ofness* não é claramente definido. Ademais, não se distingue suficientemente a ideia de *aboutness* em relação à de *ofness*.

A família FRBR é composta por três modelos, totalizando 11 entidades. Os modelos FRAD e FRSAD foram propostos a partir das entidades dos grupos 2 e 3 do modelo FRBR. Assim, o modelo FRBR é constituído por: grupo 1 - as entidades dos produtos de trabalho intelectual ou artístico, sendo: obra, expressão, manifestação e item, em inglês, *work, expression, manifestation, item*; grupo 2 - as entidades vinculadas aos responsáveis pela produção do conteúdo, disseminação e guarda, sendo: pessoa, entidade coletiva e família,¹³⁶ em inglês, *person, corporate body, family* (as quais formam o modelo FRAD); e grupo 3 - as entidades que exercem os assuntos de uma obra, a saber: conceito, objeto, evento e lugar, em inglês, *concept, object, event, place* (que formam o modelo FRSAD) (INTERNATIONAL..., 2010).

O objetivo do modelo FRSAD, segundo o grupo de trabalho que o produziu, é estruturar conceitualmente as entidades do grupo 3 e, como essas entidades se relacionam com os dados de autoridade assunto, e as necessidades dos usuários. Além disso, o modelo também busca auxiliar na avaliação do potencial de uso e compartilhamento internacional dos

¹³⁴ “[...] appropriate analysis of aboutness – i.e., the relation between a work and its subject matter.” (INTERNATIONAL..., 2010, p. 10).

¹³⁵ “[...] recognizes that any statement on the “*ofness*” of a work can be a **subject statement** and is likely to be the target of a catalogue user’s search. The FRSAD model is therefore applicable to situations where *ofness* statements (depiction) are included in subject access. (INTERNATIONAL..., 2010, p. 11, negrito do modelo).

¹³⁶ Inserido a partir do modelo FRAD em 2009. Antes havia duas entidades nesse grupo: pessoa e entidade coletiva.

identificação vinculada ao *thema*, que possui como denominação um ou mais *nomen(s)*. Portanto, *thema* se define como qualquer entidade utilizada como assunto de uma obra e, *nomen*, como qualquer signo ou sequência de signos (caracteres alfanuméricos, símbolos, sons, etc) mediante os quais conhece, refere e aborda o *thema* (INTERNATIONAL..., 2010).

O modelo ainda considera *thema* uma superclasse de todas as entidades da família FRBR.¹³⁷ Quanto a isso, como consta no modelo, o grupo de trabalho (*working group - FRSAD WG*), responsável pelo FRSAD, “[...] precisou distinguir *thema* a partir do conceito de entidade definido no FRBR porque *thema* é uma superclasse de todas as entidades do FRBR [...]” (INTERNATIONAL..., 2010, p. 16, tradução nossa).¹³⁸

De fato, a superioridade atribuída ao *thema* permite aproximação ao ‘conceito’, como vem sendo trabalhado na Biblioteconomia e Ciência da Informação, mas que, por sua vez, apresenta significado diferente da entidade ‘conceito’ do modelo, que indica uma ideia, em oposição a coisas materiais. A entidade conceito do modelo é adotada para contemplar áreas de conhecimento, disciplinas, escolas de pensamento, teorias, dentre outros aspectos, que junto às outras três (objeto, evento e lugar) coloca em relevo o que se entende como pertinente para a representação dos documentos.¹³⁹

Em certa medida, é possível aproximar as quatro entidades do FRSAD (conceito, objeto, evento e lugar) às unidades que compõem as fórmulas de Kaiser e Ranganathan. Entretanto, as propostas destes dois teóricos eram mais especificamente voltadas à área, já que configuravam enunciados, representativos dos conteúdos dos textos, que funcionassem, portanto, para as operações de busca.

Quanto aos termos assunto, tema e conceito no modelo, são por vezes empregados como sinônimos, mas a questão não é discutida, adotando-se *thema* em todas as situações. Assim,

Para *thema*, outros possíveis termos (em inglês) são incluídos, como ‘assunto’, ‘tópico’ e ‘conceito’; no entanto, mesmo as discussões no Grupo de Trabalho provocaram a existência de visões diferentes sobre essa

¹³⁷ “[...] *thema* is a superclass of all FRBR entities [...]” (INTERNATIONAL..., 2010, p. 16).

¹³⁸ “[...] needed to distinguish *thema* from the previously defined FRBR entity concept because *thema* is a superclass of all FRBR entities [...]” (INTERNATIONAL..., 2010, p. 16).

¹³⁹ A entidade ‘tempo’ não está prevista no FRSAD. Interessante se figurasse com o objetivo de representar diferentes abordagens temáticas, como por exemplo, os conceitos gregos de *chronos* (tempo cronológico) e *kairos* (tempo baseado em eventos), que veiculam diferentes conceitos. Esse ponto foi destacado na tese de: MILANI, Suellen Oliveira. **Bias na representação de assunto:** uma discussão de oposições binárias nos Functional Requirements for Subject Authority Data (FRSAD). 2014. 134 f. Tese (doutorado)-Universidade Estadual Paulista, Faculdade de Filosofia e Ciências de Marília, 2014.

granularidade (alguns veem ‘assunto’ e ‘tópico’ como sinônimos, enquanto outros veem ‘tópico’ como um componente do ‘assunto’).¹⁴⁰ (INTERNATIONAL..., 2010, p. 16, tradução nossa).

A escolha pelo latim para nomear as duas entidades ‘superiores’ é indicada no modelo do seguinte modo: adota-se “[...] *thema* (plural *themata* ou *themas*) e *nomen* (plural *nomina* ou *nomens*), porque não possuem em nosso contexto significado pré-existente, são culturalmente neutros e não requerem tradução.”¹⁴¹ (INTERNATIONAL..., 2010, p. 16, tradução nossa).

No entanto, não cabe falar em neutralidade quando se faz menção à cultura; a expressão ‘culturalmente neutro’ é em si contraditória. De qualquer modo, algumas palavras latinas são adotadas em vários idiomas, ou seja, elas são usadas nestes idiomas, portanto, apresentam significados. Na base desta proposta de uso de *thema* e de *nomen*, no entanto, está o problema da ausência de exploração da terminologia desenvolvida pela área, que poderia contribuir para fundamentar o modelo.

Diante da aparente neutralidade que o modelo FRSAD sinaliza, como citado acima, alguns pontos são propostos nesta pesquisa como possível caminho para sua fundamentação. Deste modo, o modelo:

- não define qual estrutura teórica é aporte da sua criação;
- não explora conceitualmente a questão do assunto;
- sinaliza a existência de vínculo do assunto com *thema* e *nomen*, entretanto não fundamenta teoricamente a relação;
- emprega os termos assunto, tema e conceito como sinônimos de *thema*, sem buscar elaborar um escopo teórico a respeito.

Os pontos elencados não são novos. Pode-se dizer que o modelo conceitual FRSAD retoma antigos problemas, sem indicá-los como tal e desenvolvê-los.

Žumer, Salaba e Zeng (2007) apontam a demanda de um modelo conceitual que desse conta do aspecto temático. Destaca-se que prevaleceu um entendimento empírico para a

¹⁴⁰ “For *thema*, other possible (English) terms include ‘subject’, ‘topic’, and ‘concept’; however, even discussions within the Working Group proved that there are very different views on granularity (some see ‘subject’ and ‘topic’ as synonyms, while others see ‘topic’ as a component of ‘subject’).” (INTERNATIONAL..., 2010, p. 16).

¹⁴¹ “[...] *thema* (plural *themata* or *themas*) and *nomen* (plural *nomina* or *nomens*), because they have no pre-existing meaning in our context, are culturally neutral and do not require translation.” (INTERNATIONAL..., 2010, p. 16).

fundamentação, ou seja, o aspecto do uso, quer de quem trabalha profissionalmente, quer de quem pesquisa autoridade assunto, em diferentes espaços, conforme é possível constatar:

O Grupo de Trabalho FRSAR, sentiu necessidade de definir as tarefas do usuário sendo necessário um estudo real e dois foram conduzidos. O primeiro foi um estudo piloto em 2006 durante a *Semantic Technologies Conference* (San Jose, Califórnia, EUA). A maioria dos participantes do estudo eram ou criadores de ferramentas semânticas, incluindo vocabulários controlados, taxonomias e ontologias, ou desenvolvedores e gestores de sistemas de tecnologia semântica. O segundo estudo foi uma pesquisa internacional enviada para profissionais da informação em todo o mundo durante os meses de maio a setembro de 2007. Entre os participantes foram incluídos criadores de registro de autoridade, criadores e gestores de vocabulários, catalogadores, bibliotecários de metadados e de referência, entre outros. Os participantes foram solicitados a descrever seu trabalho e a utilização de dados de autoridade assunto em diferentes contextos, incluindo catalogação/criação de metadados, o trabalho de autoridade assunto, e pesquisas ou auxílios a outras pessoas por buscas de informações bibliográficas. Os resultados destes estudos enriqueceram nossa compreensão do uso de dados de autoridade assunto informando e confirmando as tarefas do usuário no FRSAR.¹⁴² (ŽUMER; SALABA; ZENG, 2007, p. 490, tradução nossa).

As autoras também desenvolveram e divulgaram pesquisas sobre o FRSAD, algumas anteriores à versão do relatório de 2010, como segue. Zeng e Žumer (2009a) realizaram o mapeamento do modelo FRSAD e outros modelos abstratos, apresentando-o na *ISKO UK Conference*, em Londres. Já Zeng e Žumer (2009b) realizaram introdução ao FRSAD e o mapearam com uso de *Simple Knowledge Organization System* (SKOS, em português, Sistema de Organização do Conhecimento) e Žumer, Zeng e Salaba (2015) indicaram aplicação do FRBR e FRSAD especificamente na Classificação Decimal Universal (CDU). Menos no primeiro caso, as iniciativas realizam, em alguma medida, a modelagem de assuntos, mas na perspectiva de instrumentos específicos adotados para tal.

Mesmo não sendo foco do modelo considerar a perspectiva do vocabulário, é explicitado que “até certo ponto, a granularidade de um *thema* também depende do

¹⁴² The FRSAR Working Group felt strongly that, in order to define user tasks, an actual user study was necessary, and two studies were therefore conducted. The first was a pilot study at the 2006 Semantic Technologies Conference (San Jose, California, USA). Most study participants were either creators of semantic tools, including controlled vocabularies, taxonomies and ontologies, or developers and managers of semantic technology systems. The second study was an international survey sent to information professionals throughout the world during the months of may-september 2007. Participants included authority record creators, vocabulary creators and managers, catalogers, metadata librarians, and reference librarians among others. Participants were asked to describe their work and their use of subject authority data in different contexts, including cataloging/metadata creation, subject authority work, and searching or helping others search bibliographic information. The results of these studies enriched our understanding of subject authority data use and informed and further confirmed the FRSAR user tasks. (ŽUMER; SALABA; ZENG, 2007, p. 490).

vocabulário controlado usado para sua(s) denominação(ões).” (INTERNATIONAL..., 2010, p. 18, tradução nossa).¹⁴³ Mas, nesse aspecto, nota-se ser o contrário, pois o vocabulário controlado deve ser construído a partir do modo como se quer trabalhar um *thema* (ou seja, um conceito). Essa abordagem do modelo reflete o pensamento de uma linguagem documentária como uma lista de denominações, usada apenas para padronizar, a exemplo das incipientes listas de cabeçalhos de assunto. Essa situação ocorre quando uma linguagem documentária pronta é adotada, portanto, sem contemplar as especificidades do sistema de informação em questão.

Delsey (2005) aponta que os principais aspectos dos modelos FRBR e FRANAR devem ser reexaminados como parte de uma análise mais intensa em relação ao acesso por assunto e sugere maneiras de abordar o refinamento e a extensão dos modelos. Traz questões sobre o atual âmbito de cobertura representado pelas entidades definidas nos modelos, a necessidade de esclarecer os atributos de entidades adicionais e a representação tanto das relações semânticas, quanto sintáticas refletidas em tesouros, listas de cabeçalhos de assunto, esquemas de classificação e outros instrumentos de indexação. Cabe salientar que o artigo foi publicado por Delsey no mesmo ano de formação do grupo de trabalho da IFLA sobre o FRSAD. Assim, é interessante que, já no momento do início da criação do grupo de trabalho sobre o FRSAD, o autor tenha mencionado questões que se mostram em aberto no modelo, como as relações semânticas e sintáticas em instrumentos, como listas de cabeçalhos de assunto, tesouros e esquemas de classificação, haja vista a diferença entre eles. Ao mesmo tempo, chama a atenção a qualificação, pelo autor, de todos estes instrumentos como voltados à indexação.

Por sua vez, Gemberling (2010) afirma que o tratamento de assunto na família FRBR tem atraído menos atenção em relação a outros aspectos. Destaca que proposições para a elaboração (do que ele chama) de categorias – conceito, objeto, evento e lugar – são uma forma de corroborar o aumento da complexidade semântica e que propostas para trabalhos promissores devem ultrapassar a simplicidade da entidade *thema*. Nessa direção, o autor aborda e discute o problema da ambiguidade na criação de cabeçalhos de assunto.

Crítico do FRSAD, Furner (2012) argumenta que não existe constatação na literatura de pressupostos ontológicos subjacentes à criação do modelo e que não há clareza acerca do raciocínio por trás das escolhas feitas. O autor propõe discussão sobre os pressupostos ocultos ao desenho do modelo FRSAD, bem como alternativas para analisar as implicações das

¹⁴³ “To some extent the granularity of a *thema* also depends upon the controlled vocabulary used for its appellation(s).” (INTERNATIONAL..., 2010, p. 18).

escolhas do grupo de trabalho da IFLA. Aponta que o modelo foca nos requisitos de produção e organização das descrições dos assuntos e sua construção foi orientada pela compilação de resoluções da prática, por razões pragmáticas, e não por questões filosóficas básicas sobre assuntos de obras e fenômenos relativos.

Corrobora-se com Furner (2012) que o FRSAD não apresenta uma posição filosófica sobre o *aboutness* na modelagem de dados. Segundo ele: “Em última análise, o Grupo de Trabalho FRSAR não assume uma posição de natureza filosófica sobre o *aboutness*, mas olha o problema do ponto de vista do usuário.”¹⁴⁴ (FURNER, 2012, p. 495; INTERNATIONAL..., 2010, p. 11, tradução nossa).

Todavia, ainda de acordo com Furner (2012), a decisão de indicar o ponto de vista para o olhar do usuário já é o resultado de uma posição filosófica. Além disso, indica dois conceitos interpretados e aplicados frequentemente sem reflexão: *aboutness* e subjetividade.

Gemberling (2016) resgata Furner (2012) para apontar a ausência de pressupostos subjacentes ao desenho do modelo FRSAD. Furner (2012) indica o debate na natureza dos conceitos gerais entre os nominalistas e realistas, argumentando que o FRSAD implica uma posição realista e, Gemberling (2016) preconiza a defesa do realismo, pois argumenta que a decisão sobre o debate filosófico não foi essencial para o grupo de trabalho criar e estruturar o FRSAD.¹⁴⁵

Melo e Bräscher (2014) apresentam o modelo conceitual FRSAD como uma tendência a seguir para a atividade de representação do assunto, discutindo o conceito e os relacionamentos semânticos no contexto da organização da informação. Analisam o modelo quanto aos seus objetivos e a forma que ele aborda o conceito (*thema*) e os relacionamentos semânticos (*thema-to-thema*). Portanto, as autoras entendem *thema* como conceito, confirmando o ponto de vista já exposto. Concluem que o FRSAD amplia as possibilidades de representação temática de uma obra, ao considerar como assunto todas as entidades da família FRBR.

Quanto à abordagem do modelo, o grupo de trabalho estabelece no relatório três termos de referência em relação à noção de assunto, aproximando-a do termo *aboutness*. O primeiro termo de referência é indicado da seguinte forma: “construir um modelo conceitual

¹⁴⁴ “Ultimately, the FRSAR Working Group does not take a philosophical position on the nature of aboutness; rather, it looks at the problem from the user’s point of view.” (FURNER, 2012, p. 495; INTERNATIONAL..., 2010, p. 11).

¹⁴⁵ “The FRSAD working group claimed not to take any philosophical position on the nature of subjects.” (GEMBERLING, 2016, p. 136).

de entidades do Grupo 3 dentro da estrutura FRBR como elas se relacionam ao *aboutness* das obras.”¹⁴⁶ (INTERNATIONAL..., 2010, p. 7, tradução nossa).

Os outros dois termos de referência indicados no modelo são: proporcionar estruturação para relacionar dados, que englobem registros de autoridade de assunto com as necessidades dos usuários e, ajudar a evolução do potencial de utilização do intercâmbio internacional de dados de autoridade de assunto, tanto em bibliotecas quanto em outros setores (INTERNATIONAL..., 2010).

É fato que o modelo referencia o *aboutness*, porém não fundamenta a noção de assunto e não realiza qualquer estruturação teórica sobre o entendimento de ambos, apontando-os como uma relação entre a obra e a matéria do assunto (*subject matter*). Sem desenvolver uma definição teórica, remete ao assunto principal (*subject main*) ou assunto nuclear (*subject core*) de uma obra.

Conforme apresentado, o modelo pauta duas supra entidades, sendo *thema* para indicação de qualquer entidade usada como assunto de uma obra e, *nomen*, para forma, como qualquer sinal ou sequência de sinais que um *thema* é conhecido, referido ou endereçado. Nesse mesmo caminho, no contexto do controle de autoridade assunto, um *nomen* deve ser a denominação de um *thema*, o que é apresentado no modelo pela figura 8 a seguir.

Figura 8 – Relacionamentos FRSAD



Fonte: *International Federation of Library Associations and Institutions* (2010, p. 15).

As setas com pontas duplas estipulam relação de vários para vários, pois qualquer obra pode possuir mais de um *thema* e qualquer *thema* pode ser o assunto de várias obras. O modelo toma como exemplo, o livro “Uma breve história do tempo: do *big bang* aos buracos negros”, de Stephen W. Hawking, e aponta que a obra possui vários *themas*, como: cosmologia, espaço e tempo, unificação da física, buracos negros, *big bang*, história do tempo, universo, etc. Nessa direção, outras obras também podem sustentar esses *themas* e,

¹⁴⁶ “To build a conceptual model of Group 3 entities within the FRBR framework as they relate to the *aboutness* of works.” (INTERNATIONAL..., 2010, p. 7).

para qualquer um dos apontados, existem outros possíveis *nomens* em outros idiomas e diferentes vocabulários controlados (INTERNATIONAL..., 2010).

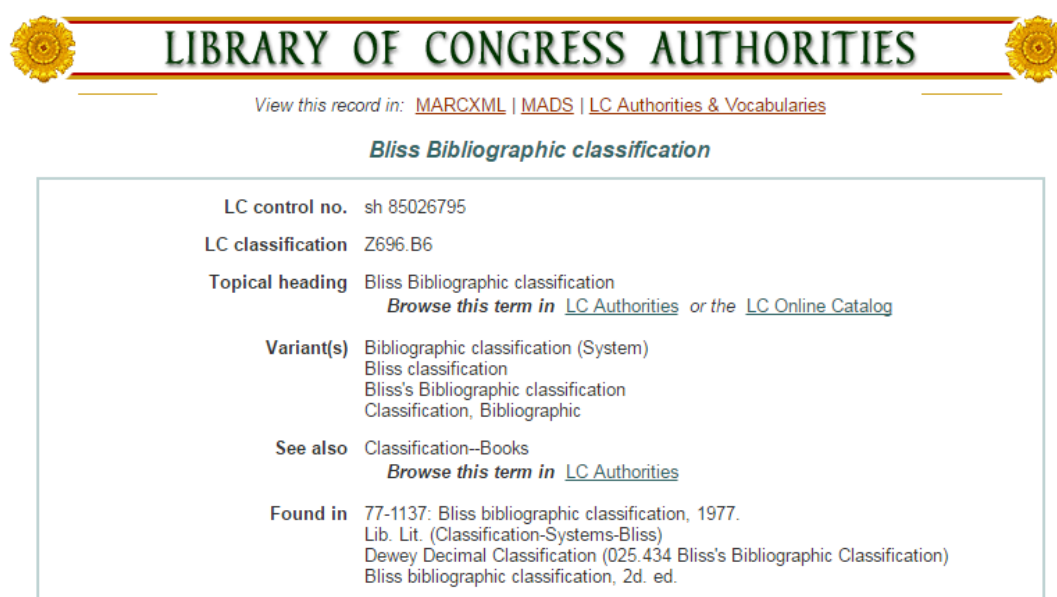
Há dois atributos gerais para *thema*, sendo tipo (*type*) e nota de escopo (*scope note*), os quais dependem do âmbito de aplicação segundo o modelo conceitual FRSAD. Encontra-se no referido modelo um conjunto mais amplo de atributos para *nomens*, como tipo de *nomen* (*type of nomen*), esquema (*scheme*), fonte de referência do *nomen* (*reference source of nomen*), representação do *nomen* (*representation of nomen*), linguagem do *nomen* (*language of nomen*), escrita do *nomen* (*script of nomen*), conversão da escrita (*script conversion*), forma do *nomen* (*form of nomen*), tempo de validade do *nomen* (*time of validity of nomen*), público (*audience*) e estado do *nomen* (*status of nomen*).

A fim de ilustrar os referidos atributos, seguem explicações: tipo de *nomen* estrutura a forma identificadora ao nome controlado; esquema é o próprio vocabulário controlado, como a Lista de Cabeçalhos de Assunto da *Library of Congress* e a Classificação Decimal de Dewey; fonte de referência do *nomen* refere-se a algum tipo de material especializado ou geral, como por exemplo, uma enciclopédia, portanto a fonte de consulta para a escolha; representação do *nomen* destina-se à forma de apresentação, como a alfanumérica, a visual, a sonora; linguagem do *nomen* é o idioma, se inglês, espanhol, português, entre outros; escrita do *nomen* baseia-se, por exemplo, se em braile, chinês, cirílica, dentre outras; conversão da escrita vai de encontro à mudança na escrita, como exemplo, documento escrito em chinês e convertido para o português; forma do *nomen* indica, por exemplo, uma abreviação ou uma fórmula; tempo de validade do *nomen* vincula-se com um período, uma data; público são os grupos de pessoas destinatárias do material, como crianças, universitários, estudantes de idiomas, dentre outros; e, por fim, estado do *nomen* é a forma como o referido é adotado, por exemplo, forma oficial ou provisória.

Alguns desses atributos para *nomen* podem ser encontrados na figura 9, que retrata o cabeçalho “*Bliss Bibliographic classification*” no catálogo de autoridade de assunto da *Library of Congress* (LC). Os atributos podem ser observados no cabeçalho tópico (*topical heading*) que possui símbolos ou modos de representação, como por exemplo, o número *Dewey Decimal Classification* (025.434 *Bliss's Bibliographic Classification*). Variantes (*variant(s)*) podem ser considerados tópicos que complementam o significado para o *nomen* e, a remissiva ver também (*see also*), é complemento para o *nomen*. Encontra-se a fonte (*Found in*), no caso fonte de referência, e o número de controle (*LC control no.*) como sendo o estado do *nomen*.

Caso se considere como *thema*, *classification bibliographic*, as outras formas como *bibliographic classification (system)*, *Bliss classification*, *Bliss's bibliographic classification* e *classification, bibliographic* também podem formar outros *themas*, que, elencados em sequência de *nomens*, tem valor em um catálogo ou base de dados, controlado pelo vocabulário.

Figura 9 – Cabeçalho “*Bliss Bibliographic classification*”



The image shows the header of the Library of Congress Authorities page for 'Bliss Bibliographic classification'. It features the library's logo and navigation links. Below the header, there is a table with the following information:

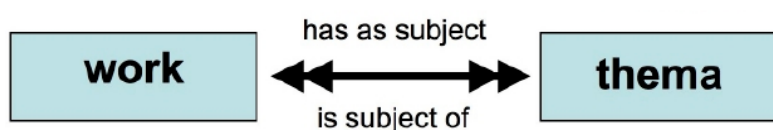
LC control no.	sh 85026795
LC classification	Z696.B6
Topical heading	Bliss Bibliographic classification <i>Browse this term in</i> LC Authorities <i>or the</i> LC Online Catalog
Variant(s)	Bibliographic classification (System) Bliss classification Bliss's Bibliographic classification Classification, Bibliographic
See also	Classification--Books <i>Browse this term in</i> LC Authorities
Found in	77-1137: Bliss bibliographic classification, 1977. Lib. Lit. (Classification-Systems-Bliss) Dewey Decimal Classification (025.434 Bliss's Bibliographic Classification) Bliss bibliographic classification, 2d. ed.

Fonte: *Library of Congress Authorities* (<https://lcn.loc.gov/sh85026795>).

Outros atributos adicionais podem ser definidos para implementação do modelo FRASAD, segundo dois tipos de relações: as relativas a entidades do mesmo tipo, como *nomen-to-nomen* e *thema-to-thema*, e as que reúnem diferentes tipos de entidades, como *work-to-thema* e *thema-to-nomen*. O segundo tipo, com diferentes entidades, demandam as especificações ressaltadas a seguir.

A relação *work-to-thema* ocorre ao se considerar que a obra tem como assunto o *thema* e o *thema* ser assunto da obra, conforme exemplificado na figura 10.

Figura 10 – Relação *work-to-thema*



Fonte: *International Federation of Library Associations and Institutions* (2010, p. 16).

Nesse caso, mesmo *thema* sendo considerado uma supra entidade em relação a todas as outras dos grupos 1, 2 e 3 (FRBR, FRAD e FRSAD), o relacionamento busca permitir a modelização das relações em um nível geral e abstrato. Isso também deve permitir que as entidades específicas do grupo 3 (conceito, objeto, evento e lugar) sejam universalmente aplicadas; o modelo elenca, como exemplo, o emprego realizado pelo instrumento *Faceted Application of Subject Terminology – FAST*.¹⁴⁷ Entretanto, dependendo do instrumento, a unidade constituinte representativa da noção de assunto pode ser a palavra, como no caso das listas de cabeçalhos, ou o termo, no caso dos tesouros. Além disso, não se pode misturar momentos distintos: a construção do instrumento e o processo de indexação.

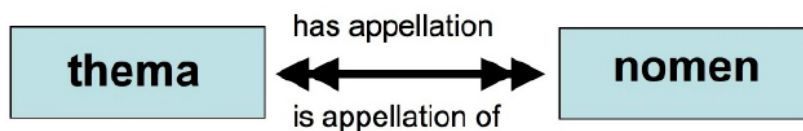
O modelo considera ainda que os *themas* interferem na variação do *aboutness* em nível de complexidade ou simplicidade, dependendo das circunstâncias da obra que pode ser expressa na relação um para um, considerando a forma simples, ou, na relação um para muitos, que transborda o *aboutness* a partir de vários *themas*. Então há comprometimento também neste aspecto, pois o vocabulário não pode distorcer o conteúdo identificado de um documento.

Isso é percebido no modelo:

Frequentemente, a complexidade de um *thema* está associada à complexidade do *nomen* pelo qual ele é representado. Uma vez que o modelo proposto introduz uma divisão clara entre *thema* (“a coisa”) e o *nomen* (“o rótulo” usado para se referir a ele) [...] ¹⁴⁸ (INTERNATIONAL..., 2010, p. 18, tradução nossa).

No outro tipo de relacionamento, *thema-to-nomen*, afirma-se que o *thema* tem como denominação um *nomen* e *nomen* ser denominação de um *thema*, retratado conforme a figura 11.

Figura 11 – Relação *thema-to-nomen*



Fonte: *International Federation of Library Associations and Institutions* (2010, p. 18).

¹⁴⁷ Derivado das Listas de Cabeçalhos de Assunto da *Library of Congress* (*Library of Congress Subject Headings - LCSH*), FAST é um esquema de terminologia de assunto amplamente utilizado por bibliotecas. Seu desenvolvimento é oriundo da colaboração entre a *Online Computer Library Center* (OCLC) e a *Library of Congress*, iniciado em 1998.

¹⁴⁸ Often the complexity of a *thema* is associated with the complexity of the *nomen* by which it is represented. Since the proposed model introduces a clear split between the *thema* (“the thing”) and the *nomen* (“the label” used to refer to it) [...] (INTERNATIONAL..., 2010, p. 18).

Nomen é indicado como uma superclasse para as entidades nome (*name*), identificador (*identifier*) e ponto de acesso controlado (*controlled access point*) do FRAD. Os “exemplos incluem ‘amor’, ‘∞’ ou ‘595.733’”. Um *nomen* pode ser legível por humanos ou por máquina”.¹⁴⁹ (INTERNATIONAL..., 2010, p. 18, tradução nossa).

De maneira geral, o FRSAD compreende que a linguagem natural traça diferentes nuances, como a estruturação “tem denominação” e “é denominado”, formando relacionamento de muitos para muitos, pois um *thema* detém um ou mais *nomen(s)*. Também pode incidir um *nomen* referente a mais de um *thema*, mas em um vocabulário controlado, um *nomen* é denominado por apenas um *thema*.

No inglês, é sabido que dentre as formas de uso do verbo *to have* (usado na terceira pessoa como *has*), uma delas serve para significar posse. Assim, “*has appellation*” relaciona a posse de um *thema* para um *nomen* e vice-versa. Já o verbo *to be* (usado na terceira pessoa como *is*) equivale ao ser/estar do português. Com isso, nota-se um *thema* para um *nomen* e vice-versa ‘ser’ ou ‘estar’; em analogia: o ser enquanto algo essencial e fixo, e o estar atrelado a volatilidade, a algo que muda.

A relação *thema-to-thema* envolve os relacionamentos hierárquicos, que segundo o modelo, definem a desambiguação dos termos através dos níveis de superordenação e subordinação, facilitando a tarefa de navegação do usuário, especificamente na identificação, seleção e exploração da busca. Ainda de acordo com o modelo, *thema-to-thema* envolve a estruturação de relações do tipo: genérica (*generic relationship*), todo-parte (*whole-part relationship*) e de equivalência (*instance relationship*) seguindo, assim, a normatividade da ISO 25964 (INTERNATIONAL STANDARD ORGANIZATION, 2009). O modelo considera também as relações poli-hierárquicas (*polyhierarchical relationship*) baseadas na ANSI/NISO Z39.19 (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2005).

O relatório do FRSAD finaliza indicando que, para *nomen-to-nomen*, o modelo estabelece a relação de equivalência (*equivalence relationship*) e todo-parte (*whole-part relationship*), também embasada na ISO 25964 e na ANSI/NISO Z39.19. Na primeira relação (de equivalência) para o *nomen*, é ressaltado o acesso ao assunto em situações como: a) *nomens* serem sinônimos; ou b) quase sinônimos; c) terem variantes lexicais; d) ser específico, mas estar representado por outro *nomen* com maior alcance; e) ser específico, mas

¹⁴⁹ “Examples include ‘love’, ‘∞’ or ‘595.733’. A *nomen* can be human readable or machine-readable.” (INTERNATIONAL..., 2010, p. 18).

estar representado por combinação de dois ou mais termos conhecido como equivalência composta.

Ainda na relação de equivalência entre *nomens*, em diferentes idiomas, o modelo destaca:

É óbvio que as relações de equivalência não assumem uma equivalência exata. A equivalência inexistente e parcial é encontrada frequentemente em vocabulários controlados. Na realidade, os *nomens* nas situações anteriores d) e e) representam diferentes *themas*. Mas como em alguns vocabulários controlados, esses *nomens* estão conectados como termos preferidos e alternativos, pode-se interpretar que esses *nomens* são assumidos como representando o mesmo *thema* em um vocabulário controlado.¹⁵⁰ (INTERNATIONAL..., 2010, p. 31, tradução nossa).

A exposição permite constatar que os parâmetros indicados no FRSAD tratam da estruturação de um vocabulário controlado por *themas* em *nomens* e, vice versa, mas não explicita como a noção de assunto se desenvolve. As formas de relacionamentos devem ser empregadas para função particular, como por exemplo, organizar termos em um sistema de significação, pois as palavras isoladas, ou o(s) *nomen(s)* isolado(s), não tem significado, ou então, podem ter muitos significados.

Ademais, nota-se que o modelo direciona quatro tarefas para a comunidade usuária utilizar os dados de autoridade assunto, a saber: encontrar, identificar, selecionar e explorar. A última tarefa busca cobrir um domínio de assunto e sua terminologia com as respectivas relações entre os *themas* e, o FRSAD, ainda direciona que o usuário poderá explorar a correspondência do(s) *nomen(s)* de um *thema* em um sistema de autoridade de assunto com o(s) respectivo(s) *nomen(s)* do mesmo *thema* em outro sistema de autoridade com a mesma finalidade. Essas tarefas não são inéditas, pois remontam aos objetivos do catálogo propostos por Cutter (1904), conforme já demonstrado na subseção 4.2.

Estabelece ainda que as tarefas de encontrar, identificar e selecionar sustentam o uso dos dados para encontrar, identificar e selecionar um ou vários assuntos e, respectivas denominações, que correspondam aos critérios estabelecidos pelos usuários utilizando atributos e relacionamentos.

Assim sendo, segundo o modelo, é possível encontrar um *thema* ou conjunto de *themas* e distinguir entre *themas* similares, e encontrar *nomens* que confirmem e distingam

¹⁵⁰ It is obvious that equivalence relationships do not assume exact equivalence. Inexact and partial equivalence are often found in controlled vocabularies. In reality, the *nomens* in the above d) and e) situations represent different *themas*. But since in some controlled vocabularies these *nomens* are connected as preferred and alternative terms, it may be interpreted that these *nomens* are assumed to represent the same *thema* in a controlled vocabulary. (INTERNATIONAL..., 2010, p. 31).

nomens similares. A mesma simetria, segundo o modelo, decorre em identificar um *thema* para confirmar se o encontrado pode se distinguir de outros *themas* similares e identificar um *nomen* a fim de confirmar que o *nomen* encontrado se distingue de *nomens* similares. Ao seleccionar, utilizam-se dados de autoridade de assunto para um *thema* e o conjunto de *themas* encontrados e, o mesmo decorre, com um *nomen* ou conjunto de *nomens* recuperados.

Diante da apresentação do modelo e de alguns estudos, nessa subsecção foi possível constatar que há lacunas conceituais que sustentem a noção de assunto no FRSAD.

Com o propósito de alcançar o entendimento sobre a noção de assunto no modelo FRSAD, objeto empírico dessa pesquisa, na próxima parte da análise busca-se observar convergências e diferenças entre os momentos analisados, incluindo os desenvolvimentos entre eles, e o modelo FRSAD.

O modelo foi pensado pela e para a área de Biblioteconomia e Ciência da Informação e, portanto, a primeira parte da análise (subsecção 7.1) complementa essa segunda na busca de perspectivas que possam embasar o que foi construído teoricamente pela área ao longo de tantos anos.

7.3 ANÁLISE SOBRE A NOÇÃO DE ASSUNTO NO FRSAD

Dando continuidade ao que a pesquisa propõe, nessa segunda parte da análise busca-se compreender se o modelo FRSAD reflete o acúmulo teórico da área de Biblioteconomia e Ciência da Informação quanto à ideia de assunto no processo de representação temática e seus métodos de realização, ou, se seria uma nova proposta, com pouca aderência com a área. Se o FRSAD não refletir o conhecimento da área, isso pode ter significados diferentes: o modelo pode apresentar conteúdos inconsistentes; e/ou pode apresentar outros conteúdos, não aderentes à área, ainda que consistentes.

É importante ressaltar que não foram e não serão assinaladas toda a estrutura do FRSAD, mas sim as evidências que emergem da construção da noção em foco.

Foi desenvolvido na área o ‘conceito’ como modo de fazer a abstração sobre os atributos temáticos do documento. Seja uma ideia, um objeto material, ou um lugar, trata-se estes atributos como conceitos por serem unidades linguísticas passíveis de operação, portanto, de desestruturação e de estruturação, na forma de uma mensagem documentária. Mas, quando se fala em conceito - objeto - evento - lugar no modelo, usa-se a palavra conceito no sentido mais restrito de ideia. Para uma possível diferenciação, no modelo houve

indicação do *thema* como uma entidade de nível elevado. Ao que parece, *thema* aproxima-se do que se chamou de conceito na área até então. Nesse cenário, que envolve a ênfase do conceito orientado a conferir valor à noção de assunto, faz-se aproximação com o segundo momento teórico, quando ocorre o desdobramento do pensamento anglo-saxão (quadro 7), a partir da primeira categoria (estruturação do assunto), caracterizado pela exposição de operações em nível conceitual do assunto. Porém, isso não significa entender o momento como precursor, mas como propulsor de uma perspectiva que seguiu e segue explorada pela área.

Partindo para o contexto das duas entidades de identificação do assunto, apontadas no FRSAD como de nível elevado, ou seja, consideradas entidades superiores – *thema* e *nomen* – ambas são caracterizadas por determinar o conteúdo informativo. Žumer, Zeng e Salaba (2012) apontam que para o compartilhamento de uma infraestrutura bibliográfica, referente aos dados de autoridade assunto, o *thema* é denominado por um ou mais *nomen(s)*. Nesse contexto, é considerada aproximação com o primeiro momento teórico (quadro 6), pela segunda categoria (o controle do assunto para a entrada), pois no contexto influenciado por Cutter, a efetividade das palavras pelos cabeçalhos, quer seja por uma classe ou não, se dá por nomes. Nesse aspecto, palavras isoladas não tem significado, ou então, possuem muitos significados, mas no FRSAD, o assunto autorizado tem como significado o *thema* que é denominado por um ou mais *nomen(s)*.

Apesar de Cutter não ter feito a proposição consciente de uma perspectiva linguística, houve mesmo que inconscientemente a percepção dessa natureza. Analisando o fato de o modelo considerar *thema* uma entidade superior em relação a todas as demais da família FRBR, Melo e Bräscher (2014) relacionam *thema* à abordagem conceitual. Assim, a partir de Melo e Bräscher (2014), no que tange à aproximação entre *thema* e conceito, ressalta-se o vínculo com o segundo momento teórico (quadro 7), pela segunda categoria, o controle do assunto para a entrada.

Se um *thema* pode ser trabalhado, segundo o modelo, enquanto *nomen(s)*, um dentre os vários atributos que compõem o primeiro é a perspectiva do vocabulário, mas, nesse aspecto, deve ser construído a partir do modo como se quer trabalhar um *thema* (ou seja, um conceito). Conforme já indicado, essa abordagem do modelo reflete o pensamento de uma linguagem documentária como uma lista de denominações. Por outro lado, constata-se no modelo “os casos de um *thema* sem um *nomen* estarem além do escopo deste modelo” (INTERNATIONAL..., 2010, p. 16, tradução nossa), sendo possível remeter a relação de um significado sem significante, ou, conceito sem termo. Conforme foi observado no

levantamento dos três momentos sobre a construção da noção de assunto, a área desenvolveu questões teóricas e metodológicas pautadas pela Linguística (signo como significante e significado) e pela Terminologia (a unidade terminológica é o termo-conceito), com foco na construção de linguagens documentárias. Entretanto, a proposta do modelo parece não reconhecer esses avanços que se realizaram, em especial, na vertente europeia, mas que se aproxima com o desdobramento da vertente anglo-saxã, inicial no tocante à construção de tesouros.

Nesse sentido, destaca-se aproximação, de forma concomitante, ao primeiro e segundo momentos teóricos, que se encontra respectivamente, nos quadros 6 e 7, pela terceira categoria (instrumento adotado).

Pelo quadro 9, é possível ilustrar que o modelo, em alguma medida, tem aproximação com perspectivas da área quanto à ideia de assunto no processo de representação temática e seus métodos de realização. Porém, de outro modo, há pontos com fragilidades que demonstram lacunas e há ausência de acúmulo teórico da área da Biblioteconomia e Ciência da Informação.

Quadro 9 – O estudo e o modelo FRSAD com as categorias resultantes dos momentos teóricos de construção da noção de assunto: convergências

FRSAD: apresentação e estudos	Categorias	Primeiro momento de construção da noção de assunto <i>Quadro 6</i>	Segundo momento de construção da noção de assunto <i>Quadro 7</i>	Terceiro momento de construção da noção de assunto <i>Quadro 8</i>
As entidades do FRSAD (conceito, objeto, evento e lugar) são diferentes em termos de gradação, mas orientam, em alguma medida, dar conta de uma representação do assunto de maneira especializada.			Estruturação do assunto	
De acordo com Žumer, Zeng e Salaba (2012), o FRSAD estabelece possibilidades para compartilhamento de infraestrutura bibliográfica, referente aos dados de autoridade assunto, sendo essa identificação vinculada ao <i>thema</i> , que possui como denominação um ou mais <i>nomen</i> .		O controle do assunto para a entrada		
Melo e Bräscher (2014) relacionam <i>thema</i> a abordagem conceitual.			O controle do assunto para a entrada	
Mesmo não sendo foco do modelo considerar a perspectiva do vocabulário, é explicitado que “até certo ponto, a granularidade de um <i>thema</i> também depende do vocabulário controlado usado para sua(s) denominação(ões).”		Instrumento adotado		

<p>(INTERNATIONAL..., 2010, p. 18, tradução nossa). Mas, nesse aspecto, nota-se ser o contrário, pois o vocabulário controlado deve ser construído a partir do modo como se quer trabalhar um <i>thema</i> (ou seja, um conceito). Essa abordagem do modelo reflete o pensamento de uma linguagem documentária como uma lista de denominações, usadas apenas para padronizar, a exemplo das incipientes listas de cabeçalhos de assunto.</p>			
<p>A proposta do modelo parece não reconhecer os avanços que se realizaram, em especial, na vertente europeia. Para ilustrar essa ausência, no modelo é feita a seguinte indicação: “os casos de um <i>thema</i> sem um <i>nomen</i> estarem além do escopo deste modelo.” (INTERNATIONAL..., 2010, p. 16, tradução nossa). Em analogia com a literatura da área é possível relacionar um significado sem significante, ou, conceito sem termo.</p>		<p>Instrumento adotado</p>	

Fonte: elaborado pela autora.

No terceiro momento, pela vertente europeia, a área é caracterizada por questões teóricas e metodológicas pautadas pela Linguística (signo como significante e significado) e pela Terminologia (a unidade terminológica é o termo-conceito), entretanto a proposta do modelo parece não reconhecer esses avanços que se realizaram.

No tocante à entidade conceito, a primeira do modelo FRSAD, a proposta deixa muitas dúvidas. Ao pensar nas entidades pela abstração, seria preciso falar em ideia, objeto material, ação e lugar. O conceito, como até o momento a área vem usando, aproxima-se ao *thema* do modelo.

Como já foi mencionado, nota-se fragilidade no uso dos termos no modelo. Neste sentido, a imprecisão terminológica da literatura foi retratada na subseção 5.6, referente ao segundo momento teórico da noção de assunto pelo desdobramento anglo-saxão. O modelo FRSAD faz uso de forma ampla da palavra *subject* (assunto), mesmo porque o inglês é o idioma preferencial da IFLA e das suas publicações. Também é o idioma da literatura de vertente anglo-saxã.

Apesar do *subject* ser o significante central do modelo, outros termos aparecem com frequência, como *aboutness* e *topic* e ainda entra em jogo o *thema* e o *nomen*. Conforme exposto, o modelo não supera a divergência conceitual oriunda do desdobramento da vertente anglo-saxã e, os dois últimos termos em latim, de alguma forma, buscam fornecer sentido à representação pelo *subject*.

Entre os esforços para alcançar o compartilhamento global e o uso de dados de autoridade de assunto, no modelo há indicação de concentração no *nomen*, como por exemplo, termo ou descritor encontrado em um vocabulário controlado, mas isso demanda compreensão dos assuntos em termos de seus significados, bem como as relações entre esses assuntos. Os fundamentos da Linguística Documentária, a partir da vertente europeia, seria útil para dar consistência ao que se observa.

Mesmo sem deter o papel de norma e, de forma ampla, o objetivo do modelo colabora na abstração do assunto quanto a construção do registro de autoridade de assunto, não fica nítido como o assunto pode se manifestar através das 11 entidades.

As relações entre essas 11 entidades da família FRBR (obra, expressão, manifestação, item, pessoa, entidade coletividade, família, conceito, objeto, evento e lugar) são priorizadas no FRSAD para abstração do assunto e essa indicação relaciona-se com o *nomen*, pelo qual um dado *thema* é conhecido. Desse modo, *thema* e *nomen* reportam o objetivo de vincular os registros de autoridade de assuntos em catálogos ou bases de dados.

Apesar das convergências com as categorias construídas a partir do referencial teórico trabalhado, o FRSAD tem lacuna quanto ao conhecimento teórico da área de Biblioteconomia e Ciência da Informação. Ademais, em nenhuma parte do modelo houve aproximação de entendimento do assunto enquanto algo a ser construído e interpretado a partir do texto do documento; o assunto não está dado no documento. Ainda cabe ressaltar que a abstração fornecida pelo modelo concentra atenção no registro de autoridade de assunto e não na noção desse assunto. Por vezes, o modelo mais confunde do que esclarece ao trabalhar com termos no contexto da representação temática.

Por essa análise, torna-se evidente que não há uma escolha teórica explícita que nutra o que seja a noção de assunto no modelo, mas pode-se dizer que ele se pauta pela vertente anglo-saxã. De outro modo, pode ser ressaltado que o modelo não trabalha com a análise e síntese de textos até o registro da informação documentária, aspecto caracterizador da vertente europeia.

Há uma preocupação latente com os nomes de autoridades de assunto, mas não com a caracterização da construção da representação temática, como por exemplo, em diferentes produtos e aspectos. Dessa forma, a construção da noção de assunto no modelo tem lacunas em relação ao acúmulo da literatura da área.

8 CONSIDERAÇÕES FINAIS

Nas pesquisas iniciais realizadas para esta investigação, percebeu-se que a palavra assunto deflagra diferentes nuances da representação temática, fazendo perceber outros termos também adotados para nomear este tipo de representação. Através do amadurecimento das leituras, foi constatada a existência de abordagens diversas, construídas em momentos históricos e culturais distintos, e que são presentes ainda hoje, de algum modo, na literatura da área. De acordo com a exploração da literatura realizada, vertentes constituíram-se com suas respectivas abordagens, mas elas influenciam-se umas às outras, pois apesar dos fundamentos que as distinguem, elas se aproximam por um propósito: a representação do assunto.

Por este motivo, nessa investigação, a exploração realizada sobre a noção de assunto baseou-se em diferentes abordagens que marcaram a Biblioteconomia e Ciência da Informação para este tema. Examinar as vertentes em uma perspectiva geográfica e cronológica foi pensada como uma forma preliminar de categorização, uma vez que não há escassez de literatura, o que evidencia a necessidade de sistematização de diferentes abordagens que permita algum tipo de contribuição teórica.

Cotidianamente, a palavra assunto pode referir-se a diferentes aspectos, mas na Biblioteconomia e Ciência da Informação é abarcada pela construção, desencadeada pela interpretação do texto do documento, para posterior representação via registro documentário. O entendimento de como o assunto foi se constituindo na área foi estruturado em: precedentes da construção da noção de assunto; e, primeiro, segundo e terceiro momentos de construção da noção de assunto.

A seção sobre os ‘precedentes da construção da noção de assunto’ resgatou as propostas iniciais de estudos no continente europeu, envolvendo marcas históricas do produto bibliográfico até o trabalho de Adrien Baillet (1649-1706). Essa abordagem torna-se interessante à medida que recupera o nome de bibliógrafos e seus trabalhos, além de algumas bibliotecas e catálogos, e a compilação de regras que marcaram a representação temática na Catalogação.

Em seguida, o ‘primeiro momento de construção da noção de assunto’ é marcado pelo contexto anglo-saxão no século XIX, pela vertente denominada Catalogação de Assunto (*Subject Cataloguing*), que se desenvolveu com ênfase na elaboração de catálogos e da forte influência dos princípios alfabéticos e das listas de cabeçalhos de assunto. Foi um século que marcou a Biblioteconomia pelas contribuições que recebeu de teóricos da Catalogação, como por exemplo, de Anthony Panizzi (1797-1879) e Charles Jewett (1816-1868), que se

envolveram com a formulação de princípios para as práticas realizadas. Foi nesse contexto que Charles Cutter (1837-1903) se destacou por investir esforços na abordagem dos assuntos e na tradição do catálogo dicionário.

O ‘segundo momento de construção da noção de assunto’ centra-se em um quadro teórico estabelecido por forte influência anglo-saxã. Anteriormente sistematizado por Cutter na forma de um cabeçalho, nesse momento, o assunto passa a ser realizado pela decomposição. O fato principal é que essas relações, assim como outras, não aconteceram de forma linear, ou seja, ocorreram com o envolvimento de influências variadas. Esse momento é marcado pela presença de diferentes teóricos, partindo de Julius Kaiser (1868-1927) e Shiyali Ranganathan (1892-1972), passando por alguns pesquisadores do *Classification Research Group*, além de Calvin Mooers, alcançando avanços sobre a Indexação e o Índice, o Conceito e a Análise Conceitual, a Normalização da Indexação. A seção é finalizada pela explicitação da Dispersão Terminológica que pode ser evidenciada pela análise do conjunto desses estudos, haja vista o modo como foram realizados e consolidados no século XX na vertente anglo-saxã.

Já no último e ‘terceiro momento de construção da noção de assunto’ trabalhado, buscou-se pela articulação de ideias que permitiram a formação de um novo cenário que, diferente do contexto anglo-saxão, desenvolveu-se por outro modo de dispor as informações, em especial, para uso da informação científica. As discussões ocasionadas nesse momento pela vertente europeia favoreceram um cenário de pesquisa que, também de modo distinto da vertente anterior, não se caracterizou pela dispersão da produção intelectual. Para tanto, as ideias propostas por Paul Otlet (1868-1944) e, mais tarde, as de Jean-Claude Gardin (1925-2013) alcançaram, sobretudo, a Espanha e o Brasil pela necessidade de se desenvolver meios e linguagens fundamentadas teoricamente para o acesso de documentos.

Ao longo da investigação, foi possível constatar que a noção de assunto é construída pela coexistência de elementos diversos, os quais foram evidenciados pelo referencial e categorias traçadas. Com isso, percebeu-se que a noção de assunto é pensada por uma forma de estruturação, do controle para a entrada e do instrumento adotado. O assunto, nas três vertentes, foi pensado para o universo do objeto que se pretendeu representar: na primeira, as palavras dos textos dos livros foram transpostas para as listas de cabeçalhos de assunto; na segunda, houve avanço na preocupação com a especialização dos assuntos, sobretudo dos periódicos, o que acarretou na necessidade do desenvolvimento de instrumento que desse conta de sistemas conceituais, estruturados a partir de conceitos e suas características, como ocorreu com o tesouro; e, na terceira, ocorreu ampliação de fundamentos linguísticos e

terminológicos que desencadeou contiguidade entre o instrumento (na forma de um vocabulário) e o universo de conceitos presentes nos textos.

Na tentativa de determinar as características que envolvem a noção de assunto, as categorias foram desenvolvidas de modo a instalarem uma tensão no ato de pensar elementos para sua concretização, o que influi tanto no pensar de quem trabalha o assunto quanto no resultado da representação para uso de quem utiliza.

Profissionais de ambientes informacionais, independente de acervos convencionais ou virtuais, manipulam o assunto para a representação em bases de dados. Apesar de ser atividade corriqueira da prática de bibliotecários(as), na construção de um registro documentário não se pode perder de vista que há um conjunto teórico de sustentação para esse fazer.

Assim, o assunto de um documento em um registro não é uma palavra ‘solta’, pois essa pode não ter, ou, ter diferentes significados. Os assuntos dos documentos são representações construídas e, conforme demonstrado ao longo do referencial trabalhado, eles foram elaborados e aplicados em diferentes tempos e lugares. Resgatar esse cenário colabora para o fortalecimento da área ao demonstrar que a representação temática não pode ser relegada somente a trabalhos empíricos. Há necessidade contínua de trabalhos teóricos que permitam fundamentar devidamente as atividades práticas. Esse movimento, que não privilegia a busca por teorias e metodologias rigorosamente fundamentadas, levou a que a organização da informação fosse rotulada de ‘processo técnico’, entendido como operação mecânica, desprovida de base intelectual.

O objetivo geral da pesquisa foi o de investigar e analisar criticamente a base teórica e conceitual que envolve a fundamentação da noção de assunto no modelo FRSAD. O desenvolvimento dos três momentos teóricos levou a perceber que o modelo FRSAD desconhece as diferentes vertentes da Biblioteconomia e Ciência da Informação. Ele reflete, em alguma medida, aspectos da vertente anglo-saxã, conforme demonstrado na análise.

Na aproximação das convergências para pensar a noção em foco, relacionada à estruturação do assunto, ao controle do assunto para a entrada e ao instrumento adotado, o modelo demonstra que a sua validade está nos esforços para alcançar o compartilhamento global e o uso de dados de autoridade de assunto controlado. Considerando que isso é comum à área, pois sem controle o assunto enquanto palavra não contribui para a construção de um sistema simbólico, o modelo não se desenvolve quanto aos aspectos relacionados ao conceito, desenvolvidos na literatura para uma compreensão e representação do assunto.

Portanto, se o modelo não trabalha com a sistematização e o agrupamento por critérios pré-determinados de natureza teórica, não colabora para o movimento de abstração sobre a noção de assuntos a partir de um sistema de significação próprio, como por exemplo, decorre dos conceitos metodológicos de linguagem advindos da vertente europeia.

Evidencia-se que há relação entre a noção de assunto no modelo FRSAD e o desenvolvimento teórico, mas por outro lado, as discussões implicaram a ausência dos aspectos linguísticos em torno do assunto no FRSAD. O assunto enquanto um atributo é uma unidade linguística passível de operação, demandando, portanto, processos de desestruturação e de estruturação, na forma de uma mensagem documentária. O registro documentário deve ser portador da mensagem documentária, situação que permite o acesso por parte do público que faz uso dos catálogos ou bases de dados.

O modelo não tem como função ser um instrumento normalizador para a representação temática ou para a construção de vocabulário controlado. Ao sinalizar uma possível isenção na escolha de dois termos em latim, *thema* e *nomen*, transmite abordagem cristalizada da relação significante e significado. Entretanto, não demonstrou um princípio consistente para essa e nem para a escolha das quatro entidades (conceito, objeto, evento e lugar) que buscam subsidiar a noção de assunto.

O modelo FRSAD é apresentado como uma referência internacional e um marco que proporcione o entendimento dos registros bibliográficos e as respectivas autoridades de assunto com o objetivo de atender as buscas dos usuários. Entretanto, deixa lacuna nessa perspectiva pelo motivo de não ter acumulado a literatura da área no que tange os fundamentos para o assunto.

Nos limites da pesquisa, é possível considerar que as escolhas dos referenciais, ao mesmo tempo em que nutriram diferentes perspectivas sobre o assunto, avançaram na sistematização de ideias que, por vezes, são confundidas entre si frente às diferentes abordagens teóricas existente na Biblioteconomia e Ciência da Informação. Como exemplo, tem o emprego, por vezes conflituoso, dos termos análise documentária e análise de assunto, assim como, aqueles termos que são foco desta pesquisa: assunto, conceito, conteúdo, *aboutness*, dentre outros.

Por outro lado, explorar a literatura com o objetivo de análise de um modelo empírico contribuiu, também, para a expansão de olhar crítico sobre as propostas da IFLA, oferecidas à comunidade bibliotecária.

Compreender fundamentos teóricos que subsidiem o entendimento da noção de assunto é necessária à consolidação da representação temática, situação que permite fortalecer

as interlocuções entre a teoria, o ensino e a prática profissional da área em Biblioteconomia e Ciência da Informação.

Deseja-se que a apropriação realizada nessa pesquisa sobre a noção de assunto, visando à compreensão inicial de uma perspectiva teórica e, posteriormente, da análise do modelo conceitual FRSAD, contribua com a área, possibilitando discussões que alimentem o cenário da representação temática. Sendo assim, almeja-se que essa pesquisa abra portas para o aprofundamento de críticas às questões em torno da noção de assunto, integrando as agendas de investigação de pesquisadores que venham a desenvolver literatura sobre o tema, na busca pela compreensão do panorama da representação temática.

REFERÊNCIAS

- ABBAGNANO, Nicola. **Dicionário de filosofia**. 5. ed. rev. e ampl. São Paulo: Martins Fontes, 2007.
- AMERICAN LIBRARY ASSOCIATION. **Glosario ALA de bibliotecología y ciencias de la información**. Madrid: Diaz de Santos, c1988.
- ARAÚJO, Andre Vieira de Freitas. Pioneirismo bibliográfico em um polímata do séc. XVI: Conrad Gesner. **Informação & Informação**, Londrina, v. 20, n. 2, p. 118-142, 2015.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12676**: métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992.
- AULETE Digital. Disponível em: <http://www.aulete.com.br/assunto>. Acesso em: 6 mar. 2018.
- BARBOSA, Alice Príncipe. Classificações facetadas. **Ciência da Informação**, Brasília, v. 1, n. 2, p. 73-81, 1972.
- BARDIN, Laurence. **Análise de conteúdo**. Lisboa: Edições 70, 2018.
- BEGHTOL, Claire. Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. **Journal of Documentation**, London, v. 42, n. 2, p. 84-113, june. 1986.
- BELY, Nathalie *et al.* **Procédures d'analyse sémantique appliquées a la documentation scientifique**. Paris: Gauthier-Villars Éditeur, 1970.
- BUCKLAND, Michael. Information schools: a monk, library science and the information age. *In*: HAUKE, Petra. **Bibliothekswissenschaft - quo vadis? / Library Science - quo vadis**. München: K. G. Saur, 2005. p. 19-32.
- BURKE, Peter. **Uma história social do conhecimento**: de Gutenberg a Diderot. Rio de Janeiro: J. Zahar, 2003.
- BUONOCORE, Domingo. **Diccionario de Bibliotecologia**. 2. ed. aum. Buenos Aires: Marymar, 1976.
- CAMPOS, Maria Luiza de Almeida; GOMES, Hagar Espanha. Organização de domínios de conhecimento e os princípios ranganathianos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 8, n. 2, p. 150-163, jul./dez. 2003.
- CAPURRO, Rafael; HJØRLAND, Birger. O conceito de informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 12, n. 1, p. 148-207, jan./abr. 2007.
- CERVO, Amado Luiz; BERVIAN, Pedro Alcino; SILVA, Roberto da. **Metodologia científica**. 6. ed. São Paulo: Pearson Prentice Hall, 2007.

CESARINO, Maria Augusta da Nóbrega; PINTO, Maria Cristina Mello Ferreira. Cabeçalho de assunto como linguagem de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 7, n. 2, p. 268-88, set. 1978.

CHAUMIER, Jacques. Indexação: conceito, etapas, instrumentos. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 21, n. 1/2, p. 63-79, jan./jun. 1988.

CHEN, Peter P. **Modelagem de dados**: a abordagem entidade-relacionamento para projeto lógico. São Paulo: Makron Books, 1990.

CINTRA, Anna Maria Marques; TÁLAMO, Maria de Fátima Gonçalves Moreira; LARA, Marilda Lopes Ginez de; KOBASHI, Nair Yumiko. Do termo ao descritor. **Revista de Comunicação e Artes**, São Paulo, v. 17, n. 28, p. 75-82, jan./abr. 1994.

COLLISON, Robert L. **Índices e indexação**. São Paulo: Polígono, 1971.

CRIPPA, Giulia. Narrativa como gesto bibliográfico: Gabriel Naudé entre erudição e política. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 22, número especial, p. 21-35, jul. 2017.

CUNHA, Celso. Gramática básica do português contemporâneo. *In*: AULETE Digital. **Gramática básica**. Disponível em: <http://www.aulete.com.br/site.php?mdl=gramatica>. Acesso em: 4 abr. 2018.

CUNHA, Isabel M. R. Ferin. Análise documentária. *In*: SMIT, Johanna W. (coord.). **Análise documentária**: a análise da síntese. 2. ed. Brasília: IBICT, 1987.

CUTTER, Charles Ammi. **Rules for a printed dictionary catalogue**. Washington: Librarian of the Boston Athenaeum, 1876. Disponível em: <https://archive.org/details/cu31924029518978>. Acesso em: 2 nov. 2016.

CUTTER, Charles Ammi. **Rules for a printed dictionary catalogue**. 4. ed. Washington: Government Printing Office, 1904.

DAHLBERG, Ingetraut. Teoria do conceito. **Ciência da Informação**, Brasília, v. 7, n. 2, 1978.

DAHLBERG, Ingetraut. Knowledge organization: its scope and possibilities. **Knowledge Organization**, Würzburg, v. 20, n. 4, p. 211-222. 1993.

DELSEY, Tom. Modeling subject access: extending the FRBR and FRANAR conceptual models. **Cataloging & Classification Quarterly**, New York, v. 39, issue 3/4, p. 49-61. 2005.

DIAS, Antonio Caetano. **Elementos de catalogação**. Rio de Janeiro: Associação Brasileira de Bibliotecários, 1967.

DIAS, Eduardo Wense; NAVES, Madalena Martins Lopes. **Análise de assunto**: teoria e prática. Brasília: Briquet de Lemos, 2013.

DICIONÁRIO da língua portuguesa: dicionários editora. Porto: Porto Editora, 2004.

DOUSA, Thomas M. **Julius Otto Kaiser's systematic indexing: a study of its theoretical content in its historical context.** 2010. Doctorate (degree Philosophy in Library and Information Science)-University of Illinois, Urbana-Champaign, 2010.

DUTTA, Bidyarthi; DUTTA, Chaitali. Concept of 'subject' in library and information science from a new angle. **Annals of Library and Information Studies**, Indian, v. 60, n. 2, p. 78-87, june. 2013.

EL HADI, Widad Mustafa (ed.). Fondements épistémologiques et théoriques de la science de l'information-documentation: hommage aux pionniers francophones. *In: COLLOQUE ISKO FRANCE*, 11., 2017, Paris. **Actes** [...]. Paris: ISTE Editions, 2018.

FARIA, Maria Isabel; PERICÃO, Maria da Graça. **Dicionário do livro: da escrita ao livro electrónico.** Coimbra: Almedina, 2008.

FERRAZ, Iraneuda Maria Cardinali. Uso do catálogo de biblioteca: uma abordagem histórica. **Trans-in-formação**, Campinas, v. 3, n. 1/2/3, p. 90-114, jan./dez. 1991.

FERRAZ, Wanda. **Relação de assuntos para cabeçalhos de fichas.** 5. ed. Rio de Janeiro: Freitas Bastos, 1977.

FERREIRA, António Gomes. **Dicionário de latim-português.** Porto: Porto Codex, 1987.

FIÚZA, Marysia Malheiros. Funções e desenvolvimento do catálogo: uma visão retrospectiva. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 9, n. 2, p. 139-158, set. 1980.

FIÚZA, Marysia Malheiros. A catalogação bibliográfica até o advento das novas tecnologias. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 16, n. 1, p. 43-53, mar. 1987.

FOSKETT, Antony Charles. **A abordagem temática da informação.** São Paulo; Brasília: Editora da Universidade de Brasília; Editora Polígono, 1973.

FRÍAS MONTOYA, José Antonio. **La descripción bibliográfica y sus puntos de acceso en el catálogo de la biblioteca: evolución histórica y problemática actual.** Volume I. 1995. 299 f. Tese (Doutorado em Biblioteconomía y Documentación)-Universidad Complutense de Madrid, Madrid, 1995.

FUJITA, Mariângela Spotti Lopes. Sistema de indexação precis. I - PRECIS: perspectiva histórica e técnica do seu desenvolvimento e aplicação (continua). **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 21, n. 1/2, p. 21-45. 1988.

FURNER, Jonathan. FRISAD and the ontology of subjects of works. **Cataloging & Classification Quarterly**, New York, v. 50, issue 5-7, p. 494-516, jun. 2012. (Special Issue The FRBR Family of Models).

GARCÍA GUTIÉRREZ, Antonio Luis. **Estructura lingüística de la documentación: teoría y método.** Murcia: Universidad de Murcia, 1990.

GARCÍA-TALAVERA, Miguel Diaz. Forma. *In*: DICIONÁRIO Santillana: para estudantes. 2. ed. São Paulo: Moderna, 2011.

GARDIN, Jean-Claude. Chapitre premier: fonctions du centre. *In*: ASSOCIATION NATIONALE D'ÉTUDES POUR LA DOCUMENTATION AUTOMATIQUE (A.N.E.D.A.). **L'organisation de la documentation scientifique**. Paris: Gauthier-Villars, 1964. cap. 1, p. 9-21.

GARDIN, Jean-Claude. Recherches sur l'indexation automatique. **Revue Française d'Informatique et de Recherche Opérationnelle**, le anée, n. 6, p. 27-46, nov./déc. 1967.

GARDIN, Jean-Claude *et al.* **L'automatisation des recherches documentaires: un modèle général 'Le SYNTOL'**. Paris: Gauthier-Villars, 1968.

GARRIDO ARILLA, María Rosa. **Teoría e historia de la catalogación de documentos**. Madrid: Síntesis, 1999.

GEMBERLING, Ted. Thema and FRBR's third group. **Cataloging & Classification Quarterly**, New York, v. 48, issue 5, p. 445-449, may. 2010.

GEMBERLING, Ted. FRISAD, Semiotics, and FRBR-LRM. **Cataloging & Classification Quarterly**, New York, v. 54, issue 2, p. 136-144, febr. 2016.

GIL LEIVA, Isidoro. **Manual de indización: teoría y práctica**. Gijón: Trea, 2008.

GOMES, Hagar Espanha; MARINHO, Marcílio Teixeira. **Introdução ao estudo do cabeçalho de assunto**. Rio de Janeiro, 1984. Disponível em: http://www.conexao.org/bit/cabecalho/cab_ass.htm. Acesso em: 20 mar. 2018.

GUEDES, Emanuel Guedson Ferrerira. **O conceito de aboutness na organização e representação do conhecimento**. 2009. 90 f. Dissertação (Mestrado em Ciência da Informação)–Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2009.

GUIMARÃES, José Augusto Chaves. A Análise documentária no âmbito do tratamento da informação: elementos históricos e conceituais. *In*: RODRIGUES, G. M.; LOPES, I. L. (org.). **Organização e representação do conhecimento na perspectiva da Ciência da Informação**. Brasília: Thesaurus, 2003. p. 100-117.

GUIMARÃES, José Augusto Chaves; NASCIMENTO, Lúcia Maria Barbosa do; MORAES, João Batista Ernesto de. A diplomática como perspectiva metodológica para o tratamento de conteúdo de documentos técnicos. *In*: VALENTIM, Marta Lúcia Pomin (org.). **Métodos qualitativos de pesquisa em ciência da informação**. São Paulo: Polis, 2005. p. 135- 160.

GUIMARÃES, José Augusto Chaves. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-Americana de Ciência da Informação**, Brasília, v. 1, n. 1, p. 77-99, jan./jun. 2008.

GUIMARÃES, José Augusto Chaves. Abordagens teóricas de tratamento temático da informação: catalogação de assunto, indexação e análise documental. **Ibersid**, Zaragoza, v. 3, p. 105-117. 2009.

GUIMARÃES, José Augusto Chaves. Prefácio. *In*: ALMEIDA, Carlos Cândido de. **Elementos de linguística e semiologia na organização da informação**. São Paulo: Cultura Acadêmica, 2011. 9-12. *E-book*. Disponível em: <https://repositorio.unesp.br/handle/11449/109182>. Acesso em: 5 jan. 2018.

HENDERSON, Madeline M. Calvin N. Mooers, october 24, 1919 - december 1, 1994. **Journal of the American Society for Information Science**, Maryland, n. 47, n. 9, p. 659-661. 1996.

HJØRLAND, Birger. The concept of “subject” in information science. **Journal of Documentation**, London, v. 48, n. 2, p. 172-200. 1992.

HJØRLAND, Birger. Towards a theory of aboutness, subject, topicality, theme, domain, field, content... and relevance. **Journal of the American Society for Information Science and Technology**, Maryland, v. 52, n. 9, p. 774-778, jul. 2001.

HJØRLAND, Birger. Subject (of Documents). **Knowledge Organization**, Würzburg, v. 44, n. 1, p. 55-64. 2017.

HOUAISS, Antônio. **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2001.

HUDON, Michèle. Epistemological and theoretical foundations of Information - Documentation Science: a tribute to francophone pioneers 11th ISKO-France Conference, 2017. **Knowledge Organization**, Würzburg, v. 45, n. 3, p. 255-267. 2018.

HUTCHINS, J. W. On the problem of “aboutness” in document analysis. **Journal of Informatics**, v. 1, n. 1, p. 17-35, april. 1977.

HUTCHINS, J. W. The concept of ‘aboutness’ in subject indexing. **Aslib Proceedings**, v. 30, n. 5, p. 172-181. 1978.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Functional Requirements for Bibliographic Records**: final report. 1998. Disponível em: <https://www.ifla.org/publications/functional-requirements-for-bibliographic-records>. Acesso em: 3 mar. 2017.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Functional Requirements for Bibliographic Records**: final report: amended and corrected through february 2009. 2009. Disponível em: <https://www.ifla.org/publications/functional-requirements-for-bibliographic-records>. Acesso em: 3 mar. 2017.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Functional Requirements for Subject Authority Data (FRSAD)**: a conceptual model. 2010. Disponível em: <http://www.ifla.org/files/assets/classification-and->

indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf. Acesso em: 1 mar. 2016.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. Section Cataloguing. FRBR Review Group. **World-wide review of the FRBR-Library Reference Model, a consolidation of the FRBR, FRAD and FRSAD conceptual models**. 28 february 2016a. Disponível em: <http://www.ifla.org/node/10280?og=54>. Acesso em: 29 set. 2016.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. Section Subject Analysis and Access. Subject Analysis and Access: the new name of our Section. **News**, 21 september 2016b. Disponível em: <https://www.ifla.org/node/10875>. Acesso em: 30 set. 2016.

INTERNATIONAL STANDARD ORGANIZATION. **ISO 25964: thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval**. Geneve: International Standard Organization, 2009.

KAULA, Prithvi N. **Repensando os conceitos no estudo da classificação**. 1982. Disponível em: <http://www.conexao.org/bit/kaula/index.htm>. Acesso em: 4 fev. 2019.

KOBASHI, Nair Yumiko. **A elaboração de informações documentárias: em busca de uma metodologia**. 1994. 195 f. Tese (Doutorado em Ciências da Comunicação)–Programa de Pós-Graduação em Ciências da Comunicação, Universidade de São Paulo, São Paulo, 1994.

KOBASHI, Nair Yumiko. Análise documentária e representação da informação. **INFORMARE: cadernos do Programa de Pós-Graduação em Ciência da Informação**, Rio de Janeiro, v. 2, n. 2, p. 5-27, jul./dez. 1996.

KOBASHI, Nair Yumiko. Linguística textual e elaboração de informações documentárias: algumas reflexões. In: GASPAR, N. R.; ROMÃO, L. M. S. (orgs.). **Discurso e texto: multiplicidade de sentidos na Ciência da Informação**. São Carlos: EdUFSCar, 2008. p. 47-66.

LANCASTER, Frederick Wilfrid. **Indexação e resumos: teoria e prática**. 2. ed. Brasília: Briquet de Lemos, 2004.

LANGRIDGE, Derek. Cabeçalhos de assunto. In: KNIGHT, G. Norman (coord.). **Treinamento em indexação: um curso da Society of Indexers**. Rio de Janeiro: FGV, 1974. cap. 5. p. 75-85.

LARA, Marilda Lopes Ginez de. **A representação documentária: em jogo a significação**. 1993. 133 f. Dissertação (Mestrado em Ciências da Comunicação)–Programa de Pós-Graduação em Ciências da Comunicação, Universidade de São Paulo, São Paulo, 1993.

LARA, Marilda Lopes Ginez de. **Dos sistemas de classificação bibliográfica às search engines (II)**. São Paulo: Associação Paulista de Bibliotecários, 2001. (Ensaio APB, n. 91).

LARA, Marilda Lopes Ginez de; TÁLAMO, Maria de Fátima Gonçalves Moreira. Linguística documentária e terminologia: experiência didática na interface das disciplinas. In:

ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 7., 2006, Marília. **Anais** [...]. Marília: Unesp; ANCIB, 2006.

LARA, Marilda Lopes Ginez de; TÁLAMO, Maria de Fátima Gonçalves Moreira. Uma experiência na interface Linguística Documentária e Terminologia. **Datagramazero**, Rio de Janeiro, v. 8, n. 5, out. 2007.

LARA, Marilda Lopes Ginez de. Conceitos de organização e representação do conhecimento na ótica das reflexões do grupo Temma. **Informação & Informação**, Londrina, v. 16, n. 3, p. 92-121, jan./jun. 2011.

LITTON, Gaston. **Classificação e catálogos**. São Paulo: Mcgraw-hill, 1971.

LÓPEZ GUILLAMÓN, Ignacio. **De la “notitia librorum” a los metadatos. Itinerario de la teoría sobre catalogación bibliográfica**. 2011. 1031 f. Tese (Doutorado em Biblioteconomía y Documentación)–Universidad de Extremadura, Badajoz, 2011.

MALCLÈS, Louise Nöelle. **La bibliografía**. Buenos Aires: EUDEBA, 1960.

MANN, Margaret. **Catalogação e classificação de livros**. Rio de Janeiro: Editôra Fundo de Cultura, 1962.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. 5. ed. São Paulo: Atlas, 2003.

MARTÍNEZ DE SOUSA, José. **Diccionario de bibliologia y ciencias afines**. Salamanca: Fundación Germán Sánchez Ruipérez, 1989.

MARTINS, Gilberto de Andrade; THEÓPHILO, Carlos Renato. **Metodologia da investigação científica para Ciências Sociais aplicadas**. 3. ed. São Paulo: Atlas, 2016.

MELO, Maria Antônia Fonseca; BRÄSCHER, Marisa. Requisitos Funcionais para Dados de Autoridade Assunto (FRSAD): entidades, atributos e relacionamentos. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 12, n. 2, p. 102-119, maio/ago. 2014.

METCALFE, John. When is a Subject not a Subject? *In*: RAWSKI, Conrad H. (ed.). **Towards a theory of Librarianship**. New York: Scarecrow Press, 1973. p. 303-338.

MIKSA, Francis Louis. **Charles Ammi Cutter: library systematizer**. Colorado: Libraries Unlimited, 1977.

MIKSA, Francis Louis. The legacy of the library catalogue for the present. **Library Trends**, Champaign, v. 61, n. 1, p. 7-34. 2012.

MOOERS, Calvin N. Zatocoding applied to mechanical organization of knowledge. **American Documentation**, v. 2, p. 20-32, jan. 1951. Disponível em: <https://courses.engr.illinois.edu/cs473/fa2013/misc/zatocoding.pdf>. Acesso em: 13 set. 2018.

MOOERS, Calvin N. The indexing language of an information retrieval system. *In*: CHAN, Lois Mai; RICHMOND, Phyllis A.; SVENONIUS, Elaine (eds.). **Theory of subjects analysis: a sourcebook**. Colorado: Libraries Unlimited, 1985. p. 247-261.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **ANSI/NISO Z39.19-2005**: guidelines for the construction, format, and management of monolingual thesauri. 2005.

NEEDHAM, Christopher Donald. **Organizing knowledge in libraries**: an introduction to information retrieval. 2. th. rev. London: André Deutsch, 1971. Disponível em: <https://archive.org/details/in.ernet.dli.2015.110959>. Acesso em: 25 abr. 2018.

ORTEGA, Cristina Dotta. Fundamentos da organização da informação frente à produção de documentos. **Transinformação**, Campinas, v. 20, n. 1, p. 7-15, jan./abr. 2008.

ORTEGA, Cristina Dotta. Aspectos terminológicos em Ciência da Informação: a função documentária em jogo. *In*: ENCONTRO DE REPRESENTAÇÃO DOCUMENTAL, 1., 2017, São Carlos. **Anais [...]**. São Carlos: UFSCar, 2017. Disponível em: <http://www.telescopium.ufscar.br/index.php/enredo/enredo/schedConf/presentations>. Acesso em: 20 jul. 2017.

OTLET, Paul. Les sciences bibliographiques et la documentation. **IIB Bulletin**, n. 8, 1903.

OTLET, Paul. L'organisation rationnelle de l'information et de la documentation en matière économique. **IIB Bulletin**, n. 10, 1905.

OTLET, Paul. **Tratado de documentação**: o livro sobre o livro teoria e prática. Brasília: Briquet de Lemos/Livros, 2018.

PANIZZI, A. Rules for the Compilation of the Catalogue. *In*: BRITISH MUSEUM. **Catalogue of Printed Books in the British Museum**: volume I. London: J. B. Nichols and Son Printers, 1841. Disponível em: <https://babel.hathitrust.org/cgi/pt?id=umn.31951001998306w;view=1up;seq=9>. Acesso em: 25 jul. 2018.

PEIXOTO, Jorge. **Técnica bibliográfica**: subsídio para a bibliografia portuguesa. II parte – catalogação. Coimbra: Atlântida, 1962.

PENNA, Carlos Victor. **Catalogacion y clasificacion de libros**. Buenos Aires: ACME Agency, 1945.

PINTO, Maria Cristina Mello Ferreira. Análise e representação de assuntos em sistemas de recuperação da informação; linguagens de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 2, p. 169-186, set. 1985.

PINTO MOLINA, María. **Análisis documental**: fundamentos y procedimientos. 2. ed. rev. aum. Madrid: EUDEMA, 1993.

PLACER, Xavier. **A bibliografia e sua técnica**. Rio de Janeiro: Departamento de Imprensa Nacional, 1955.

QUIVY, Raymond; CAMPENHOUDT, Luc Van. **Manual de investigação em ciências sociais**. 4. ed. Lisboa: Gradiva, 2005.

REITZ, Joan M. **Online Dictionary for Library and Information Science**. Califórnia: ABC-CLIO Corporate, 2014. Disponível em: https://www.abc-clio.com/ODLIS/odlis_s.aspx. Acesso em: 3 abr. 2017.

RIAZ, Muhammad. **Advanced indexing and abstracting practices**. New Delhi: Atlantic Publishers & Distributors, 1989.

ROVIRA, Carmen. **Los epigrafs en el catalogo diccionario**. 2. ed. rev. Washington: Union Panamericana, 1966.

SAGREDO FERNÁNDEZ, Félix; IZQUIERDO ARROYO, José Maria. **Concepción lógico-lingüística de la documentación**. Madrid: IBERCOM, 1983.

SALES, Rodrigo de. **A presença de Kaiser no quadro teórico do Tratamento Temático da Informação (TTI)**. 2012. 190 f. Tese (Doutorado em Ciência da Informação)-Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2012.

SANTOS, Paola. Paul Otlet: um pioneiro da organização das redes mundiais de tratamento e difusão da informação registrada. **Ciência da Informação**, Brasília, v. 36, n. 2, p. 54-63, ago. 2007.

SERRAI, Alfredo. História da biblioteca como evolução de uma idéia e de um sistema. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 4, n. 2, p. 141-161, set. 1975.

SHARP, Henry Alexander. **Cataloging: a textbook for use in libraries**. London: Allied Publishers Private, 1964. Disponível em: <https://archive.org/stream/in.ernet.dli.2015.533483/2015.533483.cataloguing-a#page/n5/mode/2up>. Acesso em: 2 abr. 2018.

SHERA, Jesse Hauk; EGAN, Margaret E. Exame atual da Biblioteconomia e da Documentação. *In*: BRADFORD, S. C. **Documentação**. Rio de Janeiro: Editôra Fundo de Cultura, 1961. p. 15-61.

SHERA, Jesse Hauk; EGAN, Margaret E. **Catálogo sistemático: princípios básicos e utilização**. Brasília: Editôra Universidade de Brasília, 1969.

SIMÕES, Maria Da Graça De Melo. Do conceito e dos princípios metodológicos de indexação. Subsídios para uma análise crítica. *In*: JORENTE, Maria José Vicentini; PADRÓN, Dunia Ilanes (orgs.). **Una mirada a la ciencia de la información desde los nuevos contextos paradigmáticos de la posmodernidad**. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2017. p. 151-176.

STATEMENT of International Cataloguing Principles. [S. l.]: IFLA, 2009. Disponível em: https://www.ifla.org/files/assets/cataloguing/icp/icp_2009-en.pdf. Acesso em: 4 mar. 2019.

STATEMENT of International Cataloguing Principles. [S. l.]: IFLA, 2016. Disponível em:

<https://www.ifla.org/publications/statement-of-international-cataloguing-principles>. Acesso em: 4 mar. 2019.

STROUT, Ruth French. The development of the catalog and cataloging codes. **The Library Quarterly**, Chicago, v. 26, n. 4, p. 254-275, oct. 1956.

TÁLAMO, Maria de Fátima Gonçalves Moreira *et al.* Otlet, o criador de estruturas informacionais pela paz mundial. *In*: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 20., 2002, Fortaleza. **Anais** [...]. Fortaleza, 2002.

TÁLAMO, Maria de Fátima Gonçalves Moreira; LARA, Marilda Lara Lopes Ginez de. O campo da Lingüística Documentária. **Transinformação**, Campinas, v. 18, n. 3, p. 203-211, set./dez. 2006.

TÁLAMO, Maria de Fátima Gonçalves Moreira; LARA, Marilda Lara Lopes Ginez de. Interface entre linguística, terminologia e documentação. **Brazilian Journal of Information Science**, Marília, v. 3, n. 2, p. 58-74, jul./dez. 2009.

TAYLOR, Arlene G. On the subject of subjects. **The Journal of Academic Librarianship**, v. 21, n. 6, p. 484-491, nov. 1995.

TAYLOR, Margaret S. **Fundamentals of practical cataloguing**. London: George Allen and Unwin, 1948.

TEIXEIRA, José Carlos A. **Cabeçalhos de assunto**: (manual para estudantes). Niterói: Universidade Federal Fluminense, 1979.

TILLET, Barbara B. Bibliographic structures: the evolution of catalog entries, references, and tracings. *In*: SVENONIUS, Elaine (org.). **Conceptual foundations of descriptive cataloging**. San Diego: Academic Press, 1989. p. 149-165.

TÔRRES, Lecy Maria Caldas. Sistematização da sintaxe de cabeçalho de assunto. **Conexão Rio**. [2000?]. Disponível em: <http://www.conexaorio.com/bitl/lecy/lecy.htm>. Acesso em: 23 abr. 2018.

UNITED NATIONS INTERNATIONAL SYSTEM FOR INFORMATION IN SCIENCE AND TECHNOLOGY (UNISIST). **Principes d'indexation**. Paris: UNESCO, 1975. Disponível em: <http://unesdoc.unesco.org/images/0001/000164/016443FB.pdf>. Acesso em: 11 set. 2018.

UNITED NATIONS INTERNATIONAL SYSTEM FOR INFORMATION IN SCIENCE AND TECHNOLOGY (UNISIST). Princípios de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 10, n. 1, p. 83-94, mar. 1981.

VERNER, Mathilde. Adrien Baillet (1649-1706) and his Rules for an alphabetic subject catalog. **The Library Quarterly**, Chicago, v. 38, n. 3, jul. 1968.

WHEATLEY, Henry Benjamin. **How to make an index**. London: Eliot Stock, 1902. Disponível em: <https://archive.org/details/howtomakeindex00wheauoft>. Acesso em: 4 jun. 2018.

WORLD ENCYCLOPEDIA OF LIBRARY AND INFORMATION SERVICES. 3. ed. Chicago: American Library Association, 1993.

ZENG, Marcia Lei; ŽUMER, Maja. Mapping FRSAD model and other abstract models. *In*: ISKO UK CONFERENCE, 2009a, London. **Proceedings** [...]. London: ISKO UK, 2009a. Disponível em: <http://www.iskouk.org/content/mapping-frsad-model-and-other-abstract-models>. Acesso em: 3 mar. 2016.

ZENG, Marcia Lei; ŽUMER, Maja. Introducing FRSAD and Mapping it with SKOS and other models. *In*: WORLD LIBRARY AND INFORMATION CONGRESS, IFLA GENERAL CONFERENCE AND COUNCIL, 75., 2009b, Milan. **Proceedings** [...]. Milan: IFLA, 2009b. Disponível em: <http://www.ifla.org/past-wlic/2009/>. Acesso em: 3 mar. 2016.

ZENG, Marcia Lei; ŽUMER, Maja. **Introducing FRSAD and mapping it with SKOS and other models**. [2010?]. Based on the work of the FRSAR Working Group. Disponível em: <http://nkos.slis.kent.edu/FRSAR/introducingFRSAD.pdf>. Acesso em: 2 set. 2016.

ŽUMER, Maja; ROE, Sandy. News from the classification & indexing section. **Metadata Newsletter**, v. 2, n. 1, p. 21-23, June. 2016. Disponível em: https://www.ifla.org/files/assets/classification-and-indexing/metadata_newsletter-20160803.pdf. Acesso em: 13 jul. 2016.

ŽUMER, Maja; SALABA, Athena; ZENG, Marcia Lei. Functional Requirements for Subject Authority Records (FRSAR): a conceptual model of aboutness. *In*: INTERNATIONAL CONFERENCE ON ASIAN DIGITAL LIBRARIES, 10., 2007, Hanoi, Vietman. **Proceedings** [...]. Germany: Springer-Verlag Berlin Heidelberg, 2007. p. 487-492. Disponível em: <https://www.researchgate.net/publication/225622866>. Acesso em: 30 mar. 2016.

ŽUMER, Maja; ZENG, Marcia Lei; SALABA, Athena. **FRSAD: conceptual modeling of aboutness**. California: Libraries Unlimited, 2012.

ŽUMER, Maja; ZENG, Marcia Lei; SALABA, Athena. Application of FRBR and FRSAD to classification systems. *In*: INTERNATIONAL UDC SEMINAR, 2015, Lisbon. **Proceedings** [...]. Würzburg: Ergon-Verlag, 2015. Edited by Aida Slavic e Maria Inês Cordeiro.

ANEXO – MODELO *FUNCTIONAL REQUIREMENTS FOR SUBJECT AUTHORITY DATA (FRSAD)*



International Federation of Library Associations and Institutions

**Functional Requirements for Subject Authority Data (FRSAD)
A Conceptual Model**

IFLA Working Group on the
Functional Requirements for Subject Authority Records (FRSAR)

Editors: Marcia Lei Zeng, Maja Žumer, Athena Salaba

Approved by the Standing Committee of the IFLA Section on Classification and
Indexing

June 2010

CONTENTS

MEMBERS OF THE IFLA WORKING GROUP ON THE FUNCTIONAL REQUIREMENTS FOR SUBJECT AUTHORITY RECORDS (FRSAR)	4
1. BACKGROUND.....	6
2. PURPOSE AND SCOPE.....	8
2.1 PURPOSE.....	8
2.2 SCOPE.....	9
2.3 ABOUTNESS AND OFNESS.....	10
2.4 METHOD.....	12
2.5 COMPONENTS OF THE STUDY.....	12
3. ENTITIES.....	14
3.1 DIAGRAMMING CONVENTIONS.....	14
3.2 GENERAL FRAMEWORK.....	14
3.3 CHOICE OF TERMS FOR FRSAD ENTITIES.....	16
3.4 THEMA.....	16
3.5 NOMEN.....	18
4. ATTRIBUTES.....	19
4.1 ATTRIBUTES OF A THEMA.....	19
4.2 ATTRIBUTES OF A NOMEN.....	21
5. RELATIONSHIPS.....	25
5.1 WORK-TO-THEMA RELATIONSHIP.....	25
5.2 THEMA-TO-NOMEN RELATIONSHIP.....	25
5.3 THEMA-TO-THEMA RELATIONSHIPS.....	26
5.4 NOMEN-TO-NOMEN RELATIONSHIPS.....	31
6. USER TASKS.....	33
6.1 USERS AND USE.....	33
6.2 USER TASKS.....	34
6.3 ASSESSING VALUES RELATIVE TO USER TASKS.....	35
6.4 MAPPING OF ATTRIBUTES, RELATIONSHIPS, AND USER TASKS.....	37
7. CONCLUSION.....	38
APPENDIX A. MODELING ABOUTNESS.....	39
A.1 SUBJECT RELATIONSHIP AND GROUP 3 ENTITIES INTRODUCED IN FRBR.....	39
A.2 POSSIBLE APPROACHES TO THE MODEL OF ABOUTNESS.....	40
APPENDIX B. RELATIONSHIP OF FRSAD WITH FRBR AND FRAD.....	45
B.1 RELATIONSHIP OF FRSAD WITH FRBR.....	45
B.2 RELATIONSHIP OF FRSAD WITH FRAD.....	46
APPENDIX C. FRSAD MODEL AND OTHER MODELS.....	48
C.1 THE IMPORTANCE OF THE THEMA-NOMEN MODEL.....	48
C.2 MAPPING THE FRSAD MODEL TO OTHER MODELS.....	49
C.3 CONCLUSION.....	50
APPENDIX D. EXAMPLES FROM SUBJECT AUTHORITY SYSTEMS.....	51
D.1 EXISTING MODELS OF THEMA TYPES.....	51
D.2 THEMA-THEMA RELATIONSHIPS PRESENTED IN SUBJECT AUTHORITY DATA.....	55

D.3	SAME THEMA REPRESENTED BY NOMENS FROM DIFFERENT SCHEMES.....	68
D.4	EXAMPLES OF DISPLAY RECORDS FROM CONTROLLED VOCABULARIES OR SUBJECT AUTHORITY FILES	68
REFERENCES	73

**MEMBERS OF THE IFLA WORKING GROUP ON THE
FUNCTIONAL REQUIREMENTS FOR SUBJECT AUTHORITY
RECORDS (FRSAR)**

Working Group

Leda Bultrini
ARPA Lazio (Regional Environment Protection Agency), Italy

Lois Mai Chan
University of Kentucky, USA

Jonathan Furner
University of California Los Angeles, USA

Edward O'Neill
OCLC, USA

Gerhard Riesthuis
University of Amsterdam, The Netherlands

Athena Salaba (Co-Chair and Secretary)
Kent State University, USA

Diane Vizine-Goetz
OCLC, USA

Ekaterina Zaytseva
Russian National Public Library for Science and Technology, Russia

Marcia Lei Zeng (Chair)
Kent State University, USA

Maja Žumer (Co-Chair)
University of Ljubljana, Slovenia

Advisory Group

Victoria Francu (2005-06)
Central University, Bucharest, Romania

Hemalata Iyer (2008)
Albany State University of New York, USA

Dorothy McGarry
University of California, Los Angeles, USA

David Miller
Curry College, USA

Päivi Pekkarinen
National Library of Health Sciences, Finland

Barbara Tillett
Library of Congress, USA

This document was prepared by:

Marcia Lei Zeng, Maja Žumer, Athena Salaba, Jonathan Furner, Lois Mai Chan, Ed
O'Neill, Diane Vizine-Goetz

1. BACKGROUND

The IFLA Study Group on the Functional Requirements for Bibliographic Records (FRBR) developed a conceptual model showing the entities and relationships of the bibliographic universe in 1997. The purpose of the FRBR model is to identify the functional requirements of information in bibliographic records to facilitate the specified user tasks.¹ The basic entities of the FRBR model are the result of a logical analysis of the data typically represented in bibliographic records. The entities are divided into three groups:

Group 1 entities are defined as the products of intellectual or artistic endeavours that are named or described in bibliographic records: *work*, *expression*, *manifestation*, and *item*.

Group 2 entities are those responsible for the intellectual or artistic content, the physical production and dissemination, or the custodianship of the Group 1 entities: *person*, *corporate body*, and *family*².

Group 3 entities represent an additional set of entities that serve as the subjects of works: *concept*, *object*, *event*, and *place*.

The FRBR final report presents the entity-relationship model, identifies entities and their attributes, and defines relationships among entities. Although in the FRBR model the entities of all three groups are defined, the main focus is on the first group. The developers of FRBR envisioned that its extensions would cover the additional data that are normally recorded in authority records.

The Working Group on Functional Requirements and Numbering of Authority Records (FRANAR) was established in April 1999. It was charged to continue the work of FRBR by developing a conceptual model for entities described in authority records. Authority data in the context of their work is defined as “the aggregate of information about a *person*, *family*, *corporate body*, or *work* whose name is used as the basis for a controlled access point for bibliographic citations or records in a library catalogue or bibliographic file”³. The primary purpose of the Functional Requirements for Authority Data (FRAD) conceptual model is “to provide a framework for the analysis of functional requirements for the kind of authority data that is required to support authority control and for the international sharing of authority data. The model focuses on data, regardless of how they may be packaged (e.g., in authority records)”⁴. While the FRANAR Working Group has

¹ *Functional Requirements for Bibliographic Records: Final Report*. (1998). IFLA Study Group on the Functional Requirements for Bibliographic Records. München: KG Saur

² "Family" entity was added in *Functional Requirements for Authority Data - A Conceptual Model*. (2009).

³ *Functional Requirements for Authority Data - A Conceptual Model*. (2009). IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR), ed. by Glenn E. Patton. München: K.G. Saur. p.15.

⁴ *Ibid.*, p. 13.

included some aspects of subject data in their model, they have not undertaken the full analysis of the entities and relationships relevant to subject authorities⁵.

As a result, the IFLA Working Group on the Functional Requirements for Subject Authority Records (FRSAR) was formed in 2005 to address subject authority data issues and to investigate the direct and indirect uses of subject authority data by a wide range of users. The FRSAR Working Group (2005 to present) and the FRANAR Working Group (1999 to 2009) both worked in parallel to develop models within the FRBR framework. By the time FRANAR released its final report in June 2009, FRSAR also had released its first draft report of the *Functional Requirements for Subject Authority Data* (FRSAD) for world-wide review. Since the two reports were developed independently, the relationship between the FRSAD and FRAD models is explained in Appendix B of this report.

⁵ *Functional Requirements for Authority Data - A Conceptual Model*. (2009). IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR), ed. by Glenn E. Patton. München: K.G. Saur. p. 8.

2. PURPOSE AND SCOPE

2.1 Purpose

Subject access to information has been a significant approach of users to satisfy their information needs. Research results have demonstrated that the integration of controlled vocabulary information with an information retrieval system helps users perform more effective subject searches. This integration becomes possible when subject authority data (information about subjects from authority files) are linked to bibliographic files and are made available to users.

The purpose of authority control is to ensure consistency in representing a value—a name of a person, a place name, or a term or code representing a subject—in the elements used as access points in information retrieval. For example, “World War, 1939-1945” has been established as an authorised subject heading in the *Library of Congress Subject Headings* (LCSH). When using LCSH, in cataloguing or indexing, all publications about World War II are assigned the established heading regardless of whether a publication refers to the war as the “European War, 1939-1945”, “Second World War”, “World War 2”, “World War II”, “WWII”, “World War Two”, or “2nd World War”. The synonymous expressions lead to the authorised heading. This ensures that all publications about World War II can be retrieved by and displayed under the same subject heading, either in local catalogue or database or in a union catalogue.

In almost all large bibliographic databases, authority control is achieved manually or semi-automatically by means of an authority file. The file contains data about access points – names, titles, or subject terms – that have been authorised for use in bibliographic records. In addition to ensuring consistency in subject representation, a subject authority system may also record established semantic relationships among subject concepts and/or their labels. Data in a subject authority system are connected through semantic relationships, which may be expressed in subject authority records or generated according to specific needs (e.g., presenting the broader and narrower concepts) in printed or online displays of thesauri, subject headings lists, classification schemes, and other subject authority systems. Such systems have been referred to as “controlled vocabularies”, “structured vocabularies”, “concept schemes”, “encoding schemes”, and “knowledge organization systems” interchangeably depending on their function and structure, as well as according to the communities that use them. Given the purpose of this report, the discussions about subject authority data apply to all systems and structures referred to by these terms. The study follows FRBR’s approach in that it makes no *priori* assumption about the physical structure or storage of authority data.

2.2 Scope

The primary purpose of this study is to produce a framework that will provide a clearly stated and commonly shared understanding of what the subject authority data/record/file aims to provide information about, and the expectation of what such data should achieve in terms of answering user needs. The role of the FRBR Working Group was defined in the following terms of reference:

- To build a conceptual model of Group 3 entities within the FRBR framework as they relate to the *aboutness* of works;
- To provide a clearly defined, structured frame of reference for relating the data that are recorded in subject authority records to the needs of the users of that data;
- To assist in an assessment of the potential for international sharing and use of subject authority data both within the library sector and beyond.

To fulfil these terms of reference, the FRBR Working Group established two sub-groups: User Tasks Sub-Group and Subject Entities Sub-Group.

The User Tasks Sub-Group focused on user studies and the definition of user tasks. For the purposes of this study, the users of subject authority data include information professionals who create and maintain subject authority data, information professionals who create and maintain metadata, intermediaries and end users who search for information to fulfil information needs. The functional requirements for subject authority data are defined in relation to the following general tasks that are performed by these users:

Find one or more subjects and/or their appellations, that correspond(s) to the user's stated criteria, using attributes and relationships;

Identify a subject and/or its appellation based on its attributes or relationships (i.e., to distinguish between two or more subjects or appellations with similar characteristics and to confirm that the appropriate subject or appellation has been found);

Select a subject and/or its appellation appropriate to the user's needs (i.e., to choose or reject based on the user's requirements and needs);

Explore relationships between subjects and/or their appellations (e.g., to explore relationships in order to understand the structure of a subject domain and its terminology).

The Subject Entities Sub-Group focused on the Group 3 entities including the study of current FRBR Group 3 entities and alternatives in order to define:

- a) entities that can serve as subjects of a *work* (the “has as subject” relationship);
- b) possible sub-entities in the Group 3 cluster; and
- c) additional entities related to the Group 3 cluster.

The FRSAR Working Group is aware that some controlled vocabularies provide terminology to express other aspects of *works* in addition to subject (such as form, genre, and target audience of resources). While very important and the focus of many user queries, these aspects describe **isness** or what class the *work* belongs to based on form or genre (e.g., novel, play, poem, essay, biography, symphony, concerto, sonata, map, drawing, painting, photograph, etc.) rather than what the *work* is about. Some of these aspects are explicitly covered by the FRBR model, for example, “form of work,” “intended audience,” etc. as attributes of *work*. While the Group acknowledges that there are cases where a vocabulary provides terminology, or has been used, also for **isness**, the focus of the FRSAD model is on **aboutness** (the FRBR-defined relationship *work* “has as subject ...”). On the other hand, any case of a *work* about a form or genre (e.g. about romance novels, about dictionaries) clearly falls within the aboutness category.

2.3 Aboutness and Ofness

When modelling the fundamental classes of bibliographic entities we necessarily face the challenge of carrying out the most appropriate analysis of aboutness—i.e., the relation between a work and its subject matter. Aboutness is a concept that is central to the field of knowledge organization, and many authors have made significant contributions to our understanding of the nature of work–subject relations. Some of these contributions appear in the literature of library and information science (LIS),^{6,7,8} while others have been made by philosophers of logic and language^{9,10,11}. A review of these literatures will show that there is not as much consensus on the nature of aboutness as one might hope to find: in fact, there are a wide variety of views.

At the risk of oversimplifying what is undoubtedly a complex situation, we might consider that it is possible to place views on aboutness on a spectrum whose poles represent the two extremes of “nominalism” (or, following Hjørland,¹² “idealism”) and “realism.” For the thoroughgoing nominalist, it does not make sense to talk of works “having” or “being about” subjects—aboutness should be conceived not as a property of

⁶ Wilson, P. (1968). *Two kinds of power: An essay on bibliographic control*. Berkeley, CA: University of California Press.

⁷ Hutchins, W. J. (1977). On the problem of “aboutness” in document analysis. *Journal of Informatics*, 1 (1): 17-35.

⁸ Hjørland, B. (1992). The concept of “subject” in information science. *Journal of Documentation*, 48 (2): 172-200.

⁹ Ryle, G. (1933). “About.” *Analysis*, 1 (1): 10-12.

¹⁰ Putnam, H. (1958). Formalization of the concept “about.” *Philosophy of Science*, 25 (2): 125-130.

¹¹ Goodman, N. (1961). About. *Mind*, 70 (277): 1-24.

¹² Hjørland, *op. cit.*

works but rather as a relation, constructed by a particular person at a particular time, between a particular set of works and a particular linguistic expression (i.e., a name or label). The realist, on the other hand, is content to proceed on the assumption that subjects are real things that exist separately from the linguistic expressions that we use to name them, and that it is possible to determine “the” subject(s) of any given work. Of course, there are other points of view on aboutness that may be located either at intermediate points between these two poles, or on different spectrums. But it is probably fair to say that most people who are actively engaged in the tasks of designing bibliographic classification schemes, indexing documents in accordance with such schemes, and using those schemes as tools for finding the kinds of documents we want, act in accordance with assumptions that are consistent with some version of the realist viewpoint.

Ultimately, the FRSAR Working Group does not take a philosophical position on the nature of aboutness, rather, it looks at the problem from the user’s point of view. When confronted with an information need that can potentially be met by finding and using a document about a certain subject, the user both expects to be able to formulate a search statement specifying the subject, and expects that the tools and services at hand are capable of comparing such search statements with the subject statements generated by cataloguers and indexers.

Those LIS authors who have focused on the subjects of visual resources, such as artworks and photographs, have often been concerned with how to distinguish between the “aboutness” and the “ofness” (both specific and generic depiction or representation) of such works.¹³ In this sense, “aboutness” has a narrower meaning than that used above. A painting of a sunset over San Francisco, for instance, might be analyzed as being (generically) “of” sunsets and (specifically) “of” San Francisco, but also “about” the passage of time. Standard metadata schemas for cultural objects correspondingly allow for distinctions to be made among (a) **description** of the **kinds** of things depicted in *works*, (b) **identification** of the **particular** people, objects, events, and places depicted, and (c) **interpretation** of the **meanings** of *works*.¹⁴ The FRSAR Working Group recognizes that any statement on the “ofness” of a work can be a **subject statement** and is likely to be the target of a catalogue user’s search. The FRSAD model is therefore applicable to situations where ofness statements (depiction) are included in subject access.

¹³ Shatford, S. (1986). Analyzing the subject of a picture: A theoretical approach. *Cataloging & Classification Quarterly*, 6 (3): 39-62.

¹⁴ *Categories for the Description of Works of Art* (CDWA). (2000). Eds. Baca, M. and Harpring, P. The J. Paul Getty Trust and College Art Association. Los Angeles, CA: Getty Research Institute. Available at: http://www.getty.edu/research/conducting_research/standards/cdwa/index.html (accessed 2010-01-20).

2.4 Method

The method used to build this conceptual model is the entity analysis technique used in FRBR. The development of such a conceptual model consists of the following steps¹⁵:

- 1) The analysis starts with the **user tasks** as well as the key objects that are of interest to users in a specific domain.
- 2) The attention is not on individual data but on the “things” the data describe. Each of the objects of interest or **entities** defined for the model serves as the focal point for a cluster of data.
- 3) At a higher level, an entity diagram depicts the **relationships** that normally hold between one type of entity and another type of entity.
- 4) Important characteristics or **attributes** of each entity are then identified.
- 5) Each attribute and relationship is mapped to the user tasks. Relative values of importance are assigned to each attribute and relationship with specific reference to the task performed and the entity that is the object of the user’s interest.

These steps were followed in the development of this model. The background of the initial conceptual analysis of user tasks and entities performed by the Group is explained in detail in Appendix A.

2.5 Components of the Study

The remainder of the report is divided into two major segments: the main body of the report follows the FRBR report structure and presents the entity-relationship model; the second segment contains four appendices that explain the methodology and implementation considerations.

The remaining body of the report consists of four chapters:

- Chapter 3 of the study identifies and defines the entities used in the model.
- Chapter 4 analyses the attributes associated with each of the entities defined in the model.
- Chapter 5 analyses the relationships used in the model, including the relationships operating both at the general level and between specific instances of entities.
- Chapter 6 presents the user tasks and then maps the attributes associated with each entity to the four generic user tasks the subject authority data is intended to support, showing the relevance of each attribute or relationship to each of the user tasks.

The appendices contain additional materials:

¹⁵ *Functional Requirements for Bibliographic Records: Final Report*. (1998). IFLA Study Group on the Functional Requirements for Bibliographic Records. München: KG Saur. p.9. Section 2.3 Methodology.

- Appendix A discusses the initial analysis used as basis in building the conceptual model for aboutness. It analyses the subject relationship and Group 3 entities introduced in FRBR as well as possible approaches to the model of aboutness.
- Appendix B explains the relationship of the FRSAD model with the FRBR and the FRAD models, given the fact that both FRAD and FRSAD are based on FRBR but were developed in parallel.
- Appendix C furthers the discussion of the importance of the FRSAD model and maps it with related developments including the ISO standard for thesauri, the W3C's Simple Knowledge Organization System (SKOS) and OWL Web Ontology Language, and the Dublin Core Abstract Model.
- Appendix D contains examples from existing subject authority systems through the perspective of the FRSAD model.

3. ENTITIES

3.1 Diagramming Conventions

FRSAD follows the conventions used in both FRBR and FRAD:



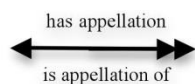
A rectangle represents an entity.



A single-headed arrow on a line represents a relationship in which any given instance of the entity at the beginning of the line may be associated with only one instance of the entity to which the arrow is pointing.



A double-headed arrow on a line represents a relationship in which any given instance of the entity at the beginning of the line may be associated with one or more instances of the entity to which the arrow is pointing.



A relationship above the line indicates left-to-right direction; a relationship below the line indicates a right-to-left direction.

3.2 General Framework

The FRSAR Working Group proposes a generalisation of FRBR, as seen in Figure 3.1. This diagram is based on the original FRBR Figure 3.3 that depicts the “subject” relationships between works and entities in Group 1, Group 2, and Group 3. A new entity, *family*, added by FRAD in Group 2, is also reflected in this general framework. The entities in the third group represent an additional set of entities that serve as the subjects of *works*. The group includes, according to the FRBR model, *concept* (an abstract notion or idea), *object* (a material thing), *event* (an action or occurrence), and *place* (a location).

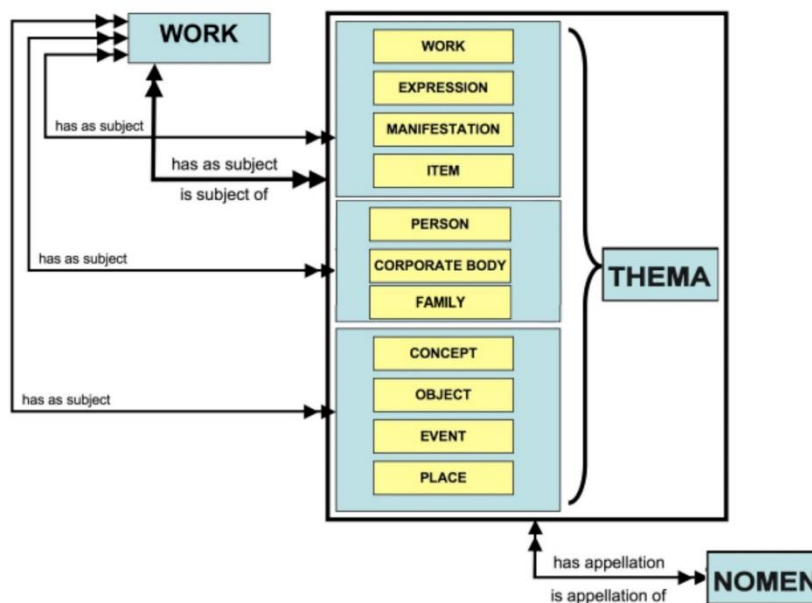


Figure 3.1 FRSAD's relation to FRBR (with the addition of FRAD entity family)

The FRSAR Working Group introduced the following two entities:

Thema: any entity used as a subject of a *work*

Nomen: any sign or sequence of signs (alphanumeric characters, symbols, sound, etc.) that a *thema* is known by, referred to, or addressed as.

The Functional Requirements for Subject Authority Data (FRSAD) model is presented as:



Figure 3.2 FRSAD conceptual model

Both “has as subject/is subject of” and “has appellation/is appellation of” relationships are many-to-many relationships. Any *work* can have more than one *thema* and any *thema* can be the subject of more than one *work*. We can take “A brief history of time: from the big bang to black holes” by Stephen W. Hawking as an example. The *work* has several *themas*: “cosmology”, “space and time”, “unification of physics”, “black holes”, “big bang”, “history of time”, “universe”, etc. There are many other *works* about any of these

themas. For any of the *themas* in this list (presented here as terms in English) there are other possible *nomens* in other languages and in different controlled vocabularies.

Some *works* are perceived as having no *thema* as subject (such as certain musical works or abstract artwork), and no subject access is provided to them. These cases are not covered by FRSAD. The cases of a *thema* without a *nomen* are also beyond the scope of this model.

3.3 Choice of Terms for FRSAD Entities

The Working Group chose Latin terms, *thema* (plural *themata* or *themas*) and *nomen* (plural *nomina* or *nomens*), because they have no pre-existing meaning in our context, are culturally neutral and do not require translation. For *thema*, other possible (English) terms include “subject”, “topic”, and “concept”; however, even discussions within the Working Group proved that there are very different views on granularity (some see “subject” and “topic” as synonyms, while others see “topic” as a component of “subject”). The Working Group needed to distinguish *thema* from the previously defined FRBR entity *concept* because *thema* is a superclass of all FRBR entities (to be explained in the next section). For *nomen*, it is the case that the term “name” is often considered synonymous with proper name. In addition, the Working Group needed to distinguish *nomen* from the FRAD entity *name* because *nomen* includes FRAD entities *name*, *identifier*, and *controlled access point*.

3.4 THEMA

Thema is defined as “any entity used as a subject of a *work*”. Therefore this model confirms one of the basic relationships defined in FRBR: **WORK has as subject THEMA / THEMA is subject of WORK**.

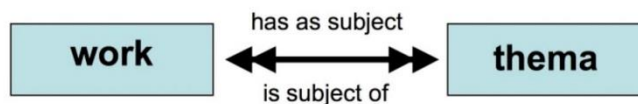


Figure 3.3 *Work-Thema* relationship

According to Delsey, the first broad objective of FRSAD is to ensure that the scope of the entities defined is sufficient to cover everything that a user of a library catalogue might view as a “subject”¹⁶. We may therefore see *thema* from different points of view. From the point of view of end-users and intermediaries, *thema* comprises the aboutness of the

¹⁶ Delsey, T. (2005). Modeling subject access: Extending the FRBR and FRANAR conceptual models. *Cataloging & Classification Quarterly* 39 (3/4): 49–61.

(possibly unknown) resources that will satisfy the information need. From the point of view of information professionals who create metadata, one or more *themas* capture the aboutness of a particular resource.

Within the FRBR framework, *thema*, which can be viewed as an entity on its own and also as a super-entity or superclass, includes existing Group 1 and Group 2 entities, and additionally, all others that serve as the subjects of *works* (i.e., Group 3). In other words, *thema* is a superclass of all FRBR entities (Figure 3.4). Defining *thema* as a super entity enables modelling of relationships and attributes on a more general and abstract level.

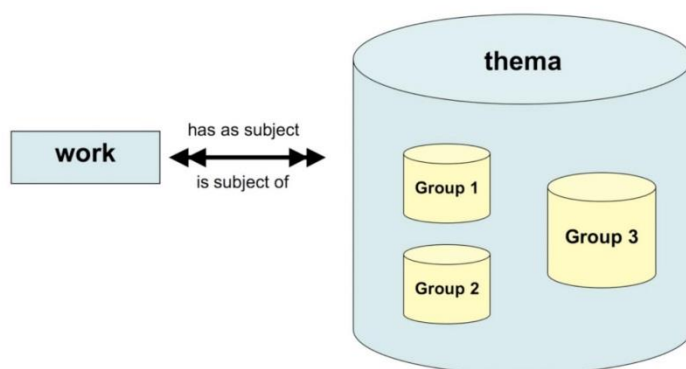


Figure 3.4 Within the FRBR framework, *thema* includes existing Group 1 and Group 2 entities and, in addition, all others that serve as the subjects of *works* (i.e., Group 3)

While the original FRBR Group 3 entities (*object, concept, event, place*) may be used in a given implementation, the Working Group does not propose them as a universally applicable set of Group 3 entities. In a particular application *themas* would normally have implementation-specific types; but, based on the pilot study, as explained in Appendix A, there seems to be no generally applicable categorization of *thema*. Appendix D provides examples of existing implementations through the perspective of the FRSAD model.

Themas can vary substantially in complexity or simplicity. Depending on the circumstances (the subject authority system, user needs, the nature of the *work*, etc.) the aboutness of a *work* can be expressed as a one-to-one relationship between the *work* and the *thema*; this means that the totality of the aboutness is encompassed in a single *thema*. In other circumstances the relationship is one-to-many, meaning that the aboutness of the *work* is captured in two or more *themas*. It is virtually impossible to define what the universal “atomic” level of a *thema* might be, because any *thema* can be fragmented further. The argument can be reversed: simple *themas* may be combined or aggregated, resulting in more complex *thema(s)*. In each particular implementation the atomic level is specified and rules guide the creation of *nomens* for complex *themas*.

To some extent the granularity of a *thema* also depends upon the controlled vocabulary used for its appellation(s). Often the complexity of a *thema* is associated with the complexity of the *nomen* by which it is represented. Since the proposed model introduces a clear split between the *thema* (“the thing”) and the *nomen* (“the label” used to refer to it), the complexity of the semantic and syntactic rules for creating or establishing a *nomen* is not directly reflected in the complexity of the *thema*, nor is it completely independent. Some types of controlled vocabularies (such as subject headings systems) enable the establishment of complex *themas* (e.g., by creating pre-coordinated strings), while others (such as thesauri) are mainly conceived for the use of more atomic *themas*.

3.5 NOMEN

The FRASAD model proposes a new relationship: **THEMA has appellation NOMEN/NOMEN is appellation of THEMA** (Figure 3.5).

Nomen is defined as “any sign or sequence of signs (alphanumeric characters, symbols, sound, etc.) that a *thema* is known by, referred to, or addressed as. Examples include “love,” “∞,” or “595.733.” A *Nomen* can be human-readable or machine-readable. *Nomen* is a superclass of the FRAD entities *name*, *identifier*, and *controlled access point*.

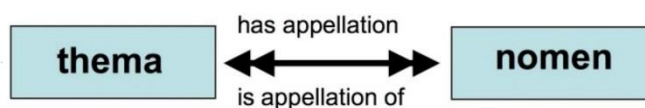


Figure 3.5 *Thema-Nomen* relationship

In general (i.e., in natural language or when mapping different vocabularies) the “has-appellation/is appellation of” relationship is a many-to-many relationship. A *thema* has one or more *nomens* and there may be a *nomen* referring to more than one *thema*. In a given controlled vocabulary, however, a *nomen* should be an appellation of only one *thema*, as illustrated in Figure 3.6. See Appendix D for examples from subject authority systems.

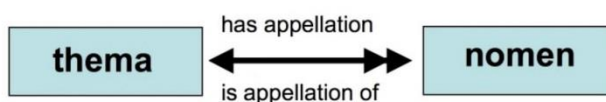


Figure 3.6 *Thema-Nomen* relationship within a controlled vocabulary

4. ATTRIBUTES

4.1 Attributes of a *THEMA*

In the FRSAD model, the entity *thema* is defined in a very abstract and general way. Attributes of a *thema* are implementation-dependent and will vary. “Type” and “scope note” can be considered general attributes, but particular values of “type” are, again, implementation-dependent. In any implementation there will normally be additional attributes of a *thema* other than “type” and “scope note”. Those attributes will be dependent on both the type of *themas* and the application domain.

4.1.1 Type of *thema*

The category to which a *thema* belongs in the context of a particular knowledge organisation system.

In an implementation *themas* can be organised based on category, kind, or type. This report does not suggest specific types, because they may differ depending on the implementation (see Appendix A).

For example, in some implementations the original FRBR entities *work*, *expression*, *manifestation*, *item*, *person*, *family*, *corporate body*, *concept*, *object*, *event*, and *place* can be used as types, perhaps even adding *time* as has been suggested (see Appendix A.2). In general, any entity defined in FRBR and/or FRAD may become a type of *thema*. And their attributes (also defined in FRBR and FRAD) will equally apply.

In other implementations, a different set of types may be defined. Two examples from existing implementations, UMLS and AAT, are presented below. Detailed explanations of these subject authority systems can be found in Appendix D.

1) *Unified Medical Language System* (UMLS) semantic types¹⁷

Entities

Physical Object
Organism
Anatomical Structure

¹⁷ National Library of Medicine. (2003-). *Unified Medical Language System. Current Semantic Types*. UMLS 2004AB Documentation. Last updated: 21 March 2008. Available at: http://www.nlm.nih.gov/research/umls/META3_current_semantic_types.html (accessed 2010-01-20).

Manufactured Object
 Substance
 Conceptual Entity
 Idea or Concept
 Finding
 Organism Attribute
 Intellectual Product
 Language
 Occupation or Discipline
 Organization
 Group Attribute
 Group

Events

Activity
 Phenomenon or Process

2) *Art and Architecture Thesaurus* (AAT) facets¹⁸

[Top of the AAT hierarchies]
 Associated Concepts
 Physical Attributes
 Styles and Periods
 Agents
 Activities
 Materials
 Objects

These examples show very different approaches to defining types of *themas*. In the UMLS, *themas* are first differentiated as “Entity” or “Event.” The types of UMLS “Entity” are “Physical Object” or “Conceptual Entity.” The types of “Events” are grouped into “Activity” and “Phenomenon or Process.” In AAT, all *themas* are categorized into seven types: “Associated Concepts,” “Physical Attributes,” “Styles and Periods,” “Agents,” “Activities,” “Materials,” and “Objects.”

Clearly since *themas* are very different, they will also necessarily have different attributes. In the first example “Substance” (a physical object) will have very different attributes from “Organization” (a conceptual entity) in the UMLS. The same is true for the attributes of “Styles and Periods,” “Agents,” and “Materials” in the AAT.

¹⁸ Art and Architecture Thesaurus Online Hierarchy Display. In: *Art and Architecture Thesaurus*. (2000-). Los Angeles: J. Paul Getty Trust, Vocabulary Program. Available at: <http://www.getty.edu/vow/AATHierarchy?find=&logic=AND¬e=&english=N&subjectid=300000000> (accessed 2010-01-20).

Another possible distinction can be made at the *thema* level between **Classes** and **Instances**. These two types of *thema* are fundamental and many subject authority systems recognize them. The **Class/Instance** distinction is essentially equivalent to the universal/particular distinction, typically made on the basis of instantiability (and hence is sometimes characterized in philosophical literature as a kind/instance distinction).

For example:

<u>Class</u>	<u>Instance</u>
Palaces	Buckingham Palace
Ships	The Lusitania
Battles	The Battle of Hastings

4.1.2 Scope Note

A text describing and/or defining the *thema* or specifying its scope within the particular subject authority system.

4.2 Attributes of a *NOMEN*

The attributes of a *nomen* listed below represent the most common (general) attributes. Not all are applicable in every case, and the list is not comprehensive. While all listed attributes are applicable to individual instances of *nomens*, some may also be used for an entire subject authority system and declared on that level. The examples of attribute values are illustrative only and should not be seen as prescriptive. In any particular implementation the actual values of an attribute are selected from a controlled list and/or are coded.

In addition to the proposed general attributes, presented here, there may be additional, implementation-specific attributes.

Current subject authority records typically include other elements such as administrative data. Also, current authority system may allow merging data -- which describe both *thema(s)* and *nomen(s)* -- into one record. As this is a conceptual model, such aspects of implementation are not discussed.

4.2.1 Type of *nomen*

Category to which the *nomen* belongs.

In addition to other implementation-specific types, there are two important values of this attribute:

- identifier – the sign or sequence of signs assigned to an entity that is persistent and unique within a domain
- controlled name – the name constructed during the authority control or vocabulary maintenance process that usually serves as an access point (note: labelled as *controlled access point* in FRAD).

If needed, the values of the type attribute may be further refined; for example, additional refinement may include different kinds/formats of identifiers (e.g., URI, ISBN).

4.2.2 Scheme

The scheme in which the *nomen* is established, including value encoding schemes (subject heading lists, thesauri, classification systems, name authority lists, etc.) and syntax encoding schemes (standards for encoding dates, etc.).

Examples of attribute values:

- LCSH
- DDC
- UDC
- ULAN
- ISO 8601

4.2.3 Reference Source of nomen

The source in which the *nomen* is found. It may also be modelled as a relationship with the appropriate Group 1 entity.

Examples of attribute values:

- Encyclopaedia Britannica
- Webster's Third New International Dictionary (1961)
- Columbia Gazetteer

4.2.4 Representation of nomen

The data type in which the *nomen* is expressed.

Examples of values:

- alphanumeric
- sound
- graphic

4.2.5 Language of nomen

The language in which the *nomen* is expressed.

Examples of values:

- English
- Greek
- Chinese
- Slovenian

4.2.6 Script of nomen

The script in which the *nomen* is expressed.

Examples of values:

- Cyrillic
- Thai
- Chinese (Simplified)
- Chinese (Traditional)

4.2.7 Script conversion

The rule, system, or standard used to render the *nomen* in a different representation.

Examples of values:

- Pinyin
- ISO 3602, 1998, Romanisation of Japanese (kana script)

4.2.8 Form of nomen

Any additional information that helps to interpret the *nomen*.

Examples of attribute values:

- Full name
- Abbreviation
- Formula

4.2.9 Time of validity of nomen

The time period, in which the *nomen* is/was used or is/was valid within a subject vocabulary system.

This should not be confused with the temporal aspect of a *thema*.

Examples of values:

- until May 11, 1949
- after 1945
- 1945 - 1967

4.2.10 Audience

The community or user group for which the *nomen* is the preferred form.

In the global environment it is usually impossible to declare one *nomen* of a *thema* to be the preferred form. The notion of “preferred” form can, in general, be tied only to a particular community, defined by name, rule, or convention.

Examples of values:

- English-speaking users
- Scientists
- Children

4.2.11 Status of nomen

The status of the *nomen* in a subject authority system.

This should not be confused with the management of a subject authority system (e.g., including or excluding a *thema*).

Examples:

- Proposed
- Accepted
- Obsolete

5. RELATIONSHIPS

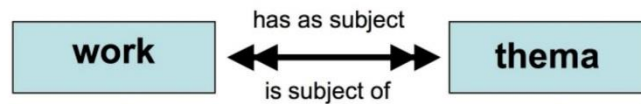
The FRSAD model establishes two sets of relationships:

- 1) Relationships between different types of entities: *WORK*-to-*THEMA* and *THEMA*-to-*NOMEN*. These are the primary relationships and are also illustrated in Chapter 3 where the entities are presented.
- 2) Relationships between entities of the same type: *THEMA*-to-*THEMA* and *NOMEN*-to-*NOMEN*. These are presented in detail in this chapter.

5.1 *WORK*-to-*THEMA* Relationship

The *Work*-to-*Thema* relationship is discussed in Section 3.2. In the FRSAD model, *thema* includes existing Group 1 and Group 2 entities and all entities that serve as the subjects of *works*. Their relationships can be illustrated as:

Work has as subject Thema / Thema is subject of Work



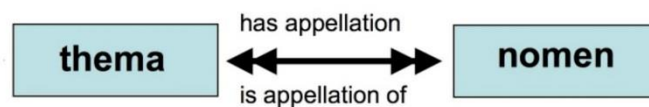
(From Figure 3.3 *Work-Thema* relationship)

Thema refers to anything that can be the subject of a *work*. Presented within the entity-relationship model, the *WORK*-to-*THEMA* relationship is many-to-many: any *work* can have one or more *themas*, and any *thema* may be the subject of one or more *works*.

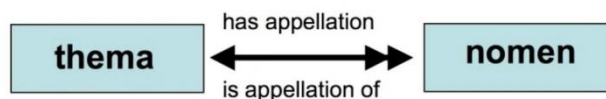
5.2 *THEMA*-to-*NOMEN* Relationship

The *THEMA*-to-*NOMEN* relationship is specified with the following statements:

Thema has appellation Nomen / Nomen is appellation of Thema



(From Figure 3.5 *Thema-Nomen* relationship)



(From Figure 3.6 *Thema-Nomen* relationship within a controlled vocabulary)

As stated in Section 3.4, in general (i.e., in natural language or when mapping different vocabularies) the relationship “has appellation/is appellation of” is a many-to-many relationship. Any *thema* may have more than one *nomen* (see Figure 3.5); and any *nomen* may be the appellation of more than one *thema*. In subject authority-related processes such as integrating or mapping controlled vocabularies, this condition may result in ambiguity regarding the meaning, scope, and definition of *themas* represented by the same *nomen*. Therefore, in a given controlled vocabulary, a *nomen* normally is an appellation of only one *thema* but a *thema* can have more than one *nomen* (see Figure 3.6). Based on this general principle, a more complex *nomen*, for example, with added qualifiers, must be constructed when necessary in order to eliminate ambiguity.

5.3 *THEMA*-to-*THEMA* Relationships

Only relationships directly applicable for subject access are analyzed here. The FRBR and FRAD models cover additional entity-to-entity relationships such as relationships between *works*.

In order to ensure that (1) the attributes relevant to the construction and use of subject authority data are adequately covered, and (2) the model provides a clear and pragmatic representation of the relationships that are “reflected through subject access points in bibliographic records as well as those reflected in the syndetic structure of thesauri, subject headings lists, and classification schemes and in the syntactic structure of indexing strings”¹⁹, the *thema*-to-*thema* relationship types are discussed in the context of subject authority systems.

5.3.1 Hierarchical Relationships

Hierarchical structures show relationships between and among concepts and classes of concepts. Hierarchical relationships reveal degrees or levels of superordination and subordination, where the superordinate term represents a class or a whole, and subordinate terms refer to its members or parts. *Hierarchical structures* are found in classification schemes, subject heading systems, thesauri, and other knowledge organization systems. Used in the bibliographic universe, hierarchical relationships provide disambiguation functions to assist with the *identify* user task. Yet they are the

¹⁹ Delsey, T. (2005). Modeling subject access: Extending the FRBR and FRANAR conceptual models. *Cataloging & Classification Quarterly* 39 (3/4): 49–61.

most effective in furthering linking and navigation objectives, and satisfying the *select*, and especially, the *explore* user tasks. They are of particular aid to users with undefined or very broad information needs and they also allow users to improve their searching.

Typically, a hierarchical relationship may be one of three types: *the generic relationship*, *the hierarchical whole-part relationship*, and *the instance relationship*²⁰. Some concepts can belong to more than one superordinate concept simultaneously. They are considered to have *polyhierarchical relationships*. Other perspective hierarchical relationships also exist (as explained in Section 5.3.1.5).

5.3.1.1 The Generic Relationship

The *generic relationship* is the logical relationship of inclusion. The primary function of the hierarchical relationship is to convey the same concept, but at different levels of specificity²¹. “Of limited domain and range, it is strictly defined in terms of the properties of reflexivity, anti-symmetry, and transitivity”²². It is sometimes represented as the “all-some” relationship. For example, all parrots are birds, and some birds are parrots. But not all parrots are pets therefore the genus-species relationship between parrots and pets does not exist in logic²³. In the computer science literature and formal ontology construction, the characteristic of “inheritance” of genus-species relationships is also widely presumed. This “hierarchical force” assumes that what is true of a given class (e.g., furniture) is true of all member-classes it subsumes (chairs, tables, and so on.)

5.3.1.2 The Whole-Part Relationship

The *whole-part relationship* covers situations where one concept is inherently included in another, regardless of context, so that concepts can be organized into hierarchies (with the “whole” treated as a broader term). For example, blood vessels are part of the cardiovascular system in anatomy.

In addition to physical component part relationships, “whole and part” can be applied to several common types of situations such as geographical regions, hierarchical organizational structures, disciplines or fields of discourse. Because such relationships, being synthetic rather than analytic, are not necessarily or logically true in subject authority systems they may be differentiated as special hierarchical relationships (rather than genus-species and perspective hierarchies) or as associative relationships.

²⁰ ISO (2009). *ISO/CD 25964-1, Information and documentation — Thesauri and interoperability with other vocabularies — Part 1: Thesauri for information retrieval*. ISO/TC 46 / SC 9 ISO 25964 Working Group.

²¹ Clarke, S.G. (2001). Thesaural relationships. In: *Relationships in Knowledge Organization*. Eds. Bean, C.A. and Green, R. Dordrecht: Kluwer. p. 42.

²² Svenonius, E. (2000). *The Intellectual Foundation of Information Organization*. Cambridge, Mass.: MIT Press, p.151.

²³ Svenonius, *op. cit.*

5.3.1.3 The Instance Relationship

The *instance relationship* identifies the link between a general class of things or events, expressed by a common noun, and an individual instance of that category, expressed by a proper noun. For example, “Mydoom” and “ILOVEYOU” are two instances of “computer worms” that are expressed by proper nouns.

5.3.1.4 Polyhierarchical Relationship

Some concepts can belong to more than one superordinate concept and are therefore considered to possess *polyhierarchical relationships*. These relationships can be (a) generic, e.g., music instrument “organ” belongs under both the “wind instrument” hierarchy and the “keyboard instrument” hierarchy; (b) whole-part, e.g., “biochemistry”, is part of “biology” and is also part of “chemistry”; or (c) more than one type, e.g., “skull”, belongs under the “bones” (kind-of), and also under the “head” (part-of), hierarchies²⁴.

5.3.1.5 Other Hierarchical Relationship

Other hierarchical relationships, which do not have the logical properties of the above hierarchies, are seen often in subject authority systems. This may be partially due to the requirements of literary warrant (the natural language used to describe content objects), user warrant (the language of users), and sometimes, organizational warrant (the needs and priorities of the organization)²⁵. Their value is that they provide points of view about a concept and the aspect under which it is considered. For instance, although an insect can belong to only one genus-species hierarchy (e.g., Arthropoda), it can belong to as many perspective hierarchies as there are aspects of insects to be studied. In a classification scheme, an insect can be looked at, or studied, from the point of view of agricultural pests, disease carriers, food, and control technology²⁶. Other reasons to employ such hierarchies are that concepts and terms like “happiness” are poly-semantic, vague, or ambiguous. Hence there might be no agreement as to what genus (class) such concepts belong to.

5.3.2 Associative Relationships

Associative relationships cover affiliations between pairs of *themas* that are not related hierarchically yet are semantically or conceptually connected and co-occurring. Associative relationships between *themas* are made explicit in some of the subject authority systems.

²⁴ Examples are from ISO (2009). *ISO/CD 25964-1, Information and documentation — Thesauri and interoperability with other vocabularies — Part 1: Thesauri for information retrieval*. ISO/TC 46 / SC 9 ISO 25964 Working Group.

²⁵ NISO. (2005). *ANSI/NISO Z39.19-2005. Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*. Bethesda, Maryland: NISO Press.

²⁶ Svenonius, *op. cit.*

In general, associative relation links are established among *themas* belonging to different hierarchies, or among overlapping *themas* within the same array on a particular level of the hierarchy. Most commonly considered associative relationships fall into these categories^{27, 28, 29}:

Associative Relationships	Examples
Cause / Effect	accident / injury
Process / Agent	velocity measurement / speedometer
Action / Product of the action	weaving/cloth
Action / Patient or Target	teaching / student
Concept or Thing / Properties	steel alloy / corrosion resistance
Thing or Action / Counter-agent	pest / pesticide
Thing/ Its parts (if it does not qualify for the hierarchical whole-part relationship)	car / engine
Raw material / Product	grapes / wine
Action / Property	communication/communication skills
Field of study/ Objects or phenomena studied	forestry/forests

In each particular implementations, a decision would be made about whether to include associative relationships and if so, which ones to include and at what level of specificity.

5.3.3 Other Approaches to Semantic Relationships

In the literature and in practice, other approaches to differentiate semantic relation types have been used. A taxonomy of subject relationships, compiled in 1996 and shared at an American Libraries Association (ALA) conference, listed over a hundred associative relationships and 26 hierarchical relationships³⁰. Over 40 in the associative group and 20 in the hierarchical group have been verified by other sources³¹.

²⁷ Lancaster, F. W. (1986). *Vocabulary Control for Information Retrieval*. 2nd ed. Arlington, Virginia: Information Resources Press.

²⁸ NISO. *op.cit.*

²⁹ Aitchison, J., A. Gilchrist, and D. Bawden (2000). *Thesaurus Construction and Use: A Practical Manual*. 4th ed. London: Fitzroy Dearborn.

³⁰ Michel, D. (1996). Taxonomy of Subject Relationships. Appendix B (Part 2), Final Report to the ALCTS/CCS Subject Analysis Committee, submitted by Subcommittee on Subject Relationships/Reference Structures. ALA Association for Library Collections and Technical Services (ALCTS) Cataloging and Classification Section (CCS) Subject Analysis Committee (SAC). Available at: <http://www.ala.org/ala/mgrps/divs/alcts/mgrps/ccs/cmtes/subjectanalysis/subjectrelations/msrscu2.pdf> Also available is the hierarchical display of this Taxonomy, at: <http://www.ala.org/ala/mgrps/divs/alcts/mgrps/ccs/cmtes/sac/inact/subjectrelations/appendixbpartii.cfm> (accessed 2010-03-31).

³¹ Subcommittee on Subject Relationships/Reference Structures. (1997). Final Report to the ALCTS/CCS Subject Analysis Committee. ALA Association for Library Collections and Technical Services (ALCTS) Cataloging and Classification Section (CCS) Subject Analysis Committee (SAC). Available at: <http://www.ala.org/ala/mgrps/divs/alcts/mgrps/ccs/cmtes/sac/inact/subjectrelations/finalreport.cfm>

The *Unified Medical Language System* (UMLS)³² classified semantic relationship types into two main groups and a number of sub-groups:

- isa
- associated_with
 - physically_related_to
 - spatially_related_to
 - functionally_related_to
 - temporally_related_to
 - conceptually_related_to

Spatial relationship types in UMLS include location_of, adjacent_to, surrounds, and traverses.

Whereas in another case, such relationship types for geographical regions only are identified as³³:

- Inherently spatial
 - Containment
 - Overlap
 - Proximity
 - Directional
- Explicitly stated
 - PartOf
 - AdministrativePartOf
 - AdministrativePartitionMemberOf
 - AdministrativeSeatOf
 - ConventionallyQualifiedBy
 - SubfeatureOf
 - GeophysicalPartitionMemberOf
- PhysicallyConnectedTo
 - FlowsInto

These examples illustrate implementation-dependent relationship typing.

³² National Library of Medicine. (2004) Unified Medical Language System. Current relations in the semantic network. In: NLM. *Unified Medical Language System-Semantic Network Documentation, Section 3. Semantic Networks*. Available at: http://www.nlm.nih.gov/research/umls/META3_current_relations.html (accessed 2010-01-20).

³³ Hill, L. (1999). Content standards for digital gazetteers. Presentation at the *JCDL2002 NKOS Workshop "Digital gazetteers--Integration into distributed digital library services"*, July 18, 2002, Portland, Oregon. Available at: <http://nkos.slis.kent.edu/DL02workshop.htm> (accessed 2010-01-20).

5.4 *NOMEN-to-NOMEN* Relationships

Only the equivalence and whole-part relationships are discussed below. Other *nomen-to-nomen* relationships may also be established.

5.4.1 Equivalence Relationship

Equivalence of *nomen* is a very important notion in subject access. Two *nomens* are equivalent if they are appellations of the same *thema*. The equivalence relationships in a monolingual controlled vocabulary can be found in five general situations.^{34,35}

- a) the *nomens* are synonyms
- b) the *nomens* are near or quasi-synonyms
- c) the *nomens* have lexical variants
- d) a *nomen* is regarded as unnecessarily specific and it is represented by another *nomen* with broader scope
- e) a *nomen* is regarded as unnecessarily specific and it is represented by a combination of two or more terms (known as “compound equivalence”).

It is obvious that equivalence relationships do not assume exact equivalence. Inexact and partial equivalence are often found in controlled vocabularies. In reality, the *nomens* in the above d) and e) situations represent different *themas*. But since in some controlled vocabularies these *nomens* are connected as preferred and alternative terms, it may be interpreted that these *nomens* are assumed to represent the same *thema* in a controlled vocabulary.

In addition, equivalence relationships exist between *nomens* in different languages and across schemes. For example, “iron” (a term in English), “železo” (a term in Slovenian), and “Fe” (a chemical symbol) are all *nomens* for the same metal and are therefore considered equivalent.

The equivalence relationships of *nomens* can be specified further. For example:

- Replaces/Is replaced by
[e.g., “integrated plant control” is replaced by “centralized control”]
- Has variant form/Is variant form
 - Has acronym/is acronym for
[e.g., “VS” is acronym for “virtual storage”]
 - Has abbreviation/Is abbreviation of
 - Has transliterated form/Is transliteration of

5.4.2 The Whole-Part Relationship

The *whole-part relationship* also exists between *nomens*. A *nomen* may have components (parts). These components may or may not be a *nomen* on their own. The composition of

³⁴ ISO. (2009). ISO/CD 25964-1. *op. cit.*

³⁵ NISO. (2005). Z39.19-2005. *op. cit.*

such *nomen* may be governed by rules, for example, the citation order in faceted classification schemes or the order of subdivisions in a subject heading.

6. USER TASKS

6.1 Users and Use

During the early stages of developing the entity-relationship conceptual model of subject authority records, the FRSAR Working Group considered it essential to analyze the users of subject authority data, to identify the contexts in which the data is used, and to characterize different usage scenarios.

Potential user groups include:

- a) information professionals who create and maintain subject authority data, including cataloguers and controlled vocabulary creators;
- b) information professionals who create and maintain metadata;
- c) reference services librarians and other information professionals who search for information as intermediaries; and
- d) end-users who search for information to fulfil their information needs.

Intermediaries (group c) act on behalf of end-users (group d). They interact with the bibliographic data in a similar way as end-users (although at a higher level of expertise). Therefore, for the purpose of this use analysis the intermediaries and end-users are considered belonging to the same end-user group.

Figure 6.1 illustrates the three-point perspective of subject authority data users.

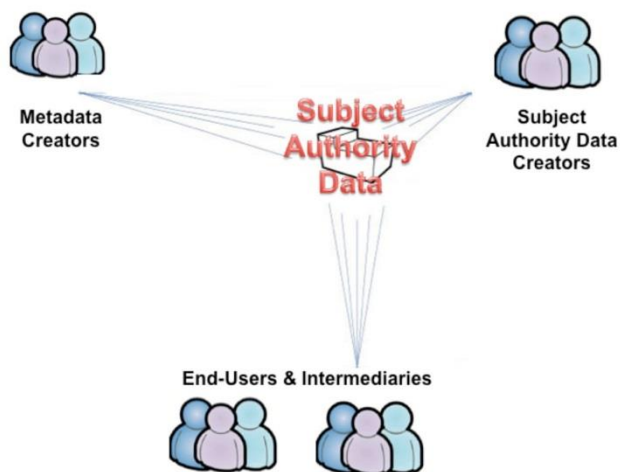


Figure 6.1 Users of subject authority data

6.2 User Tasks

When using subject authority data, a user may need to *find*, *identify*, and *select* a subject entity or entities. A user may also choose to *explore* a subject domain and its terminology as well as the relationships that exist among the *themas*. In addition, the user may explore the correlation of the *nomen(s)* of a *thema* in one subject authority system to the respective *nomen(s)* of the same *thema* in another subject authority system.

Based on the results from the two user studies conducted by the User Tasks Sub-Group, four tasks for subject authority data have been defined:

- **Find** one or more subjects and/or their appellations, that correspond(s) to the user's stated criteria, using attributes and relationships;
- **Identify** a subject and/or its appellation based on their attributes or relationships (i.e., to distinguish between two or more subjects or appellations with similar characteristics and to confirm that the appropriate subject or appellation has been found);
- **Select** a subject and/or its appellation appropriate to the user's needs (i.e., to choose or reject based on the user's requirements and needs);
- **Explore** relationships between subjects and/or their appellations (e.g., to explore relationships in order to understand the structure of a subject domain and its terminology).

The *explore* user task is a new task introduced in FRSAD; whereas the *find*, *identify*, and *select* user tasks have been previously introduced in the FRBR and/or FRAD conceptual models. The subject authority data use survey conducted by FRSAR's User Tasks Sub-Group indicates that a large number of participants (69%) use subject authority data to explore relationships among terms during cataloguing and metadata creation. In addition, 62% of participants use subject authority data to explore relationships while searching for bibliographic resources, and 64% use these data to navigate and browse bibliographic descriptions. These numbers reflect a major use of subject authority data for a task that was not present in the FRAD and FRBR models; therefore, the group deemed it important to add the *explore* user task.

Figure 6.2 shows a comparison of user tasks as defined in FRBR, FRAD, and FRSAD.

User Tasks

FRBR	FRAD	FRSAD
Find Identify Select Obtain	Find Identify Contextualize Justify	Find Identify Select Explore

Figure 6.2 Comparison of user tasks as defined in FRBR, FRAD, and FRSAD

6.3 Assessing Values Relative to User Tasks

While in some cases the users' information needs are limited to authority data only, in most cases the users will utilise subject authority data to *find*, *identify*, *select*, and/or *obtain* Group 1 entities as specified by the FRBR user tasks.

The following is a list of tasks that result from placing the primary subject authority data user tasks (*find*, *identify*, *select*, and *explore*) in the context of different user groups as it relates to interacting only with subject authority data. The activities of using subject authority data to access bibliographic data are covered by FRBR. In the following examples, text within double quotation marks represents a nomen; a thema is referred by English text within curly brackets. Single quotation marks are used for thema types.

FIND: using the data to find one or more subjects and/or their appellations, that correspond(s) to the user's stated criteria, using attributes and relationships

- 1) Using subject authority data to **find** a *thema* or a set of *themas* based on the user's search criteria. For example:
 - A user is looking for a *thema* of the type 'substance' within the medical field.
 - A user is looking for a *thema* that is a 'medical condition(s)' for which {chloromadinone acetate} is used for treatment.
 - A user is looking for *themas* that are particular artistic styles (for example, {modernism}) using *thema-to-thema* hierarchical relationships.
- 2) Using subject authority data to **find** a *nomen* or a set of *nomens* for a *thema*. For example:
 - A user is looking for the *Dewey Decimal Classification* number (*nomen*) for the *thema* {dragonflies} (as it is referred to in English).

- A user is looking for the preferred *nomen* in the *Library of Congress Subject Headings* for the *thema* {lilac flower} (as commonly referred to in English).

IDENTIFY: using the data to identify a subject and/or its appellation based on their attributes or relationships

- 3) Using subject authority data to **identify** a *thema*, i.e., to confirm that the *thema* found is the one sought by the user, or to distinguish between two similar *themas*.
For example:
 - A user is employing subject authority data to identify whether the *thema* {clothing} or the *thema* {costume} is more appropriate for a specific information need.
- 4) Using subject authority data to **identify** a *nomen*, i.e., to confirm that the *nomen* found is the one sought by the user, or to distinguish between two similar *nomens*.
For example:
 - A user employs subject authority data to verify whether the *nomen* “craftsman style” is the appropriate *nomen* in a particular system.

SELECT: using the data to select a subject and/or its appellation appropriate to the user’s needs

- 5) Using subject authority data to **select** a *thema* from the set of *themas* found.
For example:
 - Select a *thema* at the appropriate level of specificity from a hierarchy of related *themas*: A user is using subject authority data to select the *thema* {volley ball} as a more appropriate subject access point in a bibliographic record rather than the broader *thema* {ball games}.
- 6) Using subject authority data to **select** a *nomen* from the set of *nomens* found.
For example:
 - Select the preferred *nomen* for a *thema* within a subject authority system to use in searching or in assigning access points: A user is using subject authority data to select “ale glasses” among the *nomens* “ale glasses,” “glass, beer,” and “malt-beverage glass” found in the *Art and Architecture Thesaurus*.

EXPLORE: using the data in order to explore relationships between subjects and/or their appellations

- 7) Using subject authority data to **explore** the relationships between two or more *themas* within the same subject authority system.
For example:
 - A user is using subject authority data to explore associative relationships of the *thema* {digital libraries} and other *themas*.
- 8) Using subject authority data to **explore** the relationships between two or more *nomens* within the same subject authority system.
For example:
 - A user is using subject authority data to explore the relationship of the *nomen* “ladybugs” and the *nomen* “ladybirds” in LCSH.

- 9) Using subject authority data to **explore** the correlation of *themas* between two or more subject authority systems.

For example:

- A user is using subject authority data to explore the correlation of the *thema* {domestic cats} between the LCSH and the *Sears List of Subject Headings*.

- 10) Using subject authority data to **explore** the correlation of *nomens* between two or more subject authority systems.

For example:

- A user is using subject authority data to explore whether the *Medical Subject Headings* (MeSH) *nomen* “cataract” correlates to the *National Library of Medicine (NLM) Classification’s nomen* “WW 260.”

- 11) Using subject authority data to **explore** the structure of a subject domain within a subject authority system.

For example:

- A user is using subject authority data to explore how the domain {computer science} is represented within the *American Society for Information Science and Technology (ASIS&T) Thesaurus*.

6.4 Mapping of Attributes, Relationships, and User Tasks

Tables 6.1 and 6.2 map the attributes and relationships defined in chapters 4 and 5 to the defined set of user tasks. The mapping is intended to clarify which attributes and relationships are required to support each particular user task. Only attributes and relationships specified in the FRSAD model are mapped. The decision as to which attributes and relationships to include or indicate as mandatory and, to some extent, the determination of importance, are application- or implementation-specific. The degree of importance included in Table 6.1 followed the FRBR report and is based on an analysis of common library and other information agencies’ practice.

THEMA

Attributes	Find	Identify	Select	Explore
Type of thema	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Scope note		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Thema-to-Thema Relationships				
Hierarchical relationship	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Associative relationship	<input type="checkbox"/>		<input type="checkbox"/>	<input checked="" type="checkbox"/>

■ = strong importance □ = moderate importance Blank = not important

Table 6.1 Mapping for *Thema*

NOMEN

Attributes	Find	Identify	Select	Explore
Type of nomen	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Scheme	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Reference Source of nomen		<input type="checkbox"/>		
Representation of nomen	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Language of nomen	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Script of nomen	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Script conversion	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
Form of nomen	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Time of validity of nomen	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Audience	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Status of nomen	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
Nomen-to-Nomen Relationships				
Equivalence relationship	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Whole-part relationship	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

= strong importance
 = moderate importance
Blank = not important

Table 6.2 Mapping for *Nomen***7. CONCLUSION**

In this report, the FRSAR Working Group has presented a conceptual model within the FRBR framework as it relates to the aboutness of *works*. The report defines a structured frame of reference for relating the data that are recorded in subject authority records to the needs of the users of these data. The FRSAD model is developed with the goal of assisting in an assessment of the potential for international sharing and use of subject authority data both within the library sector and beyond. It enhances consideration for the functional requirements for subject authority data at a level that is independent of any implementation, system, or specific context.

APPENDIX A. MODELING ABOUTNESS

A.1 Subject Relationship and Group 3 Entities Introduced in FRBR

The subject relationship introduced in the FRBR model is illustrated in Figure A.1:

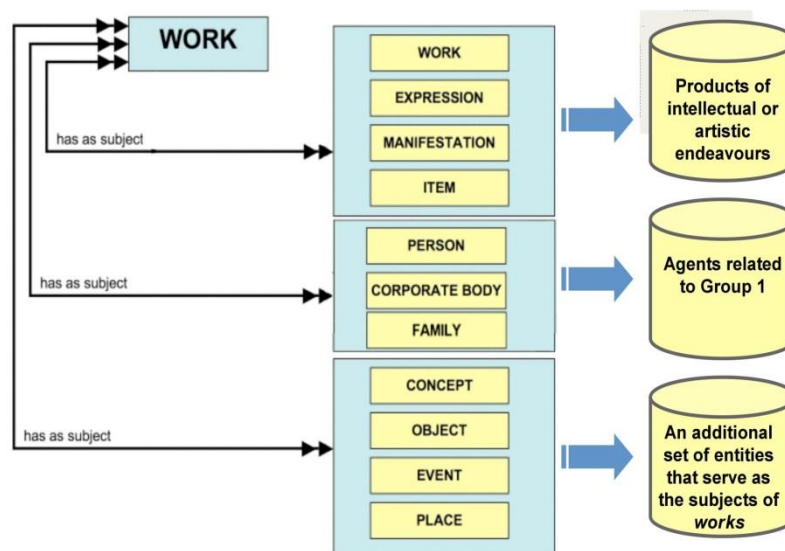


Figure A.1 Extension of FRBR Figure 3.3 "Group 3 entities and 'subject' relationships"

The diagram in FRBR Figure 3.3 depicts the "subject" relationships between *works* and entities in Group 1, Group 2, and Group 3. These three groups are represented as the components on the right side of the above figure. The left and centre components in the figure are based on the FRBR Figure, with the *Family* entity added in Group 2 according to the FRAD model.

The entities in Group 3 represent an additional set of entities that serve as the subjects of *works*. The FRBR report specified Group 3 entities under Figure 3.3 as³⁶:

3.1.3 Group 3 Entities: Concept, Object, Event, Place

³⁶ *Functional Requirements for Bibliographic Records: Final Report*. (1998). IFLA Study Group on the Functional Requirements for Bibliographic Records. München: KG Saur, p. 17.

The entities in the third group (outlined in bold in Figure 3.3) represent an additional set of entities that serve as the subjects of *works*. The group includes *concept* (an abstract notion or idea), *object* (a material thing), *event* (an action or occurrence), and *place* (a location).

The diagram depicts the “subject” relationships between entities in the third group and the *work* entity in the first group. The diagram indicates that a *work* may have as its subject one or more than one *concept*, *object*, *event*, and/or *place*. Conversely, a *concept*, *object*, *event*, and/or *place* may be the subject of one or more than one *work*.

The diagram also depicts the “subject” relationships between *work* and the entities in the first and second groups. The diagram indicates that a work may have as its subject one or more than one *work*, *expression*, *manifestation*, *item*, *person*, and/or *corporate body*.

A.2 Possible Approaches to the Model of Aboutness

The FRBR Working Group has, as the central part of its terms of reference, the goal of building a conceptual model of Group 3 entities within the FRBR framework as they relate to the **aboutness** of *works*.

It is mentioned in the FRBR study that “further analysis is needed of the entities that are the centre of focus for subject authorities, thesauri, and classification schemes, and of the relationships between those entities”³⁷. In the years following the publication of the FRBR model, some researchers focused on Group 3 entities, particularly on the fact that *time* is not included.³⁸ Consequently, time and space are not treated symmetrically. Some discussions brought attention to the lack of coverage of activities and processes.

Tom Delsey, in his paper published in *Cataloging & Classification Quarterly* in 2005, highlighted the aspects of the FRBR model that “will need to be re-examined as part of a more intensive analysis of subject access”³⁹. Delsey followed up with a presentation of a paper at the IFLA satellite meeting in Järvenpää, Finland, before the IFLA General Conference in Oslo in August 2005. His presentation has provoked much discussion among the members of the FRBR Working Group.

Delsey identified three “broad objectives” to be met by re-examination of the ways in which the FRBR model analyzes data relevant to subject access:

- 1) “to ensure that the scope of the *entities* defined in the [FRBR and FRAD] models is sufficient to cover everything that a user of a library catalogue might view as a ‘subject’”;

³⁷*Ibid.*, p. 7.

³⁸ Heaney, M. (1997). Time is of the essence. Available at: <http://www.bodley.ox.ac.uk/users/mh/time978a.htm> (accessed 2010-01-20).

³⁹ Delsey, T. (2005). Modeling subject access: Extending the FRBR and FRANAR conceptual models. *Cataloging & Classification Quarterly* 39 (3/4): 49–61.

- 2) “to ensure that the *attributes* that come into play in the construction and use of subject access points and subject authority records are adequately covered”; and
- 3) “to ensure that the models provide a clear and pragmatic representation of the *relationships* that are reflected through subject access points in bibliographic records as well as those reflected in the syndetic structure of thesauri, subject heading lists, and classification schemes and in the syntactic structure of indexing strings⁴⁰ (emphases added)”.

Delsey identified two “key questions” related to entities: “The first [of the key questions] is whether the entities are defined in sufficiently broad terms to cover fully what we might characterize as the “subject” universe. The second is whether the categorizations represented by the entities defined in the models are appropriate and meaningful for the purposes of clarifying the bibliographic conventions through which that “subject” universe is reflected”⁴¹. In other words:

1. Are the entity classes *collectively exhaustive*? Does the model cover the whole universe of subject-related entity classes?
2. Are the entity classes *individually appropriate*? Does the model carve up the universe of subject-related entity classes in the “right” way⁴²?

As a first step, the FRSAR Entities Sub-Group performed a pilot study, in which four students and faculty members at the Kent State University School of Library and Information Science classified existing subject terms used by the NSDL (National Science Digital Library) contributors. These included about 3000 terms assigned based on a variety of subject vocabularies and free keywords. They classified terms into six categories: concrete stuff, abstract stuff, event, time, place, and others. The same method was also applied by one of the Working Group members to another set of subject terms from controlled vocabularies used in two library science textbooks. The results show that there is a blurred distinction between concrete and abstract concepts; for example, the distinction between a particular chair as a physical object and the concept of chairs. In addition, there were difficulties in classifying named instances (proper names), which resulted in many terms being put into the “others” category. The results of this test indicate that it would be difficult for any user (end user, librarian, or vocabulary developer) to conduct such a task when using subject authority data. These categories do not seem helpful or necessary to the end users either.

Following the pilot study, the Working Group discussed several possible previously identified approaches to the development of a theoretical framework of aboutness.

⁴⁰ Delsey, *op. cit.*, p. 50.

⁴¹ Delsey, *op. cit.*, p. 50.

⁴² See also: Furner, J. (2006). The ontology of subjects of works. Paper presented at ASIS&T 2006: Annual Meeting of the American Society for Information Science and Technology, Austin, TX, November 3–8, 2006.

Scenario 1

Keep FRBR Group 3 entities (*concept*, *object*, *event*, and *place*) and only analyze attributes and relationships. The advantage is that the Working Group uses an existing framework. However, as demonstrated in the pilot study of the FR SAR Entities Sub-Group, the Group 3 entities need to be revisited. Adding *time* to the FRBR list solves one part of the problem, but the resulting model still does not cover processes, activities, or situations.

An additional argument for rejecting this scenario is that the original categorisation of Group 3 entities into four classes goes too far towards prescribing a particular way of structuring the subject languages that are used to provide access to *works*. Any subject authority system that lacks a faceted structure to distinguish clearly between *concepts*, *objects*, *events*, and *places* can be modelled only with difficulty. Rather than taking a stand on exactly which aspects to identify for the entire information community, the Working Group felt it was important to provide a higher level, more theoretical approach and not to impose any constraint on the forms that subject authority systems take in particular implementations. This modelling does not limit any community from implementing the original FRBR Group 3 entities; on the contrary, it allows for more flexibility.

Scenario 2

Take Ranganathan's facets as the basis of the new framework. The facets would become entities:

- Personality
- Matter
- Energy
- Space
- Time

The advantage is that this approach is well known in the library community, has been justified theoretically, and covers all areas of aboutness quite well. The issues are whether we would still have problems defining some of the entities, and whether librarians and end users would have trouble understanding and applying them.

Scenario 3

Take the <indecs>⁴³ model as the basis of the new framework. The main focus of the <indecs> model is intellectual property and rights management, but it also overlaps significantly with FRBR. The basic <indecs> entities are defined as:

- **Percept**: an entity that is perceived directly with at least one of the five senses.

⁴³ Rust, G. and Bide, M. (2000). The <indecs> metadata framework: Principles, Model and Data Dictionary. Version2. Indecs Framework Ltd. Available at: http://www.doi.org/topics/indecs/indecs_framework_2000.pdf (accessed 2010-01-20).

- **Being:** an entity that has characteristics of animate life; anything which lives and dies
- **Thing:** an entity without the characteristics of animate life
- **Concept:** an entity that cannot be perceived directly through the mode of one of the five senses; an abstract entity, a notion or idea; an abstract noun; an unobservable proposition, which exists independently of time and space
- **Relation:** the interaction of percepts and/or concepts; a connection between two or more entities
 - **Event:** a dynamic relation involving two or more entities; something that happens; a relation through which an attribute of an entity is changed, added or removed
 - **Situation:** a static relation involving two or more entities; something that continues to be the case; a relation in which the attributes of entities remain unchanged

Being and **Thing** together correspond to a supertype of the FRBR entity *object*; **Concept** roughly corresponds to the FRBR entity *concept*; and **Event** corresponds to the FRBR entity *event*. Thus, the three main differences between the <indec> model and the FRBR model are (a) the subtyping of **Percept** in the <indec> model into **Being** and **Thing**, and in the FRBR model into *item*, *person*, and *object*, (b) the absence of an FRBR entity that directly corresponds to the <indec> entity **Situation**, and (c) the absence of an <indec> entity that directly corresponds to the FRBR entity *place*.

As Delsey⁴⁴ notes, these differences raise corresponding questions about the possibility of making changes to the set of Group 3 entities defined in the original FRBR model: (a) Should the original entity *Object* be subtyped into two entities—e.g., **Inanimate object**, and **Animate object**? (b) Should **Situation** be added as an entity? (c) Should the FRBR entity *place* be removed? Note that in the FRBR report *places* are treated as entities only to the extent that they are the subject of a *work*.

Scenario 4

Make a pragmatic list of entities. Buizza and Guerrini created one example of such a list⁴⁵ for the Italian project Nuovo soggettario. Two logical entities, the *subject* (the topic, the basic theme of the work, the summarisation of its main contents) and *concept* (a unit of thought, each of the single elements which make up the subject), were defined. The list shows, as an example, what can be a concept in a specific implementation and draws on categories, roles and relationships from the report of the project:

- Object (material thing)
- Abstraction
- Living organism
- Person
- Corporate body

⁴⁴ Delsey, op. cit., p. 51 - 52

⁴⁵ Buizza, P. and Guerrini, M. 2002. A conceptual model for the New Soggettario: Subject indexing in the light of FRBR. *Cataloging & Classification Quarterly*, 34(4): 31–45.

- Work
- Matter/material
- Property/quality
- Action
- Process
- Event
- Place
- Time

The problem with such lists is that the entities are not mutually exclusive, have overlaps, and rely on individual common everyday definitions of the entities. The authors' original purpose was to show the compatibility of those categories with the model. It is also a warning that making a pragmatic list of entities would be a disadvantage for a theoretical model.

Scenario 5

Do not make any recommendation on categorisation of subjects. This approach is a more abstract view and does not pose restrictions on any implementations. It also allows a more abstract, general view.

This last scenario (5) was the decision taken by the Working Group, based on comparative analysis of all scenarios and the pilot user study. None of scenarios 1-5 are ideal for all situations, while each may be a good solution for particular implementations. Any further categorization of Group 3 entities would prescribe a particular way of structuring the subject authority systems that are used to provide access to works. A good model should allow for any multiple domain-specific structures and should be flexible enough to accommodate different implementations. This can be achieved only by a more abstract theoretical model, completely independent of any implementation that enables the treatment of attributes and relationships on a more general level.

APPENDIX B. RELATIONSHIP OF FRSAD WITH FRBR AND FRAD

Two models, FRAD (by FRANAR Working Group) and FRSAD (by FRSAR Working Group), complement and further develop some aspects of the original model of FRBR. The three models together have been labelled the 'FRBR family', suggesting that they are all considered parts of a larger general model. There exist some differences among them, though; the respective FRBR, FRANAR and FRSAR working groups have made different modelling decisions during their independent model development. Eventually FRBR, FRAD and FRSAD will have to be harmonised and a resultant new consolidated model will be developed. In order to facilitate this process the FRSAR Working Group takes this opportunity to list the most important differences below.

B.1 Relationship of FRSAD with FRBR

The FRSAR Working Group follows FRBR in the methodology, specification, and presentation of entities and relationships. The "has as subject" (many-to-many) relationship, established between the *work* and the entity(ies) representing the aboutness of the *work*, is kept in its entirety in FRSAD. As in FRBR, the FRSAD model also starts with a user tasks analysis and follows with the establishment of appropriate entities and relationships. The four areas where some differences were introduced in FRSAD are:

- The addition of the "Explore" task;
- *Thema* is introduced as a superclass of all entities that can be subjects of a *work*. Attributes and relationships of *thema* are presented;
- No entities are explicitly predefined in Group 3;
- *Nomen* is introduced (including attributes and relationships) and is defined as a separate entity instead of an attribute.

The inclusion of the "explore" task is based on the findings of the user study conducted by the FRSAR Working Group. Users of subject authority data also use these data to explore a domain, to get acquainted with the terminology, and to identify semantic relationships. The FRSAR Working Group is confident that the same is true for bibliographic information in general, and recommends that the 'explore' task be added to the general model.

In Fig. 3.3 of the FRBR report, the depicted "subject" relationship has three boxes representing all three groups of entities respectively, on the right side of the 'has as subject' relationship. FRSAD has developed this further by creating a superclass (*thema*), thus enabling the modelling of the "has as subject" relationship on a more general level. *Thema* includes Group 1, Group 2 and all other entities that can be the subjects of a *work*. Therefore, the subject relationship can easily be modelled as "*work* has as subject *thema*."

FRBR defines four entities in Group 3: *concept*, *object*, *event*, and *place*. The FRSAR Working Group, based on the pilot user study, literature review, and independent analysis, decided to avoid any predefined subclasses. There seems to be no universal categorisation of *themas* and any attempt to declare one would necessarily limit the usability of a general model. Each particular implementation will need to define the categories or types of *themas*. The original FRBR Group 3 entities are, therefore, only one possible scenario. (Please refer to Section 4.1.1.)

FRSAD introduces a differentiation between a thing itself and its appellation. The appellation (name, label, etc.) is often modelled as an attribute of the entity it refers to (also in FRBR). While simpler, this approach makes it impossible to introduce the attributes (e.g. language) and relationships (e.g. the relationship between a former and current name) of the appellation itself, because in an E-R model one may not have attributes of an attribute. *Nomen* is therefore introduced in FRSAD as an entity, rather than an attribute, to enable appropriate modelling.

B.2 Relationship of FRSAD with FRAD

The FRANAR Working Group was established in 1999 with the mandate of developing FRBR further in the area of authority files. Later the decision was made by FRANAR to focus on Group 2 entities and *work* only. As a consequence, the FRSAR Working Group was established to cover the ‘has as subject’ relationship and the appropriate entities. The FRAD and FRSAD models were therefore developed independently. The working groups, although both following FRBR and its modelling approach, have made several different decisions. The most significant ones are:

- User tasks: “Contextualise” and “Justify” in FRAD vs. “Explore” in FRSAD;
- *Name* in FRAD vs. *Nomen* in FRSAD;
- *Name*, *Identifier* and *Controlled access point* as separate entities in FRAD vs. values of the attribute “Type of *Nomen*” in FRSAD;
- *Rules* and *Agency* as new entities in FRAD and not explicitly modelled in FRSAD.

B.2.1 User tasks

The Working Group believes that “explore” is a generalisation of “contextualise” and expresses better the user task of browsing, getting acquainted, becoming familiar with, and discovering.

“Justify,” on the other hand, is a task of information professionals and not end-users. It is an important task on its own, but falls within metadata creation and not metadata use. Since FRSAD follows the FRBR approach that has not extended its model to cover such tasks, it is not included in the FRSAD model.

B.2.2 Name and Nomen

Although similar at first glance, the two entities are different: FRAD *name* is defined as “a character or group of words and/or characters by which an entity is known in the real

world". The FRSAD *nomen* is a more general entity, comprising any (textual or other) appellation both in the real world and in artificial systems. In relation to FRAD, *nomen* is a superclass of FRAD *name*, *identifier* and *controlled access point*.

B.2.3 Separate appellation entities in FRAD vs. *nomen* entity and "type" attribute in FRSAD

Nomen is the general appellation entity in FRSAD with specific types. This allows the introduction of any type that will be necessary for an implementation. In addition, some possible general values of the attribute "type" are already suggested, such as "identifier". FRAD "name," "identifier" and "controlled access point" are therefore possible types of *nomens*. This approach allows flexibility; even particular kinds of identifiers (URL, ISBN, etc.) can be defined as values of "type" of a *nomen*.

B.2.4 Rules and Agency

Rules and *agency* are not specifically modelled in FRSAD. The position of the Working Group is that the focus of the model is not on the cataloguing process and it is not necessary to include that level of detail. If needed, *rules* (which are applied in all phases of cataloguing, not only in creation of controlled access points) should be considered instances of *work*. *Agencies*, which apply the rules, should be considered instances of *corporate body*. If modelled, they are in a relationship with the attribute assignment event.

APPENDIX C. FRISAD MODEL AND OTHER MODELS

C.1 The Importance of the THEMA-NOMEN Model

As early as 1923, Ogden and Richards⁴⁶ published a famous triangle of meaning that illustrated the relationship between language, thought content, and referent. The graph (Figure D.1) implies that the referent of an expression (a word or another sign or symbol) is relative to different language users. The theoretical foundation of it can be traced back to Aristotle, who distinguished objects, and the words that refer to them, and the corresponding experiences in the *psyche*. Equally, Frege distinguished between two types of meaning: thought content and referent, in his essay *Über Sinn und Bedeutung*. It is not enough to try to understand what a thing is, based on its name, because it may have been named in ancient times, and the name reflects only what the name-givers thought was the nature of reality then. Therefore multiple terms may refer to the same object or idea, a single term may refer ambiguously to more than one object or idea, and outdated terms may be confusing⁴⁷.

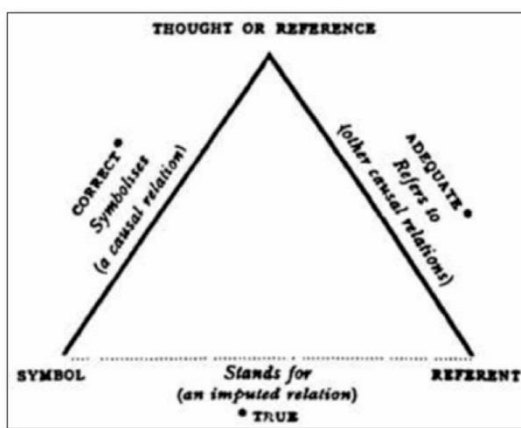


Figure D.1 Ogden's Semiotic Triangle. (Ogden and Richards, 1923,⁴⁸ p.11)

Ogden's model was also adopted by researchers in library and information science as the basis for building subject authority systems^{49,50}.

⁴⁶ Ogden, C. K., and Richards, I. A. (1923). *The Meaning of Meaning: A Study of the Influence of Language Upon Thought and of the Science of Symbolism*. London: Routledge & Kegan Paul.

⁴⁷ Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, NF 100. 1892, S. 25–50. Available at: <http://www.gavagai.de/HHP31.htm> (accessed 2009-05-22).

⁴⁸ Ogden and Richards. *op. cit.*

The importance of the *thema-nomen* model for subject authority data is to separate *subjects* from what they are known as, referred to, or addressed as. Among the efforts to achieve global sharing and use of subject authority data, some efforts have focused on *nomen*, e.g., translated metadata vocabulary, a symmetrical multilingual thesaurus, or a multi-access index to a vocabulary. However, most efforts have focused on the conceptual level, e.g., mappings between two thesauri or between a classification scheme and a thesaurus. Such efforts usually encounter much greater challenges because they are concerned with the subject mappings in terms of their meaning as well as the relationships among the subjects.

C.2 Mapping the FRSAD Model to Other Models

This *thema-nomen* conceptual model matches well with encoding schemas such as SKOS Simple Knowledge Organization System (SKOS) and OWL Web Ontology Language (OWL), which provide models for expressing the basic structure and content of knowledge organization systems (KOS) such as thesauri, classification schemes, subject heading lists, taxonomies and other similar types of controlled vocabularies, as well as ontologies. SKOS defines classes and properties sufficiently for representing the common features found in a standard thesaurus and other KOS structures. The SKOS model is based on a concept-centric view of vocabulary, where primitive objects are not labels; rather, they are concepts represented by labels. As an application of the RDF (Resource Description Framework), SKOS allows concepts to be composed and published on the World Wide Web, linked with data on the Web and integrated into other concept schemes. Each SKOS concept is defined as an *RDF resource* and each concept can have *RDF properties* attached. These include: one or more preferred terms (at most one in each natural language); alternative terms or synonyms; and, definitions and notes with specification of their language⁵¹. Each of these can be matched to what have been defined in the FRSAD model in terms of *thema*, *nomen*, and their attributes. SKOS also has specific properties to represent all of the semantic relationships that are described in Chapter 5.

Regarding issues of complexity and granularity of the *themas* and comprehensive semantic relationships between and among *themas* that FRSAD attempts to cover, OWL has even better matches. OWL ontologies provide classes, properties, individuals, and data values and are stored as Semantic Web documents⁵². OWL 1 mainly focused on constructs for expressing information about classes and individuals. OWL 2, the newest

⁴⁹ Dahlberg, I. (1992). Knowledge organization and terminology: philosophical and linguistic bases. *International Classification*. 19(2):65-71.

⁵⁰ Campbell et al., *op. cit.*

⁵¹ *SKOS Simple Knowledge Organization System Reference* (2009). Eds. Miles, A. and Bechhofer, S. W3C Candidate Recommendation 17 March 2009. Available at: <http://www.w3.org/TR/skos-reference/> (accessed 2010-01-20).

⁵² *OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax*. (2009). Eds. Motik, B, Patel-Schneider, P.F. and Parsia, B. W3C Working Draft 21 April 2009. Available at: <http://www.w3.org/TR/owl2-syntax/> (accessed 2010-01-20).

W3C working draft, offers new constructs for expressing additional restrictions on properties, new characteristics of properties, incompatibility of properties, properties chains, and key properties⁵³. OWL 2 provides axioms (statements that say what is true in the domain) that allow relationships to be established between class expressions, including: SubClassOf, EquivalentClasses, DisjointClasses, and DisjointUnion. More importantly, in OWL 2, classes and property expressions are used to construct class expressions, sometimes also called *descriptions*, and, in the description logic literature, *complex concepts*. It provides for enumeration of individuals and all standard Boolean connectives: *AND*, *OR*, and *NOT*. The ObjectIntersectionOf, ObjectUnionOf, and ObjectComplementOf class expressions provide for the standard set-theoretic operations on class expressions. The ObjectOneOf class expression contains exactly the specified individuals.

When the DCMI Abstract Model became a DCMI Recommendation in 2007, its one-to-one principle (i.e., each DC metadata description describes one, and only one, resource) was recognized or followed by other metadata standards. According to the DCMI model, a record can contain *description sets*, which may contain *descriptions* composed of *statements*, which use *property-value* pairs.⁵⁴ This results in information that can be processed, exchanged, referred to, and linked to at the statement level. When a record contains descriptions of the resource, the individual descriptions also can be linked to the authority data that manages the values associated with those properties (e.g., the subject authority data, the property name authority data, or the geographic authority data). Such an information model is independent of any particular encoding syntax and facilitates the development of better mappings and cross-syntax translations⁵⁵. The conceptual model proposed by the FRSAR Working Group corresponds to this abstract model by allowing any *thema* to be independent of any *nomen*, including any syntax that a *nomen* may use. Therefore this conceptual model will facilitate the sharing and reuse of subject authority data amongst not only the subject authority systems themselves, but also metadata resources.

C.3 Conclusion

Putting the subject authority data within the context of the Semantic Web developments, especially from the perspective of the Web of Data, subject authority data that are modelled based on FRSAD and encoded in SKOS and OWL will be able to become part of linked open data and contribute to the further development of the Semantic Web.

⁵³ *OWL 2 Web Ontology Language New Features and Rationale*. (2009). Eds. Golbreich, C. and Wallace, Evan K. W3C Working Draft 21 April 2009. Available at: <http://www.w3.org/TR/owl2-new-features/> (accessed 2010-01-20).

⁵⁴ *DCMI Abstract Model*. (2007). Eds. Powell, A., Nilsson, M., Naeve, A. Johnston, P. and Baker, T. Available at: <http://dublincore.org/documents/abstract-model/> (accessed 2010-01-20).

⁵⁵ *Ibid.*

APPENDIX D. EXAMPLES FROM SUBJECT AUTHORITY SYSTEMS

This appendix provides examples found in implementations of existing subject authority systems through the perspective of the FRSAD model, presented in four parts: 1) existing models of *thema* types; 2) *thema-thema* relationships presented in subject authority data (both in individual vocabularies and cross-schemes); 3) same *thema* represented by *nomens* from different schemes; and 4) examples of display records from controlled vocabularies or subject authority systems.

D.1 Existing Models of THEMA Types

In Chapter 4 Attributes, “type” is defined as a general attribute of *thema* because other attributes are usually implementation-dependent. In any particular application, *themas* would normally have particular implementation-specific types. Based on our preliminary study, there seems to be no generally applicable categorization of *themas*. This is also supported by the following examples, ranging from general (*Faceted Application of Subject Terminology*) to more specialized subject domains such as biomedical and health sciences (*Unified Medical Language System* and *The Foundational Model of Anatomy Ontology*) and art and architecture (*Art and Architecture Thesaurus*).

Example D.1.1 *Faceted Application of Subject Terminology (FAST) subject facets*⁵⁶:

Faceted Application of Subject Terminology (FAST) is an adaptation of the *Library of Congress Subject Headings (LCSH)* with a simplified syntax. LCSH headings form the basis for FAST authority file. FAST employs a faceted approach by defining headings according to their functions and categorizes all headings into eight facets. Seven of them are subject facets and one is form (genre) facet. The subject facets include:

- Topical
- Personal Names (as Subjects)
- Corporate Names (as Subjects)
- Geographics
- Periods
- Titles
- Events

Headings in the FAST database include both single-concept and multiple-concept headings. Each FAST heading or heading-string belongs to a single facet.

⁵⁶ FAST: *Faceted Application of Subject Terminology*. [2001-]. OCLC Online Computer Library Center. Available at: <http://www.oclc.org/research/activities/fast/default.htm> (accessed 2010-01-26).

Example D.1.2. Unified Medical Language System® (UMLS) semantic types^{57, 58}

The *Unified Medical Language System*® (UMLS), developed, maintained, and distributed by the National Library of Medicine of the United States, provides a unified system for correlating a large number of biomedical terminologies and facilitates the development of computer systems that behave as if they “understand” the meaning of the language of biomedicine and health. In order to facilitate the establishment of correspondences in the meanings of terms, the same concepts occurring in different constituent vocabularies are assigned to high level semantic types encompassed within the *UMLS Semantic Network*. It consists of: (a) a set of broad subject categories, or **Semantic Types**, that provide a consistent categorization of all concepts represented in the *UMLS Metathesaurus*®, and (b) a set of useful and important relationships, or **Semantic Relations**, which exist between Semantic Types. More than 130 semantic types and 50 semantic relationships defined by the UMLS can be found in the UMLS 2004 AB Documentation⁵⁹. The following are the high level semantic types:

Entities

- Physical Object
 - Organism
 - Anatomical Structure
 - Manufactured Object
 - Substance
- Conceptual Entity
 - Idea or Concept
 - Finding
 - Organism Attribute
 - Intellectual Product
 - Language
 - Occupation or Discipline
 - Organization
 - Group Attribute
 - Group

Events

- Activity
- Phenomenon or Process

The scope of the *UMLS Semantic Network* is broad, allowing for the semantic categorization of a wide range of terminology in multiple domains. The top level types are **Entities** (including “Physical Object” and “Conceptual Entity”) and **Events** (including “Activity” and “Phenomenon or Process”). Looking at its major groupings of semantic types (such as organisms, anatomical structures, biologic function, chemicals, events, physical objects, and concepts or ideas) it is obvious that they are designed to be especially applicable in the domain of biomedical and health areas.

⁵⁷ National Library of Medicine. (2003-). Unified Medical Language System. Current Semantic Types. UMLS 2004AB Documentation. Last updated: 21 March 2008. Available at:

http://www.nlm.nih.gov/research/umls/META3_current_semantic_types.html (accessed May 22, 2009).

⁵⁸ UMLS Factsheet. Available at: <http://www.nlm.nih.gov/pubs/factsheets/umls.html> (accessed May 22, 2009).

⁵⁹ *ibid.*

Example D.1.3. *The Foundational Model of Anatomy Ontology semantic types*⁶⁰

The Foundational Model of Anatomy (FMA) initially developed as an enhancement of the anatomical content of UMLS, is a domain ontology of the concepts and relationships that pertain to the structural organization of the human body. It is found that while there is considerable correspondence in the meaning of anatomical terms in the UMLS sources, there is very little similarity in the arrangement of anatomical terms among the source schemas. It is important that the underlying semantic structure of these abstractions must also be aligned. The top-level semantic types are **Anatomical Entity**, **Attribute Entity**, and **Dimensional Entity**:

- Anatomical Entity**
 - Non-physical anatomical entity
 - Physical anatomical entity
- Attribute Entity**
 - Cell morphology
 - Cell shape type
 - Cell surface feature
 - Concept name
 - Miscellaneous term
 - Organ part phenotype
 - Physical attribute relationship
 - Physical state
 - Structural relationship value
- Dimensional Entity**
 - Line
 - Point
 - Surface
 - Volume

As a domain ontology, the FMA represents deep knowledge of the structure of the human body. Its emphasis is on the highest level of granularity of the concepts. Meanwhile it also presents a great number of specific structural relationships between the references of these concepts. According to project documentation⁶¹, the FMA consists of approximately 75,000 anatomical classes, 130,000 unique terms, 205,000 frames, and 170 unique slots showing different types of relations, attributes, and attributed relationships. FMA is a typical example of modeling that shows how **semantic types** for a concept scheme can be defined. It not only encompasses the diverse entities that make up the human body but is also capable of modeling a great deal of knowledge relating these entities.

Example D.1.4. *Art and Architecture Thesaurus (AAT) facets*⁶²

Art and Architecture Thesaurus (AAT) is a controlled vocabulary for fine art, architecture, decorative arts, archival materials, and material culture for the purposes of

⁶⁰ *The Foundational Model of Anatomy ontology* (FMA). 2006-- School of Medicine, University of Washington. Available at: <http://sig.biostr.washington.edu/projects/fm/index.html> (accessed 2010-01-26).

⁶¹ About FMA. [2006]. School of Medicine, University of Washington. Available at: <http://sig.biostr.washington.edu/projects/fm/AboutFM.html> (accessed 2010-01-26).

⁶² *Art and Architecture Thesaurus Online. Hierarchy Display. op. cit.*

indexing, cataloging, and searching, as well as research tools. It was developed for literature about art and architecture and for records describing works of art and architecture. The facets in AAT are conceptually organized in a scheme that proceeds from abstract concepts to concrete, physical artifacts. These facets are: “Associated Concepts”, “Physical Attributes”, “Styles and Periods”, “Agents”, “Activities”, “Materials”, and “Objects”. Homogeneous groupings of terminology, or hierarchies, are arranged within the seven facets of the AAT:

- Top of the AAT hierarchies
 - **Associated Concepts Facet**
 - Associated Concepts
 - **Physical Attributes Facet**
 - Attributes and Properties
 - Conditions and Effects
 - Design Elements
 - Color
 - **Styles and Periods Facet**
 - Styles and Periods
 - **Agents Facet**
 - People
 - Organizations
 - Living Organisms
 - **Activities Facet**
 - Disciplines
 - Functions
 - Events
 - Physical and Mental Activities
 - Processes and Techniques
 - **Materials Facet**
 - Materials
 - **Objects Facet**
 - Object Groupings and Systems
 - Object Genres (Hierarchy Name)
 - Components (Hierarchy Name)
 - Built Environment (Hierarchy Name)
 - Furnishings and Equipment
 - Visual and Verbal Communication

The conceptual framework of facets is not subject-specific. One example is the subject “Renaissance painting”. Terms to describe Renaissance paintings will be found in many locations in the AAT hierarchies rather than a defined portion that is specific only for Renaissance painting⁶³.

In summary, all examples in this section indicate that in actual implementations there are always attempts to define some fundamental facets or atoms to accommodate all types of

⁶³About the AAT. Los Angeles: J. Paul Getty Trust, Vocabulary Program. Revised 12 November 2008. Available at: http://www.getty.edu/research/conducting_research/vocabularies/aat/about.html (accessed May 22, 2009).

themas. However, the resulting *themas* “types” differ from implementation to implementation.

D.2 THEMA-THEMA Relationships presented in Subject Authority Data

Authority **records** can be stored and displayed differently within a system, and they may also have various combinations of components when displayed to:

- information professionals who create and maintain subject authority data, including cataloguers and controlled vocabulary creators;
- information professionals who create and maintain metadata;
- reference services librarians and other information professionals who search for information as intermediaries; and
- end-users who search for information to fulfil their information needs.

Therefore, it is the authority **data**, not the **records**, which will be the focus in the examples presented in the following sections.

D.2.1 *Thema-Thema* relationships presented by individual vocabularies

The emphasis of this section is on the semantic relations presented in vocabularies. The following examples demonstrate how *thema-to-thema* relationships are presented in different vocabularies for the same *thema*, “mercury” (as a liquid metal and/or as an element). The same object can be viewed from different perspectives and therefore it may belong under different hierarchies (polyhierarchical relationship). Webster’s definition of mercury is: “a heavy silvery toxic univalent and bivalent metallic element; the only metal that is liquid at ordinary temperatures”⁶⁴.

[Note: In the figures in this section, an oval shaped node is used to represent a *thema*.]

⁶⁴ *Webster's Online Dictionary*. Definition: Mercury. Available at: <http://www.websters-online-dictionary.org/definition/mercury> (accessed July 2008).

Example D.2.1.1. LC Subject Authority*Thema:* **mercury (as a liquid metal)**

[Note: in the following entry, MARC21 coding is used:

010 = Library of Congress control number

040 = Cataloging source

053 = LC Classification Number

\$c = Explanatory term (specifying topic)

150 = Heading--Topical term

450 = *See from Tracing*--Topical term (unauthorized form/variant of term)550 = *See Also From Tracing*--Topical Term;

\$a = Topical term or geographic name entry element

\$w = Control subfield; g - Broader term.]

LC Control Number: sh 85083794**HEADING:** Mercury

000 00558cz a2200217n 450

001 4734282

005 19900221112154.6

008 860211ij anannbavn |a ana

035 __ |a (DLC)sh 85083794

906 __ |t 8528 |u fk03 |v 0

010 __ |a sh 85083794

040 __ |a DLC |c DLC |d DLC

053 _0 |a QD181.H6 |c Chemistry

053 _0 |a TA480.M4 |c Engineering materials

053 _0 |a TN271.M4 |c Prospecting

053 _0 |a TP245.M5 |c Chemical technology

150 __ |a Mercury

450 __ |a Hydrargyrum

450 __ |a Quicksilver

550 __ |w g |a Liquid metals

953 __ |a xx00 |b fg07

[Note: in this captured screenshot, subfield signs are displayed as a vertical bar.]

Figure D.1: A record from the LC Subject Authority File

Several semantic relationships are indicated in this record. There is a semantic relationship between this *thema*, which has a *nomen* “Mercury”, and another *thema*, which has a *nomen* “Liquid metals” (see illustration below). This can be recognized by the field tag 550, which means “see also”. (Inter-system relationships will be explained later in section D.2.2.)

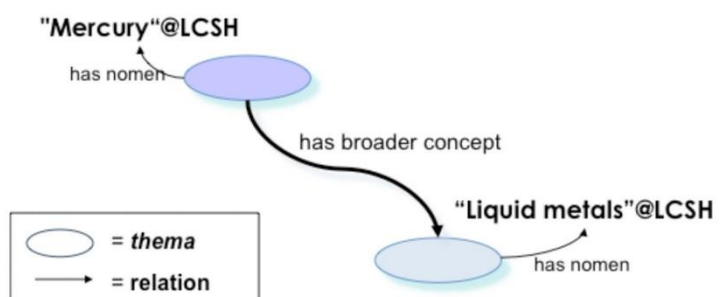


Figure D.2 Illustration of the semantic relations between two *themas* represented in Figure D.1

Example D.2.1.2. *Art and Architecture Thesaurus*:

Thema: **mercury (as a liquid metal and as an element)**

ID: 300011026

Record Type: concept

mercury (<mercury and amalgam>, nonferrous metal, ... Materials)

Note: Pure metallic element having symbol Hg and atomic number 80; a lustrous silvery metal that is liquid at ordinary temperatures. Use also for this metal as processed and formed, usually in combination with other substances, to make various objects and materials.

Terms:

mercury (preferred, C, D, U, LC, English-P)

Hg (C, UF, U, A, English)

quicksilver (C, UF, U, English)

argento vivo (C, D, U, Italian-P)

Facet/Hierarchy Code: M.MT

Hierarchical Position:

```

Materials Facet
... Materials
..... materials
..... <materials by composition>
..... inorganic material
..... <metal and metal products>
..... metal
..... <metal by composition or origin>
..... nonferrous metal
..... <mercury and amalgam>
..... mercury ←

```

Additional Parents:

```

Materials Facet ←
... Materials
..... materials
..... <materials by form>
..... <materials by chemical form>
..... elements (chemical substances)
..... mercury ←

```

Figure D.3 An online display record of the AAT concept “Mercury”

Figure D.3 shows a screen captured from the *Art and Architecture Thesaurus* (AAT) online version. Hierarchical relationships of the *themas* represented by *nomens* “mercury”, “elements (chemical substances)”, and “nonferrous metal” are presented in the hierarchies. Such semantic relationships are illustrated in the following figure (Figure D.4).

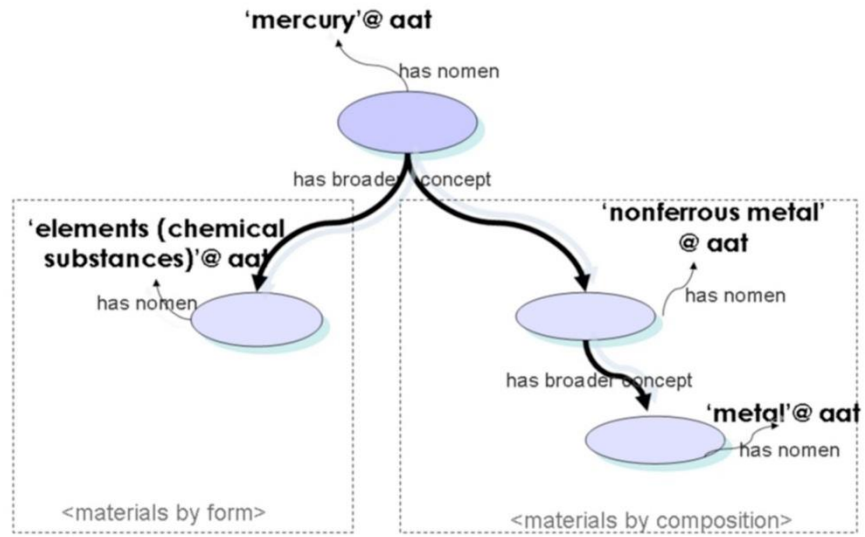


Figure D.4 Illustration of the semantic relations between the *themas* presented in Figure D.3

Example D.2.1.3. Medical Subject Headings (MeSH): Standard Display

Thema: mercury (as a liquid metal and as an element):

MeSH Heading	Mercury
Tree Number	D01.268.556.504
Tree Number	D01.268.956.437
Tree Number	D01.552.544.504
<p>Inorganic Chemicals [D01]</p> <p>Elements [D01.268]</p> <p>Metals, Heavy [D01.268.556]</p> <p>Mercury [D01.268.556.504] ←</p>	
<p>Inorganic Chemicals [D01]</p> <p>Elements [D01.268]</p> <p>Transition Elements [D01.268.956]</p> <p>Mercury [D01.268.956.437] ←</p>	
<p>Inorganic Chemicals [D01]</p> <p>Metals [D01.552]</p> <p>Metals, Heavy [D01.552.544]</p> <p>Mercury [D01.552.544.504] ←</p>	
See Also	Mercury Isotopes
See Also	Mercury Radioisotopes
See Also	Organomercury Compounds
Allowable Qualifiers	AD AE AG AI AN BL CF CH CL CT DF DU EC HI IM IP ME PD PH PK RE SD ST TO TU UR

Figure D.5 Extracted portion from a MeSH record indicating semantic relations

Figure D.5 shows data derived from a Standard Display of a MeSH record found through the MeSH Browser. It can be viewed from three segments:

a) The hierarchical relationships can be traced following the “Tree Numbers”. Analysis reveals two immediate hierarchical relationships (see Figure D.6; notational form of *nomens* are not included): (1) between *themas* represented by *nomens* “Mercury” and “Transition Elements”; (2) between *themas* represented by the *nomens* “Mercury” and “Metals, Heavy”. The latter can be traced up to two upper classes”.

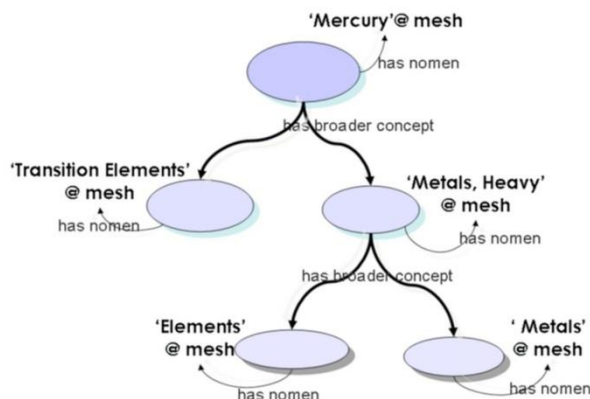


Figure D.6 Illustration of the hierarchical relationships (through the tree structure) from the extracted MeSH Heading record shown in Figure D.5

b) The information indicates that the *thema* represented by a *nomen*, “Mercury”, has associative relationships (“see also”) with *themas* represented by *nomens* “Mercury Isotopes”, “Mercury Radioisotopes”, and “Organomercury Compounds”, as illustrated in Figure D.7:

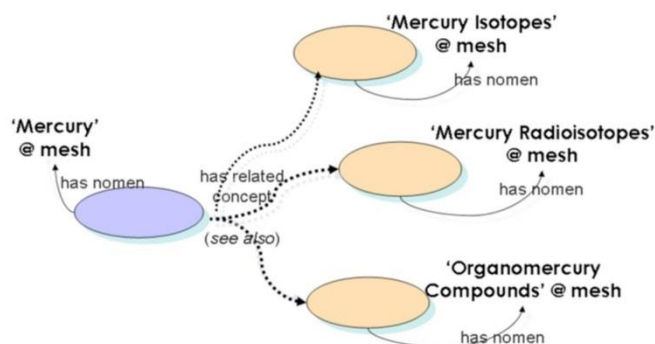


Figure D.7 Illustration of the associative relationships (“see also”) from the extracted MeSH record shown in Figure D.5

c) The MeSH record also provides allowable qualifiers to enable the forming of more complex concepts. In this example, the concept can be further limited to specific aspects: “administration & dosage (AD)”, “isolation & purification (IP)”, “toxicity (TO)”, etc. These facilitate the forming of specific subject headings (e.g., “Mercury – TO”, or “Mercury – IP”) to represent different *themas*.

Example D.2.1.4 . Dewey Decimal Classification

Thema: mercury (as a metal)

Class Number: 669.71
Segmented Number: 669/.71
Caption: Mercury

Main Classes

600	<u>Technology</u>
660	<u>Chemical engineering</u>
669	<u>Metallurgy</u>
669.1-669.7	<u>Metallurgy of specific metals and their alloys</u>
669.2-669.7	<u>Nonferrous metals</u>
669.7	<u>Other nonferrous metals</u>
669.71	Mercury

Figure D.8a. Screen captured from OCLC Connexion WebDewey for classes related to “mercury (as a metal)”

Thema: mercury (as an element)

Class Number: 546.663
Segmented Number: 546/.663
Caption: *Mercury

Main Classes

500	<u>Science</u>
540	<u>Chemistry</u>
541-547	<u>Chemistry</u>
546	<u>Inorganic chemistry</u>
546.6	<u>Groups 8, 9, 10, 11, 12, 13, 14</u>
546.66	<u>Group 12</u>
546.663	*Mercury
546.6635	<u>Mercury (Element)--physical chemistry</u>

Figure D.8b. Screen captured from OCLC Connexion WebDewey for classes related to “mercury (as an element)”

It should be noted that although the relationships are similar to what is presented in other thesauri (shown before), in a classification scheme such relationships are presented through the notational codes associated with *themas*, which reflect the conceptual hierarchy of a scheme. Hence it is the **notations** (669.71 and 546.663), not the **captions**, that represent the *themas*, as one can find from the above figures where both captions are “Mercury” although they are affiliated with two different classes in DDC. The two pairs

of hierarchical relationships are illustrated in the following figures: Figure D.9a is for *thema* “mercury as a metal” and Figure D.9b is for *thema* “mercury as an element”.

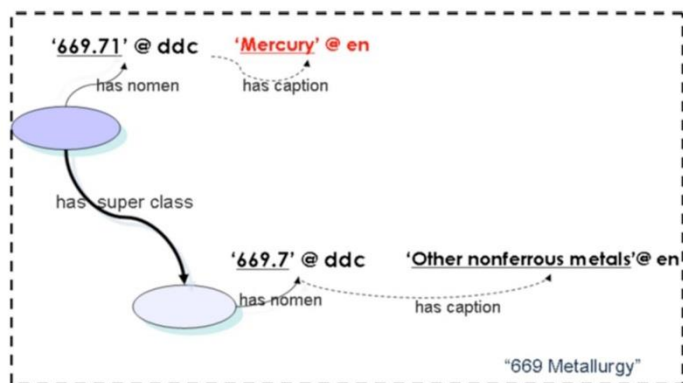


Figure D.9a Illustration of the hierarchical relationships (through the classificatory structure) between the DDC classes shown in Figure D.8a

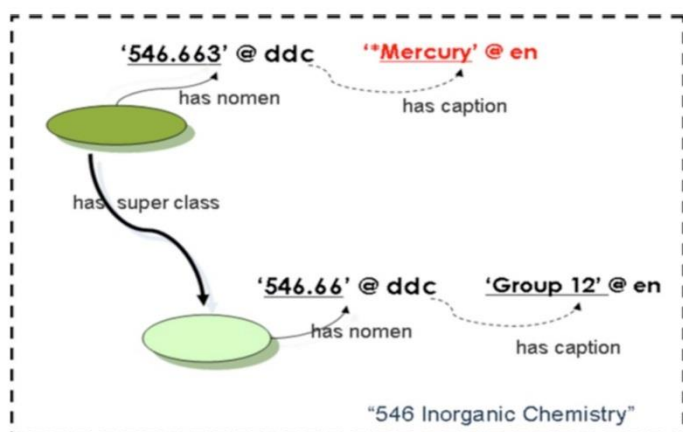


Figure D.9b Illustration of the hierarchical relationships (through the classificatory structure) between the DDC classes shown in Figure D.8b

D.2.2 Inter-system *THEMA* crosswalking through *NOMENS*

Example D.2.2.1 *INSPEC Thesaurus* and *INSPEC Classification*

Thema: **mercury (planet)**

Note: Although the term “Mercury” has multiple meanings and is a good example of homographs, the focus of this section is **not** on homograph control.

From *INSPEC Thesaurus* (2004, pg. h76):

[Note: CC= Classification Code]

	Mercury (planet)	←
	BT planets	
	TT planets	
	RT transits	
→	CC A9630D	
	DI January 1971	
	PT planets	

From *INSPEC Classification* (2004 pg. 84):

	A9630	Planets and satellites	
		<i>(exc. the Moon)</i>	
		<i>for Earth, see A91... for celestial mechanics, see A9510...</i>	
→	A9630D	Mercury	←

Figure D.10 Extracted entries from *INSPEC Thesaurus* (top) and *INSPEC Classification* (bottom) showing inter-system *thema* crosswalking

Example D.2.2.1 demonstrates that a *thema*, “planet Mercury”, can be crosswalked through the *nomens* in two different authority systems, where “Mercury (planet)” is a *nomen* (in a form of a thesaurus term) from the *INSPEC Thesaurus* and “A9630D” is a *nomen* (in a form of a notation in a classification) from the *INSPEC Classification*. This is illustrated in Figure D.11.

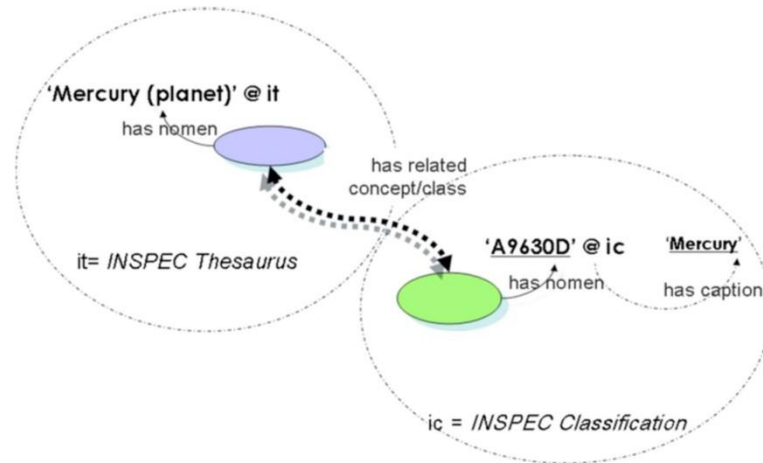


Figure D.11 Illustration of the inter-system *themas'* crosswalking between *INSPEC Thesaurus* and *INSPEC Classification* shown in Figure D.10

Example D.2.2.2. LCSH and Library of Congress Classification (LCC)

Thema: **“Mercury” (as a metal and an element)**

Example taken from Library of Congress Subject Authority File:

[Note: in the following entry, MARC21 coding is used:

010 = Library of Congress control number

040 = Cataloging source

053 = LC Classification Number

\$c = Explanatory term (specifying topic)

150 = Heading--Topical term

450 = *See* from Tracing--Topical term (unauthorized form/variant of term)

550 = *See Also* From Tracing--Topical Term;

\$a = Topical term or geographic name entry element

\$w = Control subfield; g = Broader term.]

This same record is also used in a previous section (D.2.1) when semantic relationships between *themas* from the same scheme are presented. In the following example, the relationships of *themas* from different schemes are further explored.

LC Control Number: sh 85083794

HEADING: Mercury

000 00558cz a2200217n 450

001 4734282

← 005 19900221112154 6

008 860211i| anannbabn |a ana

035 __ |a (DLC)sh 85083794

906 __ |t 8528 |u fk03 |v 0

010 __ |a sh 85083794

040 __ |a DLC |c DLC |d DLC

053 _0 |a QD181.H6 |c Chemistry

053 _0 |a TA480.M4 |c Engineering materials

053 _0 |a TN271.M4 |c Prospecting

053 _0 |a TP245.M5 |c Chemical technology

150 __ |a Mercury

450 __ |a Hydrargyrum

450 __ |a Quicksilver

550 __ |w g |a Liquid metals

953 __ |a xx00 |b fg07

[Note: in this captured screenshot, subfield signs are displayed as a vertical bar.]

Figure D.12. A record from the LC Subject Authority File

In this example, the *thema* “mercury” (as a metal and an element), represented by the *nomen* “Mercury” in LCSH, is crosswalked to the *Library of Congress Classification* (LCC) where the *thema* is placed in different classes that have the *nomens* “QD181.H6” (in Chemistry), “TA480.M4” (in Engineering materials), “TN271.M4” (in Prospecting), and “TP245.M5” (in Chemical technology). Figure D.13 illustrates such relationships.

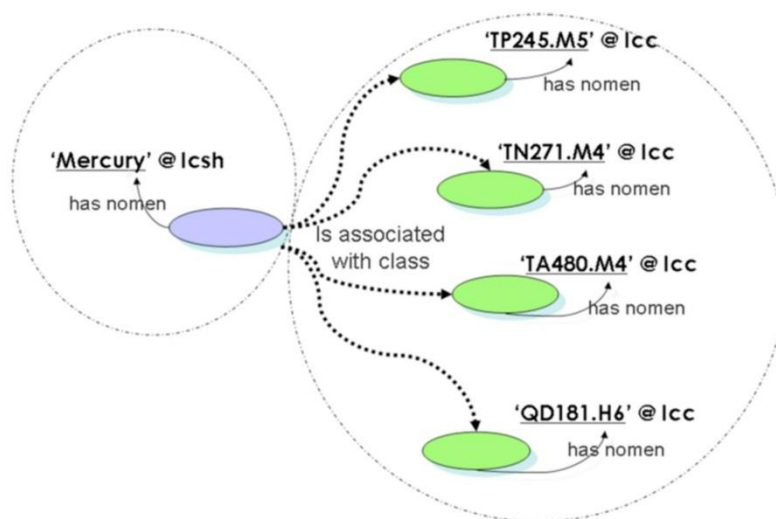


Figure D.13. Illustration of the inter-system *themas'* crosswalking between LCSH and LCC showing in Figure D.12

D.3 Same THEMA Represented by NOMENs from Different Schemes

The following case demonstrates that, to some extent, the granularity of a *thema* is also dependent on its appellations in a particular scheme.

For example, a resource is about “academic library labor unions in Germany”. The *thema* will be represented by the *nomens* established in different schemes such as:

DDC: “331.881102770943”

Constructed/combined from:

331.8811 – labor unions in industries and occupations other than
extractive, manufacturing, construction
-027.7 – academic libraries
-0943 – Germany

LCSH: “Library employees--Labor unions—Germany”

“Universities and colleges--Employees--Labor unions—Germany”

“Collective bargaining--Academic librarians--Germany”

“Libraries and labor unions--Germany”

FAST:

“Library employees--Labor unions”

“Universities and colleges--Employees--Labor unions”

“Collective bargaining--Academic librarians”

“Libraries and labor unions”

“Germany”

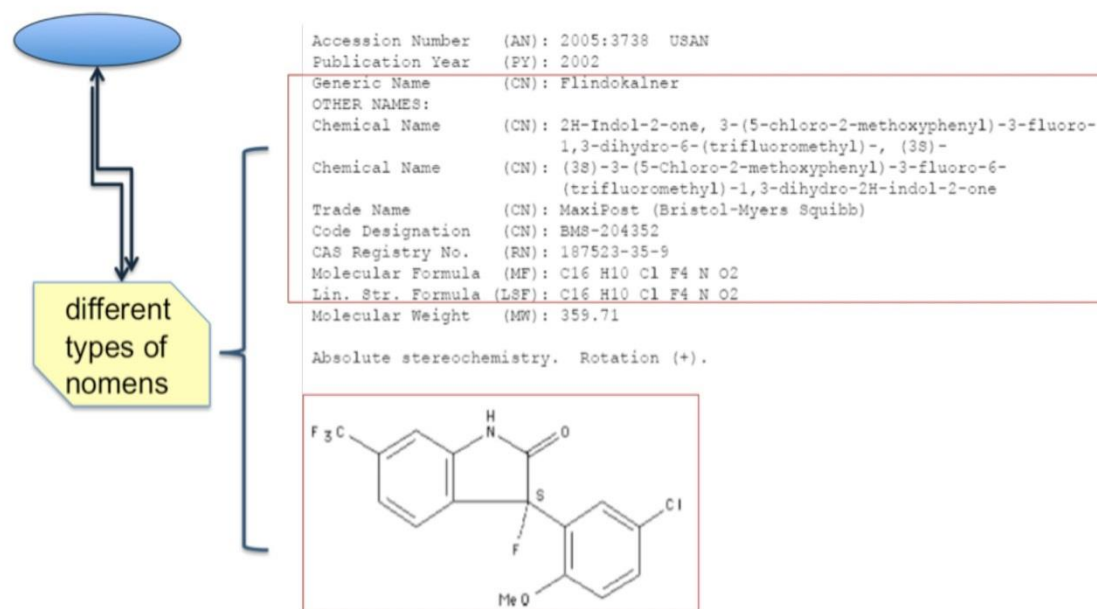
As this example demonstrates, schemes may allow the representation of *themas* at different levels of specificity through the structure and syntax of the *nomens* they have established.

D.4 Examples of Display Records from Controlled Vocabularies or Subject Authority Files

As shown in section D.2, authority **records** can be displayed differently within a particular system; furthermore, they can also have various combinations of authority **data** when displayed to different users (e.g., subject authority data creators and maintainers, metadata creators and end-users). Following are captured screens of records displayed online. They contain mixed information regarding *thema*, *nomen*, relationships between a *thema* and its *nomens*, as well as among different *themas*. In addition, they demonstrate that *thema* types are implementation-dependent and vary in different domains.

Example D.4.1. A chemical substance and its NOMEN -- A display record from *The USP Dictionary of U.S. Adopted Names and International Drug Names*

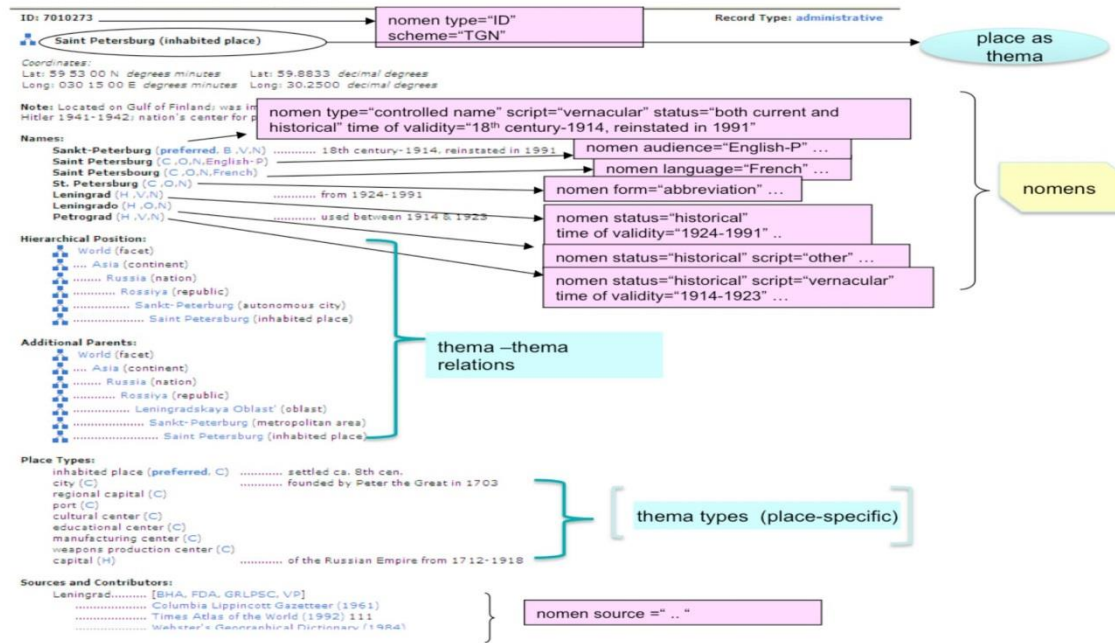
The figure below demonstrates how a *thema* could have various *nomens* in the context of specific systems. The forms of the *nomens* for this chemical compound are not only in various names represented in natural language, but also those represented in artificial languages such as codes, formulas and a graph.



Source: STN Database Summary Sheet: USAN (The USP Dictionary of U.S. Adopted Names and International Drug Names)
<http://www.cas.org/ASSETS/773D56DEC03E4769BF0E1BC206BB371E/usan.pdf>, p.5. Record reprinted with permission.

Example D.4.2. A place as a thema – A display record from *Getty Thesaurus of Geographic Names (TGN)*

This example presents: (1) the hierarchical relationships of a *thema* (in this case a place) with other *themas*, i.e., the “whole-part” relationships; (2) various *nomens*, to be chosen as preferred terms in different contexts, with attributes regarding the form, time of validity, status, audience, and source of a particular *nomen*; and (3) *thema* types that are place-specific.



Source: *Getty Thesaurus of Geographic Names Online*. http://www.getty.edu/research/conducting_research/vocabularies/tgn/
Record reprinted with permission.

Example D.4.3. A display record (Extensive Concept View) from *Medical Subject Headings (MeSH)*

Thema-thema relationships presented in the *Medical Subject Headings (MeSH)* have been explained in a previous section with Example D.2.1.3 and Figure D.6 and D.7. The following Expanded Concept View displays an additional component for “Concept 1: Mercury.” The summary of the semantic relationships displayed in this record is presented below the figure.

MeSH Heading	Mercury	→	nomen type="constructed name" Scheme="MeSH"
Tree Number	D01.268.556.504	}	nomens, used to represent hierarchical relations between themas in tree structures
Tree Number	D01.268.956.437		
Tree Number	D01.552.544.504		
Annotation	Hg-202	→	nomen type="constructed name" form="code" Scheme="..."
Concept 1 (Preferred)	Mercury	→	thema represented by nomen
	Concept UI	M0013448	→
Scope Note	A silver metallic element that exists as a liquid at room temperature. It has the atomic symbol Hg (from hydrargyrum, liquid silver), atomic number 80, and atomic weight 200.59. Mercury is used in many industrial applications and its salts have been employed therapeutically as purgatives, antisyphilitics, disinfectants, and astringents. It can be absorbed through the skin and mucous membranes which leads to MERCURY POISONING . Because of its toxicity, the clinical use of mercury and mercurials is diminishing.		
Semantic Type	T131 (Hazardous or Poisonous Substance)	}	thema - thema relations
Semantic Type	T196 (Element, Ion, or Isotope)		
CAS Type 1 Name	Mercury	→	nomen type="constructed name" Scheme="UMLS"
Registry Number	7439-97-6	→	nomen type="ID" Scheme="UMLS"
Term (Preferred)	Mercury	→	nomen type="constructed name" Scheme="NLM(1966)"
Term UI	T025687	→	nomen type="ID" Scheme="NLM(1966)"
Date	01-JAN-1999	→	nomen time of validity = "19990101"
Lexical Tag	NON	→	nomen transcription/transliteration="none"
Thesaurus	NLM (1966)	→	nomen source="NLM(1966)"
See Also	Mercury Isotopes	}	thema - thema relations
See Also	Mercury Radioisotopes		
See Also	Organomercury Compounds		
Allowable Qualifiers	AD AE AG AI AN BL CF CH CL CT DF DU EC HI IM IP ME PD PH PK RE SD SI TO TU UR		
Entry Combination	poisoning:Mercury Poisoning	→	nomen-nomen relation
Date of Entry	19990101	→	nomen time of validity="19990101"
Unique ID	D008628		

Source: Medical Subject Headings on MeSH Browser (2008 MeSH); <http://www.nlm.nih.gov/mesh/2008/MBrowser.html>

This Expanded Concept View presents various types of semantic relationships among *themas*:

- a) Two immediate hierarchical relationships: (1) between *themas* represented by *nomens* “Mercury” and “Transition Elements”. The same is true for these *themas* and their *nomens* with notational forms; (2) between *themas* represented by the *nomens* “Mercury” and “Metals, Heavy”. The latter can be traced up to two upper classes.
- b) Associative relationships between “Mercury” (as a liquid metal and as an element) and other *themas* represented by *nomens* “Mercury Isotopes”, “Mercury Radioisotopes”, and “Organomercury Compounds”.
- c) Allowable qualifiers enable the concept to be further limited to specific perspectives (e.g., “administration & dosage (AD)”, “isolation & purification (IP)”, and “toxicity (TO)”). These facilitate the forming of specific subject headings (e.g., “Mercury – TO”, or “Mercury – IP”) to represent different *themas*.
- d) The semantic types of this *thema*: “T131 (Hazardous or Poisonous Substance)” and “T196 (Element, Ion, or Isotope)” as defined by UMLS.

Thema-nomen relationships are clearly presented in the record, including the *nomens* in natural languages and as specific identification numbers. Various attributes of *nomens* are also presented.

REFERENCES

- Aitchison, J., Gilchrist, A. and Bawden, D. (2000). *Thesaurus Construction and Use: A Practical Manual*. 4th ed. London: Fitzroy Dearborn.
- Art and Architecture Thesaurus*. (2000-). Los Angeles: J. Paul Getty Trust, Vocabulary Program. Available at: http://www.getty.edu/research/conducting_research/vocabularies/aat/ (accessed 2010-01-20).
- Buizza, P. and Guerrini, M. 2002. A conceptual model for the New Soggetario: Subject indexing in the light of FRBR. *Cataloging & Classification Quarterly*, 34(4): 31–45.
- Campbell, K. E., Oliver, D.E., Spackman, K.A. and Shortliffe, E.H. (1998). Representing thoughts, words, and things in the UMLS. *Journal of the American Medical Informatics Association*, 5(5): 421–431.
- Categories for the Description of Works of Art (CDWA)*. (2000). Eds. Baca, M. and Harpring, P. The J. Paul Getty Trust and College Art Association. Los Angeles, CA: Getty Research Institute.
Available at:
http://www.getty.edu/research/conducting_research/standards/cdwa/index.html (accessed 2010-01-20)
- Clarke, S.G. (2001). Thesaural relationships. In: *Relationships in Knowledge Organization*. Eds. Bean, C.A. and Green, R. Dordrecht: Kluwer.
- Dahlberg, I. (1992). Knowledge organization and terminology: philosophical and linguistic bases. *International Classification*. 19(2):65-71.
- DCMI Abstract Model*. (2007). Eds. Powell, A., Nilsson, M., Naeve, A. Johnston, P. and Baker, T. Available at: <http://dublincore.org/documents/abstract-model/> (accessed 2010-01-20).
- Delsey, T. (2005). Modeling subject access: Extending the FRBR and FRANAR conceptual models. *Cataloging & Classification Quarterly* 39 (3/4): 49–61.
- Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, NF 100. 1892, S. 25–50. Available at: <http://www.gavagai.de/HHP31.htm> (accessed 2009-05-22).
- Functional Requirements for Authority Data - A Conceptual Model*. (2009). IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR), ed. by Glenn E. Patton. München: K.G. Saur.

- Functional Requirements for Bibliographic Records: Final Report*. (1998). IFLA Study Group on the Functional Requirements for Bibliographic Records. München: KG Saur. Available at: <http://www.ifla.org/files/cataloguing/frbr/frbr.pdf> (accessed 2010-01-20).
- Furner, J. (2006). The ontology of subjects of works. Paper presented at ASIS&T 2006: Annual Meeting of the American Society for Information Science and Technology, Austin, TX, November 3–8, 2006.
- Goodman, N. (1961). About. *Mind*, 70 (277): 1-24.
- Heaney, M. (1997). Time is of the essence. Available at: <http://www.bodley.ox.ac.uk/users/mh/time978a.htm> (accessed 2010-01-20).
- Hill, L. (1999). Content standards for digital gazetteers. Presentation at the JCDL2002 NKOS Workshop "Digital gazetteers--Integration into distributed digital library services", July 18, 2002, Portland, Oregon. Available at: <http://nkos.slis.kent.edu/DL02workshop.htm> (accessed 2010-01-20).
- Hjørland, B. (1992). The concept of "subject" in information science. *Journal of Documentation*, 48 (2): 172-200.
- Hutchins, W. J. (1977). On the problem of "aboutness" in document analysis. *Journal of Informatics*, 1 (1): 17-35.
- ISO (2009). *ISO/CD 25964-1, Information and documentation — Thesauri and interoperability with other vocabularies — Part 1: Thesauri for information retrieval*. ISO/TC 46 / SC 9 ISO 25964 Working Group.
- Lancaster, F.W. (1986). *Vocabulary Control for Information Retrieval*. 2nd ed. Arlington, Virginia: Information Resources Press.
- Michel, D. (1996). Taxonomy of Subject Relationships. Appendix B (Part 2), Final Report to the ALCTS/CCS Subject Analysis Committee, submitted by Subcommittee on Subject Relationships/Reference Structures. ALA Association for Library Collections and Technical Services (ALCTS) Cataloging and Classification Section (CCS) Subject Analysis Committee (SAC). Available at: <http://www.ala.org/ala/mgrps/divs/alcts/mgrps/ccs/cmtes/subjectanalysis/subjectrelations/mrsrscu2.pdf>. Also available is the hierarchical display of this Taxonomy, at: <http://www.ala.org/ala/mgrps/divs/alcts/mgrps/ccs/cmtes/sac/inact/subjectrelations/appendixbpartii.cfm> (accessed 2010-03-31).
- National Library of Medicine. (2003-). *Unified Medical Language System*. Available at: <http://www.nlm.nih.gov/research/umls/> (accessed 2010-01-20).
- NISO. (2005). ANSI/NISO Z39.19-2005. *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*. Bethesda, Maryland: NISO Press.

- Ogden, C. K., and Richards, I. A. (1923). *The Meaning of Meaning: A Study of the Influence of Language Upon Thought and of the Science of Symbolism*. London: Routledge & Kegan Paul.
- OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax*. (2009). Eds. Motik, B, Patel-Schneider, P.F. and Parsia, B. W3C Working Draft 21 April 2009. Available at: <http://www.w3.org/TR/owl2-syntax/> (accessed 2010-01-20).
- OWL 2 Web Ontology Language New Features and Rationale*. (2009). Eds. Golbreich, C. and Wallace, Evan K. W3C Working Draft 21 April 2009. Available at: <http://www.w3.org/TR/owl2-new-features/> (accessed 2010-01-20).
- Putnam, H. (1958). Formalization of the concept "about." *Philosophy of Science*, 25 (2): 125-130.
- Rust, G. and Bide, M. (2000). The <indecs> metadata framework: Principles, Model and Data Dictionary. Version2. Indecs Framework Ltd. Available at: http://www.doi.org/topics/indecs/indecs_framework_2000.pdf (accessed 2010-01-20).
- Ryle, G. (1933). "About." *Analysis*, 1 (1): 10-12.
- Shatford, S. (1986). Analyzing the subject of a picture: A theoretical approach. *Cataloging & Classification Quarterly*, 6 (3): 39-62.
- SKOS Simple Knowledge Organization System Reference* (2009). Eds. Miles, A. and Bechhofer, S. W3C Candidate Recommendation 17 March 2009. Available at: <http://www.w3.org/TR/skos-reference/> (accessed 2010-01-20).
- Subcommittee on Subject Relationships/Reference Structures. (1997). Final Report to the ALCTS/CCS Subject Analysis Committee. ALA Association for Library Collections and Technical Services (ALCTS) Cataloging and Classification Section (CCS) Subject Analysis Committee (SAC). Available at: <http://www.ala.org/ala/mgrps/divs/alcts/mgrps/ccs/cmtes/sac/inact/subjectrelations/finalreport.cfm> (accessed 2010-01-20).
- Svenonius, E. (2000). *The Intellectual Foundation of Information Organization*. Cambridge, Mass.: MIT Press.
- Wilson, P. (1968). *Two kinds of power: An essay on bibliographic control*. Berkeley, CA: University of California Press.