

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO

Kátia Soares Braga

**METADADOS PARA REPRESENTAÇÃO TEMÁTICA
DE CONTEÚDOS INFORMACIONAIS
EM SISTEMAS DE INFORMAÇÃO:
UMA CONTRIBUIÇÃO DO TESAURO DA CÂMARA DOS DEPUTADOS (TECAD)**

BRASÍLIA/DF

2013

Kátia Soares Braga

**METADADOS PARA REPRESENTAÇÃO TEMÁTICA
DE CONTEÚDOS INFORMACIONAIS
EM SISTEMAS DE INFORMAÇÃO:
UMA CONTRIBUIÇÃO DO TESAURO DA CÂMARA DOS DEPUTADOS (TECAD)**

Monografia apresentada ao programa de Pós-graduação em Ciência da Informação, no curso Arquitetura e Organização da Informação da Escola de Ciência da Informação, da Universidade Federal de Minas Gerais, como requisito para a obtenção do certificado de Especialista

Linha de Pesquisa: Gestão da Informação e do Conhecimento

Orientadora: Prof. Dra. Maria Aparecida Moura

BRASÍLIA

2013

Dados Internacionais de Catalogação da Publicação (CIP)

Kátia Soares Braga (CRB/DF 1522)

Braga, Kátia Soares.

Metadados para representação temática de conteúdos informacionais em sistemas de informação: uma contribuição do Tesouro da Câmara dos Deputados (TECAD) / Kátia Soares Braga. Brasília: [s.n.], 2013. 97f.p.

Orientadora: Prof. Dra. Maria Aparecida Moura

Monografia (Especialização em Arquitetura e Organização da Informação) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Programa de Pós-Graduação em Ciência da Informação

1. Metadados. 2. Representação temática. 3. Sistema de Organização do Conhecimento (SOC). 4. Tesouro. 5. Gestão da Informação e do Conhecimento – Câmara dos Deputados (Brasil). 6. Arquitetura e Organização da Informação – Câmara dos Deputados (Brasil). I. Título: uma contribuição do Tesouro da Câmara dos Deputados (TECAD).

CDU 001.103.2: 025.4



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

Monografia apresentada no Curso de Especialização em Arquitetura e Organização da Informação, de autoria de Kátia Soares Braga, aprovada pela banca examinadora constituída pelos seguintes professores:

Prof. Dra. Maria Aparecida Moura
Universidade Federal de Minas Gerais

Prof. Dra. Gercina Ângela Borém de Oliveira Lima
Universidade Federal de Minas Gerais

Prof. Ms. Raphael da Silva Cavalcante
Câmara dos Deputados

Prof. Dr. Jorge Tadeu de Ramos Neves
Coordenador(a) do curso de Especialização em Arquitetura e Organização da Informação da
Escola de Ciência da Informação da UFMG.

Data de aprovação: Brasília, 29 de abril de 2013

Via N3 - Projeção "L" - Setor de Garagens Ministeriais Norte
Garagem da Câmara dos Deputados - Complexo Avançado - Prédio do CEFOR –
Brasília, DF – Brasil
CEP: 70160-900
Telefone: (61) 3216-7500 / 3216-7501
Fax: (61) 3216-7515

Para Pedro,

Para Hilda,

Para Marcelo,

com o afeto de sempre.

Agradeço aos meus colegas Bruno, Fabíola e Márcia que, junto a mim, vêm partilhando de momentos profissionais no Tesouro da Câmara dos Deputados;

Aos colegas do curso, que me fizeram aprender mais e melhor sobre as atividades de gestão e informação e conhecimento da Câmara dos Deputados;

Aos professores, por terem tornado o curso tão estimulante e desafiador; em especial a professora Cida Moura que me orientou;

Aos gestores e demais colegas do CEDI, pela oportunidade e apoio que me deram para a participação nesse curso.

"Um objeto encontra sua imagem, um objeto encontra seu nome. Acontece que a imagem e o nome deste objeto se encontram"

"Uma forma qualquer pode substituir a imagem de um objeto"

"As figuras vagas têm uma significação tão necessária quanto perfeita como as figuras precisas"

"Um objeto nunca executa a mesma atividade que seu nome ou que sua imagem".

As palavras e as imagens, René Magritte, 1928.

"Na linguagem, tudo são idiossincrasias¹".

Kátia Soares Braga

¹**Idiossincrasia**, temperamento peculiar, é uma característica comportamental ou estrutural peculiar a um indivíduo ou grupo. Símbolos idiossincrásicos são símbolos que podem significar alguma coisa para uma pessoa em particular. Por isso, linguistas preconizam que palavras não são apenas arbitrárias, mas também importantes sinais idiossincrásicos.

RESUMO

Essa é uma pesquisa social qualitativa do tipo descritiva onde foi realizado um estudo de caso com o objetivo de analisar a contribuição dos Sistemas de Organização do Conhecimento (SOC) para a modelagem de metadados para a representação temática de conteúdos informacionais nos sistemas de informação da Câmara dos Deputados. A falta de campos estruturados e a ausência de definição de metadados nos sistemas de informação da Câmara dos Deputados comprometem o tratamento, a recuperação e a visibilidade da informação; assim como o desenvolvimento e uso do próprio tesouro, dos vocabulários controlados e das linguagens documentárias. Os SOC devem prover o léxico e a sintaxe para que a gramática de suas terminologias contribua para uma comunicação eficaz e de qualidade entre os usuários e os sistemas de informação, principalmente em se tratando do ambiente *web* da internet tal qual o Portal da Câmara dos Deputados que visa reunir e disponibilizar as informações e notícias sobre o processo legislativo, legislação e orçamento da União. Para coleta e análise dos dados foram utilizadas as técnicas de pesquisas bibliográfica e documental para levantamento dos referenciais teórico e para identificar as características dos tesouros, dos vocabulários documentários e das linguagens documentárias, assim como os sistemas de informação e as tipologias documentais da Câmara dos Deputados. Na fundamentação, foram utilizados referenciais teóricos sobre os SOC; as teorias que influenciam o uso e desenvolvimento desses sistemas, tais como teoria do conceito, classificação, indexação e outros aspectos relacionados à representação temática; além de padrões, perfil e esquema de metadados. Os resultados apresentaram uma aplicação da modelagem de metadados para representação temática dos conteúdos informacionais nos sistemas de informação da Câmara dos Deputados. Sem pretender ser uma solução definitiva, essa modelagem apontou a importância da organização e da representação temática dos conteúdos informacionais nos sistemas de informação para a gestão da informação do conhecimento e o desenho da arquitetura da informação, principalmente o papel a ser desempenhado pelos SOC nesse processo.

Palavras-chave: Metadados, Representação temática, Sistema de Organização do Conhecimento (SOC), Tesouro, Gestão da Informação e do Conhecimento – Câmara dos Deputados (Brasil), Arquitetura e Organização da Informação -- Câmara dos Deputados (Brasil).

ABSTRACT

This is a descriptive qualitative social research. It was conducted a case study in order to analyze the contribution of Knowledge Organization Systems (KOS) for modeling metadata for the thematic representation of information content in the information systems of the House of Representatives, Brazil. The lack of structured fields and the undefinition of metadata in the information systems disturb the treatment, retrieval, and visibility of information; as well as the development and use of thesaurus, controlled vocabularies and indexing languages. The Knowledge Organization Systems (KOS) should provide lexicon, and also syntax for the terminology's grammar to establish effective communication between users and information systems. Especially in the internet environment as the House of Representatives' homepage that provides information and news on the legislative process, legislation and Union's budget. For collecting and analyzing data were used bibliographical and documentary research techniques. These techniques revealed characteristics of thesauri, controlled vocabularies and documentary languages, as well as the information systems and document typologies. As theoretical framework, it was used theories of Knowledge Organization Systems (KOS), such as the theory of the concept, classification, indexing, and other aspects related to the thematic representation issue; so standards, profile and metadata schema. The results showed an application of modeling metadata for thematic representation of the informational content in the information systems of the House of Representatives, Brazil. This model intends to reveal the importance of organization and thematic representation of the information content in information systems for information and knowledge management, and organization and information architecture, especially the role to be played by the KOS.

Keywords: Metadata, thematic representation, Knowledge Organization System (KOS), thesaurus, Information and Knowledge Management – House of Representatives (Brazil), Organization and Information Architecture – House of Representatives (Brazil).

LISTA DE ABREVIATURAS

AACR	<i>Anglo-American Cataloguing Rule</i>
CDD	Código Decimal de Dewey
CDDORIS	Classificação Decimal de Direito
CDU	Classificação Decimal Universal
CEDI	Centro de Documentação e Informação
Celeg	Coordenação de Estudos Legislativos
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
DCMI	Dublin Core Metadata Initiative
FRAD	Requisitos Funcionais Para Dados de Autoridades
FRBR	Requisitos Funcionais para Registros Bibliográficos
GT	Grupo de Trabalho
ISBD	<i>International Standard Bibliographic Description</i>
ISKO	<i>International Society for Knowledge Organization</i>
KO	<i>Knowledge Organization</i>
KOS	<i>Knowledge Organization System</i>
LCSH	<i>Library of Congress Subject Headings</i>
LeXML	Rede de Informação Legislativa e Jurídica
MARC	<i>Machine Readable Cataloging</i>

OC	Organização do Conhecimento
PCF	Plano de Classificação Funcional
RDA	<i>Resource Description and Access</i>
RDF	Resource Description Framework
RVBI	Rede Virtual de Bibliotecas – Congresso Nacional
Selid	Seção de Linguagens Documentárias
SKOS-Core	<i>Simple Knowledge Organization System</i>
SOC	Sistema de Organização do Conhecimento
SRI	Sistema de Recuperação da Informação
TECAD	Tesouro da Câmara dos Deputados
TGT	Teoria Geral da Terminologia
TI Controle	Comunidade de Tecnologia da Informação Aplicada ao Controle
VCGE	Vocabulário Controlado do Governo Eletrônico
W3C	World Wide Web Consortium
XML	<i>eXtensible Markup Language</i>

SUMÁRIO

1. INTRODUÇÃO.....	15
1.1 PROBLEMA	16
1.2 OBJETIVOS	17
1.2.1 Objetivo geral	17
1.2.2 Objetivos específicos.....	17
1.3 JUSTIFICATIVA	18
1.4 ESTRUTURA DA MONOGRAFIA	20
2. ORGANIZAÇÃO DO CONHECIMENTO	22
2.1 SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO (SOC).....	24
3. ORGANIZAÇÃO E REPRESENTAÇÃO TEMÁTICA DA INFORMAÇÃO	35
3.1 TEORIAS DO CONCEITO	36
3.2 TEORIAS DA CLASSIFICAÇÃO	40
3.3 REPRESENTAÇÃO DO CONHECIMENTO: INDEXAÇÃO E CLASSIFICAÇÃO	43
4. RECUPERAÇÃO E PADRONIZAÇÃO DA DESCRIÇÃO DA INFORMAÇÃO	49
4.1 METADADOS.....	50
4.2 PADRÕES DE METADADOS	54
4.3 PADRÃO DUBLIN CORE DE METADADOS	56
4.4 QUALIDADE DOS METADADOS	58
5. METODOLOGIA.....	61
6. OS SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO (SOC) DA CÂMARA DOS DEPUTADOS	62
7. CONCLUSÕES.....	77

REFERÊNCIAS.....	80
8. ANEXO A.....	89
9. ANEXO B.....	99

1. Introdução

Desde 2006, atuo no Grupo de Trabalho (GT) formado para desenvolvimento de um tesauro para a Câmara dos Deputados. A partir de 2009, assumi a coordenação desse trabalho que logo deixou de existir na forma de GT para se constituir na Seção de Linguagens Documentárias (Selid) cujo objetivo principal era o de “dotar a Câmara dos Deputados de um tesauro ou vocabulário sistêmico, considerando os pressupostos constitucionais, institucionais e organizacionais do Poder Legislativo, a ser utilizado como instrumento terminológico na gestão corporativa da informação” (MANUAL, 2013). Além de um tesauro, essa seção passou a ser responsável por outros instrumentos terminológicos, vocabulários controlados e linguagens documentárias a serem utilizadas na representação temática dos diversos conteúdos informacionais dos sistemas de informação da Câmara dos Deputados.

Apesar de existir desde o início da década de 2000, o GT passou por diversas mudanças no que diz respeito aos seus objetivos e, principalmente, à sua composição. Essas mudanças foram decorrentes não só de limitações administrativas, típicas de órgãos públicos, mas também de diferentes visões técnicas, metodológicas e gerenciais que foram se fazendo necessárias e foram sendo assimiladas pela coordenação e pela equipe responsável pelo projeto.

Dentre algumas dessas limitações, que sobrevive até os dias atuais, está a ausência de uma ferramenta tecnológica sistêmica para a gestão do tesauro, dos demais vocabulários controlados e das linguagens documentárias, que atenda não só às necessidades desses instrumentos terminológicos, mas que seja capaz de acompanhar todo fluxo do estudo dos termos, desde a sugestão até a homologação do termo candidato. Quanto às mudanças, tal como atesta o objetivo principal reformulado ainda no âmbito do GT no ano de 2008, o projeto passou a incluir como objeto de tratamento todos os tipos de informação e documentação, ou seja, administrativa, técnica, legislativa ou política; arquivística, bibliográfica ou jornalística; corrente ou histórica; impressa, digital ou multimídia, dentre outras. Além disso, passou a visar o ambiente *web*, a ser disponibilizado no Portal da Câmara dos Deputados para diversos tipos de usuários, desde aqueles mais especializados formados por parlamentares, consultores legislativos e pesquisadores; até a

sociedade em geral, incluindo o cidadão e grupos organizados oriundos de movimentos sociais ou partidos políticos.

Para atender a todas essas novas demandas, o Tesouro da Câmara dos Deputados (TECAD), bem como os vocabulários controlados e linguagens documentárias a ele relacionadas, precisam se integrar aos mecanismos de indexação e busca nos sistemas de informação da Câmara dos Deputados. Para tanto, é necessário que o desenho de uma arquitetura de informação, que se destinará a otimizar a gestão da informação e do conhecimento na Casa, aceite o desafio de incluir a modelagem de metadados, principalmente no tocante à representação temática tendo em vista o tratamento e, portanto, recuperação dos conteúdos informacionais. Sendo assim, esse trabalho pretende contribuir, se não para uma proposta definitiva, mas para uma reflexão a respeito da necessidade, importância e urgência da modelagem de metadados para a representação temática nos sistemas de informação da Câmara dos Deputados. Nesse sentido, esse trabalho pretende dar sua contribuição, principalmente à luz da experiência que obteve por meio do privilégio de coordenar a Seção de Linguagens Documentárias (Selid) da Coordenação de Estudos Legislativos (Celeg) do Centro de Documentação e Informação (CEDI), principalmente no que diz respeito ao desenvolvimento e manutenção do TECAD.

1.1 Problema

A fim de realizar o levantamento de requisitos e a modelagem de metadados que atendessem às demandas de organização dos sistemas de informação da Câmara dos Deputados, seria necessário identificar e caracterizar esses sistemas, as tipologias documentais e o perfil dos usuários. Porém, eles serão identificados apenas de forma genérica e sem maiores detalhes no intuito de contextualizar e subsidiar a análise.

Essa análise será do ponto de vista dos Sistemas de Organização do Conhecimento (SOC), tais como o tesouro e os demais vocabulários controlados e linguagens documentárias, que serão melhor conceituados, caracterizados e exemplificados na revisão de literatura. Do mesmo modo, a análise se restringirá à

representação temática dos conteúdos informacionais, que incluem os processos de indexação e classificação, ambos diretamente relacionados ao desenvolvimento e uso dos SOC. A representação descritiva, muito embora seja outra etapa da catalogação a qual também fazem parte a indexação e a classificação, ficará de fora desse estudo haja vista que têm menor relação com os SOC.

O problema que ora se coloca é, portanto, o de analisar a contribuição dos Sistemas de Organização do Conhecimento (SOC) na modelagem de metadados para a representação temática dos conteúdos informacionais nos sistemas de informação da Câmara dos Deputados. Ou, por meio da seguinte questão posta: Como os Sistemas de Organização do Conhecimento (SOC), a exemplo do tesouro, dos demais vocabulários controlados e linguagens documentárias, podem contribuir para a modelagem de metadados para a representação temática de conteúdos informacionais nos sistemas de informação da Câmara dos Deputados?

1.2 Objetivos

1.2.1 Objetivo geral

Tendo em vista o problema de pesquisa acima, esse estudo tem o objetivo geral de analisar a contribuição dos Sistemas de Organização do Conhecimento (SOC) na modelagem de metadados para a representação temática dos conteúdos informacionais nos sistemas de informação da Câmara dos Deputados.

1.2.2 Objetivos específicos

São objetivos específicos desse trabalho:

- Analisar os Sistemas de Organização do Conhecimento (SOC) utilizados ou necessários para a representação temática de conteúdos informacionais nos sistemas de informação da Câmara dos Deputados;

- Identificar os sistemas de informação e as tipologias documentais da Câmara dos Deputados;

- Analisar a modelagem de metadados aplicada à representação temática dos conteúdos informacionais dos sistemas de informação da Câmara dos Deputados.

Nesse trabalho, as análises consistem no levantamento, conceituação e caracterização dos Sistemas de Organização do Conhecimento (SOC) e dos metadados à luz das teorias, metodologias, padrões e normas existentes. A modelagem de metadados analisada, no entanto, não pretende ser uma prescrição, mas uma sugestão a ser aplicada às demandas dos sistemas de informação, das tipologias documentais e dos segmentos de usuários da Câmara dos Deputados identificados apenas de forma genérica nesse trabalho.

1.3 Justificativa

À exceção de boa parte dos sistemas de recuperação gerenciados e alimentados pela biblioteca, os demais sistemas de informação da Câmara dos Deputados ainda não possuem campos estruturados, tampouco definição de metadados, principalmente no que tange à representação temática de seus diferentes tipos documentais. Isso por si só, compromete o tratamento e, conseqüentemente, a recuperação de toda e qualquer informação especializada. De outro lado, abandona alguns tipos de conteúdo à sua própria sorte, uma vez que recebem pouco ou nenhum tipo de tratamento. Como exemplos, pode-se mencionar a informação legislativa composta pelas proposições legislativas e a legislação, que são produtos das atividades de uma das funções principais da Câmara dos Deputados que é legislar; assim como a informação jornalística que tem o papel de informar o cidadão e imprimir maior transparência a todas as atividades dos parlamentares no exercício daquela e das demais funções legislativas. Dentre outras, essas são informações que carecem de campos estruturados ou de um tratamento adequado que atendam às demandas específicas de representação e, portanto, de recuperação temática.

A falta de estruturação e a indefinição de metadados para a representação temática dos conteúdos informacionais comprometem a interoperabilidade dos sistemas e a visibilidade da informação. Não bastasse isso, comprometem o desenvolvimento e uso do próprio tesouro, dos vocabulários controlados e das linguagens documentárias envolvidos na representação temática dos conteúdos informacionais haja vista que os Sistemas de Organização do Conhecimento (SOC) devem prover não apenas o léxico, mas também a sintaxe para que a gramática de suas terminologias possa estabelecer uma comunicação eficaz e de qualidade entre os usuários e os sistemas de informação. Principalmente em se tratando do ambiente *web* da internet tal qual o Portal da Câmara dos Deputados que visa reunir e disponibilizar nessa rede mundial as informações e notícias sobre o processo legislativo, legislação e orçamento da União.

Entretanto, atualmente ainda não há um setor na Câmara dos Deputados que centralize ou que seja responsável em executar e refletir sobre essas questões de estruturação de campos ou definição de metadados, principalmente no que diz respeito à representação temática, pois o senso comum que prevalece na Casa é o de que o tesouro, os vocabulários controlados e as linguagens documentárias irão por si só resolver os problemas de tratamento e visibilidade da informação, assim como de interoperabilidade entre os sistemas. Alguns profissionais individual ou setorialmente têm tido essa preocupação, mas a responsabilidade acaba por recair sobre os clientes ou os desenvolvedores dos sistemas que, apesar de toda a boa vontade, muitas vezes não têm competências técnicas de avaliar e decidir sobre essas questões, ainda mais de forma isolada.

As questões acima levantadas, que justificam a necessidade e importância desse trabalho, são alguns dos fatores impeditivos para uma boa gestão da informação e do conhecimento e para o desenho de uma arquitetura e organização da informação voltadas especificamente para atender às necessidades do Poder Legislativo. Aliás, essa é uma das limitações dessa pesquisa, pois faltam padrões e normas de metadados que atendam a toda essa complexidade da informação produzida e gerenciada pelo Legislativo, apesar de existirem algumas iniciativas e esforços nesse sentido, inclusive no Brasil, a exemplo da Rede de Informação Legislativa e Jurídica (LexML).

Diante disso, essa pesquisa pretende contribuir para alguns dos aspectos mais relevantes para a gestão da informação e do conhecimento e da arquitetura e organização da informação no âmbito da Câmara dos Deputados, que é a organização e representação temática dos conteúdos informacionais. Para isso, pretende reunir elementos para que os tesauros, os vocabulários controlados e as linguagens documentárias possam não só atender às demandas informacionais específicas, mas que possam se desenvolver de modo a integrar os sistemas de informação da Casa. Apesar disso, essa pesquisa não se aterá a nenhum sistema de informação em particular, nem a nenhum tipo documental ou segmento específico de usuários, pois a intenção é dar o devido distanciamento para melhor analisar o objeto, assim como trazer à tona e ressaltar as experiências práticas, teóricas, metodológicas e gerenciais acumuladas na coordenação do Tesouro da Câmara dos Deputados (TECAD).

1.4 Estrutura da monografia

Essa monografia é composta de sete capítulos, distribuídos da seguinte forma:

- Capítulo 1: introdução;
- Capítulo 2: referencial teórico sobre os Sistemas de Organização do Conhecimento (SOC), conceituação e caracterização;
- Capítulo 3: referencial teórico das teorias que influenciam o uso e desenvolvimento dos Sistemas de Organização do Conhecimento (SOC), tais como teoria do conceito, classificação, indexação e outros aspectos relacionados à organização e à representação temática;
- Capítulo 4: referencial teórico com padrões, perfil e esquema de metadados para a representação temática dos conteúdos informacionais nos sistemas de informação a partir dos tesauros,
- Capítulo 5: metodologia utilizada no levantamento e análise dos dados;

- Capítulo 6: características dos vocabulários controlados e das linguagens documentárias que compõem os Sistemas de Organização do Conhecimento (SOC) da Câmara dos Deputados e aplicação de perfil de metadados para representação temática dos conteúdos informacionais nos sistemas de informação da Casa;
- Capítulo final: traz as conclusões.

As páginas finais trazem as referências bibliográficas das fontes de informação utilizadas na fundamentação teórica desse trabalho.

O capítulo 2 aborda os Sistemas de Organização do Conhecimento (SOC), sua conceituação, tipologia e caracterização.

2. Organização do Conhecimento

Para Dahlberg (1993, p. 211), a origem da Organização do Conhecimento (OC) reside nas preocupações de Paul Otlet e Henri La Fontaine, no início do século XX, às quais foram se juntando documentalistas e cientistas da informação e, “desde a obra de Wüster², os terminólogos se envolveram também com o assunto.”

Para Fujita (2008) e Martínéz Tamayo e Valdez (2008), apesar das origens remotas na Teoria do Conhecimento, desde a Antiguidade, o reconhecimento da identidade e da conformação do campo científico denominado *Knowledge Organization* (KO) ganharam destaque e fortalecimento com a criação e início dos trabalhos da *International Society for Knowledge Organization* (ISKO), a partir de 1993.

No entanto, Barité (2000) afirma que a organização do conhecimento é uma disciplina de formação recente, que estuda as leis, os princípios e os procedimentos pelos quais se estruturam o conhecimento especializado em qualquer disciplina. O objeto de estudo da organização do conhecimento é o conhecimento socializado e registrado, e, no que diz respeito à Biblioteconomia e Documentação, abrange o desenvolvimento teórico e prático para a construção, gestão e o uso de classificações científicas, taxonomias, nomenclaturas e linguagens documentais.

Para Dahlberg (1993; 2006), organização do conhecimento é a ciência que ordena, estrutura e organiza sistematicamente as unidades de conhecimento (conceitos) de acordo com os seus elementos de conhecimento inerentes (características) e a aplicação desses conceitos e classes de conceitos ordenados para melhor identificar e representar os conteúdos de objetos ou indivíduos. A partir

² Eugène Wüster (1898-1977) foi um austríaco que, nos anos 1930, realizou estudos científicos na área da terminologia com foco nos termos em suas áreas de especialidade para uma comunicação precisa e sem ambiguidade entre os cientistas e, com isso, estabeleceu as bases da futura Teoria Geral da Terminologia (TGT).

dessa organização do conhecimento criam-se ferramentas que apresentam a interpretação organizada e estruturada do objeto, chamados de Sistemas de Organização do Conhecimento (SOC).

Enquanto ciência, a organização do conhecimento se ocupa dos seguintes aspectos: dos fundamentos teóricos e aplicações práticas do desenho e construção de Sistemas de Organização do Conhecimento (SOC), tanto gerais quanto especializados, assim como das técnicas de organização de conhecimento, incluindo aquelas que usam a linguagem natural; da influência de outras disciplinas na organização do conhecimento, tais como a linguística e a terminologia; assim como da influência da organização do conhecimento em outras disciplinas e, por fim, da pesquisa, educação, treinamento, normalização, questões éticas e legais, gestão e tecnologia referentes à organização do conhecimento, seus especialistas e organizações (DAHLBERG, 2006).

Dahlberg (2006) traça uma analogia entre a estrutura da matéria e o grau de complexidade de representação do conhecimento, aí incluídos os elementos do conhecimento (características de conceitos), unidades de conhecimentos (os próprios conceitos), unidades mais amplas de conhecimento (combinação de conceitos) e sistemas de conhecimento (entidades compostas de unidades de conhecimento arranjadas numa estrutura coesiva, adequadamente planejada). Esses diferentes níveis do conceito têm, portanto, impacto na organização e representação do conhecimento.

Apoiadas nessa visão de conceito de Dahlberg, Brascher e Café (2008) afirmam que organização do conhecimento é o processo de modelagem do conhecimento que visa a construção de representações do conhecimento. Esse processo tem por base a análise do conceito e de suas características para o estabelecimento da posição que cada conceito ocupa num determinado domínio, bem como das suas relações com os demais conceitos que compõem esse sistema nocional. Brascher e Café (2008) descrevem a organização do conhecimento também como um processo cognitivo, pois a representação construída não se restringe ao conhecimento expresso por um autor, mas é fruto de um processo de

análise de domínio que procura refletir uma visão consensual sobre a realidade que se pretende representar.

2.1 Sistemas de Organização do Conhecimento (SOC)

O termo Sistema de Organização do Conhecimento (SOC) é uma tradução para o português do original inglês *Knowledge Organization System* (KOS). O termo foi proposto pelo *Networked Knowledge Organization Systems Working Group* na primeira Conferência da *ACM Digital Libraries* em 1998, Pittsburgh, Pennsylvania. Assim como a sigla KOS, utilizada com frequência na literatura, adotou-se o correspondente SOC em português. Os SOC ou esquemas de representação do conhecimento, como alguns autores preferem denominar, são encontrados na literatura das áreas de Ciência da Informação, Biblioteconomia e Documentação para designar instrumentos que fazem a tradução dos conteúdos dos documentos originais e completos para esquemas estruturados sistematicamente, que representam esse conteúdo, com a finalidade principal de organizar a informação e o conhecimento e, conseqüentemente, facilitar a recuperação das informações contidas nos documentos (CARLAN, 2010).

Para Hodge (2000), os Sistemas de Organização do Conhecimento (SOC) englobam todos os tipos de instrumentos usados para organizar a informação e promover o gerenciamento do conhecimento e constituem o “coração” dos Sistemas de Recuperação da Informação (SRI) das bibliotecas, museus e arquivos, no ambiente físico, e, principalmente, no ambiente web.

Hjorland (2007 *apud* CARLAN, 2010) refere-se aos Sistemas de Organização do Conhecimento (SOC) como ferramentas que apresentam a interpretação organizada de estruturas do conhecimento, também chamadas de ferramentas semânticas. Os SOC são estruturas sistemáticas que visam a construção de modelos abstratos do mundo real, representando os conceitos de um domínio. Essas ferramentas semânticas são utilizadas para o tratamento da informação viabilizando a recuperação da informação, tanto no ambiente informatizado como no tradicional.

Segundo Boccato (2011), os sistemas de recuperação da informação de bases de dados eletrônicas, catálogos coletivos informatizados, bancos de dados, repositórios, portais, gerenciadores de publicações eletrônicas, bem como as páginas na web requerem Sistemas de Organização do Conhecimento (SOC) que permitem a representação temática de recursos informacionais nos meios impresso e eletrônico, armazenados em acervos locais de bibliotecas tradicionais ou em bibliotecas digitais ou virtuais disponíveis na internet. Tais SOC são instrumentos que possibilitam a organização de estruturas semânticas de conhecimento e são compostos pelos sistemas de classificação, listas de cabeçalhos de assunto, taxonomias, tesouros, ontologias, entre outras linguagens documentárias já bastante conhecidas e utilizadas pelos bibliotecários em sua prática profissional.

Segundo Lara (2001), para que a linguagem documentária atue como um modo de organização e desempenhe o papel de instrumento de comunicação é preciso reunir algumas qualidades, tais como: a) funcionar como código inteligível e fonte para a interpretação do sentido e; b) caracterizar-se como metalinguagem e incorporar o usuário como integrante do processo. Essas características dependem do rigor metodológico utilizado na sua construção, principalmente quanto à normalização semântica. "Para funcionar como metalinguagem e integrar o usuário como participante do processo, a linguagem documentária deve utilizar referências de linguagem e de significado que sejam razoavelmente compartilhadas", em referência às garantias literária e de usuário prescritas para o desenvolvimento desses tipos de sistemas (LARA, 2001, p.234).

Moreiro González (2011, p. 41) aponta diversos fins para os Sistemas de Organização do Conhecimento (SOC), que são:

- permitir categorizar recursos com esquemas de organização determinados, para facilitar sua posterior recuperação,
- permitir ao usuário interpretar estruturas de conhecimento de maneira organizada,
- contribuir para melhorar a eficiência dos serviços de informação,
- buscar a terminologia específica das instituições,
- muitas vezes, provêm do trabalho cooperativo em âmbito empresarial.

Hodge (2000) aponta algumas características comuns dos Sistemas de Organização do Conhecimento (SOC) usadas em organização de bibliotecas digitais:

- cada SOC impõe uma visão particular do mundo, de uma coleção e de seus itens;
- a mesma entidade pode ser caracterizada de diferentes maneiras, dependendo do SOC que é usado;
- deve haver identificação suficiente entre o conceito expresso no SOC e o objeto do mundo real, ao qual aquele conceito se refere, pois quando uma pessoa procura algo sobre determinado objeto, o SOC deve ser capaz de conectar o conceito do objeto com sua representação no sistema.

Porém, os estudos de Hodge (2000) e do *Networked Knowledge Organization Systems/Services* (2010) apontam que os Sistemas de Organização do Conhecimento (SOC) possuem características distintas para determinados contextos de aplicação. A partir da classificação de Hodge (2000), Boccato (2011) considera que as normas ANSI/NISO Z39.19-2005, BS 8723 e a ISO 25.964 demonstram que existem atuações e tratamentos diferenciados para os diversos sistemas de organização do conhecimento, tal como aqueles reunidos e sistematizados a seguir:

1. Classificação e Categorização:

a) Anéis de Sinônimos: conjunto de termos considerados equivalentes para fins de recuperação. Eles são aplicáveis a interfaces de sistemas automatizados de informação no acesso aos conteúdos informacionais que utilizam a linguagem natural como instrumento de recuperação.

b) Sistemas de Classificação: linguagem documentária hierárquica, pré-coordenada, formada por signos artificiais que apresentam uma relação de superordenação e subordinação entre si, em que cada conceito é superior ao conceito seguinte, apresentando-se, dessa maneira, numa estrutura normativa.

(BOCCATO, 2008, p. 273). Sua aplicação dá-se, normalmente, para fins de indicar sua localização nas estantes, nos catálogos e nos próprios sistemas de classificação.

Segundo Chowdhury e Chowdhury (2007), há classificações dos seguintes tipos, seguidas de alguns exemplos:

- Classificações enumerativas: todas as classes e assuntos possíveis, desde os mais simples até os mais complexos, são enumerados. O esquema das classes superordenadas, coordenadas e subordinadas é apresentado em forma de notação enumerativa e hierárquica e, por isso, pode ser facilmente consultado e localizado pelos usuários, assim como permite fazer um mapa mental do assunto. Por ser uma classificação pré-definida, no entanto, é também mais rígida e limitada às classes enumeradas. Um exemplo é a Classificação da Biblioteca do Congresso Americano (LCC).
- Classificações analítico-sintéticas: o assunto de um dado documento será dividido entre seus elementos constituintes. O esquema de classificação será usado para localizar notações para cada um desses elementos que, por sua vez, devem ser reunidos, em observância a determinadas regras, para formar o número final. Proveem várias tabelas auxiliares de assuntos e regras que evitam a necessidade de uma longa lista de classes. Essas características dão maior flexibilidade, mas também tornam o trabalho dos classificadores mais complexo uma vez que eles precisam construir números em alguns assuntos específicos e não apenas selecionar de uma lista. Tanto a Classificação Decimal de Dewey (CDD) quanto a Classificação Decimal Universal (CDU) são tipicamente classificações analítico-sintética.
- Classificações facetadas: lista as várias facetas de cada assunto ou a classe principal e provê um conjunto de regras para a construção de números de classe. No entanto, qualquer componente ou faceta de um assunto pode se encaixar em uma ou mais das cinco categorias

fundamentais delineadas por Ranganathan³. O classificador precisa, então, descobrir as várias facetas do assunto e construir, ele mesmo, um número de classificação, o que lhe permite representar de forma flexível tanto assuntos simples, quanto complexos. Um exemplo é a Classificação de Colon (CC), desenvolvido por Ranganathan, sistema este utilizado em algumas bibliotecas da Índia, do Sudão e do Reino Unido. As ideias ranganathianas contribuíram significativamente no desenvolvimento de tesouros facetados, a exemplo do *Thesaurofacet*, publicado em 1965.

c) Sistemas de classificação de negócios para gerenciamento de registros: conjunto de termos organizados hierarquicamente, a partir da superordenação e subordinação existente entre si, para tratar, classificar e controlar informações geradas por uma corporação para sua recuperação, manutenção, eliminação de registros, entre outras informações corporativas. Existem vários tipos de sistemas de classificação: numérico, cronológico, alfabético (assuntos gerais), tipo do negócio ou atividade.

d) Listas de cabeçalhos de assunto: linguagens pré-coordenadas, controladas, de estruturas associativa ou combinatória de palavras ou expressões de diversas áreas do conhecimento, cuja finalidade é a de representar para recuperar a informação.

Segundo explicam em maiores detalhes Chowdury e Chowdury (2007), a lista de cabeçalho de assuntos é uma lista alfabética, de termos e expressões com referências cruzadas e notas, que podem ser usados para representar os assuntos de um conteúdo informacional. Apesar de serem arranjados alfabeticamente, os termos ou expressões são semanticamente relacionados. Essas listas foram desenvolvidas para complementar a classificação bibliográfica e também podem ser usadas na busca ou navegação em uma coleção. Para os autores, o exemplo mais proeminente é a lista de Cabeçalho de Assuntos da Biblioteca do Congresso

³ Shialy Rammarita Ranganathan (1897-1972), foi um matemático e bibliotecário indiano, considerado o pai da teoria da classificação facetada, que será melhor detalhada no capítulo seguinte dessa monografia.

Americano, denominada na língua original de *Library of Congress Subject Headings* (LCSH), que se baseia nos seguintes princípios fundamentais: na uniformização, unicidade conceitual, especificidade e consistência quanto à forma e estrutura dos cabeçalhos, termos e expressões, principalmente aqueles similares, além de atender às necessidades dos usuários, orientar-se pela a garantia literária e manter a estabilidade nas alterações que são realizadas estritamente quando necessárias e não sem antes medir o seu impacto nos sistemas de informação.

e) Taxonomias: vocabulários controlados formados por termos preferenciais, organizados em estruturas hierárquicas ou poli hierárquicas.

Currás (2010, p. 55 a 78) reúne diversos conceitos de taxonomia com diferentes propósitos, todos importantes para a sua compreensão, quais sejam:

- um campo da ciência que classifica a vida, isto é, toda ideia e atividade humana,
- uma linguagem controlada, ou seja, uma lista organizada de palavras e frases, ou sistemas de notação que se usam para iniciar um processo de indexação e recuperação de informação,
- um esquema de navegação ordenado hierarquicamente,
- trata de estruturas conceituais adequadas para seu uso na web semântica,
- organiza de forma ordenada documentos de temática variada, criando níveis taxonômios apropriados para o processamento informático,
- cria ordem no caos de unidades conceituais, advindas de documentos variados, e, por fim,
- uma aplicação web para gestão de linguagens documentárias orientada especificamente para o desenvolvimento de tesouros hierárquicos que pode ser usado no desenvolvimento de estruturas de navegação da web ou como complemento articulador num gestor de conteúdos.

Currás (2010, p.68) ainda dá destaque para taxonomia corporativa como eixo principal de uma arquitetura de informação empresarial cujo foco é a gestão do conhecimento, sua organização e suas peculiaridades. Com isso concordam Campos e Gomes (2008), para quem as taxonomias são estruturas classificatórias que servem de instrumento para a organização e recuperação de informação em empresas e instituições. Estão sendo vistas como meios de acesso atuando como mapas conceituais dos tópicos explorados em um serviço de recuperação. O desenvolvimento de taxonomias para o negócio da empresa tem sido um dos pilares da gestão da informação e do conhecimento.

Bravo (2011) afirma que as taxonomias são hierarquias semânticas que priorizam a navegação e, para tanto, sua aplicação em meio digital. Quanto à estrutura, podem classificar uma área temática ou o conteúdo de um sítio na web com o fim de organizá-los a partir de determinadas propriedades ou características. Para isso, requer-se uma análise conceitual que diferencie as propriedades e características dessa área temática ou sítio na web, assim como de suas áreas menores. Entretanto, em sua concepção clássica, a taxonomia aplica um critério monohierárquico no estabelecimento do sistema de classificação, ou seja, cada uma das classes que a compõem só pode ocupar um único lugar na estrutura hierárquica.

Na avaliação de Bravo (2011), a maior parte das definições e abordagens sobre taxonomia restringe sua aplicação aos contextos digitais e, mais especificamente, ao desenvolvimento de sítios na web. Nesse caso, consideram que são sistemas para organizar o conteúdo em sítios na web, intranets ou portais com o fim de desenvolver sistemas de busca, facilitar a navegação e exploração, assim como o descobrimento de recursos de informação. As taxonomias, entretanto, mais que nos conteúdos, se focalizam nos usuários. O conteúdo em um sítio na web de uma organização é bastante complexo, pois pode ser classificado pelas funções, pelos produtos ou serviços, pelos departamentos, pelos lugares, pelas pessoas, dentre outras formas. Assim mesmo, as taxonomias se caracterizam por sua estrutura dinâmica e flexível.

2. Modelos de Relacionamentos:

a) Ontologias: Moreira, Alvarenga e Oliveira (2004) reúnem algumas das definições mais comuns sobre ontologia, abaixo:

- é um sistema conceitual informal a partir de uma base de conhecimento concreta,
- é uma representação de um sistema conceitual que utiliza uma teoria lógica,
- é um vocabulário utilizado por uma teoria lógica,
- é uma especificação de uma conceitualização.

Apesar disso, ainda é a definição de Gruber (1993, p. 193) que é a mais disseminada entre os especialistas até os dias de hoje. Para o autor, uma ontologia é uma especificação explícita de uma conceitualização que, por sua vez, é composta por objetos, conceitos e outras entidades que existem em uma determinada área e as relações que ocorrem entre elas. Desse modo, permitem organizar, classificar, recuperar, compartilhar e reutilizar informações em ambientes digitais. Tal qual os tesauros, as ontologias também são constituídas por um conjunto de termos estruturados hierarquicamente, formando uma estrutura de classes e subclasses.

Chowdhury e Chowdhury (2007) afirmam que, no contexto da web, ontologias proveem um conhecimento compartilhado de um domínio que é necessário para compreender diferenças e conotações dos termos, assim como para facilitar a interoperabilidade no processamento de dados pelos computadores. Com isso, concorda Bravo (2011), para quem, dadas as características das ontologias, elas se constituem na ferramenta fundamental para o desenvolvimento da web semântica, dado que permitem capturar o conhecimento de um modo genérico e formal de maneira que possa ser compartilhado e reutilizado.

b) Redes semânticas: conjunto de termos que representam conceitos, modelados a partir de uma notação gráfica composta por nós interconectados. As redes semânticas podem ser utilizadas para a representação do conhecimento ou como ferramenta de apoio para sistemas automatizados de inferências sobre o conhecimento.

c) Tesouros: sistemas de organização do conhecimento que permitem a organização de termos descritores a partir de relações lógico-semânticas de ordens hierárquicas, equivalentes e associativas.

Toutain (2006, p. 22) define o termo tesouro como “vocabulário controlado, compreendido como estruturas terminológicas, que visa padronizar a linguagem, em serviços de informação, cobrindo um domínio específico do conhecimento, traduzido de uma linguagem natural para uma linguagem de máquina”.

Os tesouros são formados por estruturas conceituais complexas de relações semânticas que permitem ao usuário acessar e associar conceitos a partir de descritores que possam representar sua necessidade de informação. As estruturas caracterizadas nos tesouros só permitem a recuperação de documentos que possuem tais relações descritas em seu contexto inicial de representação. Ressalta-se que o incremento das relações de equivalência e associativas no repertório terminológico possibilita maior aproximação entre o tesouro e a linguagem de busca do usuário, condição essa de suma importância para o adequado desempenho do instrumento terminológico em sistemas automatizados de informação (BOCCATO, 2011).

Dentre as principais funções de um tesouro, segundo Lancaster (2004), destacam-se o controle de sinônimos e quase sinônimos e a distinção de homógrafos. O tesouro possibilita, ainda, a facilidade na condução da busca por intermédio dos termos relacionados e das referências cruzadas de maneira a permitir uma consistência melhor da indexação e do transporte da linguagem de busca para a linguagem de indexação. Dessa maneira, o tesouro pode reduzir o tempo e aumentar a eficiência nas atividades de indexação e recuperação da informação.

Na internet presencia-se a evolução dos tesouros a partir do uso de linguagens de marcação, como o XML (*eXtensible Markup Language*), na construção de tesouros navegáveis, permitindo a migração de tesouros tradicionais a formatos inteligíveis por máquinas, que propicia o intercâmbio e a interoperabilidade entre sistemas de organização do conhecimento. O SKOS-Core

(*Simple Knowledge Organization System*), iniciativa do World Wide Web Consortium (W3C) é um esquema suportado no RDF (*Resource Description Framework*) que atende tais finalidades a partir do desenvolvimento de especificações e normas que proporcionam um modelo para representar a estrutura conceitual básica e o conteúdo de sistemas de organização do conhecimento, tais como os tesouros, sistemas de classificação, listas de cabeçalhos de assunto, taxonomias e outros vocabulários controlados similares na *Web Semântica* (GRUPO DE TRABAJO DEL W3C, 2010).

Além dos tipos descritos acima, os estudos de Hodge (2000) e Moreiro González (2011, p. 42) apontam para as listas de termos, tais como:

- a) **Lista de autoridades:** lista de termos usada para controlar a variedade de nomes para entidades. Em geral, esse tipo de Sistema de Organização do Conhecimento (SOC) não inclui uma organização nem uma estrutura complexa, a apresentação pode ser alfabética ou organizada por um esquema de classificação superficial.
- b) **Glossários:** lista de termos, em geral, com definições, relativos a um assunto específico ou um trabalho particular.
- c) **Dicionários:** listas de palavras em ordem alfabética e suas definições. Seu escopo é mais geral que os glossários. Podem prover informações sobre a origem de uma palavra, variações quanto à morfologia e a escrita da palavra, bem como, os múltiplos significados dentre as disciplinas.
- d) **Índices toponímicos:** dicionários de termos relativos a lugares geográficos.

Boccatto (2011) ressalta a presença das folksonomias como um sistema social de organização do conhecimento aplicável também no contexto da internet. Moreiro González (2011) acrescenta que é um sistema de indexação colaborativo e social, mediante a utilização de rótulos simples ou palavras-chave livres sem quaisquer hierarquias ou relações entre elas, sendo algumas de suas principais funções o fato dos usuários poderem ordenar sua informação para que, posteriormente, possam recuperá-la ou compartilhá-la, assim como agilizar a representação da informação por usuários leigos que são inexperientes quanto ao uso de complexos Sistemas de

Organização do Conhecimento (SOC). No entanto, entre as suas maiores desvantagens, citam-se a carência de controle das palavras-chaves, principalmente dos sinônimos, a ausência de hierarquias ou associações entre elas e, principalmente, a ambiguidade na representação da informação e conteúdos informacionais.

Não obstante a diversidade de tipos, as linguagens documentárias e, analogamente, boa parte dos demais Sistemas de Organização do Conhecimento (SOC), possuem ao menos três elementos básicos, segundo Gardin et al. (*apud* Cintra, 2002, p. 35-36):

- um léxico, identificado como uma lista de elementos descritores, devidamente filtrados e depurados;
- uma rede paradigmática para traduzir certas relações essenciais e, geralmente estáveis, entre descritores; e
- uma rede sintagmática destinada a expressar as relações contingentes entre os descritores, relações que são válidas no contexto particular onde aparecem.

Nesse sentido, os Sistemas de Organização do Conhecimento (SOC) podem possibilitar a comunicação sistema-usuário, pois funcionam como instrumentos auxiliares no processo da representação da informação e, por essa razão, são instrumentos essenciais na atividade de indexar e recuperar documentos (OLIVEIRA, 2012).

Esse capítulo aborda as teorias que influenciam o uso e desenvolvimento dos Sistemas de Organização do Conhecimento (SOC), tais como as teorias do conceito e da classificação, a indexação e outros aspectos relacionados à organização e à representação temática da informação.

3. Organização e representação temática da informação

Campos (2004) analisa as principais teorias que fundamentam a organização da informação e a representação do conhecimento nas áreas de ciência da informação, ciência da computação e da terminologia. Da ciência da informação, a autora analisa as teorias ligadas à representação de sistemas de conceitos, como a teoria da classificação facetada de Ranganathan e a teoria do conceito de Dahlberg, que possibilitam a representação de domínios de conhecimento. Da ciência da computação, investiga os modelos representacionais associados à modelagem de sistemas de banco de dados, mais especificamente o modelo orientado a objetos, além de um campo mais recente que repensa as possibilidades representacionais e de organização de domínios de conhecimento que é a ontologia formal. Da teoria da terminologia, os princípios estabelecidos para a determinação de conceitos e suas relações apresentadas por Wüster e seus seguidores.

Segundo Campos (2004), os mecanismos de representação de conhecimento permitem que processos de formalização sobre os objetos e suas relações, em contextos predefinidos, possam ser facilmente representados. No âmbito da ciência da informação, possibilitam a elaboração de linguagens documentárias verbais e notacionais, visando à recuperação de informação e à organização dos conteúdos informacionais de documentos. No âmbito da ciência da computação, servem para auxiliar a implementação de estruturas computáveis. No âmbito da terminologia, esses mesmos mecanismos permitem a sistematização dos conceitos e, conseqüentemente, a elaboração de definições consistentes.

Para Francelin e Kobashi (2011), na área de organização da informação e do conhecimento, existe um conjunto variado de tendências teórico-epistemológicas e teórico-metodológicas sobre o conceito, que são, em suma, as seguintes:

a) o conceito que é abordado majoritariamente para fins operacionais para a construção ou análise de instrumentos, tais como classificações, tesouros e ontologias,

b) a Teoria do Conceito de Dahlberg, a Teoria Geral da Terminologia e a Teoria da Classificação Facetada de Ranganathan,

c) as abordagens lógicas, voltadas para a organização de sistemas de conceitos fundamentados na lógica aristotélica, mais bem representada pelas *categorias*, e pelo princípio de fixação da linguagem ou conceitos por meio de juízos e proposições verdadeiras, características básicas da filosofia analítica e do positivismo lógico,

d) as principais linhas de força teórico-metodológicas estão centradas em Dahlberg, Ranganathan e Wüster, com forte influência da lógica filosófica.

A seguir, serão detalhadas algumas dessas aplicações teóricas, metodológicas e técnicas que fundamentam a organização e representação do conhecimento, principalmente no que diz respeito ao desenvolvimento e uso dos Sistemas de Organização do Conhecimento (SOC).

3.1 Teorias do Conceito

Na Teoria Geral da Terminologia (TGT), os conceitos, segundo Wüster, são o ponto de partida para o trabalho terminológico. A terminologia, cujo objetivo é o de estabelecer delimitações claras entre os conceitos, considera que termos e conceitos podem ser independentes um do outro. Nesse sentido, a concepção de termo de Wüster caracteriza-se pela possibilidade de descrição de um conceito por meio de uma definição como passo anterior à própria relação conceito-termo. Por isso, pode-se identificar um conjunto de conceitos de um domínio especializado, organizá-los em um sistema estruturado e defini-los sem mesmo identificar com precisão os termos que os designam. Wüster afirma que a terminologia parte do conceito para buscar sua denominação, porém não descarta totalmente a ideia de denominação, por ser indispensável contar com ela ou com algum outro signo para

identificar e fixar um conceito. Nesse sentido, a TGT se interessa praticamente apenas pelos termos, dissociando o léxico da gramática, do contexto e do discurso, vendo-os como unidades que existem e têm vida independente com o devido cuidado para que não existam termos polissêmicos, sinônimos ou homônimos. A teoria do conceito de Wüster tem, portanto, função estritamente normativa e operacional e privilegia a fixação de estruturas conceituais das áreas de especialidade bem definidas (BARROS, 2004).

A definição de conceito como unidade de pensamento foi adotada, segundo Campos (2001,p.101), por Wüster e pela ISO/TC-37. Porém, Dahlberg (1978, p.5) reformula a definição, observando que o termo pensamento pode ser subjetivo e impreciso, sendo mais apropriado caracterizá-lo como unidade de conhecimento.

Essas diferentes definições sobre o conceito são adotadas, inclusive, na normalização sobre terminologia, tal como na ISO 1087 (2000, p. 2), onde o conceito é definido como “[...] unidade de conhecimento criada pela combinação única de características”; na ANSI/ISO Z39.19-2005 (2005, p.4), definido como unidade de pensamento, formada pela combinação mental de algumas ou todas as características de um objeto concreto ou abstrato, real ou imaginário, de modo que conceitos existem na mente como entidades abstratas independente dos termos usados para expressá-los e na ISO 704 (2000, p.2) que apresenta, inicialmente, os conceitos como unidades de pensamento e, posteriormente, como unidade de conhecimento.

Dahlberg (1978) apresenta o conceito como unidade mínima e o define como uma tríade: referente, características e nome. A autora afirma que os conceitos são a reunião e compilação de enunciados verdadeiros a respeito de determinado objeto, denominados elementos dos conceitos. Esses elementos são as chamadas características dos conceitos e traduzem os atributos das coisas designadas. As características dos conceitos possuem as funções de definição dos conceitos, formação dos nomes dos conceitos e ordenação classificatória dos conceitos e índices.

Os elementos dos conceitos são obtidos por meio do método analítico-sintético que consiste na coleta dos enunciados verdadeiros e na decomposição de seus atributos essenciais e necessários ou acidentais e possíveis. Cada enunciado apresenta um atributo que é uma característica do conceito que, às vezes, pode ser

tão geral que se torna uma categoria. Os conceitos podem ser categorizados formalmente em objetos, fenômenos, processos, propriedades, relações e dimensões. No entanto, vale ressaltar, que existem inúmeras possibilidades de combinações dessas categorias (DAHLBERG, 1978).

Quando dois conceitos diferentes possuem uma ou duas características em comum, então há relações do tipo semânticas entre os conceitos, que podem ser (DAHLBERG, 1978):

- a) Relações hierárquicas (ou de gênero e espécie) entre conceitos com características idênticas, mas um possui uma característica a mais do que o outro, ou ainda, entre conceitos específicos do mesmo gênero. A relação hierárquica é de dois tipos, a relação de abstração de gênero e espécie, que forma cadeias de conceitos, e a relação lateral, ou seja, que forma renques de conceitos. A relação hierárquica baseia-se, entretanto, em uma relação lógica de implicação, ou seja, nela os conceitos devem ser da mesma natureza, o que não ocorre com o todo e suas partes, que, em muitos casos, são de natureza diferentes.

Na teoria da terminologia, o que Dalhberg denomina hierárquica, Wüster classifica de relação lógica e específica muito mais os seus tipos: relação de comparação (subordinação lógica, coordenação lógica, interseção lógica e diagonal) e relação de combinação (determinação, conjunção, disjunção). Para a terminologia, esta relação é fundamental para a elaboração de definições, pois permite a compreensão do conceito, isto é, especifica o que a “coisa é”. No entanto, para a terminologia, o objetivo de evidenciar as relações entre os conceitos não é para especificar uma estrutura sistemática, mas para o entendimento do próprio conceito (CAMPOS, 2004).

A relação hierárquica é uma das relações principais em qualquer estrutura classificatória. Segundo Campos (2004), ela é a que forma a espinha dorsal de uma estrutura, pois é a partir dela que se estabelece o primeiro elemento de uma definição.

- b) Relação partitiva: existente entre o todo e sua parte, ou entre um produto e os elementos que o constituem, sem, no entanto, especificar o que é todo ou o que é a parte.

Segundo Campos (2004), no entanto, podem ser apresentadas como metacategorias informacionais partitivas:

- objeto integral/componente: o que caracteriza este tipo de partição é a possibilidade de a estrutura do componente ser separável e possuir uma funcionalidade específica;
- membro/coleção: o componente não possui nenhum papel funcional em relação ao todo, isto é, são partes do todo, mas podem ser individualizados;
- massa/porção: o todo é considerado como um agregado homogêneo, e sua parte/porção é similar e separável;
- objeto/matéria: expressa a parte/matéria na qual a coisa é feita e que não é separável do objeto, pois não tem nenhuma função separada do objeto;
- caráter/atividade: designa uma parte de uma atividade;
- área/lugar: é um tipo partitivo de relação espacial entre uma região ocupada por diferentes objetos.

c) Relação de oposição, que podem ser de contradição ou contrariedade,

d) Relações funcionais aplicam-se a conceitos que expressam processos.

Campos (2004) ressalta a importância ainda de um dado tipo de relação que não mais se constitui entre conceitos, mas entre a forma de expressar os conceitos, ou seja, dá-se no âmbito da língua, a chamada relação de equivalência.

Dahlberg (1978) alerta também sobre a importância da definição dos conceitos. A autora explica que uma definição é a explanação do sentido da palavra ou a simples descrição de um objeto, é ainda uma limitação ou uma colocação de limites, seja de um conceito ou uma ideia. As definições podem ainda ser reais ou nominais. A definição real se relaciona com o conhecimento do objeto contido em determinado conceito e a nominal procura fixar o uso de determinada palavra.

Para Campos (2004), tanto a teoria do conceito quanto a da terminologia permitem a sistematização de conhecimentos e possuem diretrizes para a elaboração de definições e, por isso, trabalham no âmbito ontológico do conhecimento. Entretanto, a teoria da terminologia utiliza o método indutivo para pensar um dado contexto. Para Wüster, os conceitos (objeto/entidade/instância) associam-se um em relação aos outros, formando um sistema de conceitos. Entretanto, o sistema é formado a partir da análise do próprio conceito e não do contexto em que ele está inserido. O contexto é visto como um *a priori* que só é possível identificar a partir do próprio conteúdo conceitual. Na terminologia, não se fala sequer em categorias ou universais. Por sua vez, a teoria do conceito introduz uma metodologia que poderia ser denominada híbrida, de modo a agregar o método dedutivo e indutivo para pensar o particular como um todo e o todo possuindo particulares. Dahlberg, em sua teoria, apresenta categorias para representar contextos e, logo depois, analisa os conceitos de um contexto na perspectiva de ordená-los no interior dessas categorias. A princípio, sabe-se da existência das categorias (universais); entretanto, chega-se a elas, a partir de um processo indutivo de análise do conceito.

3.2 Teorias da Classificação

Henry Evelelyn Bliss (1870-1955) foi o primeiro bibliotecário a se preocupar com os fundamentos filosóficos sobre classificação. Em seu sistema final de classificação, publicado primeiramente em 1935, e revisto e ampliado em 1940-1953, Bliss mostrou os diferentes aspectos de cada área também de forma diagramática e de acordo com os pontos de vista filosófico, científico, histórico, tecnológico e artístico. Apesar disso, ele geralmente rearrumava as áreas, apresentadas bidimensionalmente, na maioria das vezes em apenas um nível hierárquico, visando à brevidade e a uma arrumação mais fácil dos livros nas estantes. Sua contribuição para a teoria da classificação foi ter posto a classificação bibliográfica novamente em contato mais estreito com os princípios filosóficos da classificação, por exemplo, com os fundamentos conceituais da formação, divisão e partição de classes. Bliss também serviu de inspiração para o classificacionista indiano Ranganathan. (DAHLBERG, 1979).

No entanto, Campos (2001, p. 32) ressalta que os sistemas anteriores ao de Ranganathan só permitiam representar o conhecimento já estabelecido, pois neles ocorre a ligação entre os assuntos dos documentos e as classificações, mas não entre o conhecimento e as classificações. Como alternativa a esta concepção, Ranganathan desenvolve um trabalho que busca compatibilizar uma classificação flexível, expansiva e prática com bases científicas, sem tornar o sistema inconsistente ou inviável e que permitirá uma única classificação capaz de atender a todos os propósitos.

Em sua teoria, Ranganathan apresenta cinco modos preliminares de formação de assuntos, que são, segundo Campos (2003) e Aranalde (2009):

a) dissecação: corte de um universo de entidades em partes que resulta em classes coordenadas ou em ordem de classes,

b) laminação: construção por superposição de faceta por faceta que resulta em classes compostas,

c) desnudação: diminuição progressiva da extensão e o aumento da intenção de um assunto básico ou de uma ideia isolada que resulta em classes subordinadas ou em cadeias de classes,

d) reunião ou agregação livre: reunião de dois ou mais assuntos básicos ou compostos e de ideias isoladas,

e) superposição: ligação de duas ou mais ideias isoladas que pertencem ao mesmo universo de ideias isoladas que resulta em combinação de classes através de reunião e agregação.

Por categoria, Ranganathan compreende um conceito de alta generalidade e de larga aplicação que serve de estrutura a um esquema de classificação, pois sob este esquema se pode reunir outros conceitos. Faceta, por sua vez, é a manifestação das categorias em classes distintas. As facetas pertencem aos assuntos dos documentos, sendo o resultado da aplicação de características que formam um conjunto cujo nome é o termo mais geral. Sob uma faceta recaem

componentes subordinados, mas com características particulares (ARANALDE, 2009).

Como melhor explica Campos (2003), faceta é um termo genérico usado para denotar algum componente de um assunto composto tendo a função de formar renques, termos e números. No interior de cada faceta, dentro do domínio de conhecimento, os conceitos podem ser organizados em renques e cadeias. Renques, por sua vez, são classes formadas a partir de uma única característica de divisão, formando séries horizontais. E cadeias são séries verticais de conceitos em que cada conceito tem uma característica a mais ou a menos conforme a cadeia seja descendente ou ascendente.

Ranganathan (*apud* CAMPOS, 2003; ARANALDE, 2009) afirma que qualquer pessoa tem total liberdade para formular o número de categorias que quiser, desde que elas sejam empiricamente testadas. No entanto, a dificuldade em determinar um número adequado de categorias que abrangessem a totalidade dos assuntos possíveis para uma classificação bibliográfica levou Ranganathan ao seu conceito de categorias fundamentais das quais todas as categorias particulares são manifestações. Com isto, Ranganathan apresenta o seu postulado das categorias indivisíveis e fundamentais, a seguir:

a) P (Personalidade): é constituída de entidades, seus tipos ou espécies e respectivas partes ou órgãos. Por esse motivo, pode-se identificar os conceitos de entidade, coisa, todo, produto final, tipos ou espécies, partes ou órgãos com essa categoria fundamental. Ela é considerada pelo autor como indefinível;

b) M (Matéria): compreende o material de que são feitas as coisas e as suas respectivas propriedades. Engloba os elementos constitutivos, estrutura, forma, matéria e propriedades. Inclui tanto propriedades materiais quanto abstratas;

c) E (Energia): indica tanto qualquer tipo de ação como também reações, processos, operações e problemas;

d) S (Espaço): traduz o aspecto espacial geográfico dos assuntos analisados que recaem sob essa categoria. Inclui divisões geográficas de continente, de país, dentre outras;

e) T (Tempo): refere-se ao aspecto cronológico dos assuntos e à limitação de períodos, séculos ou décadas.

Segundo Aranalde (2009), as categorias elaboradas por Ranganathan são ontológicas, além de empíricas, isto é, baseadas na realidade, na observação e na experiência, e buscam responder às questões fundamentais do 'por quê', 'o quê', 'como', 'onde' e 'quando'. Somado a isso, as noções de tempo e de espaço são fundamentais tanto para identificar objetos em geral quanto para a identificação dos assuntos que compõem um documento. A autora conclui que as categorias, enquanto elementos imprescindíveis para a classificação permitem identificar e falar das coisas, possibilitando a elaboração de classes mais gerais em que são ordenados os seus predicados. A formação destas classes gerais pressupõe a delimitação das características das coisas mediante as quais elas podem ser reconhecidas como semelhantes ou dessemelhantes, coincidentes ou não coincidentes, possibilitando reuni-las em uma classe similar, separando as distintas e as distribuindo em outras classes. As categorias são recortes que o pensamento humano faz da realidade, sendo um produto lógico, produto do pensamento e de sua expressão, a linguagem. Como unidades cognitivas, elas são fundamentais para os processos intelectuais de identificar, entender e organizar as coisas do mundo.

3.3 Representação do conhecimento: Indexação e Classificação

No contexto da informação e da documentação, como afirma Rodríguez Bravo (2008), a representação da informação ou do conhecimento faz referência à simbolização de documentos utilizando os elementos físicos ou de conteúdo relevantes e eliminando os supérfluos, marginais ou imprecisos.

A organização ou representação do conhecimento do conhecimento, portanto, abarca todo tipo de métodos de indexação, resumo, catalogação, classificação, gestão de registros, bibliografia e a criação de bases de dados textuais ou

bibliográficas voltados para a recuperação de informação, como lembram Martinez Tamayo e Valdez (2008) e corrobora Rodríguez Bravo (2008).

São, portanto, as principais técnicas de organização do conhecimento, segundo Martinez Tamayo e Valdez (2008):

- a) Com linguagem natural: extração de palavras-chave, nota de conteúdo, condensação ou redação de resumos,
- b) Com linguagens artificiais: indexação pré-coordenada, indexação pós-coordenada, classificação.

A indexação e a classificação, foco principal de análise nesse trabalho, têm um método que é comum a ambas e que consiste em três passos que são (MARTÍNEZ TAMAYO, VALDEZ, 2008; RODRÍGUEZ BRAVO, 2008):

- análise conceitual que consiste em determinar de que trata o item que se está processando por meio de sua leitura técnica do título - ou a audição ou a visualização - a depender do tipo documental, do prólogo, do prefácio, da capa e contracapa, do sumário e do índice de assuntos e, em alguns casos como os artigos dos periódicos, das palavras-chave do autor, dos objetivos, das conclusões, dos títulos das seções, das legendas, das ilustrações, enfim, todas as partes da publicação que possam conter informações relevantes sobre o conteúdo da obra. Nessa leitura, deve-se identificar e selecionar os conceitos ou assuntos de que tratam ou resumem o conteúdo da obra. Essas são atividades subjetivas de modo que, além de formação e treino específicos, o profissional deve conhecer bem os usuários a quem se dirigem os documentos e a terminologia utilizada por eles e, principalmente, a ser adotada no controle da representação do conteúdo.
- a representação ou tradução, quando se determinou os assuntos pelos quais irão se recuperar o item, é necessário lhe dar alguma forma antes do registro. Quando se utiliza um Sistema de Organização do Conhecimento (SOC), os assuntos expressos em linguagem natural serão traduzidos para os termos da linguagem artificial do sistema adotado, e, por fim,

- o armazenamento em registro bibliográfico.

A indexação, de acordo com Naves (2004), no campo do tratamento da informação, apresenta dois sentidos: um mais amplo, quando se refere à atividade de criar índices, seja de autor, de título, de assunto, tanto de publicações (livros, periódicos), quanto de catálogos ou de banco de dados, em bibliotecas ou centros de informação. O outro sentido, mais restrito, refere-se à indexação, classificação ou catalogação de assuntos das informações contidas em documentos.

Lancaster (2004, p. 9) manifesta sua preocupação com os usuários quando afirma que “uma indexação de assuntos eficiente implica que se tome uma decisão não somente quanto ao que é tratado num documento, mas também por que ele se reveste de provável interesse para determinado grupo de usuários”. Afirma ainda que “quanto mais especializada a clientela de um centro de informação, maior é a probabilidade de que a indexação possa e deva ser feita sob medida, ajustando-se com precisão aos interesses do grupo”. Lancaster (2004, p. 12) alerta que os “indexadores devem estar inteiramente a par dos interesses da comunidade atendida e das necessidades de informação de seus membros”.

Nesse sentido, Guinchat y Menou (1992) ressaltam que as perguntas as quais deve responder uma indexação são as que formulariam um usuário para saber qual é o objeto de um documento e para que pode servi-lo. Para isso, a indexação de qualquer documento deve ser guiada pelas questões do paradigma de Lasswell (5w) que são: quem?, o quê?, como?, quando?, onde?.

Barité (2000), que define a indexação como o processo de descrição e representação de conteúdo de um documento mediante um número limitado de conceitos extraídos do texto dos documentos (palavras-chave) ou de vocabulários controlados (descritores, termos ou cabeçalhos de assunto), faz a distinção entre a indexação pós e pré-coordenada. Na indexação pós-coordenada, os termos são combinados ou correlacionados ao momento da recuperação da informação. A indexação pré-coordenada, por sua vez, corresponde à combinação prévia, pelo indexador, dos termos que representam o conteúdo do item indexado. Nesse tipo de indexação, a sintaxe de combinação e uso dos termos é, portanto, pré-coordenada

segundo diretrizes estabelecidas pelo próprio Sistema de Organização do Conhecimento (SOC).

O número de termos será determinado pela quantidade de informação contida no item indexado, analisado com relação às necessidades dos usuários. Nessa etapa, é necessário levar em conta os princípios básicos para a atribuição de termos que são, segundo Martinez Tamaio, Valdez (2008) e Rodriguez Bravo (2008):

- Princípio da relevância que é um fator intrínseco ao termo de indexação que expressa sua representatividade isoladamente, fora de contexto,
- Princípio da pertinência em que se avalia o termo com relação ao contexto, isto é, quanto ao seu grau de adequação com relação ao documento que está sendo indexado,
- Princípio da exaustividade que mede a qualidade na seleção dos conceitos realmente significativos, isto é, que contém a informação pertinente para os usuários,
- Princípio da especificidade que consiste em atribuir termo ou termos mais específicos que representem os principais temas cuja seleção deve designar de forma precisa o conteúdo e não uma classe mais geral abaixo da qual poderá ser incluído, ou seja, que mede a qualidade na seleção dos termos de indexação que correspondem efetivamente aos conceitos incluídos no documento,
- Princípio de uso que consiste em usar a terminologia empregada pelos usuários em atendimento às suas necessidades ou das características do acervo,
- Princípio da univocidade, ou unidade de sentido, ou consistência ou coerência que trata de denominar um assunto sempre da mesma forma cujo conceito deve ser unívoco, descartando os possíveis sinônimos, homônimos e termos polissêmicos. Este é um princípio comum aos Sistemas de Organização do Conhecimento (SOC), mediante o controle do vocabulário,

- Princípio linguístico ou da ordem da língua natural em que devem ser utilizados termos pertencentes ao idioma da comunidade de usuários e respeitar a ordem natural das expressões, de modo que expressões invertidas não são aceitas como termos, ainda que possam ser incluídas como variantes.

Na indexação, sempre existe a necessidade de agregar outros assuntos, sendo os mais comuns, os nomes próprios de pessoas, famílias e dinastias, entidades corporativas, reuniões, lugares geográficos, além de nacionalidades, línguas e literaturas nacionais, guerras e batalhas, tratados, grupos étnicos, animais, vegetais e minerais, ferramentas, instrumentos musicais, jogos, órgãos e regiões do corpo, produtos químicos, dentre outros. Quando se trata de nomes próprios e títulos uniformes, geralmente se normalizam em uma lista de autoridades de nomes e títulos. Os demais termos se controlam por meio de uma lista de autoridades do assunto. O controle de autoridade é o processo de manter a coerência da forma verbal usada para representação do conteúdo (MARTÍNEZ TAMAYO, VALDEZ, 2008).

A técnica de classificação é a aplicação do método da organização do conhecimento com o fim de ordenar sistematicamente tanto a informação no catálogo, como os itens na estante, em se tratando de uma biblioteca. Via de regra, essa técnica pode ser aplicada com duas finalidades distintas, isto é, para a construção de um número de chamada para classificar os itens nas estantes e para o desenvolvimento de um catálogo de assunto classificado ou sistemático. (MARTÍNEZ TAMAYO, VALDEZ, 2008).

Deve-se ressaltar que existe um vínculo indissolúvel entre as técnicas de análise documental, organização, representação e a recuperação da informação, pois nenhuma delas faz sentido em separado. Na etapa da recuperação da informação, os usuários se dirigem ao sistema para solicitar informação. Para satisfazê-los, é necessário compatibilizar seu interesse com as características dos documentos e, para isso, é necessário que tenham sido previamente indexados. No entanto, para que essa compatibilização funcione, é preciso que os termos utilizados na indexação sejam os mesmos utilizados na busca. Esse é o grande auxílio das

linguagens documentárias, pois só assim se produzirá o correto emparelhamento entre os termos da indexação e os da busca. Mas, se a indexação não se articula com a existência de linguagens documentárias, esse circuito não funciona mais que por aproximação e sem rigor (RODRÍGUEZ BRAVO, 2008).

Entretanto, o processo de indexação, assim como a identificação e a extração dos conceitos representativos do documento vai além da existência ou não das linguagens documentárias. Existem sistemas de indexação de linguagem livre ou indexação derivada (unitermos ou palavras-chave), associados ou não com aqueles de linguagem controlada ou documental. A finalidade da análise documental, da organização ou representação da informação é a sua recuperação, o que, para sua maior eficácia, é indiscutível o valor do controle do vocabulário, de autoridades, dentre outros. (RODRÍGUEZ BRAVO, 2008).

Os sistemas com linguagens documentárias que fazem o controle do vocabulário podem favorecer tanto a exaustividade quanto a precisão na recuperação da informação. A exaustividade melhora graças às relações que se estabelecem entre o vocabulário, como as relações de equivalência que controlam a sinonímia e as relações hierárquicas e associativas. A pós-coordenação é também um mecanismo que incrementa a exaustividade. A precisão pode ser incrementada por meio do uso de qualificadores que eliminem a polissemia, por meio das relações de hierarquia que permitem o uso de termos específicos ou genéricos, a depender do caso, assim como graças à pré-coordenação e utilização de termos compostos. (RODRÍGUEZ BRAVO, 2008).

Esse capítulo aborda o referencial teórico com padrões, perfil e esquema de metadados para a representação temática dos conteúdos informacionais nos sistemas de informação. Na análise, o foco recairá sobre a contribuição dos tesauros, dos vocabulários controlados e das linguagens documentárias que compõem os Sistemas de Organização do Conhecimento (SOC) para um esquema de metadados nos sistemas de informação da Câmara dos Deputados.

4. Recuperação e padronização da descrição da informação

Recuperar a informação contida nas publicações sempre foi o principal objetivo dos instrumentos elaborados pelas bibliotecas, desde os catálogos até os Sistemas de Organização do Conhecimento (SOC). Com o desenvolvimento de tecnologias de informação e comunicação, novas formas para estruturar e disponibilizar a informação são desenvolvidas para acesso por meios eletrônicos, como por exemplo, os catálogos online e com acesso pela internet. Dessa forma, estudos resultam no desenvolvimento de métodos e padrões para a organização de recursos de informação online, designados como esquemas de metadados. Genericamente, metadados são definidos como dados sobre dados, ou informações sobre a descrição e a localização de informações com o objetivo de permitir a sua recuperação, principalmente na internet. (ROSETTO, 2003).

De acordo com Souza e Alvarenga (2004), a padronização da descrição da informação é essencial. Segundo esses autores [...] é necessária uma padronização de tecnologias, de linguagens e de metadados descritivos, de forma que todos os usuários da web obedeçam a determinadas regras comuns e compartilhadas sobre como armazenar dados e descrever a informação armazenada e que esta possa ser consumida por outros usuários humanos ou não, de maneira automática e não ambígua. (SOUZA; ALVARENGA, 2004, p. 134).

4.1 Metadados

Algumas definições sobre metadados que nos auxiliam a entender melhor sobre o seu significado, função e aplicação, principalmente, ainda que não exclusivamente no ambiente da internet, como se verá a seguir:

- uma máquina para compreensão das informações sobre os recursos *web* ou outras coisas (BERNERS-LEE, 1997);
- “dados a respeito de outros dados, ou seja, qualquer dado usado para auxiliar na identificação, descrição e localização de informações. Trata-se, em outras palavras, de dados estruturados que descrevem as características de um recurso de informação.”(TAKAHASHI, 2000, p. 172);
- “podem ser comparados a um sistema de rotulagem que descreve o recurso, seus objetivos e características, mostrando como, quando e por quem o recurso foi armazenado, e como está formatado”. Nesse sentido, “são essenciais para entender o recurso armazenado, eles descrevem informações semânticas sobre o recurso”. (DE MARCHI; COSTA, 2004, p. 2-3);
- “atributos ou dados referenciais que representam um recurso”. (ALVES; SANTOS, 2009, p. 2).

Os metadados são elementos fundamentais no processo de representação da informação, pois possibilitam que os recursos possam ser “lidos” por diferentes programas, contribuindo com o compartilhamento de dados entre sistemas. (ALVES; SOUZA, 2007). No entanto, eles representam o que um recurso contém em termos informacionais, mas que não precisam estar em meio eletrônico necessariamente (CASTRO; SANTOS, 2007).

Os principais códigos, formatos e metodologias internacionais vigentes evidenciam a afinidade e continuidade com os novos métodos de construção de metadados, a seguir arrolados (SIQUEIRA, 2001):

- A *International Standard Bibliographic Description* (ISBD), com campos obrigatórios, condicionais e operatórios, um conjunto de explicações conceituais e orientações menos prescritivas que o AACR na descrição

dos campos, mas que também não possibilitava o desenvolvimento de pontos de acesso para índices de buscas;

- O código *Anglo-American Cataloguing Rule* (AACR2) (AACR + ISBD) trazia o conceito de entrada principal, mas não fora pensado originalmente para estruturas de bases de dados nos moldes da *web*, como o *Resource Description and Access* (RDA).

Segundo Oliver (2011), a fim de solucionar algumas limitações, em 1997 aconteceu uma conferência a fim de revisar os princípios básicos do AACR com objetivo de decidir se era viável uma nova edição para o AACR2. Decidiu-se então por uma terceira edição, AACR3, que começou a ser desenvolvida em 2004. Em 2005, percebeu-se a necessidade de uma nova abordagem, em razão do que, adotou-se a ideia de um padrão estruturado para o ambiente digital, assim como um novo título que passaria a se denominar *Resource Description and Access* (RDA). O intuito é que este novo padrão fosse mais flexível e capaz de descrever diferentes tipos de recursos, desde analógicos até digitais.

Segundo Assunção e Santos (2009) e Oliver (2011), o *Resource Description and Access* (RDA) baseia-se nos modelos conceituais de dados bibliográficos e autoridade, Requisitos Funcionais para Registros Bibliográficos (FRBR) e Requisitos Funcionais Para Dados de Autoridades (FRAD); tem seu desenvolvimento baseado na Declaração dos Princípios Internacionais de Catalogação; procura satisfazer as expectativas da comunidade internacional; não obriga a utilização da ordem dos elementos e pontuação prescrita pela *International Standard Bibliographic Description* (ISBD); permite a integração dos dados criados a partir da utilização do RDA em bases de dados existentes, criadas de acordo com o *Anglo-American Cataloguing Rule* (AACR) e padrões similares, como o MARC (*Machine Readable Cataloguing*).

Segundo Siqueira (2011), em comparação às atividades anteriores de catalogação e indexação, o que caracteriza as metodologias de criação e uso de metadados na era digital é o fato de ser um processo dialógico, mais aberto, que pode ser manual ou automatizado, direto ou indireto. A convergência de processos informacionais na *web* exige protocolos, normas e formatos que os modelos

anteriores não poderiam satisfazer. Além disso, permitem a fusão de processos que antes eram compartimentalizados, tais como a catalogação e indexação. O modelo informacional da web impôs novos paradigmas, tais como a interoperabilidade, a imaterialidade e a transitividade incessante dos objetos, com os quais os metadados têm relação (SIQUEIRA, 2011).

Baptista (2007) dá mais detalhes a respeito dessa ruptura de paradigmas, a seguir:

1. A lógica da descrição não mais se baseia numa visão fixa do objeto, mas na desconstrução da própria descrição que, apesar de obedecer a padrões, permite que todo e qualquer elemento descritivo possa ser considerado como ponto de acesso ao objeto, isoladamente, como aos relacionamentos entre o objeto considerado e outros objetos, o que amplia e potencializa o conhecimento sobre determinado assunto.

2. A elaboração de registros e da representação descritiva de qualquer objeto passa a ser feita por diferentes profissionais, de acordo com seus interesses, e a partir de diversas localizações.

Segundo Tarouco, Fabre e Tamusiunas (2003), o uso de metadados traz benefícios relacionados à acessibilidade, interoperabilidade e durabilidade. Mey e Silveira (2009), entretanto, destacam as seguintes vantagens do desenvolvimento e aplicação dos metadados, que são: administram grande quantidade de dados; ajudam na descoberta, recuperação e edição efetiva dos recursos de informação na rede; permitem rapidez do registro bibliográfico; possibilitam um alto índice de relevância, como também a busca por campo. Tabosa e Paes (2012) utilizam um argumento fundamental para esse trabalho, pois segundo esses autores, a utilização dos padrões de metadados, englobando conceitos e terminologias, garante a existência de um conjunto de informações comuns, sobre um determinado tema ou área, facilitando assim, a compreensão, a integração e o uso compartilhado de informações entre usuários de diferentes formações, níveis de experiência e propósitos.

Em suma, segundo Rosetto (2003), os objetivos dos metadados são: a) localizar, identificar e recuperar dados de um recurso informacional; b) propiciar controles de ordem gerencial e administrativo permitindo conexões e remissões (links) para pontos internos e externos; c) possibilitar a interoperabilidade entre sistemas de informação, dentro de padrões; d) informar sobre as condições de acesso e uso da informação; e) ser legível tanto pelo homem como pela máquina e, por fim, f) possibilitar a elaboração de índices. Para atingir esses objetivos, os metadados devem possuir as seguintes características: a) descrição, com pormenores, das condições físicas dos componentes a fim de identificar e caracterizar o recurso de informação; b) adoção de padrões para a sintaxe e semântica da especificação do recurso de informação, c) informação sobre armazenagem, preservação, acesso e uso dos dados; d) informações administrativas e gerenciais para a devida criação e definição de responsabilidades dos metadados; e) possibilitar análises da qualidade, avaliações e formas de uso; f) autodescrever e criar documentação própria que subsidie o gerenciamento dos recursos informacionais.

Os metadados podem ser classificados de acordo com os seguintes tipos (BARRETO, 1999; MILSTEAD et al., 1999; KENNEY et al., 2001; GILLILAND-SWETLAND, 2002):

- metadados administrativos, que controlam a aquisição e acesso a cada um dos recursos informacionais o que permite auxiliar o gerenciamento dos repositórios de objetos;
- metadados descritivos, que descrevem características e conteúdo de um documento, facilitando a identificação, pesquisa e gerenciamento das informações, isto é, utilizados para o tratamento dos recursos informacionais por meio de técnicas como a catalogação, classificação e indexação;
- metadados estruturais, que descrevem a estrutura de armazenamento das fontes de dados e a forma como os objetos se interligam;
- metadados técnicos, que descrevem quais dados são necessários para a interpretação por várias ferramentas, com a necessidade de

armazenamento, manipulação, acesso, visualização ou movimentação dos recursos;

- metadados de negócio, que descrevem o contexto e o significado dos dados;
- metadados de preservação, que controlam a disponibilidade de determinado recurso informacional, assim como seu estado de conservação ou sua localização física;
- metadados de uso, que fornecem informações sobre o acesso dos recursos aos usuários.

4.2 Padrões de Metadados

Segundo Souza, et al. (1997), os padrões de metadados têm como função fornecer as definições e formar uma rede para automatizar registros de propriedades e dados cadastrais de forma padronizada e consistente. A padronização, segundo Nascimento (2008), é necessária para que haja entendimento das estruturas, da sintaxe e do significado dos metadados, e com isso melhorar a comunicação, reutilização e compartilhamento das informações, resultando em redução de trabalho e custos, de modo que alguns dos principais objetivos da utilização de padrões de metadados são, portanto:

- Auxiliar na automatização da recuperação das informações;
- Estabelecer a comunicação consistente entre bancos de dados;
- Garantir o registro de informações que sejam importantes;
- Facilitar o uso e o compartilhamento das informações e conhecimentos;
- Auxiliar a troca entre informações de diferentes bancos de dados.

De modo geral, um formato deve ter sempre uma especificação formal da estrutura e da semântica; deve definir um conjunto coerente de atributos e das terminologias adotadas na descrição (ROSETTO, 2003). Conforme Weibel *et al.* (2002), o requisito formal significa que o formato é mantido por uma agência autorizada e responsável pelas especificações conferindo credibilidade ao esquema

dos dados; o requisito estrutura possui características codificadas dos valores contidos; e o requisito semântica determina atributos cujo significado é compreendido pelo homem e pela máquina dentro de uma coerência indiscutível. Watson (ALA, 2002) identifica que um formato de metadados deve permitir atingir o objetivo dos metadados, facilitando a identificação, localização, recuperação e uso da informação pelo usuário. Cada formato é construído sob um conjunto de especificações e necessidades, e elaborados por especialistas nas áreas em que foram implementados.

Dempsey e Heery (1997), após identificar e avaliar vinte e dois formatos de metadados criaram uma tipologia com três categorias de formatos de metadados, conforme descrito a seguir:

- 1) Na banda um, estão os formatos simples, com dados não estruturados; em geral, extraídos em base automática dos recursos e indexados por motores de busca da Internet, por exemplo, Alta Vista ou Yahoo, dentre outros;
- 2) Na banda dois, estão os formatos com dados básicos estruturados, com descrições suficientes que permitem ao usuário verificar a potencialidade de sua utilidade ou o interesse por um recurso sem ter que recuperá-lo ou conectá-lo, a exemplo do Dublin Core, dentre outros;
- 3) Na banda três, encontram-se os formatos com registros altamente estruturados, que podem ser usados tanto para a localização e recuperação como para documentar os objetos uma coleção; geralmente, são padrões internacionais, a exemplo do MARC.

No entanto, há casos em que não é possível a adoção de padrões universais de metadados, por isso, a definição de um padrão de metadados pode ser realizada por um grupo de pessoas que possuem conhecimento sobre um determinado tipo de acervo, onde são definidos conjuntos de elementos descritores que formarão o padrão, que podem ser nomes padronizados, informações ou dados que dizem respeito ao acervo (NASCIMENTO, 2008). Com isso também concorda Rosetto (2003) que afirma que muitos formatos de metadados são especificações estabelecidas por consenso de determinadas comunidades a fim de atender necessidades de informação específicas.

Diante da ausência de padrões universais que atendam integralmente às necessidades de todos os sistemas, tipos de documentos e acervos, usuários e até mesmo de comunidades, são adotados determinados perfis de aplicação que consistem em tipos de esquema de metadados com elementos projetados de um ou mais esquemas de padrões de metadados, combinados por implementadores para uma aplicação particular (DZIEKANIAK, KIRINUS, 2004). Como melhor explicam Heery e Patel (2002), perfis de aplicação são elaborados sob medida para aplicações particulares, ou seja, são criados elementos que melhor descrevem os dados de uma aplicação específica de um domínio também específico e que podem ser utilizados de um ou mais esquema de padrões existentes e aplicados combinadamente em uma aplicação particular.

4.3 Padrão Dublin Core de Metadados

Muitos padrões têm sido descritos para manipulação de metadados, dentre os quais, um dos mais antigos e difundidos está o *Dublin Core*. O *Dublin Core Metadata Initiative (DCMI)* originou-se na cidade de Dublin, Ohio, Estados Unidos, no ano de 2005, onde reuniram-se profissionais multidisciplinares para seu desenvolvimento, e desde então tornaram-se responsáveis por promover, padronizar e desenvolver o conjunto de elementos do *Dublin Core* (DUBLIN CORE). Segundo Souza, Vendrusculo e Melo (2000), Rocha (2004) e Nascimento (2008), o Dublin Core é uma espécie de língua-mãe dos metadados que pode ser estendida por meio da criação de perfis de aplicação tendo em vista atender às necessidades de comunidades específicas. Ele tem sido adotado devido à sua simplicidade na descrição dos recursos, por permitir coexistir e interoperabilizar com outras normas de metadados com semânticas diferentes, por seu escopo e consenso internacionais e por sua extensibilidade que permitem adaptações às necessidades adicionais de descrição.

O Dublin Core fornece dois níveis de descrição: simplificado e qualificado. O nível simplificado possui um conjunto de 15 elementos descritores básicos planejados para facilitar a descrição de uma grande parte dos recursos eletrônicos,

de forma a facilitar a pesquisa, compartilhamento e gerenciamento das informações (DUBLIN CORE, ISO, 2009). Esses elementos, todos de uso recomendado, não sendo nenhum obrigatório, são:

1. *Contributor* (**Colaborador**): entidade responsável pela criação de contribuições ao conteúdo do recurso; identificado pelo nome de uma pessoa, organização ou serviço.

2. *Coverage* (**Cobertura**): corresponde ao escopo espacial, temporal ou a abrangência do recurso; identificado pelo local, data, período ou até mesmo entidade onde o recurso é aplicado.

3. *Creator* (**Criador**): corresponde à entidade responsável pela criação do recurso, tais como pessoas, organizações e serviços.

4. *Date* (**Data**): corresponde à data ou ao período de tempo, associado com o ciclo de vida do recurso.

5. *Description* (**Descrição**): corresponde à descrição sobre o conteúdo do recurso, como resumo, índice, gráfico, outros.

6. *Format* (**Formato**): corresponde ao suporte físico e/ou digital, ou demais dimensões do recurso, tal como tamanho.

7. *Identifier* (**Identificador**): corresponde a uma identificação do recurso, não ambígua, em um determinado contexto.

8. *Language* (**Idioma**): corresponde à linguagem ou idioma do recurso.

9. *Publisher* (**Publicador**): corresponde à entidade responsável pela disponibilização do recurso que podem ser pessoas, organizações ou serviços.

10. *Relation* (**Relação**): corresponde a relacionamentos entre os recursos, objetos ou as informações.

11. *Rights* (**Direitos Autorais**): corresponde às informações a respeito dos direitos de propriedade e uso do recurso.

12. Source (Fonte): faz referência à origem do recurso, em todo ou em parte.

13. Subject (Assunto): corresponde ao tema ou assunto do conteúdo do recurso, descrito por palavras-chaves, termos ou classificações.

14. Title (Título): corresponde ao nome, em geral, o título pelo qual o recurso é conhecido.

15. Type (Tipo): corresponde à natureza ou ao gênero do recurso, que pode ser texto, áudio, vídeo, dentre outros.

O nível qualificado do *Dublin Core* foi criado com a finalidade de representar com maior precisão os recursos descritos, tendo sido definidas duas classes de qualificadores, sendo elas (TAVARES, 2002 *apud* NASCIMENTO, 2008):

1. Elementos de refinamento que dão mais detalhes aos elementos, como por exemplo: Data – inclusão dos qualificadores de criação, validade ou modificação;

2. Esquemas de codificação que fornece a indicação de esquemas para os valores dos elementos, como por exemplo, o uso de vocabulários controlados ou outras linguagens documentárias.

Segundo Grácio (2002), a utilização dos qualificadores também não é obrigatória e possibilita o desenvolvimento de novos qualificadores de acordo com as necessidades de cada aplicação.

4.4 Qualidade dos metadados

Bentancourt e Rocha (2012) alertam para a importância da qualidade dos metadados disponibilizados para serviços de recuperação, qualquer que seja o padrão adotado pelo sistema de informação. A garantia dessa qualidade se dará por meio da definição e uso de uma política de metadados, como normas de procedimentos, treinamentos e avaliações. Entre as definições estão as escolhas quanto aos padrões de interoperabilidade e o próprio padrão de metadados a ser adotado.

A norma ISO/IEC 25012, segundo Bentancourt e Rocha (2012), aborda a qualidade dos dados de forma multidimensional, segundo critérios classificados em duas categorias, a seguir:

A primeira categoria reúne critérios relativos à qualidade inerente dos dados, isto é, a partir dos dados em si. A segunda, dependente do sistema, cuja qualidade do dado é melhorada ou dependente do sistema em que opera. Há casos em que ambas as categorias estão presentes. A seguir, as três categorias descritas:

Os 15 critérios inerentes ao sistema, denominados dimensões, são:

- a) exatidão (*accuracy*): o dado possui o valor correto do atributo de um conceito ou evento;
- b) completude (*completeness*): o dado sujeito associado com uma entidade tem valores para todos os atributos esperados e instâncias de entidades relacionadas;
- c) consistência (*consistency*): o dado tem atributos livres de contradição e coerentes com outros dados;
- d) credibilidade (*credibility*): o dado tem atributos considerados verdadeiros e fidedignos pelos usuários;
- e) atualidade (*currentness*): o dado tem atributos que são do período correto;

Por sua vez, os critérios inerentes e dependentes do sistema são:

- i. acessibilidade (*accessibility*): o dado pode ser acessado por quem necessita do suporte de tecnologia ou de uma configuração especial decorrente de alguma necessidade especial ou incapacidade;
- ii. conformidade (*compliance*): o dado tem atributos que seguem normas, convenções ou regulamentações relacionadas com a qualidade de dados;
- iii. confidencialidade (*confidentiality*): o dado tem os atributos que garantam que possa ser acessível apenas a usuários autorizados;
- iv. eficiência (*efficiency*): o dado tem os atributos que podem ser processados em níveis esperados de funcionamento e desempenho;
- v. precisão (*precision*): o dado tem os atributos exatos ou que proporcionam a discriminação;
- vi. rastreabilidade (*traceability*): o dado tem atributos que permitem rastreamento e auditoria de acesso ou de qualquer alteração;
- vii. entendibilidade (*undestandability*): os dados adotam linguagens e símbolos que podem ser lidos e interpretados por usuários.

Por fim, os critérios dependentes do sistema são:

- a) disponibilidade (*availability*): os dados podem ser recuperados por usuários autorizados e/ou aplicações;
- b) portabilidade (*portability*): os dados podem ser instalados, substituídos ou removidos de um sistema a outro mantendo a mesma qualidade;
- c) recuperabilidade (*recoverability*): os dados mantêm um nível especificado de operações e qualidade, mesmo em caso de falhas do sistema.

5. Metodologia

Essa é uma pesquisa social qualitativa do tipo descritiva, cujos métodos de coleta e análise dos dados utilizados foram (BELL, 2008; CRESWELL, 2007):

- a) Pesquisa bibliográfica para levantamento do referencial teórico utilizados na análise sobre os Sistemas de Organização do Conhecimento (SOC); as teorias relativas à representação temática e que influenciam o uso e desenvolvimento dos Sistemas de Organização do Conhecimento (SOC), tais como teoria do conceito, classificação e indexação, dentre outros; assim como os metadados;
- b) Pesquisa documental para identificar as características dos tesouros, dos vocabulários controlados e das linguagens documentárias da Câmara dos Deputados;
- c) Pesquisa documental para identificar os sistemas de informação e as tipologias documentais da Câmara dos Deputados. Os dados sobre os sistemas de informação foram fornecidos pelo Centro de Informática e sobre as tipologias documentais foram extraídos de rápido levantamento no Portal da Casa.

Esse trabalho pode ser considerado ainda um estudo de caso, pois está sendo estudado do ponto de vista dos Sistemas de Organização do Conhecimento (SOC) no âmbito da Câmara dos Deputados. Apesar da abordagem ter profundidade compatível com uma monografia de especialização, esse estudo possibilita que um determinado aspecto de um problema seja estudado, no caso, a modelagem dos metadados para representação temática sob o ponto de vista dos SOC da Câmara dos Deputados. De modo que esse estudo poderá contribuir para identificar questões-chave que indiquem a importância da organização e da representação temática dos conteúdos informacionais nos sistemas de informação para a gestão da informação do conhecimento e o desenho da arquitetura da informação, assim como os problemas e as possíveis soluções (MARTINS, 2007).

6. Os Sistemas de Organização do Conhecimento (SOC) da Câmara dos Deputados

Atualmente, a Câmara dos Deputados possui 85 (oitenta e cinco) sistemas de informação informatizados para atender às atividades administrativa, legislativa e técnica da Casa. A maior parte desses sistemas (67%) atendem às atividades administrativas que são aquelas que dão apoio à organização, gestão e funcionamento das atividades meio e fim. Em minoria, são os sistemas que atendem às atividades meio, tal como as técnicas (19%); e as atividades fim, como as legislativas (14%). Do total de sistemas, mais da maioria (88%) estão em atividade e uso. Sem muito rigor para categorizar as atividades dessa análise, uma vez que esse não foi o objetivo principal desse estudo, são exemplos das atividades desenvolvidas na Câmara dos Deputados e atendidas pelos sistemas de informação: a) atividades administrativas: Biblioteca para aplicação do Workflow de férias; b) atividades técnicas: Biblioteca Digital; c) atividades legislativas: Infoleg-Parlamentar para cadastramento de parlamentares, dados histórico-bibliográficos, movimentação parlamentar, registro dos órgãos legislativos e da organização político-partidária realizada pela Mesa Diretora.⁴

Os sistemas da biblioteca, tal como o catálogo das coleções, dentre os demais módulos de gerenciamento do acervo e seus usuários, não entrou nesse rol de sistemas da Câmara dos Deputados, pois fazem parte da Rede Virtual de Bibliotecas – Congresso Nacional (RVBI) que é uma rede cooperativa de bibliotecas, coordenada e gerenciada pelo Senado Federal que contém recursos bibliográficos de 14 (quatorze) bibliotecas da Administração Pública Federal e do governo do Distrito Federal, dos Poderes Legislativo, Executivo e Judiciário. Desde 2000, a RVBI adotou o uso do Aleph, software de gerenciamento de bibliotecas, que adota o formato internacional MARC de intercâmbio bibliográfico.

Alguns dos sistemas de informação arrolados são de acesso livre pela internet no Portal da Câmara dos Deputados, sendo a maior parte os de natureza técnica e legislativa, a exemplo da própria Biblioteca Digital e do Histórico de

⁴ Para essa análise, foram utilizados os dados de sistemas de informação da Câmara dos Deputados constantes no Anexo A.

Movimentação Parlamentar.⁵ O grande desafio desses sistemas, entretanto, não tem sido de natureza meramente tecnológica ou de infraestrutura, mas principalmente atender às demandas de organização e tratamento de diferentes tipologias informacionais, documentais e de seus respectivos suportes, assim como de usuários com perfis variados. As informações ou documentos podem ser dos tipos administrativa, arquivística, bibliográfica, fiscal, jornalística ou legislativa; disponíveis em meio digital, impresso, manuscrito - como é o caso de documentos originais ou históricos; em áudio ou imagem, dentre outros. Por dever de ofício, a Câmara dos Deputados deve estar aberta a todo e qualquer cidadão e, por isso, a informação ali produzida ou recebida deve atender livremente, na maior parte dos casos, aos usuários internos e externos que abrangem os mais especializados como os parlamentares, consultores legislativos, servidores e empregados que dão apoio às atividades técnico-legislativas; partidos políticos, frentes parlamentares e sociedade civil organizada; operadores do direito; estudantes e pesquisadores nacionais e internacionais; assim como a sociedade em geral; além daqueles que desempenham atividades administrativas e de infra-estrutura responsáveis pelo bom funcionamento da Casa.

Grosso modo, exceto pelas informações ou documentos sob a responsabilidade ou guarda das coordenações do Centro de Documentação e Informação (CEDI), pouco ou nenhum tratamento é dado a eles para a sua representação temática nos sistemas de informação da Câmara dos Deputados. Mesmo aqueles que recebem algum tipo de tratamento para representação temática, boa parte não possuem campos estruturados para isso, a exemplo de um dos sistemas de informação mais relevantes para o desempenho das atividades da Câmara dos Deputados e, por isso mesmo, um dos mais acessados pelos usuários que é o sistema que disponibiliza conteúdos sobre a Legislação Federal e Legislação Interna da Câmara dos Deputados; íntegra da Constituição, com texto atualizado, e informações históricas; fotos e publicações sobre a Assembleia Nacional Constituinte; textos das Constituições de 1967, 1946, 1937, 1934, 1891 e 1824; os dispositivos constitucionais sujeitos à regulamentação expressa e

⁵ Esses sistemas estão disponíveis nos links, a seguir: Biblioteca Digital da Câmara dos Deputados - <http://bd.camara.leg.br/bd/>; Histórico de Movimentação Parlamentar - <http://www.camara.gov.br/internet/deputado/pesquisaHistorico.asp>.

dispositivos regulamentados, sem regulamentação, com ou sem proposições apresentadas; legislação anterior à Carta de 1988; proposições apresentadas e sua tramitação no Congresso, com seus respectivos autores, dentre outras.⁶

Nem todos os sistemas da informação, mesmo aqueles que dão algum tipo de tratamento para a representação temática das informações e documentos por eles descritos, utilizam qualquer tipo de Sistema de Organização do Conhecimento (SOC). Porém, para a organização e representação temática das informações e documentos daqueles que recebem algum tipo de tratamento nos sistemas de informação da Câmara dos Deputados, são utilizados os seguintes SOC, com as características a seguir:

- a) **Sistema de Classificação:** A biblioteca da Câmara dos Deputados, para classificação das obras das coleções sob a sua responsabilidade e guarda, utiliza a Classificação Decimal Universal (CDU). Porém, nem todas as bibliotecas da RVBI, a exemplo da própria entidade gerente da rede que é a biblioteca do Senado Federal, utilizam esse mesmo código, e sim o Código Decimal de Dewey (CDD). No entanto, isso não provoca nenhum inconveniente já que os números de chamada das obras se destinam quase exclusivamente à localização do endereço das publicações nas estantes.
- b) **Taxonomias:** Atualmente, estão sendo utilizadas ao menos duas taxonomias para classificar e representar os conteúdos informacionais principais de diferentes tipos documentais em diversos sistemas de informação, que são:
 - i. as áreas temáticas do Tesouro da Câmara dos Deputados (TECAD) são utilizadas para classificar as proposições legislativas e a legislação federal e interna, assim como os discursos e outros pronunciamentos dos parlamentares quando de suas atividades legislativas nos plenários da Câmara dos Deputados. Algumas das pesquisas avançadas que a classificação segundo as áreas temáticas permite é a seleção e reunião de todas as normas jurídicas temáticas que tramitam em determinada comissão, por

⁶ Alguns dessas informações estão disponíveis no link a seguir: <http://www2.camara.leg.br/atividade-legislativa/legislacao>.

exemplo; além de permitir a geração do Anuário Estatístico das Atividades Legislativas, outro exemplo.⁷

As áreas temáticas do TECAD são aquelas pertinentes às atividades administrativas e à produção legislativa da Câmara dos Deputados, além das áreas do conhecimento. Trata-se de uma taxonomia organizacional uma vez que, para a definição das áreas temáticas e suas subáreas, foram utilizadas as seguintes fontes de informação abaixo:⁸

- Áreas das Comissões Temáticas da Câmara dos Deputados tal como dispostas no Regimento Interno,
- Áreas da Consultoria Legislativa da Câmara dos Deputados,
- Estrutura administrativa do Poder Executivo Federal,
- Vocabulário Controlado do Governo Eletrônico (VCGE)⁹,
- Classificação Decimal Universal (CDU),
- Classificação Decimal de Direito (CDDoris)¹⁰,
- Índice de assuntos da Constituição Federal de 1988,
- Áreas do conhecimento do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).¹¹

Essas áreas têm ainda a função de nortear o desenvolvimento do tesauro segundo a relevância, o interesse e a demanda das diversas unidades e sistemas de informação da Casa, sendo também utilizadas para categorizar os termos, assim como estabelecer o relacionamentos entre eles.

⁷ Os anurários estatísticos de diversos anos podem ser acessados nos seguinte link: <http://www2.camara.leg.br/atividade-legislativa/legislacao/publicacoes/anuario-estatistico-do-processo-legislativo>.

⁸ No Anexo B foi incluído o Mapa das Áreas Temáticas desenvolvido e utilizado pelo Tesauro da Câmara dos Deputados (TECAD).

⁹ Vocabulário Controlado do Governo Eletrônico (VCGE) disponível em: <http://www.governoeletronico.gov.br/acoes-e-projetos/e-ping-padroes-de-interoperabilidade/lista-de-assuntos-do-governo-lag>.

¹⁰ Classificação Decimal de Direito (Doris), disponível em: <http://legislacao.planalto.gov.br/cddir/cddir.nsf>.

¹¹ Tabela de áreas do conhecimento do CNPq, disponível em: <http://www.cnpq.br/documents/10157/186158/TabeladeAreasdoConhecimento.pdf>

Segundo explica o Manual Metodológico de Desenvolvimento do TECAD, cada área temática equivale a uma categoria. Cada categoria, por sua vez, terá tantas subcategorias quanto forem pertinentes e relevantes para estabelecer um esquema de classificação e estruturar de forma sistemática e ordenada os termos em cada uma dessas áreas temáticas. Dá-se o nome de taxonomia a esse esquema de classificação e a essa estrutura sistemática e ordenada das áreas temáticas.

A taxonomia de cada categoria do TECAD é composta, no mínimo, pelas seguintes subcategorias genéricas:

- ✓ órgãos/organização: abriga termos relativos às entidades ou instituições da estrutura administrativa e gerencial em determinada categoria;
- ✓ agentes/sujeitos: abriga termos relacionados às pessoas, indivíduos ou titulares que desenvolvem atividades, realizam trabalhos, executam processos ou ocupam cargos ou funções em determinada categoria;
- ✓ funções/competências: abriga termos relacionados às atribuições, deveres, direitos, prerrogativas, poderes ou responsabilidades atribuídos aos órgãos ou agentes em determinada categoria;
- ✓ ações/procedimentos/processos: abriga termos relacionados às atividades, atos, atuações, acontecimentos, ocorrências, operações, práticas ou trabalhos, atribuídos aos órgãos ou agentes em determinada categoria;
- ✓ objetos/serviços/produtos: abriga termos relacionados ao resultado dos trabalhos realizados, dos processos executados e das atividades desenvolvidas em determinada categoria;
- ✓ princípios/diretrizes/critérios/postulados: abriga termos relacionados aos preceitos que regem os órgãos, agentes e ações de determinada categoria.

A depender da categoria, outras subcategorias podem se fazer necessárias. Do mesmo modo, as denominações das subcategorias em cada categoria dependerão da garantia literária, do usuário e de uso de cada área temática.

A categorização auxilia no processo de sistematização e ordenação das áreas temáticas e suas subáreas, assim como a identificar conceitos, definir termos, atribuir seus relacionamentos e selecionar fontes dos quais esses termos serão extraídos.

- ii. o Plano de Classificação Funcional (PCF), que é uma taxonomia baseada nas funções, nos processos e nas atividades efetivamente desenvolvidas na Casa, que tem sido utilizado para classificar o conteúdo dos documentos recebidos, produzidos e gerenciados em demanda às atividades administrativas, financeiras, legislativas e políticas, dentre outras, realizadas na Câmara dos Deputados.
- c) **Tesouro:** Segundo o Manual Metodológico de Desenvolvimento do TECAD, seu objetivo é dotar a Câmara dos Deputados de um vocabulário controlado sistêmico que contemple os pressupostos constitucionais, institucionais e organizacionais do Poder Legislativo, a ser utilizado como instrumento terminológico na gestão corporativa da informação, principalmente na classificação, indexação e pesquisa dos conteúdos informacionais.

Com esse intuito, o TECAD tem como principais funções, tendo em vista também atender às normas internacionais de elaboração de tesouros (ISO 2788, ANSI/ISO Z39.19-2005): traduzir e converter a linguagem natural de autores, indexadores e usuários em um vocabulário controlado a ser utilizado na indexação e recuperação em sistemas de informação; promover a uniformidade e consistência terminológica na representação e recuperação dos conteúdos informacionais; prover hierarquias consistentes e claras em um sistema de navegação; servir de parâmetro para a criação e definição de metadados nos sistema de informação da Câmara dos Deputados.

A fim de atender a essas funções, o desenvolvimento do TECAD é regido pelos seguintes princípios: eliminação da ambiguidade de modo que cada termo tenha um e somente um significado; controle de sinonímia; estabelecimento de relacionamentos semânticos entre termos quando apropriado; garantia literária, organizacional, de uso e de usuário no estudo terminológico, desde a extração até a homologação; assim como o teste e validação dos termos pelos indexadores, pesquisadores e especialistas.

Portanto, o TECAD se constitui em um instrumento multidisciplinar de controle terminológico, mas com foco nas áreas jurídica, legislativa, parlamentar, política e da administração pública, onde termos são relacionados poli hierarquicamente, seja para atender à complexidade do próprio conhecimento, como do ambiente informacional da Câmara dos Deputados. Atualmente, possui aproximadamente dez mil termos entre candidatos, preferidos, não-preferidos e das listas auxiliares; porém é utilizado na indexação de apenas alguns tipos de informações e documentos, tais como as preposições legislativas, a legislação federal e interna, os discursos e outros pronunciamentos dos parlamentares. A medida que ele vem se desenvolvendo e consolidando e os primeiros produtos têm sido disponibilizados, equipes de outros sistemas de informação têm demonstrado interesse em utilizá-lo na indexação, inclusive que contêm informações e documentos em áudio e vídeo. Em breve, esse cenário deve mudar, pois com a criação e aprovação da Política de Indexação, o TECAD se constitui, por força de legislação interna, no vocabulário controlado exclusivo para uso na indexação de todos os conteúdos informacionais da Câmara dos Deputados.¹²

- d) **Listas de cabeçalhos de assunto:** O TECAD possui diversas listas auxiliares que possuem termos que individualizam determinados conceitos e particularizam certos assuntos não representados pelos descritores. Esses

¹² Criada pelo Ato da Mesa nº 80, de 31 de janeiro de 2013. Disponível em: <http://www2.camara.leg.br/legin/int/atomes/2013/atodamesa-80-31-janeiro-2013-775250-norma-cd.html>.

termos não constam na estrutura do tesauro e na indexação, podem ou não estar associados a um descritor, a depender da que estiver prescrito na Nota Aplicativa do termo. Atualmente, o TECAD conta com as seguintes listas auxiliares, a seguir: acidentes geográficos, atos Internacionais, produtos agropecuários, animais e vegetais da fauna e flora; cargos, ocupações e categorias profissionais; condecorações e concessões honoríficas; contribuições e fundos; datas comemorativas; doenças e transtornos; elementos e compostos químicos; eventos, fóruns, debates e conferências; gratificações e vantagens; línguas e dialetos; normas jurídicas; edificações e obras de engenharia; manifestações e opiniões; povos; programas, planos, projetos e sistemas; além daquelas específicas para atender algumas demandas da Casa que são as de serviços e produtos e as de unidades administrativas da Câmara dos Deputados.

e) **Listas de autoridades:** Do mesmo modo que as demais listas auxiliares do TECAD, as listas de autoridades são utilizadas de forma complementar na indexação para individualizar alguns conceitos, como: deputados federais; conselhos, entidades e organizações; partidos políticos e frentes parlamentares.

f) **Índice toponímico:** No TECAD, é composto pela lista auxiliar de nomes e localidades geográficas.

g) **Glossário:** A Câmara dos Deputados conta com um glossário que contém número restrito de termos relativos às principais atividades da Casa. Esse glossário, disponível para pesquisa no Portal da Câmara dos Deputados, já esteve integrado ao TECAD, mas atualmente não está mais sob a responsabilidade dessa equipe.¹³ No entanto, há algumas iniciativas isoladas

¹³ Em 2004, por meio da Portaria nº 120, de 18 de janeiro de 2004, foi criado o Grupo de Trabalho para estudar e propor ações relacionadas à elaboração de glossário de termos legislativos e orçamentários e à construção de Tesauro da Câmara. Esse grupo, além de apresentar um glossário em português, de termos legislativos e orçamentários e traduzi-lo para o Inglês e Espanhol, deveria também fixar diretrizes para a indexação e padronizar a linguagem documentária a ser utilizada nas bases de dados. Posteriormente, esse grupo deu origem à equipe que viria ser responsável pela elaboração do TECAD, porém, por diversas

de glossários setoriais e, cada vez mais, ganha força a discussão sobre o desenvolvimento de um glossário corporativo como uma das medidas para padronizar a terminologia utilizada nas atividades profissionais da Casa.

h) **Ontologia:** A comunidade do legislativo, composta principalmente por integrantes da Rede de Informação Legislativa e Jurídica (LeXML) e da Comunidade de Tecnologia da Informação Aplicada ao Controle (TI Controle), dentre outras casas legislativas, tem a pretensão de um dia desenvolver uma ontologia no domínio do Poder Legislativo. Por enquanto, essa comunidade tem se reunido e desenvolvido atividades de forma colaborativa por meio do intercâmbio de experiências e recursos, inclusive no que diz respeito à contabilização da equivalência dos tesouros dos órgãos envolvidos para que seja possível integrar as informações legislativas e jurídicas produzidas e disponibilizadas pelos sistemas de informação desses órgãos e, principalmente, no portal do LeXML.¹⁴ Essas atividades de compatibilização e intercâmbio são feitas no âmbito do GT Tratamento de Assunto.

O objetivo geral do LeXML é identificar e estruturar as informações legislativas e jurídicas através da integração de processos de trabalho e compartilhamento de dados utilizando padrões abertos, nas três esferas administrativas (federal, estadual e municipal) e entre os órgãos dos três poderes da República (Executivo, Judiciário e Legislativo), por meio de hiperlinks persistentes, sistemas *online* e tratamento padronizado da estrutura textual. Em suma, essa rede ambiciona ser o “Google das leis”, isto é, disponibilizar de forma rápida, fácil e precisa o rol de leis, jurisprudências, doutrinas e proposições brasileiras. Para isso, além do controle vocabular têm trabalhado na definição de padrões para armazenamento, tratamento e compartilhamento de suas informações e documentos.

i) **Folksonomia:** Boa parte das informações e documentos mais relevantes para as atividades da Casa e da sociedade já estão disponíveis

questões gerenciais e administrativas que não cabem ser discutidas nesse trabalho, o desenvolvimento do glossário foi dissociado do TECAD desde 2006, aproximadamente.

¹⁴ LeXML disponível para consulta em: <http://www.lexml.gov.br/>.

para acesso livre no Portal da Câmara dos Deputados. Apesar disso, a ideia de criar um ambiente para o desenvolvimento de uma indexação social ainda é embrionária nesse órgão; pois prevalece o consenso de que essa atividade tenha que ser desenvolvida exclusivamente por especialistas de forma centralizada. No entanto, à medida que ganha força a ideia de que a gestão da informação e do conhecimento da Câmara dos Deputados e da Administração Pública, em âmbito geral, deve ser regida pelo princípio da transparência, o desenvolvimento de uma folksonomia para tratamento da informação e documentos no portal de forma integrada a outras formas de indexação e classificação, também pode ganhar espaço.

Diante desse cenário dos Sistemas de Organização do Conhecimento (SOC) da Câmara dos Deputados, levantam-se alguns requisitos compatíveis com criação de um perfil de metadados para a representação temática das informações e documentos nos sistemas de informação da Câmara dos Deputados, a seguir:

Perfil de Metadados para Representação Temática				
Nome	Definição	Autoridade do Registro	Status Mandatório Opcional Recomendado	Requisitos
CDU	Notação numérica para representação dos assuntos principais	Indexadores da biblioteca da Câmara dos Deputados	Mandatório e exclusivo para informação bibliográfica	Classificação, por meio de notação, dos assuntos principais do conteúdo da informação ou documento Fonte: Classificação Decimal Universal (CDU)
Classificação Funcional	Arquivo da Câmara dos Deputados	Indexadores Usuários da área administrativa e técnica, especialmente arquivo	Mandatório para os documentos administrativos e arquivísticos Opcional para os demais tipos de documentos e informação	Classificação dos assuntos principais do conteúdo da informação ou documento Fonte: Plano de Classificação Funcional (PCF)

Notação	Notação alfanumérica para representação da área e subárea	Indexadores	Mandatário para todos os tipos de documentos e informação, exceto bibliográfica	Classificação, por meio de notação, da área e subárea do conteúdo da informação ou documento Fonte: Esquema de notação do TECAD (<i>versão provisória, em desenvolvimento</i>)
Área temática	Indicação da área temática principal	Indexadores	Mandatário para todos os tipos de documentos e informação Recomendado para os documentos e informações com acesso livre à internet, principalmente aqueles relativos às atividades jornalísticas e disponibilizados em áudio ou vídeo	Classificação da área principal do conteúdo da informação ou documento Fonte: Mapa de Áreas temáticas do TECAD
Subárea temática	Indicação da subárea temática principal	Indexadores	Recomendado para todos os tipos de documentos e informação	Classificação da subárea principal do conteúdo da informação ou documento Fonte: Mapa de Áreas temáticas e taxonomia do TECAD

Assunto	Assunto(s) específico(s) do conteúdo do documento ou informação	Indexadores	<p>Recomendado para todos os tipos de documentos e informação</p> <p>Obrigatório para aqueles relativos às atividades legislativas e técnicas</p>	<p>Indexação dos conteúdos informacionais por meio de seleção de termos do TECAD</p> <p>Quando o assunto estiver relacionado às atividades legislativas, recomenda-se identificar a informação legislativa pertinente, tal como proposição, lei, doutrina ou jurisprudência, dentre outros tipos.</p> <p>Recomenda-se incluir local e data, principalmente em se tratando de data ou período histórico.</p> <p>Fonte: Termos do TECAD</p>
Assunto— pessoas, entidades, eventos ou objetos	Pessoas, entidades, eventos ou objetos como assunto(s) específicos do conteúdo do documento ou informação	Indexadores	Recomendado para todos os tipos de documentos e informação	<p>Indexação dos conteúdos informacionais por meio de seleção de termos das listas auxiliares do TECAD</p> <p>Quando o assunto estiver relacionado às atividades legislativas, recomenda-se identificar a informação legislativa pertinente,</p>

				<p>tal como proposição, lei, doutrina ou jurisprudência, dentre outros tipos.</p> <p>Recomenda-se incluir local e data, principalmente em se tratando de tempo e período histórico</p> <p>Fonte: Termos do TECAD</p>
Assunto livre	Assunto(s) específico(s) do conteúdo do documento ou informação, em geral composto por neologismo	Indexadores e usuários em geral	Recomendado para todos os tipos de documentos e informação, principalmente aqueles que abordam novos tema ou assuntos	<p>Indexação dos conteúdos informacionais por meio de uso de palavras-chave ou termos extraídos de outros vocabulários controlados que não sejam o TECAD</p> <p>Fonte: Termos Livres ou extraídos de outros vocabulários controlados</p>
Assunto livre – Indexação social	Assunto(s) específico(s) do conteúdo do documento ou informação com linguagem popular e não controlada	Usuários em geral	Recomendado para todos os tipos de documentos e informação, principalmente aqueles com acesso livre à internet e estão em pauta no parlamento, na mídia ou sendo discutidos pela sociedade em geral	<p>Indexação dos conteúdos informacionais por meio de uso de termos populares utilizados pela sociedade em geral</p> <p>Fonte: Termos Livres ou extraídos de outros vocabulários controlados</p>

Resumo	Resumo informativo ou indicativo	Indexadores	<p>Opcional para todos os tipos de documentos e informação</p> <p>Recomendado para os conteúdos de documentos e informação de carácter mais especializado ou de interesse para as atividades legislativa ou técnica</p>	<p>Síntese dos conteúdos dos documentos ou informações a fim de indicar ou explicar os principais pontos em destaque e dar uma idéia geral do que se tratam.</p> <p>Quando necessário, por se tratar de informação especializada, incluir objetivos, metodologia, resultados e conclusões do conteúdo do documento ou informação</p> <p>Fonte: Texto livre</p>
Fonte: Autora				

7. Conclusões

Os Sistemas de Organização do Conhecimento (SOC) têm muito a contribuir para a representação temática dos conteúdos nos sistemas de informação da Câmara dos Deputados. No entanto, eles por si só não são capazes de resolver essa questão completamente, pois os SOC são uma espécie de léxico que fornecem os termos padronizados para a representação temática. No entanto, é necessária também a definição da sintaxe de uso desse léxico para que a gramática fique completa e a comunicação se estabeleça de forma mais próxima à perfeição entre o usuário e a informação por meio dos sistemas de informação. Estabelecer o léxico e a sintaxe dessa gramática são, portanto, maneiras de estabelecer os signos que serão utilizados para descrever cada um dos conceitos, assim como a forma que serão empregados para melhor representar os conteúdos informacionais.

Os próprios Sistemas de Organização do Conhecimento (SOC) e algumas técnicas e orientações utilizadas na indexação já são instrumentos metodológicos para resolver essa questão. Uma delas, por exemplo, é a indexação pré-coordenada, quando os termos devem ser adotados tal como prescritos no vocabulário controlado sem dar margem ao indexador de aglutinar conceitos para tornar a representação mais específica. Nesse caso, o vocabulário deve conter termos bastante específicos e, em geral, compostos e complexos para atender a toda a demanda dos indexadores. No entanto, a pré-coordenação não atende nem de longe à complexidade dos ambientes informacionais das organizações, inclusive a Câmara dos Deputados. Essa complexidade informacional, basicamente, contempla uma enorme quantidade de sistemas de informação para tratar diferentes tipos de informação e documentos, com acesso livre na internet e, por fim, para atender a uma grande diversidade de usuários com demandas e perfis específicos.

Atualmente, uma das necessidades mais elementares dos sistemas de informação da Câmara dos Deputados é a criação de campos estruturados para a representação temática de seus conteúdos, uma falha que ainda persiste em boa parte de seus sistemas, inclusive aqueles que organizam a informação legislativa. Desse modo, não é possível pensar em integração dos sistemas e em busca federada eficaz, mesmo com acesso livre na internet. Esse estudo buscou

contextualizar esse problema. Senão para apontar uma solução definitiva, ao menos para fazer uma reflexão à luz da experiência adquirida ao longo dos anos sob a coordenação e desenvolvimento das atividades do Tesouro da Câmara dos Deputados (TECAD), por excelência o instrumento terminológico de controle vocabular para a representação dos conteúdos informacionais.

Esse trabalho se constitui, portanto, em um exercício de contextualização teórica e metodológica sobre temas caros ao desenvolvimento dos Sistemas de Organização e do Conhecimento (SOC), dentre eles as teorias do conceito, da classificação e da indexação. Em segundo lugar, foi um exercício de estranhamento em relação aos sistemas de informação, tipologias documentais e usuários reais da Câmara dos Deputados, pois o objetivo não era fazer nenhuma aplicação prática, mas sim de elocubrar sobre o tema da representação temática em um universo hipotético. Por fim, a definição de um perfil de metadados à luz do Tesouro da Câmara dos Deputados (TECAD), que tomou por parâmetro o padrão Dublin Core, se constitui no principal exercício desse trabalho, ainda que se saiba incompleto e não definitivo.

Outro aspecto ainda mais relevante desse trabalho é lançar luz sobre a necessidade urgente de uma gestão de metadados, cuja responsabilidade deve ser compartilhada entre os usuários, os clientes, donos dos negócios e responsáveis pelos sistemas de informação; pelos profissionais de informação, principalmente bibliotecárias e bibliotecários, com suas técnicas de organização da informação; os profissionais da área de tecnologia da informação; assim como as equipes que desenvolvem os Sistemas da Organização e do Conhecimento (SOC). Essa gestão deve incluir desde a definição e implantação dos metadados, que nesse caso não seriam apenas para representação temática, até o estabelecimento de responsabilidades, formas de governança; além da avaliação e o controle de qualidade sobre o preenchimento dos campos definidos.

Os resultados do trabalho apontam que a definição de um perfil de metadados que atenda às necessidades dos usuários, dos diferentes documentos e informações e da diversidade dos sistemas de informação da Câmara dos Deputados, deve entrar em pauta das discussões e decisões relativas à gestão da

informação e do conhecimento, sob pena não apenas de comprometer a recuperação e a visibilidade da informação, mas de afetar também qualquer desenho ou implantação de um programa de Arquitetura e Organização da Informação para a Câmara dos Deputados.

Ao longo dos meses de aulas, apresentações e discussões entre colegas e professores durante todo o curso de especialização, ficaram claros alguns dos problemas, mas também o potencial que os setores, na figura de seus gestores e servidores, têm para desenvolver e implantar uma política de Arquitetura e Organização da Informação na Câmara dos Deputados. O curso permitiu conhecer, não só a teoria extensa, complexa e sofisticada que envolve esse tema, mas o que efetivamente é feito por cada um de nós na gestão, organização e tratamento da informação e do conhecimento. Aonde se quer chegar, assim como as maneiras para isso, no entanto, ainda está por vir e pode ter ficado menos distante a partir dos resultados de nossos trabalhos e reflexões que ora apresentamos após meses frutíferos de intercâmbio e aprendizado.

Referências¹⁵

ALVES, Rachel Cristina Vesú; SANTOS, Plácida. Metadados em ciência da informação: considerações preliminares sobre padrões para a construção normalizada de representações. In: CONGRESSO DE LA CIBERSOCIEDAD, 4., 2009. Disponível em: <<http://www.cibersociedad.net/congres2009/es/coms/metadados-em-cincia-da-informacao-consideracoes-preliminares-sobre-padroes-para-a-construsao-normalizada-de-representasoes/994/>>. Acesso em: março de 2013.

ANSI/NISO Z39.19. **Guidelines for the construction, format, and management of monolingual controlled vocabularies**. Bethesda, Maryland: ANSI/ISO, 2005.

ARANALDE, Michel Maya. Reflexões sobre os sistemas categoriais de Aristóteles, Kant e Ranganathan, **Ciência da Informação**, Brasília, v. 38, n. 1, p. 86-108, jan./abr. 2009. Disponível em: <http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1056>. Acesso em: março de 2013.

ASSUMPÇÃO, Fabrício Silva; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. **Resource Description and Access (RDA): objetivos, características e desenvolvimento do novo padrão para descrição de recursos e acesso**. 2009. Disponível em: http://prope.unesp.br/xxi_cic/27_33049772875.pdf. Acesso em: março de 2013.

BAPTISTA, Dulce Maria. O impacto dos metadados na representação descritiva. **Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis**, v.12, n.2, p. 177-190, jul./dez., 2007.

BARITÉ, Mario. **Diccionario de organización Del conocimiento: clasificación, indización, terminología**. Montevideo: el autor, 2008. Disponible en: <http://164.73.14.9/kod/espaniol/index.php>. Acesso em: março de 2013.

BARRETO, Cássia Maria. **Modelo de metadados para a descrição de documentos eletrônicos na web**. 1999. 190f. Dissertação (Mestrado em Ciências em Sistemas de Computação) - Instituto Militar de Engenharia, Rio de Janeiro, 1999.

BARROS, Lídia Almeida. **Curso básico de terminologia**. São Paulo: Editora da Universidade de São Paulo, 2004. – (Acadêmica; 54).

¹⁵ O acesso aos links das referências disponíveis online foram checados na revisão final, por isso constam a mesma data de março de 2013.

BELL, Judith. **Projeto de pesquisa**: guia para pesquisadores iniciantes em educação, saúde e ciências sociais. 4ª ed. Porto Alegre: Artmed, 2008.

BENTANCOURT, Silva Silvia Maria Puentes; ROCHA, Rafael Port da. Metadados de qualidade e visibilidade na comunicação científica. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, v. 17, n. esp. 2 – III SBCC, p.82-101, 2012. Disponível em: <http://www.periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2012v17nesp2p82/23571>. Acesso em: março de 2013.

BERNERS-LEE, T. **Metadata architecture**: documents, metadata and link. 1997. Disponível em: <http://www.w3.org/DesignIssues/Metadata.html>. Acesso em março de 2013.

BOCCATO, Vera Regina Casari. A linguagem documentária como instrumento de organização e recuperação da informação. In: HOFFMANN, W. A. M.; FURNIVAL, A. C. (Org.). **Olhar: ciência, tecnologia e sociedade**. São Carlos: Pedro e João Ed., CECH-UFSCar, 2008. p. 269-278.

BRASCHER, Marisa; CAFÉ, Lígia. Organização da Informação ou Organização do Conhecimento? **IX Enancib**. São Paulo :Enancib, 2008. Comunicação oral apresentada ao GT-02 - Organização e Representação do Conhecimento, disponível em: <http://www.ancib.org.br/media/dissertacao/1835.pdf>. Acesso em: março de 2013.

BRASIL. Congresso Nacional. Câmara dos Deputados. Ato da Mesa nº 80, de 31 de janeiro de 2013. Dispõe sobre a Política de Indexação de Conteúdos Informativos, o Tesouro da Câmara dos Deputados e dá outras providências. **Diário da Câmara dos Deputados (DCD)**, Brasília/DF, 2 de fevereiro de 2013, p. 92, suplemento A. Disponível em: <http://www2.camara.leg.br/legin/int/atomes/2013/atodamesa-80-31-janeiro-2013-775250-norma-cd.html>. Acesso em: março de 2013.

BRASIL. Congresso Nacional. Câmara dos Deputados. Portaria nº 120, de 18 de outubro de 2004. Criou Grupo de Trabalho para estudar e propor ações relacionadas à elaboração de glossário de termos legislativos e orçamentários e à construção de Tesouro da Câmara dos Deputados, **Boletim Administrativo**, Brasília/DF, v. 27, n. 198, p. 2834-2835.

BRAVO, Blanca Rodríguez. **Apuntes sobre representación y organización de la información**. Espanha: Ediciones Trea, 2011. (Biblioteconomia y Administración Cultural, 231). 206p.

CAMPOS, Maria Luiza de Almeida. **Linguagem documentária**: teorias que fundamentam sua elaboração. Niterói: EdUFF, 2001.

CAMPOS, Maria Luiza de Almeida. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 22-32, jan./abril 2004. Disponível em: <http://www.scielo.br/pdf/ci/v33n1/v33n1a03.pdf>. Acesso em: março de 2013.

CAMPOS, Maria Luiza de Almeida. Organização de domínios de conhecimento e os princípios ranganathianos. **Perspectivas em Ciência da Informação**, v. 8, n. 2, p. 150-163, jul-dez 2003. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/366>. Acesso em: março de 2013.

CAMPOS, Maria Luiza de Almeida; GOMES, Hagar Espanha. Taxonomia e classificação: o princípio da categorização. **DataGramZero**: Revista de Ciência da Informação, v.9, n.4, artigo 01, ago. 2008. Disponível em: http://dgz.org.br/ago08/F_I_art.htm. Acesso em: março de 2013.

CARLAN, Eliana. **Sistemas de Organização do Conhecimento**: uma reflexão no contexto da Ciência da Informação. Brasília: UnB, Departamento de Ciência da Informação e Documentação, 2010.

CASTRO, Fabiano Ferreira de; SANTOS, Plácida Leopoldina V. A. da Costa. Os metadados como instrumento tecnológico na padronização e potencialização dos recursos informacionais no âmbito das bibliotecas digitais na era da web semântica. **Informação & Sociedade**, João Pessoa, v. 17, n. 2, p. 13-19, maio/ ago. 2007. Disponível em: <http://www.ies.ufpb.br/ojs/index.php/ies/article/view/840/1442>. Acesso em março de 2013.

CHOWDHURY, G.G., CHOWDHURY, Sudata. **Organizing information**: from the shelf to the web. Londres Facet, 2007. 230p. p. 71-110.

CRESWELL, John W. **Projeto de pesquisa**: métodos qualitativo, quantitativo e misto. 2ª ed. Porto Alegre: Artmed, 2007.

CURRÁS, Emilia. **Ontologias, taxonomia e tesouros em teoria de sistemas e sistemática**. Tradução de Jaime Robredo. Brasília: Thesaurus, 2010. 182p. p. 55-78.

DAHLBERG, Ingetraut. Knowledge organization: a new science? **Knowledge Organization**, v. 33, n.1, p. 11-19, 2006.

DAHLBERG, Ingetraut. Knowledge organization: its scope and possibilities. **Knowledge Organization**, v. 20, n. 4, p. 211-222, 1993.

DAHLBERG, Ingetraut. Teoria da Classificação, ontem e hoje. Tradução Henry B. Cox. In: Conferência Brasileira de Classificação Bibliográfica, 1972, Rio de Janeiro. **Anais...** Brasília: IBICT/ABDF, 1979. v.1. p. 352-370. Disponível em: http://www.conexaorio.com/bitidahlbergteoria/index_teoriam.htm. Acesso em: março de 2013.

DAHLBERG, Ingetraut. Teoria do Conceito. **Ciência da Informação**, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978. Disponível em: <http://revista.ibict.br/ciinf/index.php/ciinf/article/viewFile/1680/1286>. Acesso em: março de 2013.

DE-MARCHI, Ana Carolina Bertolotti; COSTA, Antônio Carlos da Rocha. Uma proposta de padrão de metadados para objetos de aprendizagem de museus de ciências e tecnologia. **RENOTE**: Revista Novas Tecnologias da Educação, Porto Alegre, v. 2 n. 1, p.1-10, mar. 2004. Disponível em: <http://www.cinted.ufrgs.br/ciclo3/af/02-umaproposta.pdf>. Acesso em: março de 2013.

DEMPSEY, Lorcan; HEERY, Rachel. **A review of metadata**: a survey of current resource description formats (version 1.0). March 1997. [on-line]. Disponível em: http://www.ukoln.ac.uk/metadata/desire/overview/rev_toc.htm. Acesso em: março de 2013.

DUBLIN CORE. **The Dublin Core Metadata Initiative**. [online], [201?] Disponível em: <http://dublincore.org/>. Acesso em: março de 2013.

Dziekaniak, Gisele Vasconcelos; Kirinus, Josiane Boeira. Web Semântica. **Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, n.18, 2º sem. 2004. Disponível em: <http://www.periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2004v9n18p20>. Acesso em: março de 2013.

FRANCELIN, Marivalde Moacir, KOBASHI, Nair Yumiko. Concepções sobre o conceito na organização da informação e do Conhecimento. **Ciência da Informação**, Brasília, DF, v. 40 n. 2, p.207-228, maio/ago., 2011. Disponível em: <http://revista.ibict.br/ciinf/index.php/ciinf/article/viewFile/1856/1411>. Acesso em: março de 2013.

FUJITA, M. Organização e representação do conhecimento no Brasil: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. **Tendências da Pesquisa Brasileira em Ciência da Informação**: v. 1, n.1,

2008. Disponível em:
<http://inseer.ibict.br/ancib/index.php/tpbci/article/viewFile/4/1325>. Acesso em: março de 2013.

GARDIN, J.-C. et al. L'automatisation des recherches documentaires: un modèle general "Le SYNTOL". 2. ed. Revue et augmentée. Paris: Gauthier-Villars, 1968 apud CINTRA, A. M. M. et al. **Para entender as linguagens documentárias**. 2. ed. rev. e ampl. São Paulo: Polis, 2002.

GILLILAND-SWETLAND, Anne J. **Introduction to metadata: setting the stage**. [online], 2002. Disponível em:
http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.pdf. Acesso em: março de 2013.

GRÁCIO, J. C. A. **Metadados para a Descrição de Recursos da Internet: O Padrão Dublin Core, Aplicações e a Questão da Interoperabilidade**. Dissertação (Mestrado), São Paulo: Universidade Estadual Paulista, 2002.

GRUBER, Thomas R. A translating approach to portable ontology specifications. **Knowledge Acquisition**, v. 5, n. 2, p. 199-220, 1993. Disponível em:
<http://www.dbis.informatik.hu-berlin.de/dbisold/lehre/WS0203/SemWeb/lit/KSL-92-17.pdf>. Acesso em: março de 2013.

GRUPO DE TRABAJO DEL W3C. Manual de SKOS: Simple Knowledge Organization System= Sistema para la Organización del Conocimiento Simple. **Anales de Documentación**, n. 13, p. 285-320, 2010. Disponível em:
<http://revistas.um.es/analesdoc/article/view/107511/102161>. Acesso em: março de 2013.

GUINCHAT, Claire; MENOUE, Michel. **Introdução geral às ciências e técnicas da informação e documentação**. 2.ed. rev. aum. Brasília: Ibict;CNPq, 1994. Disponível em:
<http://devds-01.ibict.br:8080/rlc/handle/1/1007>. Acesso em: março de 2013.

HEERY, R; PATEL, M. Application profiles: mixing and matching metadata schemas. **Ariadne**, n. 25, 2002. Disponível em:
<http://www.ariadne.ac.uk/issue25/app-profiles>. Disponível em: março de 2013.

HJØRLAND, B. **Knowledge Organization Systems**. 2007. Disponível em:
http://www.db.dk/bh/lifeboat_ko/CONCEPTS/knowledge_organization_systems.htm. Acesso em: março de 2013 apud Carlan, Eliana. **Sistemas de Organização do Conhecimento: uma reflexão no contexto da Ciência da Informação**. Brasília: UnB, Departamento de Ciência da Informação e Documentação, 2010.

HODGE, G. **Systems of knowledge organization for digital libraries**: beyond traditional authority files. Washington: The Digital Library Federation, The Council on Library and Information Resources, 2000. Disponível em: <http://www.clir.org/pubs/abstract/pub91abst.html>. Acesso em: março de 2013.

ISO 15836. **Information and documentation**: The Dublin Core metadata element set. 2009.

ISO 1087. **Terminology work – vocabulary – Part 1: Theory and application**. Geneva: ISO, 2000.

ISO 704. **Terminology work – principles and methods**. Geneva: ISO, 2000.

KENNEY, Anne R.; RIEGER, Oya Y.; ENTLICH, Richard. **Levando la teoria a la práctica**: tutorial de digitalización de imágenes. [online], 2001. Disponível em: <http://www.library.cornell.edu/preservation/tutorial-spanish/contents.html>. Acesso em: março de 2013.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. 2. ed. Brasília: Briquet de Lemos/Livros, 2004.

LARA, Marilda Lopez Ginez de. O unicórnio (o rinoceronte, o ornitorrinco ...), a análise documentária e a linguagem documentária. **DataGramZero**: Revista de Ciência da Informação, v. 2, n. 6, dez. 2001. Disponível em: http://www.dgz.org.br/dez01/Art_03.htm. Acesso em: março de 2013.

LÉVY, P. **A conexão planetária**: o mercado, o ciberespaço, a consciência. Trad. Maria Homem e Ronaldo Entler. São Paulo: Ed. 34, 2001.

MANUAL Metodológico de Desenvolvimento da Câmara dos Deputados. Brasília/DF: Câmara dos Deputados, 2013. (não publicado).

MARTÍNEZ TAMAYO, Ana Maria, VALDEZ, Julia C. **Indización y clasificación en bibliotecas**. 1.ed. Buenos Aires: Alfagrama, 2008. 235p. (Biblioteca Alfagrama, Introducción a la biblioteconomía).

MARTINS, Gilberto de Andrade; THEÓPHILO, Carlos Renato. **Metodologia da investigação científica para ciências sociais aplicadas**. São Paulo: Atlas, 2007.

MEY, Eliane Serrão Alves; SILVEIRA, Naira Christofolletti. **Catálogo no plural**. Brasília: Briquet de Lemos, 2009.

MILSTEAD, J., FELDMAN, S. **Metadata**: Cataloging by any other name.[online], 2001. Disponível em: http://www.iicm.tugraz.at/thesis/cquetl_diss/literatur/Kapitel06/References/Milstead_et_al_1999/metadata.html. Acesso em: março de 2013.

MOREIRA, Alexandra, ALVARENGA, Lídia, OLIVEIRA, Alcione. de Paiva. Thesaurus and Ontology: a study of the definitions found in the computer and information science literature, by means of an analytical-synthetic method, **Knowledge organization**, v. 31, n. 4., p. 231-244, 2004.

MOREIRO GONZÁLEZ, José Antônio. **Linguagens documentárias e vocabulários semânticos para a web**: elementos conceituais. Salvador: EDUFBA, 2011. 310p.

NASCIMENTO, Luciana Umburanas. **Um padrão de metadados para indexação e recuperação de objetos multimídia**. Dissertação (Mestrado), Curitiba: UFPR, 2008.

NAVES, M. M. L. **Curso de indexação**: princípios e técnicas de indexação, com vistas à recuperação da informação. Disponível em: http://www.finaltec.com.br/oicd/uploads/principios_tecnicas_de_indexacao.doc. Acesso em: março de 2013.

NETWORKED Knowledge organization systems/Services.NKOS.Disponível em: <<http://nkos.slis.kent.edu/>>. Acesso em: março de 2013.

Norma Internacional de elaboração de Tesouros Monolíngues (ISO 2788) e IBICT. **Diretrizes para elaboração de tesouros monolíngues**. IBICT: Brasília, 1984.

OLIVEIRA, Dalgiza Andrade, ARAÚJO, Ronaldo Ferreira. Construção de linguagens documentárias em sistemas de recuperação da informação: a importância da garantia de usuário. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, v. 17, n. 34, p.17-30, maio./ago., 2012.

OLIVER, Chris. **Introdução à RDA**: um guia básico. Brasília: Briquet de Lemos, 2011.

ROCHA, Rafael Port da. Metadados, Web Semântica, Categorização Automática: combinando esforços humanos e computacionais para a descoberta e uso dos

recursos da web. **Em Questão**, Porto Alegre, v. 10, n. 1, p. 109-121, jan./ jun. 2004. Disponível em: <http://seer.ufrgs.br/EmQuestao/article/view/86/1127>. Acesso em: março de 2013.

RODRÍGUEZ BRAVO, Blanca. **Apuntes sobre representación y organización de la información**. .ed. Buenos Aires: Alfagrama, 2008. 235p. (Biblioteca Alfagrama, Introducción a la biblioteconomía).

ROSETTO, Márcia. **Metadados e formatos de metadados em sistemas de informação**: caracterização e definição.--São Paulo, 2003. 112 p. (Dissertação de mestrado apresentada ao Curso de Pós – Graduação da Escola e Comunicações e Artes da Universidade de São Paulo).

ROSETTO, Márcia. Metadados e recuperação da informação: padrões para bibliotecas digitais. **Anais....Ciberética: Simpósio Internacional de Propriedade Intelectual, Informação e Ética, 2; Encontro Nacional de Informação e Documentação Jurídica (Enidj); Painel Biblioteconomia em Santa Catarina, 22)**. Florianópolis: [s.n], 2003.

SIQUEIRA, Ivan Cláudio Pereira; SILVA, José Fernando Modesto. Metadados: o fio de Ariadne ou a coragem de Teseu? **Bibl. Univ.**, Belo Horizonte, v.1, n.1, p. 11-18, jan./jun. 2011. Disponível em: <https://www.bu.ufmg.br/rbu/index.php/localhost/article/view/35/15>. Acesso em: março de 2013.

SOUZA, Márcia Izabel F.; VENDRUSCULO, Laurimar G.; MELO, Geane. C. Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core. **Ciência da Informação**, Brasília, v. 29, n. 1, p. 93-102, abr. 2000. Disponível em: <http://www.scielo.br/pdf/ci/v29n1/v29n1a10.pdf>. Acesso em: março de 2013.

SOUZA, R. R.; ALVARENGA, L. A Web Semântica e suas contribuições para a ciência da informação. **Ciência da Informação**, Brasília, DF, v. 33, n. 1, p. 132-141, jan./abr. 2004. Disponível em: http://www.scielo.br/scielo.php?pid=S0100-19652004000100016&script=sci_arttext. Acesso em: março de 2013.

SOUZA, T. B.; et al. Metadados: catalogando dados na Internet. **Transinformação**, v. 9, n.2, 1997, maio/ago. Disponível em: <<http://puccamp.br/~biblio/tbsouza92.html>>. Acesso em: março de 2013.

TABOSA, Hamilton Rodrigues; PAES, Denyse Maria Borges. Ferramentas tecnológicas na representação descritiva de documentos. **Biblionline**, João Pessoa, v. 8, n. 1, p. 78-85, 2012.

TAKAHASHI, T. (Org.). **Sociedade da informação no Brasil**: livro verde. Brasília: Ministério da Ciência e Tecnologia, 2000. Disponível em: <http://www.mct.gov.br/index.php/content/view/18878.html>. Acesso em: março de 2013.

TAROUCO, Liane Margarida Rocenbach; FABRE, Marie-Christine Julie Mascarenhas; TAMUSIUNAS, Fabrício Raupp. Reusabilidade de objetos educacionais. **RENOTE**: Revista Novas Tecnologias da Educação, Porto Alegre, v. 1 n. 1, p. 1-11, fev. 2003. Disponível em: http://www.nuted.ufrgs.br/oficinas/criacao/marie_reusabilidade.pdf. Acesso em: março de 2013.

TAVARES, T.A. Estudo Comparativo de Padrões para Indexação de Vídeo. [online], 2002. apud NASCIMENTO, Luciana Umburanas. **Um padrão de metadados para indexação e recuperação de objetos multimídia**. Dissertação (Mestrado), Curitiba: UFPR, 2008.

TOUTAIN, L. M. B. B. Biblioteca digital: definição de termos. In: MARCONDES, C.H.(Org.) **Bibliotecas Digitais**: Saberes e Práticas. Salvador/Brasília, UFBA/IBICT, 2006. p. 15-24.

VAZ, M. S. M. G. **Metamídia**: Um modelo de metadados na indexação e recuperação de objetos multimídia. Tese (Doutorado), UFPE, 2000.

WEIBEL, Stuart; GODBY, Jean; MILLER, Eric. **OCLC/NCSA metadata workshop report**. [on-line]. Disponível em : <http://dublincore.org/workshops/dc1/report.shtml>. Acesso em: março de 2013.

8. ANEXO A

Sistemas de Informação da Câmara dos Deputados

	Nome do Sistema	Descrição do Sistema	Sigla Original	Sigla Atual	Status (Ativo/Inativo)
1)	Acompanhamento de OS para a Gráfica do Senado	Sistema de cadastramento e acompanhamento de ordens de serviço para a gráfica do Senado	ACOMPGRAF	ACOMPGRAF	ATIVO
2)	Agência Câmara de Notícias	Sistema de cadastro de notícias da Agência Câmara. As notícias são disponibilizadas em tempo real via web. Um Boletim é enviado por e-mail, caso o usuário se cadastre no site	AGECAN	AGECAN	ATIVO
3)	Agenda de credencial de estacionamento	Agendamento de retirada do cartão de acesso aos estacionamentos da Câmara dos Deputados	AGENDA-ESTAC	AGENDA ESTAC	ATIVO
4)	Agenda dos Plenários da Câmara dos Deputados	O sistema de Agenda dos Plenários é uma especialização ou desdobramento do Sistema Eventos. Foi criado com a finalidade de contemplar os requisitos do DECOM, muito específicos no que se refere ao agendamento de plenários para realização das reuniões de Comissões. Irá utilizar a mesma estrutura de diretórios presentes no Star Team para o sistema de Eventos, em desenvolvimento	PLENARIOS	PLENARIOS	ATIVO
5)	Aplicativos para TV Digital	Aplicação para TV Digital que complemente a transmissão das sessões plenárias, compatível com set-top boxes e aparelhos de TV digital que suportem aplicações interativas em Ginga (geralmente TVs identificadas com o selo DTVi)	APLTVD	APLIC TV DIGITAL	INATIVO
6)	Aquisição Periódicos	Sistema para controle de aquisição de periódicos. O novo sistema irá fornecer uma	SICOR	SICOR	ATIVO

		forma segura para os usuários cadastrarem as informações dos periódicos, irá fornecer pesquisas flexíveis para os usuários, e fornecerá avisos (via correio eletrônico) de vencimento de assinaturas dos periódicos			
7)	Atendimento do CEDI - SIATE	Gerenciamento das solicitações de pesquisas e atendimentos do CEDI	CSC	CSC	INATIVO
8)	Autenticador da Presidência	Sistema para autenticar os documentos da presidência da Câmara	AUTPRES	AUTPRES	ATIVO
9)	Autenticador da Primeira Vice Presidência	Autenticador de documentos da primeira vice-presidência. Mesma versão que roda na Presidência	AUT1VICEPRES	AUT1VICEPRES	ATIVO
10)	Autenticador e Repositório de documentos da Assessoria do Departamento Pessoal	Autenticador e Repositório de documentos da Assessoria do Departamento Pessoal	AUTDEPES	AUTDEPES	ATIVO
11)	Automação da Comissão Mista de Orçamento	O Sistema irá automatizar a tramitação das matérias que tramitam pela Comissão Mista de Orçamento	CMO	CMO	ATIVO
12)	Banco de Imagens	Banco de imagens usadas pela Agência e Fotografia	WEBBANCOIMAGEM	BANCO DE IMAGENS	ATIVO
13)	Banco de Notícias Seleccionadas.	Cadastro de notícias correntes e retrospectivas, a partir de agosto de 2000, extraídas dos seguintes jornais: Jornal do Brasil, O Globo, Folha de S. Paulo, O Estado de S. Paulo, Correio Braziliense, Gazeta Mercantil, e Valor Econômico (a partir de Julho/2003), além das revistas semanais Veja e IstoÉ. As informações são disponibilizadas para consulta na Intranet e Extranet.	BNS	BNS	INATIVO
14)	Banco de Talentos	Sistema de cadastro e consulta de currículos de	BT	BT	INATIVO

		Secretário Parlamentar.			
15)	Banco de Talentos - Nova Versão	Sistema de cadastro e consulta de Talentos dos Servidores da Câmara dos Deputados	TALENTOS	TALENTOS	ATIVO
16)	Biblioteca Digital	Biblioteca Digital da Câmara dos Deputados - Customização do DSpace	BD	DSPACE	ATIVO
17)	Biblioteca para aplicação do Workflow de férias	Biblioteca utilizada pela aplicação "Workflow de Férias" e responsável pela obtenção e registro dos dados no SIGESP	FERIASWF	FERIASWF	ATIVO
18)	Cadastro de Instrutores	Gerência o cadastramento de instrutores para o CEFOR	CADINS	CADINS	ATIVO
19)	Cadastro de Servidores do DEPOL	O Sistema de Cadastro de Servidores do Depol será responsável pelo registro das informações pessoais dos agentes de policia legislativa	CSD	CSD	ATIVO
20)	Cartório Digital	Controle de entrega dos certificados digitais	CERTPAR	CERTPAR	ATIVO
21)	Catálogo de Autoridades	Sistema possibilita a divulgação das informações relativas ao Catálogo de Autoridades mantido pelo DEAPA	CAAUT	CAAUT	ATIVO
22)	Catálogo de Serviços do CENIN	O Sistema de Catálogo de Serviços do CENIN mantém os serviços que o CENIN oferece aos órgãos da Casa	CTS	CTS	INATIVO
23)	Catalogo de Sistemas	Sistema para cadastro das informações dos sistemas	SICAT	SICAT	ATIVO
24)	Cautela de Armas de Fogo	O sistema registra a cautela (guarda) de arma de fogo de autoridades policiais e pessoas com porte de arma	CAF	CAF	ATIVO
25)	Central de Digitalização	Digitalização de Proposições e diários da Câmara	CEDIG	CEDIG	ATIVO
26)	Chamadas de Emergências Policiais	Sistema para registrar as chamadas de emergenciais policiais do DEPOL	CEP	CEP	INATIVO
27)	Compras Eletrônicas	Sistema de gerenciamento de licitações. Permite que os fornecedores se cadastrem e	CECD	CECD	ATIVO

		acompanhem as licitações, enviem as suas propostas e obtenham os resultados direto do seu computador, via internet			
28)	Consulta Deputados	Consulta a situação atual do deputado (movimentações na legislatura, dados cadastrais, tempo de mandato)	CONSULTADEP	CONSULTADEP	ATIVO
29)	Contagem de Aparições de Deputados na Câmara de TV	O Sistema de Contagem de Deputados é responsável pelo registro das aparições dos deputados nos programas da TV Câmara	CAD	CAD	ATIVO
30)	Controle de Acesso de Sistemas em Maker	Sistema utilizado em todos os sistemas maker para gerenciar perfil de usuário	SAM	SAM	ATIVO
31)	Controle de Audio Visual	Sistema para a gestão dos documentos de AudioVisual	COA	COA	ATIVO
32)	Controle de Carteiras do DEPOL	O Sistema de Emissão de Carteiras do DEPOL será responsável pela geração de carteira policial dos Servidores do DEPOL, bem como a inclusão e a manutenção das informações no tocante ao Tipo Sanguíneo, ser ou não doador de órgãos e se possui ou não porte de armas	CCD	CCD	ATIVO
33)	Controle de Contas Telefônicas	Sistema que gerencia as faturas telefônicas e aparelhos da Câmara dos Deputados	CTF	CTF	INATIVO
34)	Controle de Correspondências da DIRAD	Sistema de controle de correspondências recebidas pela Câmara dos Deputados, para todas áreas com exceção dos gabinetes	SCD	SCD	ATIVO
35)	Controle de Correspondências do DEAPA	O Sistema de Controle de Correspondências do DEAPA será responsável pelo registro e controle de envio e recebimento das correspondências que são destinadas aos gabinetes e seus funcionários	COR	COR	ATIVO
36)	Controle de Emissoras de Rádio Parceiras	Sistema utilizado para catalogar as emissoras parceiras da Rádio Câmara e apoiar a	CER	CER	ATIVO

		realização de campanhas publicitárias.			
37)	Controle de Estágio e Visita	Módulo de acesso via intranet para o sistema Controle de Estágio e Visita, inicialmente desenvolvido pela COREL. Agora o mesmo já se encontra com uma nova versão realizada pela CESAN em java em vez do DevMaker, porém a gestora é a Dryade ou substituto dela	SICEV	SICEV	ATIVO
38)	Controle de Estágio Visita	Sistema de controle de estágio visita do CEFOR. O estágio visita é um programa da Câmara dos Deputados para estudantes universitários indicados por Deputados. Cada a deputado tem direito a indicar dois universitários por ano	CEV	CEV	ATIVO
39)	Controle de Exposições Artísticas	O Sistema de Controle de Exposições Artísticas registra informações referentes a exposições apresentadas na Câmara dos Deputados. O sistema será responsável por manter o cadastro dos artistas, obras de artes, doações das obras e locais de exposições	CEA	CEA	ATIVO
40)	Controle de Fluxos	Controlar os fluxos de trabalho nas unidades organizacionais da Casa de forma a facilitar a administração do trâmite de demandas e de documentos eletrônicos gerados nos processos de atendimento. Viabiliza ainda a recuperação desses documentos através do conteúdo ou por metadados diversos	CFT	CONTROLE DE FLUXOS	ATIVO
41)	Controle de Gabinete do DEMED	Controla o atendimento de RPs pelo DEMED	GABDEMED	GABDEMED	INATIVO
42)	Controle de Leilão	O sistema registra todos os itens que estão autorizados para leilão na Câmara dos Deputados	CLE	CLE	ATIVO
43)	Controle de pedidos de manutenção de	Sistema de Controle de pedidos de manutenção de sistemas	PEND	PEND	ATIVO

	sistemas				
44)	Controle de Peso	Sistema de acompanhamento de dieta para servidores da Câmara, com o objetivo de perda de peso	CPO	CPO	ATIVO
45)	Controle de Processos da PROPA	O sistema permite cadastrar, gerenciar, controlar o andamento dos processos internos e externos recebidos Procuradoria Parlamentar	CJU	CJU	ATIVO
46)	Controle de Publicação de Documentos	Sistema de controle de recebimento de documentos e montagem das publicações	CPDOC	CPDOC	INATIVO
47)	Controle de Reforma dos Gabinetes	O Sistema de Controle e Reformas dos Gabinetes registra pedidos de manutenção e benfeitorias no gabinetes parlamentares	SGB	SGB	ATIVO
48)	Controle de Solicitações	Sistema de Gerenciamento e controle de solicitações do usuário	SOLICITA	SOLICITA	ATIVO
49)	Controle de Terceirizados do Cenin	Sistema utilizado para controlar faltas, substituições e férias de todos os terceirizados do Centro de Informática	CTC	CTC	ATIVO
50)	Controle de Visitas da Primeira Vice-Presidência	O Sistema será utilizado para auxiliar os colaboradores lotados na 1ª Vice-Presidência no controle e identificação das visitas ao Gabinete	CV1	CV1	INATIVO
51)	Cosev	Sistema de consulta e relatórios complementares relativos à votação e presença.	CONREL	CONREL	ATIVO
52)	Cota Postal Intitucional	Desenvolver um sistema para gerenciar a cota postal dos órgãos administrativos da casa	COTAPI	COTAPI	INATIVO
53)	Cota Postal Telefônica	O Sistema controla a gestão da Cota Postal-Telefônica, implementando como estrutura fundamental de controle uma conta corrente em que se registram todos os eventos que geram débitos ou créditos que atualizam o saldo da cota disponível.	COTAPT	COTAPT	ATIVO

54)	Cotas Passagem Aérea	O Sistema controla a gestão de Cota de Passagem Aérea, implementando como estruturar fundamental de controle uma conta corrente, que registra todos os eventos que geram débitos	COTAPA	COTAPA	ATIVO
55)	Departamento Médico da Câmara dos Deputados	Sistema desenvolvido para o controle de acesso, laboratório, arquivo médico, gabinete e marcação de consultas do departamento médico	DEMED	DEMED	INATIVO
56)	Deputado em Foco	Geração de mídias (CD, DVD) com arquivos da participação do deputado durante a legislatura (reportagens, vídeos, biografia,...)	SISDEPUTADOFOCO	DEPUTADOEMFOCO	ATIVO
57)	Diretório Nacional de Endereços - DNE	Cadastro de endereços fornecido pela Empresa Brasileira de Correios e Telégrafos	EBCTDNE	DNE	INATIVO
58)	dw aquisições	DW sobre informações de processos e contratos de aquisições	DW AQUISIÇÕES	DW AQUISIÇÕES	ATIVO
59)	dw dirad	DW sobre informações da diretoria administrativa	DW DIRAD	DW DIRAD	ATIVO
60)	dw orçamento	DW de informações sobre transferências constitucionais dos estados e municípios e de convênios firmados com municípios	DW ORÇAMENTO	DW ORÇAMENTO	ATIVO
61)	dw sigesp	DW sobre informações pessoais, funcionais e financeiras do sistema de RH	DW SIGESP	DW SIGESP	ATIVO
62)	DW Tramitações	Data Warehouse de tramitações de processos no Sidoc.	DW TRAMITAÇÕES	DW TRAMITAÇÕES	ATIVO
63)	Ementário Jurisprudencial	O Sistema de Ementário Jurisprudencial registra todas as informações de jurisprudência aplicada no âmbito da Câmara dos Deputados. O sistema será responsável por manter um cadastro contendo todas as jurisprudências utilizadas nos processos tramitam no Conselho de Ética	EJP	EJP	ATIVO

64)	Emissão de Carteira Parlamentar	O Sistema de Emissão de Carteiras do DEPOL será responsável pela geração de carteira dos Deputados e inclusão das informações necessárias para a geração da carteira bem como mantém o controle da lista dos documentos entregues pelos parlamentares necessários para a tomada de posse	ECP	ECP	ATIVO
65)	Estágio Acadêmico	Sistema para controlar os estágios de natureza acadêmica na Câmara dos Deputados	SEA	SEA	ATIVO
66)	Fiscalização de Contratos	Site estático para publicação de conhecimentos sobre fiscalização de contratos de Tecnologia da Informação	FISCON	FISCON	ATIVO
67)	Formulário Eletrônico de Cadastro Parlamentar	Permitir que os Deputados possam preencher, via formulário eletrônico, o cadastro inicial ou visualizar e solicitar alterações de suas informações já cadastradas a partir do seu gabinete ou de computador pessoal localizado fora da Câmara dos Deputados pela rede mundial de computadores a Internet. Um formulário eletrônico de cadastro parlamentar trará alternativas de preenchimento ao deputado eleito e garantirá qualidade e confiabilidade do dado cadastrado	CADPAR	CADPAR	INATIVO
68)	GERA CARTAS	Gerador de cartas/texto para envio aos Deputados/Servidores	SGC	SGC	ATIVO
69)	Gerador de Relatórios	Gerador de Relatórios para auditoria de verba parlamentar	GERAREL	GERAREL	ATIVO
70)	Gerador de Senhas	Criar senhas de acesso e emite carta ao Servidor/Deputado. Para utilização pelo componente de Autenticação (Intranet / Banco de talentos)	SGS	SGS	ATIVO
71)	Gerador de Senhas - Nova Versão	Criar senhas de acesso e emite carta ao Servidor/Deputado. Para utilização pelo	SGNNOVO	SGNNOVO	ATIVO

		componente de Autenticação (Intranet / Banco de talentos)			
72)	Gerência de Demandas de TI	Sistema para gerenciar os atendimentos de instalações de hardware e software na Câmara dos Deputados	GDT	GDT	ATIVO
73)	Gerenciamento de Solicitação de Trabalho para Lideranças	Suporte ao processo de Gerenciamento de Solicitação de Trabalho para Lideranças	STL	STL	INATIVO
74)	Gerposto	Sistema responsável pela comunicação com os postos de presença e votação para controle da autenticação biométrica, leitura e gravação dos votos e presenças	GERPOSTO	GERPOSTO	ATIVO
75)	GerpostoAdmin	"Sistema responsável pelo monitoramento e controle remoto dos postos de votação"	GERADM	GERADM	ATIVO
76)	Gestão da SEDAS	Sistema para auxiliar o gerente de projetos da SEDAS a fazer o controle das atividades dos seus funcionários	SED	SED	INATIVO
77)	Gestão de Licenças e Afastamentos de Deputados	Sistema para controlar os licenciamentos e afastamentos dos Deputados	GLA	GLA	ATIVO
78)	Gestão de Processos - BPM	Sistema para acompanhamento e interação com processos publicados no JBPM	PORTALBPM	PORTALBPM	ATIVO
79)	Gestão por Competência	Sistema utilizado pela COREH/DEPES (gestor) como uma ferramenta de suporte à política de Gestão por Competência da DRH.	COMPETENCIA	COMPETENCIA	ATIVO
80)	Guardião	Sistema de controle de Acesso aos recursos de sistema	GUARDIAO	GUARDIAO	ATIVO
81)	Guia de Serviços	Objetiva manter todos os serviços prestados pela Câmara dos Deputados para disponibilizar consultas na Intranet	GUIASERVIÇOS	GUIA DE SERVIÇOS	ATIVO
82)	Importa e Exporta Dados	Sistema responsável pela exportação e importação de dados de presença e votação.	PNLIMEX	PNLIMEX	ATIVO

		Será substituído pelo SEVExport			
83)	INFOLEG-PARLAMENTAR	A solução integrada de informações parlamentares tem o propósito de estabelecer uma integração de informações relacionados aos parlamentares, tais como: Cadastramento de parlamentares, Dados histórico-bibliográficos, Movimentação parlamentar, Registro dos órgãos legislativos e da organização político-partidária realizada pela Mesa Diretora	INFOLEG	INFOLEG	INATIVO
84)	Informações Gerenciais sobre Proposições Legislativas	Sistema de informações gerenciais sobre proposições legislativas, que podem ser analisadas segundo o tema, autor, legislatura, período, partido, etc.	DW PROPOSIÇÕES	DW PROPOSIÇÕES	ATIVO
85)	JORNAL - Jornal da Câmara	Sistema de cadastrado de notícias do Jornal da Câmara. As notícias são disponibilizadas na web	JORNALCD	JORNALCD	ATIVO

9. ANEXO B

Mapa das Áreas Temáticas do Tesouro da Câmara dos Deputados



