

**Universidade Estadual Paulista – UNESP
“Júlio de Mesquita Filho”
Faculdade de Filosofia e Ciências**

André Felipe Pereira da Silva

**AS RELAÇÕES CONCEITUAIS EM SISTEMAS DE ORGANIZAÇÃO DO
CONHECIMENTO: ANÁLISE COMPARATIVA ENTRE A CLASSE 590 DA
CLASSIFICAÇÃO DECIMAL DE DEWEY E A BBC WILDLIFE ONTOLOGY**

MARÍLIA
2019

André Felipe Pereira da Silva

**AS RELAÇÕES CONCEITUAIS EM SISTEMAS DE ORGANIZAÇÃO DO
CONHECIMENTO: ANÁLISE COMPARATIVA ENTRE A CLASSE 590 DA
CLASSIFICAÇÃO DECIMAL DE DEWEY E A BBC WILDLIFE ONTOLOGY**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Faculdade de Filosofia e Ciências, da Universidade Estadual Paulista – UNESP – Campus de Marília, para a obtenção do título de Mestre em Ciência da Informação.

Área de Concentração: Produção e Organização do Informação.

Orientador: Prof. Dr. Walter Moreira (DCI/FFC/UNESP)

MARÍLIA
2019

S586r

Silva, André Felipe Pereira da

As relações conceituais em sistemas de organização do conhecimento : análise comparativa entre a classe 590 da Classificação Decimal de Dewey e a BBC Wildlife Ontology / André Felipe Pereira da Silva. -- Marília, 2019

103 f.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Filosofia e Ciências, Marília

Orientador: Walter Moreira

1. Relações conceituais. 2. Classificação Decimal Dewey. 3. Ontologias. 4. Sistemas de organização do conhecimento. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Filosofia e Ciências, Marília. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

ANDRÉ FELIPE PEREIRA DA SILVA

AS RELAÇÕES CONCEITUAIS EM SISTEMAS DE ORGANIZAÇÃO DO
CONHECIMENTO: ANÁLISE COMPARATIVA ENTRE A CLASSE 590 DA
CLASSIFICAÇÃO DECIMAL DE DEWEY E A BBC WILDLIFE ONTOLOGY

Dissertação para obtenção do título de Mestre em Ciência da Informação, da
Faculdade de Filosofia e Ciências, da Universidade Estadual Paulista – UNESP –
Campus de Marília, na área de concentração Produção e Organização da
Informação

BANCA EXAMINADORA

Orientador: _____

Prof. Dr. Walter Moreira
Universidade Estadual Paulista “Júlio de Mesquita Filho”
Faculdade de Filosofia e Ciências, Campus de Marília

2º Examinador: _____

Prof. Dr. Edberto Ferneda
Universidade Estadual Paulista “Júlio de Mesquita Filho”
Faculdade de Filosofia e Ciências, Campus de Marília

3º Examinador: _____

Prof. Dr. Rogério Aparecido Sá Ramalho
Universidade Federal de São Carlos

Marília, de de 2019.

Para meus pais, Antonia e Flávio que me forneceram suporte para minha educação e apoiaram minhas escolhas quanto aos estudos.

Para meu orientador Prof. Dr. Walter Moreira pelas pertinentes observações e recomendações que engrandeceram construção deste trabalho.

Para os docentes que puderam contribuir para a adequação do projeto de pesquisa e sua subseqüente realização.

AGRADECIMENTOS

À Deus, pois a Ele recorri nos momentos mais críticos da realização deste trabalho para me acalmar e prosseguir com o estudo mantendo o foco.

A meu pai Flavio Candido da Silva e a minha mãe Antonia Aparecida Pereira da Silva que me aconselharam e me estimularam a prosseguir com a pesquisa quando pensei em desistir.

A meu irmão Bruno Henrique Pereira da Silva e à sua esposa Luciana Cassalatti Silvestre da Silva por acompanharem, mesmo que de longe, meu processo acadêmico e por torcerem por meu sucesso

Ao meu Prof. Dr. Walter Moreira, por me orientar no decurso do mestrado, dedicando parte de seu tempo para me auxiliar na elaboração da pesquisa, além de estar disponível sempre que necessário para sanar dúvidas e colaborar com o andamento da dissertação. Também por ser um grande exemplo de professor e pesquisador que almejo me tornar um dia.

Aos mesmos da banca de qualificação, Prof. Dr. Edberto Fereda e Prof. Dr. Rogério Aparecido Sá Ramalho, por aceitarem dispor de tempo para analisar e a contribuir com o progresso deste trabalho

Aos docentes que de alguma forma colaboraram com a realização do trabalho desde de princípios referentes à pesquisa científica a ensinamentos sobre temas específicos que auxiliaram na construção teórica deste estudo.

Aos colegas mestrados e graduandos que acreditaram em mim e me deram apoio na realização do mestrado. Compartilhando ideias, frustrações e informações importantes para a sobre turbulência vida pós-graduanda.

Por fim, aos alunos da 41ª turma de Biblioteconomia da Universidade Estadual Paulista (UNESP) que, involuntariamente, participaram da realização de meu estágio-docência. Sendo participativos e respeitosos na aula que ministrei e no decorrer das aulas que assisti.

Terei, talvez um dia, sabedoria para aprender.
Mas sei que o viver bem, é deixar bem viver.
Sonhei com mais justiça pois, da cobiça eu já me
livrei. Eu sei que esse milênio traz nova lei. Essa
paz terei! Verei a nossa vitória entrar para
história sob um lindo tom. Mesmo se um dia o
sol se for, outra cor verei. Vai ser bom! O céu
fica tão cinzento, que parece sufocar, mas, pedir
com sentimento, faz o mundo melhorar.

(Alan Menken; Stephen Schwartz, O corcunda
de Notre Dame, 1996)

RESUMO

Aborda sobre as relações conceituais que se mostram presentes em todos os sistemas de organização do conhecimento. Ao organizar o conhecimento, delimita-se classes e cada classe desempenha relações com outras classes. Neste contexto, inserem-se a Classificação Decimal de Dewey (CDD) e as ontologias como forma de organizar o conhecimento físico e digital, respectivamente. Partindo-se do pressuposto de que os sistemas de organização do conhecimento se complementam, o objetivo geral desta pesquisa é o de comparar o modelo das relações conceituais presentes nos sistemas de classificação com o modelo de relações conceituais presentes nas ontologias. Para tanto, tomou-se como objeto de análise a classe 590 da CDD sobre Animais, e a *BBC Wildlife Ontology* e, assim, foram definidos como objetivos específicos: Compreender os tipos de relações conceituais e suas características; Analisar as relações conceituais da classe 590 da CDD; Investigar a consistência das relações conceituais na *BBC Wildlife Ontology*; Realizar análise comparativa sobre os tipos de relações conceituais identificados na CDD e na *BBC Wildlife Ontology*; Criar um modelo de ontologia que sintetize os dois sistemas de organização do conhecimento analisados para fins de sistematização. Para a realização desses objetivos adotou-se, em síntese, os seguintes procedimentos metodológicos: 1) coleta de dados a partir da literatura da área; 2) análise das relações conceituais da classe 590 da CDD; 3) análise das relações conceituais na *BBC Wildlife Ontology* com a metodologia *OntoClean*; 4) Baseamento na tipologia e nas definições das relações conceituais apresentadas na norma ISO 25964 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011). Como resultados notou-se que tanto a CDD como a ontologia analisada não estão totalmente adequadas e através da metodologia *OntoClean* criou-se uma ontologia modelo que fosse mais consistente. Nas considerações finais constatou-se que o alto nível de complexidade para organizar o conhecimento dificulta consideravelmente na criação de ontologias

Palavra-chave: Relações conceituais, Classificação Decimal Dewey, ontologias, sistemas de organização do conhecimento.

ABSTRACT

It deals with the conceptual relations that are present in all Knowledge Organization Systems. When knowledge is organized, classes are delimited and each class performs relationships with other classes. The Dewey Decimal Classification (CDD) and the ontologies are inserted in this context as a way of organizing the physical and digital knowledge, respectively. Assuming that the Knowledge Systems Organization complement each other, the general objective of this research is to compare the model of the conceptual relations present in the classification systems with the model of conceptual relations present in the ontologies. Therefore, the CDD class 590 about Animals and the BBC Wildlife Ontology were taken as the object of analysis and thus were defined as specific objectives. To understand the types of conceptual relationships and their characteristics; to analyze the conceptual relationships of class 590 of the CDD; To investigate the consistency of conceptual relationships in the BBC Wildlife Ontology; To conduct a comparative analysis of the types of conceptual relationships identified in the CDD and the BBC Wildlife Ontology; To create an ontology model that synthesizes the two Knowledge Organization Systems analyzed for systematization purposes. In order to achieve these objectives, the following methodological procedures were adopted: 1) data collection from the literature of the area; 2) analysis of the conceptual relationships of class 590 from CDD; 3) analysis of conceptual relationships in BBC Wildlife Ontology by the use of OntoClean methodology 4) Basis on the typology and definitions of conceptual relationships presented in ISO 25964 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011). As results it was noticed that both the CDD and the analyzed ontology are not totally appropriate and with OntoClean methodology it was possible to create a model ontology more consistent. In the final considerations, it was observed that the high level of complexity to organize the knowledge considerably hinders the creation of ontologies

Keywords: Conceptual relationships, Dewey Decimal Classification, ontologies, Knowledge Organization Systems.

LISTA DE FIGURAS

Figura 1 – Modelo para construção do conceito	17
Figura 2 – Exemplo de aplicação de propriedades	22
Figura 3 – Complexidade dos sistemas de organização do conhecimento	32
Figura 4 – Progressão das Classes.....	60
Figura 5 – Vizinhaça de Classes na CDD.....	60
Figura 6 – Relação hierárquica da CDD	61
Figura 7 – Exemplo do recurso ontograf do Protégé 5.1.0.....	66
Figura 8 – Essência e Rigidez da propriedade	72
Figura 9 – Vizinhaça da classe 590.....	76
Figura 10 – Rede de relações conceituais na BBC Wildlife Ontology	80
Figura 11 – Hierarquia das classes da ontologia modelo	91
Figura 12 – Propriedades da classe polvos.....	92
Figura 13 – Visão geral das classes e propriedades da ontologia modelo.....	92
Figura 14 – Propriedades da instância “PolvoPaul”	94

LISTA DE QUADROS

Quadro 1 – Definições de dado, informação e conhecimento.....	15
Quadro 2 – Comparação entre classificação e categorização	28
Quadro 3 – As dimensões que influenciam a organização do conhecimento	29
Quadro 4 – Síntese das ideias de Dahlberg (1978a).....	56
Quadro 5 – Tipos de relações das ontologias.....	67
Quadro 6 – Classe <i>adaptation</i> da <i>BBC wildlife ontology</i>	81
Quadro 7 – Classe <i>Adapted to extremes</i> da <i>BBC wildlife ontology</i>	81
Quadro 8 – Propriedade <i>adaptation</i> da <i>BBC wildlife ontology</i>	82
Quadro 9 – Classe nome taxonômico da <i>BBC wildlife ontology</i>	83
Quadro 10 – Classe <i>taxonomic rank</i> da <i>BBC wildlife ontology</i>	84
Quadro 11 – Sete passos para criar uma ontologia.....	89
Quadro 12 – Classes, entes abstratos e propriedades da ontologia modelo	90

LISTA DE ABREVIATURAS E SIGLAS

BARTOC	<i>Basel Register of Thesauri, Ontologies & Classifications</i>
BBC	<i>British Broadcasting Corporation</i>
CDD	Classificação Decimal de Dewey
CDU	<i>Classificação Decimal Universal</i>
Index.	Indexação
Inf.	Informação
ISKO	<i>International Society for Knowledge Organization</i>
ISO	<i>International Organization for Standardization</i>
KOS	<i>Knowlegde Organization Systems</i>
OCLC	<i>Online Computer Library Center</i>
ONU	Organização Nações Unidas
Org.	Organização
OWL	<i>Ontology Web Language</i>
RDF	<i>Resource Description Framework</i>
SOC	Sistemas de organização do conhecimento
TCC	Trabalho de Conclusão de Curso
TE	Termo Específico
TG	Termo Geral
TR	Termo Relacionado
UP	Usado Para
USE	Usar o termo preferido
W3C	<i>The World Wide Web Consortium</i>
XML	<i>Extensible Markup Language</i>

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Tema e Problema	13
1.2	Objetivos	21
1.3	Justificativa	21
1.4	Estrutura do Trabalho.....	23
2	ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO	26
2.1	Categorias	26
2.2	Organização do conhecimento	28
2.3	Sistemas de organização do conhecimento (SOC)	29
3	AS RELAÇÕES CONCEITUAIS	37
3.1	Tipos de relações conceituais	37
3.2	Características das relações conceituais	51
3.3	Relações conceituais na Classificação Decimal de Dewey	57
3.4	Relações conceituais na ontologia.....	62
4	PROCEDIMENTOS METODOLÓGICOS	69
4.1	Materiais	69
4.2	A metodologia <i>OntoClean</i>	71
4.2.1	Essência e Rigidez.....	71
4.2.2	Identidade e unicidade	73
4.3	Procedimentos de coleta de dados e de análise de dados.....	74
5	A CLASSE 590 DA CLASSIFICAÇÃO DECIMAL DE DEWEY	76
6	A <i>BBC WILDLIFE ONTOLOGY</i>	80
7	COMPARANDO A CLASSE 590 COM A <i>BBC WILDLIFE ONTOLOGY</i>	86
8	CONSIDERAÇÕES FINAIS	96
	REFERÊNCIAS	98

1 INTRODUÇÃO

A organização do conhecimento é uma disciplina que se preocupa em compreender de modo teórico e aplicado as questões que envolvem os meios pelos quais o conhecimento pode ser organizado. Para compreender melhor essa definição, é necessário, sobretudo, esmiuçar cada um dos termos para a compreensão do todo. Para isso, surgem duas questões preliminares e essenciais: para a Ciência da informação, o que é organização? E, o que é conhecimento?

Deixando a primeira pergunta de lado por um momento, prioriza-se a princípio o entendimento do conceito de “conhecimento” para que assim, então, seja compreensível o conceito de organização.

Para tanto, a discussão iniciar-se-á buscando a compreensão de três termos bastante utilizados no âmbito Ciência da Informação, sendo eles: dado, informação e conhecimento. Estes termos serão analisados, no entanto, não de modo isolado, mas de forma integrada, demonstrando, assim, um pouco da amplitude e da forte correlação entre os conceitos.

1.1 Tema e Problema

Para a Ciência da Informação, o conceito de “conhecimento” não pode ser analisado ignorando-se o “dado” e a “informação”, pois são conceitos que estão interligados e que, além disso, não podem ser definidos separadamente.

Lima e Alvares (2012) percebem essa interligação e ressaltam que a definição de um desses conceitos está ligada à definição do outro conceito. Sendo assim, utilizam-se as características de um conceito para compreender as características do outro.

Não se pretende neste trabalho, entretanto, definir a fundo cada um destes conceitos, mas, inicialmente, nesta fase do estudo, oferecer uma visão geral deles. Assim, ainda com base em Lima e Alvares (2012), pode-se entender dado como a menor parte que compõe uma informação e que quando isolado não pode ser decifrado de maneira clara e comum ao entendimento humano.

Na mesma linha de pensamento, o dicionário de organização do conhecimento de Barité (2013, p. 46, tradução nossa) define dado como “unidade mínima de

informação que contém um valor em si, que só se torna útil e expressivo na medida em que está associado a outros dados para uma finalidade específica". Assim, um dado numérico como "1945" quando associado a um grupo como "números divisíveis por 5" ou com outros dados como o fim da Segunda Guerra Mundial, por exemplo, pode produzir informações, pois qualquer conjunto de dados organizados possui potencial para ser transformado em informação.

Setzer (1999) ao perceber a dificuldade em definir e distinguir dado, informação e conhecimento, encarrega-se de elucidar cada um destes termos. Dessa forma, define dado como "uma sequência de símbolos quantificados ou quantificáveis. Nesse contexto, Setzer (1999) aponta uma característica interessante do dado. Já que os dados são quantificáveis, eles podem ser armazenados em um computador e processados por ele. Nesse entendimento, o dado é necessariamente uma entidade matemática e puramente sintática. Isso significa que os dados podem ser descritos em representações formais.

Dentro de um computador, trechos de um texto podem ser ligados virtualmente a outros trechos, por meio de contigüidade física ou por "ponteiros", isto é, endereços da unidade de armazenamento sendo utilizada. Ponteiros podem fazer a ligação de um ponto de um texto a uma representação quantificada de uma figura, de um som, etc. (SETZER, 1999, p. 2)

Como a definição de dado oferecida por Barité (2013) sugere, a informação é construída a partir de um conjunto de dados. Setzer (1999) caracteriza a informação como uma abstração informal que representa algo significativo para alguém através de textos, imagens, sons ou animação.

Para Setzer (1999, p. 2), a frase "Paris é uma cidade fascinante", por exemplo, é uma informação, "desde que seja lida ou ouvida por alguém, desde que 'Paris' signifique a capital da França e 'fascinante' tenha a qualidade usual e intuitiva associada com aquela palavra". Nesse exemplo, Paris é um dado, cidade também é um dado e fascinante é outro dado; juntos constituem uma informação.

Finalmente, o conhecimento é uma abstração interior, pessoal de algo que foi experimentado por alguém. No exemplo de Paris, alguém tem o conhecimento de Paris apenas se a visitou. O conhecimento não pode ser descrito em sua totalidade e também não depende apenas de uma interpretação pessoal, como a informação, pois requer uma vivência do objeto do conhecimento (SETZER, 1999).

Para facilitar a visualização das definições do que são dado, informação e conhecimento observe o Quadro 1:

Quadro 1 – Definições de dado, informação e conhecimento

Dado	É a unidade mínima de informação, uma sequência de símbolos. É um elemento que compõe a informação.
Informação	É um conjunto de dados que sejam representativos o suficiente para ser compreendido por algum indivíduo.
Conhecimento	É uma abstração interior, pessoal de algo que foi experimentado por alguém. É uma interpretação da informação.

Fonte: elaborado pelo autor

Entender esses três termos no âmbito da Ciência da Informação é fundamental para entender a disciplina que se preocupe em organizar o conhecimento.

Assim, retomando o conceito posto de lado anteriormente, a organização indica, no sentido mais amplo da palavra, o ato de colocar alguma coisa em alguma ordem. Numa biblioteca, por exemplo, a organização pode ser construída estipulando locais para livros em uma aproximação lógica, seja por assuntos, por autores, etc.

Organizar é algo que o ser humano faz cotidianamente, mas, afinal, para que organizamos? De acordo com Aulete e Valente (2018), organizar é o ato de “dar ou assumir (pessoas, coisas, elementos) ordem, disposição ou estrutura adequada para que preencham certas funções, configurações, finalidades”. Isso pode ser visto na Ciência da informação, sendo esta uma das principais preocupações da área.

Em resumo, organiza-se algo para que futuramente seja encontrado. Em uma biblioteca, por exemplo, os livros estão organizados nas estantes, não apenas como uma forma de armazenamento, mas também como uma forma de localização pelos usuários, sendo isto parte do processo de recuperação da informação.

É importante lembrar, ainda, que a organização, no exemplo citado, determina as classes a que tais livros irão pertencer. Neste sentido, a organização relaciona-se diretamente com a classificação.

Uma das principais preocupações da Ciência da Informação está em garantir a recuperação da informação pelos usuários. Assim, a organização do conhecimento é peça fundamental na Ciência da informação, para garantir a recuperação da informação, pois essencialmente organiza-se o conhecimento para que este possa ser encontrado.

Observando-se a questão por uma perspectiva mais ampla, é possível pensar que é preciso organizar os objetos e as ideias para que o mundo possa ser melhor

compreendido e, conseqüentemente, para que seja possível estabelecer melhores condições de comunicação e de produção de novos conhecimentos.

Neste contexto, Café e Sales (2010, p. 117) exemplificam que

[...] quando queremos explicar a alguém o significado de determinada palavra normalmente recorremos a sua classificação, ou seja a incluímos em um grupo maior que a define. Assim, ao explicarmos o que significa a palavra cadeira diremos que é um móvel (classe) que possui um assento e um encosto. Neste momento, organizamos nosso pensamento sobre o conceito desta palavra focando a classe dos móveis e excluindo outras classes.

Neste exemplo de Café e Sales é possível observar um aspecto interessante que pode passar despercebido numa leitura mais apressada: um conceito pode ser compreendido pelo receptor da mensagem quando é inserido em uma organização de classes, isto é, numa estrutura hierárquica. Deste modo, pode-se afirmar que a estrutura classificatória auxilia na conceitualização, ou na compreensão dos conceitos.

Ao mencionar o exemplo da cadeira, fica claro que quando um termo está relacionado a uma classe ou um domínio, isso exclui outras possíveis acepções deste termo. Então, se a cadeira é um tipo de móvel, logo, ela não é uma cadeira relacionada à Academia Brasileira de Letras, por exemplo. Neste contexto a organização dos conceitos para transmitir a ideia foi essencial.

Logo, ao relacionar a cadeira dentro de uma classe mais ampla denominada moveis, é correto afirmar que o conceito de cadeira exerce uma relação conceitual com o conceito de moveis. Como poderá ser observado mais à frente, cadeira e moveis desempenham uma relação conhecida como relação hierárquica.

No entanto, para compreender as relações entre os conceitos, as quais desempenham grande relevância a esta pesquisa, é necessário entender antes o que é um conceito.

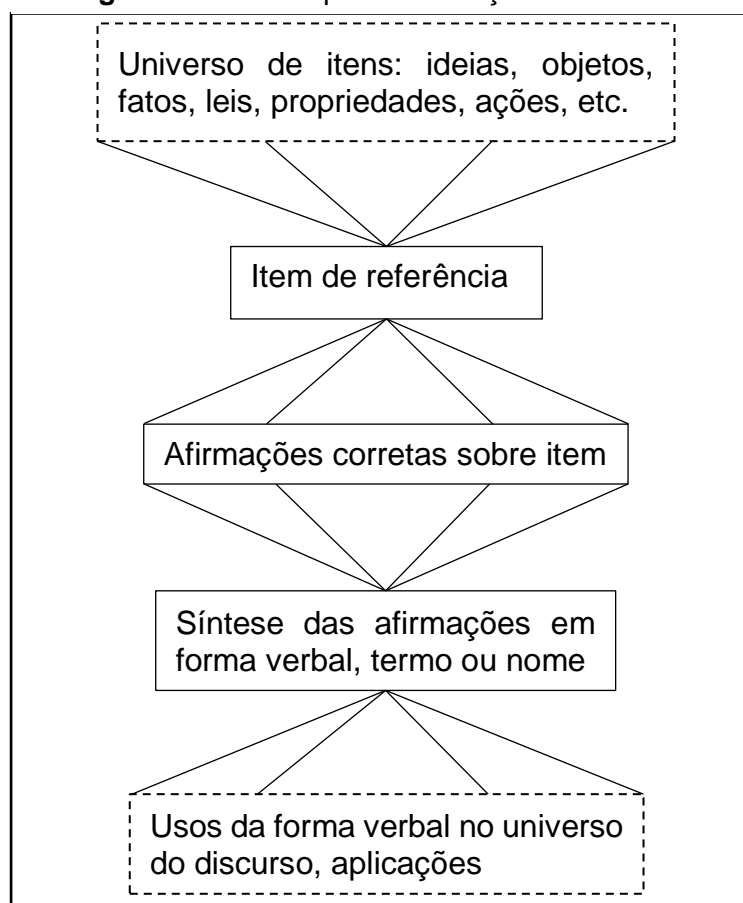
Nesse contexto, para a compreensão de alguns aspectos fundamentais do conceito, Ingetraut Dahlberg, pesquisadora da organização do conhecimento com formação original na Filosofia, desenvolveu a Teoria do Conceito.

Segundo Dahlberg (1978a), desde o momento em que os seres humanos começaram a pensar e a falar, o homem empregou símbolos para designar objetos em seu ambiente e traduzir os seus pensamentos a respeito destes objetos.

Dahlberg (1978b), considera que o conceito é uma "unidade do conhecimento". Segundo ela, o conhecimento pode ser definido como “a totalidade de proposições verdadeiras sobre o mundo, existindo - em geral - nos documentos ou nas cabeças das pessoas, então o conhecimento pode parecer existir também em todas as afirmações verdadeiras”.

Levando em conta que o ser humano é capaz de produzir afirmações corretas sobre coisas reais, Dalhberg (1978b) sugere um modelo para a construção de conceitos, conforme pode ser observado na Figura 1.

Figura 1 – Modelo para construção do conceito



Fonte: Adaptado de Dahlberg (1978b)

A Figura 1 mostra como que a partir de um universo de itens, seleciona-se um item como um item de referência ou o referente. Esses referentes podem ser “um simples objeto, um conjunto de objetos considerados como uma unidade, ou uma propriedade, uma ação, uma dimensão, etc. ou qualquer destas combinações”. As afirmações corretas sobre cada referente podem ser verificadas por meio de evidências. Essas afirmações são aceitas como verdadeiras numa forma verbal, que pode ser um termo ou um nome. Com essa forma verbal o ser humano pode, então

se comunicar verbalmente e por escrito a respeito dos conteúdos (os julgamentos, as opiniões sobre o referente, etc.) de um conceito, e também para aplicar um conceito nas afirmações, no universo do discurso, ou seja, o uso da forma verbal (DAHLBERG, 1978b).

Em outras palavras, o ser humano estabelece formas para representar seus pensamentos e, assim, reproduzir seus pensamentos e comunicar com os outros seres humanos.

A partir da necessidade do homem de se comunicar, surge a linguagem, que de acordo com Dahlberg (1978a), consiste na capacidade do homem em designar os objetos que o circundam e também permite ao homem comunicar-se com seus semelhantes.

Retomando a discussão para o entendimento do que é o conceito, é preciso lembrar que no momento em que o ser humano conhece algo, de certa forma apropria-se deste objeto conhecido, e esse conhecimento a respeito do objeto acaba transformando-o em conceito. Uma vez que se apreende um conceito, considerando-se que este conceito não é um objeto real, ele se torna uma unidade de pensamento, uma abstração, uma forma de conhecer a realidade (MIRANDA, 2005; SANTOS; MAZINI, 2011).

No entanto, para Dalhberg (1978b grifo do autor), o conceito pode sim ser considerado uma unidade, mas não uma unidade de pensamento, já que o pensamento é algo subjetivo e está apenas na cabeça do ser humano. O conceito para Dalhberg é uma unidade de conhecimento e o conhecimento é a totalidade de afirmações verdadeiras. Nas palavras de Dahlberg o conceito é “uma unidade do conhecimento, compreendendo **afirmações verdadeiras** sobre um dado **item de referência**, representado numa **forma verbal**.”

Hjørland (2009) tem um olhar semelhante a respeito do conceito de “conceito” fundamentado na ciência cognitiva. Hjørland ressalta que a função básica dos conceitos é fixar algo em um signo (ou na mente) para poder pensar sobre ele, comunicar sobre ele e agir em relação a ele.

O processo de construção de uma linguagem acontece a partir da compreensão de conceitos. Para se definir um conceito Dahlberg (1978a) constata que ele contém características que os encaixam dentro de categorias, sendo que as

categorias são entendidas, neste contexto, como o conceito em sua mais ampla extensão.

Essa extensão mencionada por Dahlberg refere-se à “soma total dos seus conceitos específicos e seus conceitos individuais, ou seja, o conjunto de conceitos para os quais a intensão de um conceito se torna verdadeira” (DAHLBERG, 1978b). Então, “a mais ampla extensão”, envolve aqueles conceitos específicos ou individuais que se aproximam.

Para Dahlberg (1978a):

[...] os elementos do conceito são obtidos pelo método analítico—sintético. Cada enunciado apresenta (no verdadeiro sentido de predicação) um atributo predicável do objeto que, no nível de conceito, se chama característica. Muitas vezes não se trata de um atributo a que corresponde uma característica, mas de uma hierarquia de características, já que o predicado de um enunciado pode tornar-se sujeito de novo enunciado e assim sucessivamente até atingirmos uma característica tão geral que possa ser considerada uma categoria.

Quando um “conceito ainda não tem nome, é possível formulá-lo pela síntese das características descobertas” neste conceito. Isso pode ser observado nos elementos químicos, os quais são muitas vezes definidos pelas características que esses elementos apresentam (DAHLBERG, 1978b).

As categorias defendidas por Dahlberg são consideradas características dos conceitos e com essas características é possível construir uma definição dos conceitos. Esta definição é feita utilizando-se de símbolos, como os termos, (ou palavras).

Categorização é o processo em que se divide o mundo em grupos ou categorias cujos membros compartilham algumas semelhanças perceptíveis dentro de um determinado contexto (JACOB, 2004).

Assim, uma categoria denominada “jogos” pode estar relacionada conceitualmente com a categoria “videogames” e também com a categoria “pôquer”, pois todos estes termos abrangem conceitos próximos e, dessa forma, podem ser relacionados conceitualmente.

Na Ciência da Informação, mais precisamente, na organização do conhecimento, os *Knowledge Organization Systems* (KOS), ou sistemas de organização do conhecimento (SOC) em português, surgem como mecanismos para facilitar a organização, a representação e recuperação da informação. Para este trabalho optou-se pela utilização da sigla em português.

Dentre os muitos SOC que surgem com o decorrer dos anos, este trabalho concentra-se, particularmente, em dois tipos: a Classificação Decimal de Dewey (CDD) e uma ontologia específica, a *BBC wildlife ontology*.

A CDD é um SOC que se utiliza de um esquema decimal para organizar o conhecimento registrado contido em bibliotecas. É um SOC feito para que os bibliotecários possam representar o conhecimento para a posterior recuperação.

As ontologias são um tipo de SOC feitas para fornecer mecanismos à máquina para conectar termos a outros e assim criar conceitos, de forma que os conceitos lá registrados possam ser associados a outros conceitos automaticamente pela máquina, por meio de mecanismos de inferência automática.

Nos SOC são utilizados símbolos (como os termos ou números) que representam conceitos. Além disso, esses símbolos que representam conceitos, também se relacionam com outros símbolos, fazendo com que, implicitamente, também representem tais conceitos.

Para compreender melhor o parágrafo acima considere o seguinte exemplo: o conceito de “casa” é um símbolo que se utiliza de três letras do alfabeto: “c”, “a”, “s” que combinadas formam uma palavra que representam um edifício habitacional. O símbolo alfabético que representa “casa” também se aproxima de outros conceitos alfabéticos como “quarto”, “banheiro”, “cozinha”, ou até mesmo o conceito de “lar”.

Essas relações podem ser associativas, hierárquicas ou equivalentes e também apresentam características que serão discutidas no capítulo 3 desta pesquisa.

Essas relações são as chamadas relações conceituais, e elas ocorrem de diversas maneiras, possuindo várias características que são hora comuns, hora distintas.

Diante de todos os tipos de relações conceituais presentes nos SOC e das características destas relações, surgem duas questões que configuram o problema desta pesquisa: Como as relações conceituais que se inserem em um SOC mais tradicional como a Classificação Decimal de Dewey podem complementar a criação de ontologias? Além disso, como a Classificação Decimal de Dewey pode ser engrandecida pelas ontologias?

Dessa forma, apresentam-se a seguir os objetivos geral e específicos desta pesquisa.

1.2 Objetivos

Tomando-se como pressuposto as relações de complementaridade entre os diferentes sistemas de organização do conhecimento, esta pesquisa tem como objetivo geral comparar as relações conceituais presentes nos Sistemas de Classificação com as relações conceituais presentes nas ontologias.

Para tanto, tomou-se como objeto de análise a Classificação Decimal de Dewey (CDD), observando-se especificamente a classe 590, referente a Animais, e a *BBC Wildlife Ontology*, referente a vida selvagem. Para tal, foram definidos os seguintes objetivos específicos:

- compreender os tipos de relações conceituais e suas características;
- analisar as relações conceituais da classe 590 da CDD;
- investigar a consistência das relações conceituais na *BBC Wildlife Ontology*;
- realizar análise comparativa sobre os tipos de relações conceituais identificados na CDD e na *BBC Wildlife Ontology*;

1.3 Justificativa

Este estudo surgiu após a realização do Trabalho de Conclusão de Curso (TCC), intitulado “Conversão de sistemas de classificação em ontologias: análise de viabilidade utilizando a CDD e o *Protégé*” (SILVA, 2017) que antecedeu esta pesquisa. Nesse TCC foi elaborada a conversão de uma classe da Classificação Decimal de Dewey (CDD) em uma ontologia.

Na criação da ontologia foi necessário compreender como a classe da CDD relacionava um termo com o outro. Notou-se então a necessidade de compreender melhor como as relações conceituais estão contidas nos sistemas de organização do conhecimento. O principal objetivo daquela pesquisa foi o de verificar se a conversão de sistemas de classificação bibliográfica em ontologias seria viável.

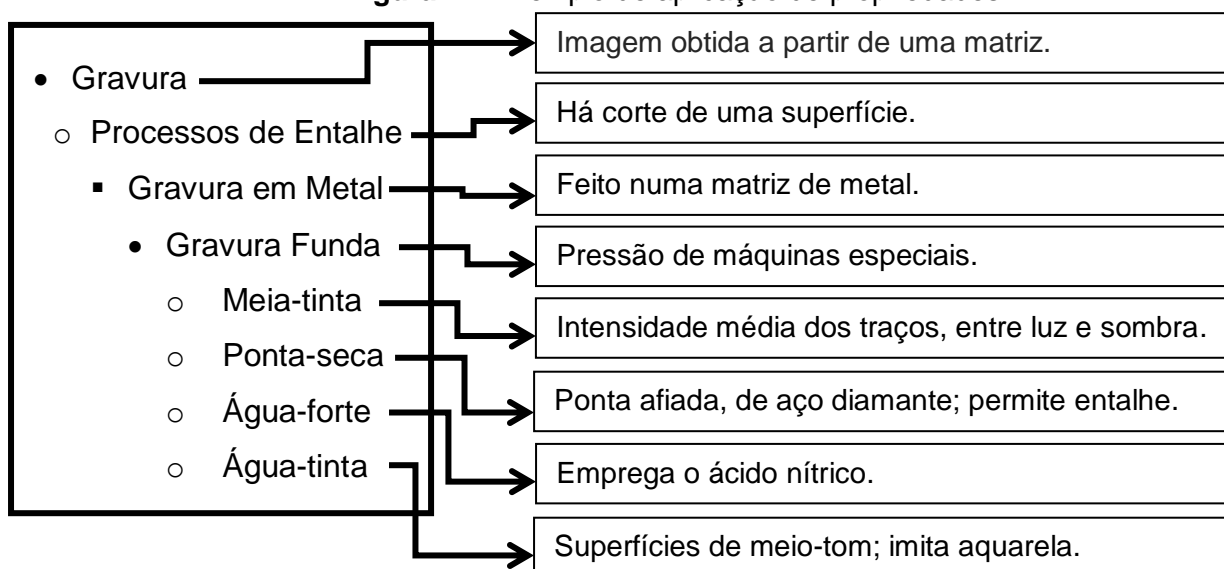
Para desenvolver aquela pesquisa, os dados coletados foram analisados adotando como procedimento de análise a comparação e sistematização das relações conceituais identificadas anteriormente na CDD com as que foram possíveis com o recurso da criação da ontologia, para os conjuntos estabelecidos. Para verificar a consistência das relações conceituais, tomou-se como base a tipologia e as definições

das relações conceituais apresentadas na norma ISO 25964 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011).

Uma ontologia necessita de propriedades para descrever um termo a fim de unificá-lo e essas propriedades são características de um termo. No TCC que guiou este estudo foi criada uma representação de algumas propriedades na classe utilizada. Dessa forma, observe que na Figura 2 cada termo apresenta uma característica. Cada um desses termos pode, obviamente, apresentar mais que uma característica, mas para fins de uma visualização menos complexa mostra-se apenas uma para entender como as propriedades funcionam.

No decorrer do trabalho houve dificuldade em definir os termos que melhor representassem um conceito e, em consequência disso, as relações conceituais não foram muito consistentes. Como, por exemplo, a gravura que aparece na Figura 2 é entendida por vezes como uma impressão, mas ao mesmo tempo a imagem obtida pelo processo da impressão. A impressão, por sua vez, também é uma classe na CDD e dessa forma precisa ser diferenciada da Gravura.

Figura 2 – Exemplo de aplicação de propriedades



Fonte: Silva (2017)

Assim, gravura foi definida como o resultado de uma impressão, enquanto que a impressão foi entendida como um processo no qual a imagem é estampada em um suporte como o papel, por exemplo.

O entendimento das relações conceituais da CDD é então importante, pois este é um sistema de organização do conhecimento que já está consolidado nas bibliotecas para organizar o acervo. Por estar tão presente na atividade organizacional do bibliotecário, é um instrumento que pode fornecer diversos benefícios para as

ontologias. As ontologias, podem aprender com a divisão decimal do conhecimento em classes elaborada por Dewey.

Por esta razão, esta pesquisa surge como uma maneira de aprofundamento quanto às relações conceituais presentes na CDD e nas ontologias como forma de dar uma base para a criação de novas conversões entre esses instrumentos, como que foi a realizada no TCC mencionado.

A fim melhorar a definição de termos na criação de ontologias, este trabalho busca por meio da estrutura de relações conceituais da CDD dar base para elaboração de ontologias.

Esta pesquisa investiga se as relações conceituais da CDD podem auxiliar a criação das ontologias e da relação entre os conceitos. Além disso, investiga se as ontologias podem ser colaborativas para estrutura classificatória da CDD.

A próxima seção apresenta-se a estrutura deste trabalho percorrendo todos os capítulos seguintes e expondo o que contém em cada um deles. Os procedimentos metodológicos serão apresentados após a fundamentação teórica, mais precisamente no capítulo 4. Optou-se por esta estruturação para que o leitor compreenda com mais profundidade a temática e o referencial teórico utilizado na pesquisa na expectativa de que isso facilite e torne mais proveitosa a compreensão dos resultados.

1.4 Estrutura do Trabalho

Esta dissertação se divide em oito capítulos sendo que os três iniciais, incluindo a introdução, possuem os pressupostos teóricos referentes aos temas de cada capítulo. Posteriormente a isso, o capítulo 4 apresenta os procedimentos utilizados na metodologia deste trabalho. Após os procedimentos metodológicos, apresentam-se dos capítulos 5 ao 7 os resultados obtidos por meio desta pesquisa. Por fim, apresentam-se as considerações finais seguidas das referências.

Assim, neste capítulo 1 de introdução, está contido o tema geral de pesquisa juntamente ao problema. Além disso, este capítulo expõe os objetivos geral e específicos da dissertação, bem como a estrutura geral do trabalho.

O capítulo 2, denominado “Organização e representação do conhecimento”, apresenta uma discussão geral sobre os conceitos de organização e representação do conhecimento, buscando oferecer uma base teórica para a compreensão das ontologias e da CDD.

O capítulo 2 se divide em três subcapítulos, sendo eles: o subcapítulo “2.1 Categorias” que mostra as definições para categoria e a diferença dela com a classificação; o subcapítulo “2.2 organização do conhecimento” que traz as principais dimensões dessa área; o subcapítulo “2.3 sistemas de organização do conhecimento” que especifica os sistemas dentro da área da organização do conhecimento demonstrando os principais objetivos dos estudos nesse tema.

Na sequência, o capítulo 3 identificado como “As relações conceituais” aborda a parte central deste trabalho demonstrando as relações conceituais em sistemas de organização do conhecimento.

Esse capítulo contém 4 subcapítulos: o subcapítulo “3.1 Tipos de relações conceituais” demonstra a variedade de relações conceituais presentes em sistemas de organização do conhecimento; o subcapítulo “3.2 Características das relações Conceituais” mostra a maneira como as relações conceituais acontecem; o subcapítulo “3.3 Relações conceituais na Classificação Decimal de Dewey” é responsável por elucidar as relações conceituais presentes na Classificação Decimal de Dewey o que será essencial para a compreensão dos resultados da pesquisa; o subcapítulo “3.4 Relações conceituais na *BBC wildlife ontology*” Assim como o anterior se propõe a demonstrar as relações conceituais na *BBC wildlife ontology* objetivando também a compreensão dos resultados da pesquisa.

O capítulo 4 é o primeiro após os pressupostos teóricos e apresenta os procedimentos metodológicos usados para desenvolver esta pesquisa. Ele é apresentado após os pressupostos teóricos para que o leitor já tenha embasamento suficiente para compreender satisfatoriamente a execução e resolução desta pesquisa.

Esse capítulo está dividido em três subcapítulos: o subcapítulo “4.1 Materiais” enumera os itens utilizados para a construção deste trabalho; o subcapítulo “4.2 Metodologia *OntoClean*” expõe a metodologia usada para avaliar a consistência das relações conceituais presentes nas ontologias estudadas neste trabalho. Esse subcapítulo se divide em mais 2 subcapítulos: “4.2.1 Essencia e Rigidez” “4.2.2 Identidade e Unicidade” que são elementos da metodologia *OntoClean*; o subcapítulo “4.3 Procedimentos de coleta de dados e de análise de dados” é responsável por mostrar como foi realizado este trabalho.

O capítulo 5 introduz aos resultados de pesquisa e nele está contida a análise das relações conceituais na classe 590 referente a animais na Classificação Decimal de Dewey.

O capítulo 6 desempenha um trabalho semelhante ao do capítulo 5 e analisa a *BBC wildlife ontology* referente a vida selvagem com auxílio da metodologia *OntoClean* usada para esse fim específico.

Após a análise dos sistemas de organização do conhecimento nos capítulos anteriores, o capítulo 7 é responsável por compará-los identificando semelhanças e diferenças em ambos.

Após a apresentação dos resultados, apresentam-se as considerações finais desta pesquisa no capítulo 8. Nas considerações finais demonstram-se como os objetivos da dissertação foram cumpridos e de que forma isso aconteceu.

Desse modo, encerra-se aqui a introdução desta pesquisa e inicia-se a apresentação da fundamentação teórica.

2 ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO

No âmbito da Ciência da Informação, mais especificamente na área da organização do conhecimento, uma das principais preocupações para o profissional está em tornar o conhecimento registrado, tanto em ambiente físico como em ambiente digital, o mais acessível possível aos usuários e de forma satisfatória, isto é, dar aos usuários mecanismos que possam facilitar a recuperação da informação.

Na prática, a organização do conhecimento preocupa-se com a melhor maneira de se representar e organizar um conhecimento para que este conhecimento possa, então, tornar-se facilmente recuperável por um usuário em um ambiente informacional.

Nesse contexto, é correto dizer que a finalidade da organização de um determinado sistema de informação é, sua efetiva compreensão e a consequente recuperação e uso dos objetos informacionais que envolve.

Por objetos informacionais, compreendem-se aqueles em que a informação foi registrada em diferentes tipos de suportes, tais como textos, imagens, registros sonoros, mapas etc. (CAFÉ; SALES, 2010).

Enfim, conforme Smit (1987) sintetiza, reúne-se e organiza-se a informação para que se possa achá-la. Além de oferecer, também, condições para a produção de novos conhecimentos.

Este capítulo sobre a organização e representação do conhecimento está dividido em subcapítulos, afim de aprofundar a discussão sobre os sistemas de organização do conhecimento.

2.1 Categorias

É compreensível que na Ciência da Informação a organização do conhecimento e, por consequência dela, a delimitação de categorias para organizar o conhecimento, é imensamente importante para o acesso e uso da informação.

A divisão do conhecimento por categorias, como lembra Jacob (2004), por exemplo, permite guiar os indivíduos, ou usuários, para descobrir uma ordenação em um ambiente que seja complexo. Ou seja, identificar, descrever e explicitar a complexidade que possa existir em um ambiente informacional.

Assim, entende-se que compreender o que é categoria é vital para esta parte do trabalho, pois é um elemento de extrema relevância para organização e para a representação do conhecimento. Além de ser um componente dos sistemas de organização do conhecimento.

Categoria é um recurso utilizado para representar o conhecimento. O termo “representação”, de acordo com Alvarenga (2003, p. 3), significa inserir “algo em lugar de”, ou seja, o termo representação está ligado ao termo substituição.

Partindo-se do mesmo princípio Lima e Alvares (2012) destacam que o ato de representar diz respeito à utilização de elementos simbólicos, tais como palavras, figuras, desenhos, imagens, mímicas, esquemas, entre outros, para substituição de um objeto, uma ideia ou um fato.

Barité *et al.* (2013, p. 127, tradução nossa) dizem que a representação do conhecimento é o “ramo da organização do conhecimento que envolve o conjunto dos processos de simbolização notacional ou conceitual do conhecimento humano no âmbito de qualquer disciplina”.

A simbolização notacional mencionada refere-se à substituição de um conceito por um símbolo. Como no caso da Classificação Decimal de Dewey que se utiliza de notações numéricas como será discutido mais à frente em seu capítulo correspondente.

Barité *et al.* (2013, p.35, tradução nossa), em seu dicionário de organização do conhecimento definem que a categoria, dentro da teoria da classificação, pode ser concebida como “cada forma ou classe fundamental de ser que podem ser agrupadas a elementos relativos a um objeto”. Portanto, pode-se observar que, em síntese, nesta definição “categoria” é entendida como uma classe.

É interessante notar que por muitos anos o conceito de “categoria” tem sido confundido com o conceito de “classe”, e isso resulta na concepção inadequada destes termos como sinônimos, conforme analisou Jacob (2004). Para ela os conceitos de classificação e de categorização referem-se a coisas distintas.

Essa distinção pode ser observada no Quadro 2.

Quadro 2 – Comparação entre classificação e categorização

Classificação	Categorização
Responsabiliza-se em separar um universo de entidades dentro de um sistema de classes que são mutuamente exclusivas e não sobrepostas, sendo elas organizadas em um contexto conceitual.	Ocupa-se em dividir grupos ou categorias nos quais seus membros sejam semelhantes dentro de um contexto imediato e por ser imediato é também um contexto variável.

Fonte: Elaborado pelo autor com base em Jacob (2004)

Entende-se, então, que a classificação é um sistema que se preocupa em eliminar ambiguidades e em separar o mundo em classes. Sendo que estas classes são específicas e inflexíveis.

A Categorização, conforme o nome sugere, envolve a delimitação de categorias. Categorias estas que agrupam indivíduos semelhantes em diferentes contextos versáteis podendo ser expansíveis.

As categorias podem ou não estar em um esquema de classificação. Sendo assim, pode-se afirmar que as classes são mais amplas que as categorias.

Dessa forma, baseando-se no que foi proposto por Jacob (2004), a categorização e a classificação não são termos que podem ser utilizados como sinônimos, mas, sim, como distintos.

2.2 Organização do conhecimento

Regressando a discussão para a área da organização do conhecimento em si, evoca-se Hjørland (2013) que ressalta que esta área se baseia nos princípios da classificação do conhecimento para organizar os conceitos e suas relações semânticas.

Para entender essa premissa proposta por Hjørland (2013), ao se definir, por exemplo, o conceito de “gato” (*Felis catus*) e sua relação com outros conceitos como “mamífero” (*Mammalia*), pode-se perceber uma relação do tipo gênero espécie, pois um gato (espécie) “é um” mamífero (gênero).

Toda organização do conhecimento, de acordo com Barité (2001), é artificial, uma vez que surge do resultado de uma construção abstrata. Além disso, a construção abstrata é criada com base em pressupostos particulares.

Neste contexto, Gnoli (2011) ressalta que na organização do conhecimento há influência de algumas dimensões (ou pressupostos). As principais dimensões

reconhecidas na área organização do conhecimento são as dimensões **ontológicas**; as dimensões **epistemológicas**; e as dimensões **pragmáticas** como pode ser observado no Quadro 3.

Quadro 3 – As dimensões que influenciam a organização do conhecimento

DIMENSÕES	Com o que lidam?
Ontológicas	Fenômenos de estudo reais
Epistemológicas	Perspectiva (propósito, alvo, cultura, disciplina, teoria, método, etc.) em que tais fenômenos são considerados e discutidos
Pragmáticas	Aspectos materiais dos documentos, suas coleções e seus usuários.

Fonte: elaborado pelo autor

Na organização do conhecimento, cuidar das informações que vão surgindo com o passar do tempo acaba por se tornar um trabalho que é mais complexo e ao mesmo tempo é também mais necessário. Em outras palavras, a vastidão de informações que surgem no decorrer dos anos faz com que a tarefa de organizá-las seja mais complicada, e por ser mais complicada cria a necessidade de encontrar formas de organizar essas informações.

Portanto, a elaboração de mecanismos que possam auxiliar os profissionais que trabalham com a organização do conhecimento mostra-se de interesse vital à Ciência da Informação.

2.3 Sistemas de organização do conhecimento (SOC)

Ao analisar a literatura da área da Ciência da Informação, é notável que sistemas de classificação, tesouros e ontologias são constantemente comparados como instrumentos que se complementam (GARCÍA JIMÉNEZ, 2004; GARCÍA MARCO, 2007; BARITÉ ROQUETA, 2011).

De forma majoritária, esses estudos buscam a compreensão da identidade de cada um desses instrumentos, isso é, buscam identificar suas características essenciais para que seja possível, também, identificar suas relações de complementaridade.

Estudos comparativos entre tipos de SOCs já são, aliás, bem antigos, considerando-se a velocidade das publicações. Soergel compara classificação com ontologias e demonstra que:

Uma classificação, por qualquer outro nome, ainda é uma classificação. O uso de um termo diferente é sintomático da falta de comunicação entre as comunidades científicas. O vasto campo de

conhecimento sobre estrutura da classificação e sobre modos de apresentação de classificações desenvolvidos em torno das classificações bibliográficas e na Ciência da Informação de modo mais amplo e o enorme capital intelectual incorporado em muitos esquemas de classificação e tesouros é largamente ignorado (SOERGEL, 1999, p. 1120, tradução nossa).

Entretanto, os trabalhos que abrangem esta temática parecem preocupar-se apenas com a simples oposição entre esses instrumentos, enfocando gerar uma concepção de que há uma certa continuidade, estando as taxonomias num extremo, passando pelas classificações e pelos tesouros em algum lugar intermediário e chegando nas ontologias num outro extremo (SOUZA; TUDHOPE; ALMEIDA, 2012; LARA, 2015).

Na literatura da ISKO (GARCÍA MARCO, 1993; 1995; HJØRLAND, 2012) tem-se agrupados estes instrumentos com uma denominação que os generaliza. Nomeiam-se estes instrumentos como já mencionado na introdução desta pesquisa como sistemas de organização do conhecimento (SOC).

Nas palavras de Hodge (2000), a expressão “sistemas de organização do conhecimento”, tem a intenção de abranger todos os tipos de instrumentos ou esquemas que sirvam para organizar a informação e para promover a gestão do conhecimento.

Além disso, os SOC também incluem

esquemas de classificação e categorização que organizam materiais em um nível geral, cabeçalhos de assunto que fornecem acesso mais detalhado e arquivos de autoridade que controlam versões variantes de informações-chave, tais como nomes geográficos e nomes pessoais. [...] também incluem vocabulários altamente estruturados, como tesouros, e esquemas menos tradicionais, tais como redes semânticas e ontologias. Como os sistemas de organização do conhecimento são mecanismos para organizar a informação, eles estão no coração de cada biblioteca, museu e arquivo. (HODGE, 2000, p. 1, tradução nossa).

A literatura brasileira da Ciência da Informação, também tem dado algumas definições. Lara (2015, p. 92), por exemplo, diz que os SOC

[...] compreendem uma série de instrumentos que tratam do ‘conteúdo’ dos documentos para fins de armazenamento e recuperação ou para promover sua gestão. Englobam os sistemas de classificação, listas de cabeçalhos de assunto, tesouros, taxonomias, terminologias, ontologias, além de listas de autoridade e anéis de sinônimos, instrumentos esses que se assemelham e se diferenciam pelo grau de estruturação semântica de seus elementos.

Na visão de Lara (2015, p. 90), as propostas de tipologias para os SOC costumam unir os objetos em grupos investigando os seus graus de profundidade semântica, níveis de estruturação dimensões e propósitos, “muitas vezes recorrendo a recursos de visualização”, este agrupamento contribui, certamente, para a melhor compreensão dos instrumentos e das relações que guardam entre si.

As tipologias abordadas por Lara (2015) revelaram a necessidade de se organizar o conhecimento para que se possa melhor compreendê-lo. Ao compreender o conhecimento, tenta-se encontrar formas de agrupamento que têm como base a identificação de traços ou atributos similares. No entanto a autora destaca que este procedimento se torna mais simples quando se refere a objetos concretos.

Os sistemas de organização do conhecimento incluem também vocabulários que são estruturados. Dentre esses SOC podem-se citar desde instrumentos que são mais tradicionais como os tesouros e instrumentos que são menos tradicionais, ou mais recentes, como acontece com as redes semânticas e as ontologias (SOUZA; TUDHOPE; ALMEIDA, 2012).

Desses estudos que agrupam os diferentes tipos de sistemas de organização do conhecimento a fim de compará-los, podem-se apontar três objetivos básicos:

- a) aprofundar a compreensão dos conceitos de sistemas de classificação, tesouros e ontologias por meio de comparações que envolvem suas naturezas e funções;
- b) compreender de que modo as ontologias podem ser beneficiadas com o conhecimento teórico-metodológico acumulado no desenvolvimento de instrumentos mais tradicionais (ou mais antigos) no âmbito da Ciência da Informação e
- c) compreender o movimento contrário ao indicado no item “b”, isto é, verificar como os sistemas de classificação e os tesouros podem ser flexibilizados e ampliados em sua estrutura relacional com o uso de ontologias.

Dessa forma, é válido recordar que com a vastidão de informações sendo produzidas diariamente tanto em suporte físico: livros, revistas, etc., como em suporte digital: artigos em revistas eletrônicas, *e-books*, etc. com o passar dos anos houve a necessidade de criar meios pelos quais toda essa informação emergente fosse organizada. Para isso, sistemas para organizar todo esse acúmulo informacional foram surgindo.

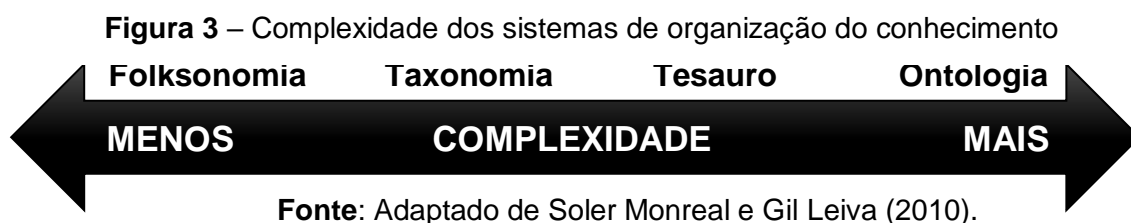
Nesse contexto, Bräscher e Carlan (2010) esclarecem que os sistemas de organização do conhecimento são instrumentos que interessam à Ciência da Informação, pois cumprem o importante papel de padronização da terminologia adotada para organização e recuperação de informações.

Além disso, os SOC são responsáveis por delimitar o uso de termos e definir os conceitos e as relações existentes em alguma área do conhecimento. Esses sistemas são provenientes do processo de se organizar o conhecimento.

Por fim, sistemas de organização do conhecimento são recursos utilizados para a representação da informação para que esta representação ocasione sua posterior recuperação. Em outras palavras, pode-se dizer que os SOC buscam fazer com o que os usuários, sejam eles humanos ou não, encontrem as informações por eles procuradas. Entende-se como usuários não humanos as máquinas e, dessa forma, os SOC voltados a elas procuram possibilitar que façam inferências lógicas simulando assim, em alguns aspectos, o raciocínio humano.

Por meio das comparações feitas nos estudos que agrupam os SOC pode-se eleger níveis de complexidade que os diferencia. Isso é perceptível em Soler Monreal e Gil Leiva (2010). Os autores relatam as diferenças de complexidade que há entre os principais instrumentos de representação do conhecimento.

Esses diferentes níveis de complexidade podem ser analisados na Figura 3 a qual mostra que a folksonomia encontra-se no extremo menos complexo dentre os SOC e a ontologia encontra-se como um instrumento com maior nível de complexidade.



Além dos instrumentos presentes na Figura 3 é apropriado acrescentar sistemas de classificação como a Classificação Decimal de Dewey (CDD) e a Classificação Decimal Universal (CDU).

A CDD é um sistema de classificação para organizar todo o conhecimento presente em bibliotecas. Serve para inserir os conteúdos dos livros em assuntos

sugeridos. CDU é um sistema que baseou-se na CDD para estabelecer as classes que, por sua vez representam assuntos. De acordo com Cintra et al. (1994), a CDD e a CDU são instrumentos que pretendem referir-se ao universo global de conhecimento tendo se baseado nas referências postuladas por Francis Bacon para a Organização do Conhecimento.

No entanto, Cintra et al. (1994, p. 42) estabelece que os tesouros partem de outros princípios: “[...] voltam-se para domínios cada vez mais particulares” que vão sendo construídos “em função de universos muito determinados. Por isso, “[...] são mais flexíveis quanto a estruturação do esquema classificatório básico e mais adequados ao atendimento das necessidades informativas de domínios especializados”.

Somado a isso pode-se inserir a CDD e a CDU no contexto da Figura 3 como sendo sistemas mais complexos que os tesouros e menos complexos que as ontologias.

Um dos SOC presentes na Figura 3 é o tesouro. A palavra tesouro tem suas origens no latim *thesaurós* e também no grego *thesauru*, tendo estes termos o significado de “tesouro” ou “repositório”, “armazenagem”. Esse termo, porém, carrega uma ampla gama de interpretações, como vocabulário, dicionário ou léxico (MOREIRA, 2003). Estes últimos, evidentemente, aproximam-se mais da organização do conhecimento, pois levam em conta o sentido, o significado das palavras.

O termo tesouro se popularizou pela criação de Peter Mark Roget, cujo dicionário de palavras, chamado *Thesaurus of english words and phrases* (Tesouro de frases e palavras inglesas, em tradução nossa), continha uma organização diferenciada que se sobressaiu em relação às que eram utilizadas pelos demais instrumentos da época, já que privilegiou uma organização em acordo ao significado das palavras, agrupando-as por significados próximos e não apenas pela ordem alfabética (DODEBEI, 2002; MOREIRA, 2003; SALES; CAFÉ, 2009; SOLER MONREAL; GIL LEIVA, 2010).

Conforme expõe Dobebei (2002), tal estrutura de agrupamento tinha o objetivo de localizar palavras que tivessem as melhores ideias expressadas em textos. As palavras, da língua inglesa, foram organizadas por Roget levando em conta seis categorias conceituais: as relações abstratas; o espaço; a matéria; o intelecto; a volição; e as afeições.

Após esta divisão das categorias, houve uma subdivisão das palavras em classes e essas, por sua vez, foram divididas em seções, subseções e conceitos isolados (DODEBEI, 2002).

O tesouro é uma ferramenta dinâmica pois se modifica continuamente. Segundo Cintra et al. (1994, p. 34), essa dinamicidade presente no tesouro pode ser percebida por sua permanente atualização, “[...] mediante operações de supressão de termos em desuso, reagrupamento de descritores em função da existência de palavras raramente utilizadas e/ou adição de termos novos”.

Esta flexibilidade dos tesouros segundo Cintra et al. (1994, p. 57) está conectada a um princípio de utilidade. Os tesouros podem ser atualizados sempre que for necessário. Dessa forma, “Cada um deles procurará organizar um dado universo nocional, de acordo com o ponto de vista que se imprime ao domínio, para responder a diferentes necessidades.”.

Isso tudo se deve ao fato de que o tesouro procura se adaptar ao usuário para que este possa usufruir de experiências mais satisfatórias. Assim, a interação que se deseja estabelecer na relação entre o usuário e o documento faz com que o fluxo da construção, manutenção e uso de uma linguagem documentária, e neste sentido específico o tesouro, obedeça a uma sequência lógica de passos.

Nesse contexto, Moreira (2003) ressalta que desde os primórdios até hoje em dia, o tesouro tem evoluído diante das construções teóricas e metodológicas. Além disso, tem também evoluído com a introdução de novos modelos cognitivos e da abordagem focada no usuário.

Outro papel importante desempenhado pelo tesouro na organização do conhecimento, é o de controlar o vocabulário com mais afincamento do que outros sistemas de classificação. Nos sistemas de classificação mais clássicos, por exemplo, não há grande preocupação com o controle vocabular

Desse modo, além de controlar o vocabulário, o tesouro permite a organização de forma hierárquica, e isso exige “[...] tantos vértices que equivalem a classes, quantos forem os aspectos escolhidos para organizar o domínio de especialidade” (CINTRA et al., 1994, p. 57). Por isso, quanto mais complexo o domínio mais classes devem ser criadas para melhorar organização.

A partir disso, pode-se definir que o tesouro é um vocabulário controlado, além de ser formado por termos-descritores que estão relacionados de maneira semântica,

isto é, estão conectados pelo sentido, ou conceito, das palavras, e que atuam como instrumentos de controle terminológico. Sua estruturação pode ocorrer de forma hierárquica ou de forma associativa, sendo que sua principal utilização está presente na indexação e recuperação de informações por meio de seu conteúdo (SALES; CAFÉ, 2009).

De certa maneira, e observando-se apenas uma de suas funções, o tesauro funciona como uma espécie de dicionário melhorado pois agrupa os termos levando em conta os assuntos. E isso remete novamente à ideia de Roget, pois pode ser usado com a finalidade de auxiliar nas consultas dos usuários em suas buscas informacionais. Portanto, pode também auxiliar os indexadores no processo de indexação.

Enquanto os tesouros vêm ao longo dos anos desempenhando um importante papel aos pesquisadores, outro instrumento destinado à organização e representação do conhecimento que vem se destacando na vertente da comunidade acadêmica, a esse instrumento é empregado o nome de a ontologia.

As Ontologias são especificações formais que têm o objetivo de descrever estruturas conceituais de domínios específicos. São aparatos desenvolvidos inicialmente pela engenharia computacional.

Então, em resumo, as ontologias podem ser tomadas como tesouros mais complexos que podem ser utilizados por máquina.

As discussões acerca das ontologias, assim como acerca da Classificação Decimal de Dewey serão retomadas de maneira mais aprofundada, neste trabalho, ao falar especificamente de suas relações conceituais no capítulo 3. Por agora, basta entender a questão da complexidade destes recursos para compreender melhor a Figura 3.

A folksonomia, também presente na Figura 3, é um instrumento que surgiu no contexto externo ao serviço do profissional do tratamento da informação, pois nela o usuário é o responsável por selecionar os termos que são convenientes aos seus propósitos.

Dessa forma as folksonomias não são instrumentos que se preocupam com a padronização dos termos, já que os usuários os definem utilizando um princípio distinto.

Assim, é bom ressaltar que

Fora do contexto profissional [...]. Todo usuário de informação tem sua maneira própria de organizá-la, o que não significa que seus métodos sejam necessariamente eficientes e eficazes, pois muitas vezes utiliza termos que não têm significação real no objeto informacional indexado. Ainda assim, na visão desse sujeito, sua forma de organização é a que lhe parece mais adequada. (VIGNOLI; ALMEIDA; CATARINO, 2014 grifo nosso).

Nesse contexto surge a folksonomia, que é a atuação do usuário na atribuição de palavras-chave empregados para organizar, representar e recuperar objetos informacionais no ciberespaço, de forma a colaborar e compartilhar informação com outros usuários em espaços virtuais, como as redes sociais e *sites* com recursos *Web 2.0* (VIGNOLI; ALMEIDA; CATARINO, 2014).

A palavra folksonomia é um neologismo que surge da junção dos termos, em inglês, “*folks*” e “*taxonomia*”. A palavra *folks* é um termo da língua inglesa que significa “pessoas”, ou “povo”. Já a palavra *taxonomia* provem de dois termos, sendo eles: *taxis* e *nomos*, ambos do grego e significam classificação e gestão, respectivamente. Dessa forma, a folksonomia pode ser literalmente traduzida como gestão da classificação feita por pessoas (MUNK; MØRK, 2007).

A taxonomia é basicamente uma classificação sistematizada, servindo, a princípio, às ciências biológicas, mas por trabalharem com termos separados em classes também se mostram importantes instrumentos para a organização e para a recuperação do conhecimento, onde as classes se manifestam em ordem lógica, apoiada igualmente em princípios classificatórios (CAMPOS; GOMES, 2007).

Assim, após a compreensão desses sistemas apresentados, o próximo capítulo abordará a respeito das relações conceituais. Relações estas que estão presentes nesses sistemas de organização do conhecimento apresentados neste capítulo.

3 AS RELAÇÕES CONCEITUAIS

Em sistemas de organização do conhecimento podem ser usados recursos que possibilitem a padronização da informação, portanto, os SOC podem fazer o uso de vocabulários controlados para permitir uma recuperação da informação bem-sucedida. Um exemplo de SOC que se utiliza de controle vocabular são os tesouros (MAZZOCCHI, 2017).

Nos SOC incluem-se diferentes tipos de relações para que estas relações possam guiar os pesquisadores e também os profissionais que trabalham com a organização do conhecimento, como os indexadores e os classificadores, a selecionar os termos que sejam mais pertinentes para expressar cada conceito (CLARKE, 2001).

Para entender melhor essas relações, serão utilizados os ensinamentos de Bean e Green (2001) as quais publicaram um livro dedicado exclusivamente ao estudo destas relações conceituais.

Na organização do conhecimento as relações são entendidas como uma associação entre duas ou mais entidades ou entre duas ou mais classes de entidades. As relações conceituais, e posteriormente suas características, serão abordadas neste trabalho, inicialmente com base nas considerações da norma ISO 25964, mas também, naturalmente, usufruindo das discussões de outros autores da área. Por essa razão foi necessário compreender no capítulo anterior o que é um tesouro.

3.1 Tipos de relações conceituais

Como este trabalho utilizará da norma ISO 25964 sobre os tesouros, para entender as relações conceituais que serão abordadas na sequência juntamente com suas características, optou-se por utilizar os tesouros como exemplo.

Assim, em um tesouro, conforme a aceção que a documentação deu a esse termo, os termos, que representam conceitos, são organizados em uma estrutura semântico-relacional hierárquica e associativa. Nessas relações, é possível observar relações que ocorrem em maneira de subordinação, de superordenação e de coordenação entre os conceitos ou termos.

Para entender melhor as maneiras pelas quais as relações ocorrem, utiliza-se como exemplo um recorte de uma relação hierárquica presente na Classificação Decimal de Dewey:

Criação de animais domésticos

Cachorros

Gatos

Antes de qualquer coisa é importante observar que “criação de animais domésticos” é a classe mais geral, enquanto “cachorros” e “gatos” são subclasses mais específicas; enquanto isso, tanto “cachorros”, como “gatos”, contêm o mesmo nível de especificidade.

Assim, “criação de animais domésticos” está superordenado (mais geral) em relação a “cachorros” e “gatos”. “cachorros” e “gatos”, por sua vez, estão subordinados (mais específicos) em relação a “criação de animais domésticos”. Por sua vez, “cachorros” e “gatos” não estão nem superordenados nem subordinados, pois empregam um mesmo nível hierárquico. A esta situação, dá-se o nome de coordenação. A coordenação é quando dois termos, ou conceitos, estão relacionados em um mesmo nível de especificidade.

Para concluir a discussão sobre os tesouros, cabe apresentar as definições deste instrumento na norma ISO 25964-1. Nessa norma, o tesouro é definido como um:

vocabulário controlado e estruturado em que os conceitos são representados por termos, organizado de modo que as relações entre os conceitos são explicitadas e os termos preferidos são acompanhados de entradas remissivas para sinônimos ou quase-sinônimos (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, p. 12, tradução nossa).

Ainda, conforme a ISO 25964-1 pode-se ler a respeito de qual o propósito do tesouro:

[...] orientar o indexador e o pesquisador a selecionarem o mesmo termo preferido ou a combinação dos termos preferidos para representar um determinado assunto. Por esta razão, um tesouro é otimizado para a navegabilidade humana e para a cobertura terminológica de um domínio (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, p. 12, tradução nossa).

O tesouro é um instrumento que age como um controlador de vocabulário, desde sua aplicação pelo indexador até o momento em que o usuário exerce a recuperação para fins de comunicação com o sistema de informação, sua necessidade de informação. Por essa razão, o tesouro pode ser entendido como um instrumento de comunicação.

A norma ISO 25964-1 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011) demonstra que os tesouros possuem três tipos principais de relações entre conceitos, como já mencionado, as relações hierárquicas, as relações associativas e as relações de equivalência; essas relações possuem características próprias que serão expostas neste trabalho.

Em primeiro lugar, definir-se-á as três relações conceituais principais dos tesouros para então, assim, esmiuçar as características e particularidades de cada uma.

A Norma ISO 25964-1 ressalta que no tocante a relações entre conceitos, esse tipo de relações:

[...] são difíceis de demonstrar, uma vez que os conceitos existem apenas na mente e, portanto, a maneira convencional de exibi-los é mostrar cada relacionamento como aplicável entre os termos preferidos que representam os conceitos em questão. (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011 tradução nossa).

Dessa forma, Green (2001) informa que as relações hierárquicas são um tipo de relação de transitividade na qual há um elemento mais geral e outro mais específico.

Na relação hierárquica de gênero-espécie deve-se pensar que qualquer atributo que esteja presente ao gênero deve estar contido na espécie (MAZZOCCHI, 2017). Isso quer dizer que para cada espécie deve incluir ao menos uma característica do gênero.

Existem duas maneiras para identificar uma hierarquia de gênero-espécie:

os termos devem pertencer à mesma categoria conceitual (é uma condição necessária, mas não suficiente porque ambas, a lógica e a perspectiva hierárquica são compatíveis), e o termo deve ser validado com o teste “todos-alguns”. (MAZZOCCHI, 2017, tradução nossa).

O teste todos-alguns consiste em colocar os dois termos que estão relacionados nas duas seguintes frases e ver como eles se comportam: Todo “termo A” é um “Termo B”. Alguns “Termo B” são “Termo A”. Caso os termos que foram inseridos para substituir o Termo A e o Termo B estejam adequados, essa relação é hierárquica. Caso os termos inseridos não estejam adequados, essa relação é não-hierárquica.

Considerando essas regras para reconhecer uma relação hierárquica pode-se pensar em alguns exemplos: considere a relação entre pássaros e canários, o canário

é um tipo de pássaro. Melhor dizendo, os canários possuem características que o identificam como um tipo pássaro.

Logo, aplicando o teste de validade todos-alguns pode-se inferir que todos os canários (espécie) são pássaros (gênero); da mesma forma que alguns pássaros são canários. As duas afirmativas se encaixam apropriadamente ao teste “todos-alguns”, portanto, a relação canários-pássaros é uma relação hierárquica.

A Norma ISO 25964-1 determina que a relação hierárquica:

[...] deve ser estabelecida entre um par de conceitos quando o escopo de um deles estiver completamente dentro do escopo do outro. Deve basear-se em graus ou níveis de superordenação e subordinação, em que o conceito superordenado representa uma classe ou o todo e os conceitos subordinados referem-se a seus membros ou partes. (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, tradução nossa)

Ainda de acordo com a Norma há um outro teste lógico para validar uma relação hierárquica. O segundo teste recomenda que “todo conceito subordinado deve pertencer à mesma categoria inerente ao seu conceito superordenado, ou seja, o termo mais amplo e mais restrito deve representar uma coisa, uma ação ou uma propriedade etc.”

Assim, observe duas situações:

- a) O termo “Metais” (uma classe de materiais) e o termo “fundição” (uma ação) representam conceitos distintos, portanto, não podem ser relacionados hierarquicamente.
- b) Por outro lado, tanto o termo “metais” como o termo “ouro” representam materiais e podem, portanto, estar hierarquicamente relacionados.

As relações hierárquicas possuem três tipos de relação aceitos no padrão internacional. Clarke, (2001) denomina essas relações como: genéricas, de instância e todo-parte. Já para Araujo (2006), as relações são denominadas como genéricas, específicas e partitivas. Para Mazzocchi (2017), estes tipos de relações são nomeados como genéricas, de instância e partitivas.

Estas denominações se referem às características semelhantes, portanto para este trabalho utilizará a seguinte denominação: genéricas, de instância e partitivas conforme a norma ISO 25964-1 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011).

Segundo Araujo (2006), nas relações genéricas os conceitos que estão superordenados têm as mesmas características daqueles que estão subordinados. Então, pode-se entender que as relações genéricas são aqueles em que os dois termos possuem características comuns.

Na relação gênero-espécie, está subentendido que a espécie herda características do gênero (MAZZOCCHI, 2017).

Clarke (2001), diz que as relações genéricas são aquelas em que o termo mais amplo nomeia uma classe ou categoria e cada um dos termos específicos indica um subconjunto dessa classe ou categoria, mas não uma instância individual dela.

Assim, percebe-se que nas relações hierárquicas genéricas os termos passam por um processo de herança na qual o termo mais específico (espécie), herda características do termo mais amplo (gênero).

A Norma ISO 25964-1 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011) determina que a relação genérica é conexão entre uma classe ou categoria e seus membros ou espécies. Além disso, a Norma estipula que deve estar em acordo ao teste todos-alguns e ao teste lógico mencionados anteriormente.

Já nas relações hierárquicas de instância, os termos específicos compartilham das mesmas características do termo que lhes é imediatamente superordenado, porém, possuem uma característica diferente a mais. Neste contexto, para Araujo (2006, p. 53), “a superordenação caminha das diferenças para as semelhanças, ou seja, da espécie para o gênero, enquanto a subordinação faz o caminho inverso, indo das semelhanças para as diferenças, do gênero para a espécie.”.

Clarke (2001) descreve que neste tipo de relação os termos mais específicos não são partes ou tipos das instâncias mais amplas e sim características individuais deles. Assim como descreveu Araujo (2006), os termos específicos possuem pelo menos uma característica que os difere dos demais.

Para este tipo de relação Clarke (2001) expõe dois exemplos que serão estruturados de acordo com a terminologia usada no tesauro. Sendo que o termo mais amplo é representado por TG (termo geral) e o termo mais específico é representado por TE (termo específico). Os exemplos mencionados são:

TG Línguas Clássicas

TE Latim

TE Sânscrito

TG Desertos

TE Deserto de Gobi

TE Deserto de Kalahari

TE Deserto de Mojave

TE Deserto do Saara

De modo a ficar mais claro, a Norma 25964-1 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, p. 6, tradução nossa) explica que “a relação de instância, vincula um conceito geral, como uma classe de coisas ou eventos e uma instância individual dessa classe, que geralmente é representada por um nome próprio”.

Então, por exemplo:

TG Regiões montanhosas

TE Alpes

TE Himalaia

O termo geral “Regiões montanhosas” representa a Classe. Já os termos específicos “Alpes” e “Himalaia” representam as instâncias, isto é, um nome, uma característica.

Por fim, as relações hierárquicas partitivas, segundo a Norma 25964-1 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, p. 60, tradução nossa) abrangem um limitado intervalo de situações. Dessa forma, as relações partitivas, partem “de uma entidade ou sistema pertence exclusivamente a um todo em particular”.

Clarke (2001) destaca que a relação partitiva é geralmente aceita como hierárquica quando acontece entre quatro tópicos específicos: entidades geográficas, entre partes do corpo, disciplinas de estudo e estruturas sociais.

Conforme Araujo (2006, p. 53), as relações partitivas se assemelham às genéricas. Porém, por estarem vinculadas aos tópicos descritos no parágrafo acima, o conceito atribuído à parte “depende daquele conferido ao todo”.

Quanto às relações partitivas a Norma ISO 25964-1 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, p. 60, tradução nossa), fornece os seguintes exemplos:

EXEMPLO 1 – Sistemas e órgãos do corpo

Sistema Cardiovascular

Vasos sanguíneos

Artérias

Veias

EXEMPLO 2 – Localização Geografia

Canadá

Ontário

Ottawa

Toronto

EXEMPLO 3 – Disciplinas e áreas do conhecimento

Ciência

Biologia

Botânica

Zoologia

EXEMPLO 4 – Estruturas sociais hierárquicas.

Exércitos

Corpo

Divisões (exército)

Batalhões

Regimentos

A maior parte dos casos das relações partitivas não é elegível para uma relação hierárquica, pois:

[...] a parte pode pertencer a mais de um todo. Por exemplo, uma relação TG / TE não deve ser estabelecida entre "bicicletas" e "rodas" porque uma roda pode fazer parte de um automóvel, um carrinho de mão ou de muitos outros artefatos. Uma busca por bicicletas recuperaria muito material indesejado se fosse estendido a todos os tipos de roda.. (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, p. 60, tradução nossa).

Além das relações que ocorrem de maneira hierárquica, existem também, as relações não-hierárquicas que “configuram-se como relacionamentos complexos que não podem ser classificados segundo os critérios de ordem e subordinação lógica”. Essas relações são chamadas de relações associativas e relações de equivalência (ARAÚJO, 2006, p. 53).

As relações associativas não-hierárquicas indicam a ligação entre descritores/termos que pertencem a campos semânticos distintos, mas próximos. As relações de equivalência remetem o conjunto dos não-descritores ou não-termos para o conjunto de termos adotados pelo sistema.

As relações associativas são aquelas em que dois termos que estão aproximados em algum contexto, não possuem uma relação hierárquica nem de equivalência (CLARKE, 2001).

Os termos “ensino” e “aluno”, por exemplo, estabelecem entre si uma relação associativa, pois “ensino” não é um tipo de “aluno” de modo a ser um termo mais geral, da mesma forma que “aluno” também não é um tipo de ensino. Assim, os termos estão relacionados de modo associativo.

A função das conexões associativas é informar os usuários do SOC sobre os termos preferidos, ou seja, os termos que querem aplicar e não o que ele pensou. As relações associativas geralmente são mostradas com a abreviação “TR” para “Termo Relacionado”. Como no exemplo abaixo:

TG Carnívoros
TR Ovíparos
TR Herbívoros
TR Onívoros

Dessa forma, o TG “Carnívoros” emprega uma relação com “ovíparos”, “herbívoros” e “onívoros”. Essa relação não ocorre de maneira hierárquica, nem de equivalência.

Por outro lado, as relações associativas são aquelas que associam pares de conceitos que são semântica ou conceitualmente associados, isto é, têm algo em comum. Essa associação precisa ocorrer de tal forma que seja necessário explicitar no tesouro, “sob o argumento de que pode sugerir termos adicionais ou alternativos para uso em indexação ou recuperação” (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, p. 60, tradução nossa).

Vale lembrar que se um termo A está associativamente relacionado a um termo B, logo, este termo B está, da mesma forma, associativamente relacionado ao termo A.

TG Mesa

TR Cadeira

TG Cadeira

TR Mesa

Outro ponto que deve ser levado em consideração quando se trata das relações associativas é que “um dos termos é frequentemente um componente necessário em qualquer explicação ou definição do outro; o termo ‘pássaros’, por exemplo, forma uma parte necessária da explicação da ‘ornitologia’.” (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, p. 60, tradução nossa).

A última relação conceitual básica dos tesouros é a relação que acontece por equivalência. Esse tipo de relação é aquele no qual um termo equivale a um outro termo.

A relação de equivalência são relações nas quais dois conceitos se assemelham de tal forma a serem considerados sinônimos. Como no caso de “docente” e “professor” que são dois termos que se referem a um mesmo conceito.

Este tipo de relação pode acontecer em dois contextos: no monolíngue, quando as relações acontecem entre a mesma língua, ou no contexto multilíngue, quando a relação acontece entre duas ou mais línguas (CLARKE, 2001). O contexto multilíngue, no entanto, não será abordado neste trabalho.

A relação de equivalência no contexto monolíngue é a relação entre um termo preferido pelo tesouro e seus correspondentes não preferidos na mesma língua natural. Esse tipo de relação, é valido ressaltar, ocorre entre termos e não entre conceitos. (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, p. 60, tradução nossa).

A norma ISO 25964-1 estipula quatro situações que pode existir equivalência:

- a) Os termos são sinônimos
- b) Os termos são quase-sinônimos
- c) O termo é considerado desnecessariamente específico e é representado por outro termo com escopo mais amplo
- d) O termo é considerado desnecessariamente específico e é representado por uma combinação de dois ou mais termos (conhecidos como "equivalência composta").

Na situação “a”, somente é necessário associar dois termos com um conceito semelhante, porém de origens linguísticas diferentes, ou seja, os sinônimos. Como por exemplo, pode-se citar a relação entre “menino” e “garoto” ou então a relação entre “cão” e “cachorro”.

Quando os termos forem sinônimos eles podem se apresentar de muitas formas diferentes, a norma ISO 25964-1 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011) designa nove formas:

1. Termos de origem linguística diferente

Como, por exemplo, os termos “abecedário” e “alfabeto”. Sendo que Abecedário é um termo criado a partir das primeiras letras do alfabeto latino (também conhecido como alfabeto romano): a + be + ce + de + o sufixo ário. E o termo Alfabeto, que é originado da união das duas primeiras letras gregas: alfa (α) + beta (β);

2. Nomes populares e nomes científicos

Nesta concepção se encaixam termos como: “*ananás comosus*” (termo científico) e “abacaxi” (nome popular); *apis mellifera scutellata*” (termo científico) *abelha* (nome popular).

3. Substantivos comuns e nomes comerciais

Nesta divisão, empregam-se aqueles termos que representam produtos que são popularmente chamados pelo nome da marca do

produto. Alguns exemplos deste caso são: “Bombril” e “lã de aço”, “Cotonete” e “hastes flexíveis com pontas de algodão”, ou então, “Velcro” e “fecho adesivo”;

4. Nomes variantes para conceitos emergentes

Englobam-se aqui os termos que são alterados conforme o conceito vai se disseminando na sociedade. O termo “caubói” por exemplo surgiu no Brasil para corresponder ao termo em inglês “cowboy”. No Brasil, há ainda os termos “vaqueiro” e “boiadeiro”.

Outro exemplo dessa situação é o termo “shampoo” do inglês e “xampu” na versão aportuguesada do Brasil.

Mais um exemplo que se encaixa nesse caso é são os termos “notebook” e “laptop”, dois termos que se referem ao computador portátil.

5. Termos atuais ou favorecidos *versus* termos desatualizados ou obsoletos

Nesse tópico pode ser citado o exemplo de “botica” que é uma palavra antiga que foi utilizada para se referir a o que é conhecido hoje como farmácia.

Outro termo, mais atual que caiu em desuso no Brasil é o “correio eletrônico” que atualmente corresponde ao termo original em inglês “*e-mail*”.

6. Grafias variantes, palavras em ordem diferente e plurais irregulares. Inclui-se aqui, também, erros ortográficos.

Os plurais irregulares e as palavras que ocorrem em ordem diferente são casos muito mais presentes na língua inglesa do que no português do Brasil. Os plurais irregulares do inglês podem muitas vezes mudar completamente a palavra. Como em “*person*”, termo que significa pessoa, tem o plural “*people*”. Ou, então, em “*mouse*” (singular) e “*mice*” (plural) do termo em português “rato”.

Em termos com ordem diferente, pode-se citar, por exemplo, uma divergência exposta por Hjørland (2012), tanto “*Information Organization*” como “*Organization of Information*” são denominações para a Organização da Informação apresentadas em dois termos diferentes.

Em português, pode-se pensar no exemplo do “creme dental” e “pasta de dente”, ambos são grafias variantes usadas para designar um produto para a higiene bucal.

7. Termos originados de diferentes culturas em um mesmo idioma

Neste aspecto em específico existem muitos exemplos que podem ser citados. Pelo fato do Brasil ser um país miscigenado entre diferentes etnias, um mesmo artefato pode mudar de nome dependendo da região.

Isso acontece com brincadeiras de rua: “Esconde-esconde” e “pique-esconde”; “bétia”, “bete”, “bets”, “tacobol”, “taco”.

Acontece também com produtos alimentícios, como no caso mencionado anteriormente, da “mandioca” “macaxeira”, “aipim”, dentre outras nomenclaturas. Outro alimento que é famoso por ter mais de uma nomenclatura é a bolacha que também é chamada de biscoito.

Em alguns casos os termos usados são específicos de uma cidade ou estado. Como, por exemplo, em Curitiba no Paraná que usam termos como “penal” para “estojo”, e “vina” para “salsicha”.

8. Abreviaturas ou siglas e nomes completos

Esta modalidade abrange os termos que são versões compactadas de outros termos sendo eles por forma de sigla como no caso siglas que compactam grandes termos, por meio do uso das primeiras letras de cada palavra, em termos mais reduzidos.

Como no caso da “ONU” que é a sigla que representa “Organização Nações Unidas”. Ou na área da organização do conhecimento as siglas “CDD” e “CDU” que significam Classificação Decimal de Dewey e Classificação Decimal Universal, respectivamente.

As abreviaturas possuem o mesmo objetivo das siglas, expor em poucas letras o que é representado com muitas. A abreviatura é feita de forma que se utiliza um ponto final para indicar que o termo está incompleto. Como por exemplo: “org.”, “inf.”, “index.”, que representam respectivamente “organização”, “informação” e “indexação”.

9. Substantivos comuns e gírias ou jargões

As gírias e jargões são palavras ou expressões que pertencem a algum grupo ou tribo. As gírias são termos usados no dia a dia enquanto os jargões podem ser considerados como aqueles usados em uma área do conhecimento.

De qualquer forma, são palavras que são entendidas. Podem pertencer a localizações geográficas como “eita pega” que podem indicar uma expressão de surpresa em alguns lugares do Brasil. Um gíria comum para expressar agradecimento é o termo “valeu”, que significa o mesmo que “obrigado”.

Os jargões podem, como dito anteriormente, representarem uma palavra ou expressão restrita a uma área. No jornalismo, por exemplo utilizam o jargão “foca” para referir-se aos novos integrantes da equipe, neste contexto, o termo “foca” equivale a “calouro em jornalismo”.

Esses são os nove tipos de sinônimos indicados na Norma para a situação “a” em que é abordado sobre os sinônimos. Na situação “b”, é entendido que os termos são quase-sinônimos. Os quase-sinônimos podem se manifestar de duas formas:

A primeira forma representa os termos que se opõem, sendo até mesmo os antônimos, considerados quase-sinônimos. Como por exemplo “consistência” e “inconsistência” ou “secura” e “umidade”.

No caso do segundo exemplo a norma ISO 25964-1 explica que

[...] o conceito buscado no tesouro é ‘a quantidade de umidade em um objeto ou material’, que pode ser expressa em termos de umidade ou de secura. Arbitrariamente, um dos termos é escolhido como o termo preferido e o outro como o termo não preferido.” (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, p. 48, tradução nossa).

Em outros casos, os conceitos de termos que não são sinônimos estão tão fortemente relacionados que isso acaba por estimular o usuário do tesouro a procurar pelo outro termo. Como exemplo pode-se citar “marreta” e “martelo” ou então “arbusto” e “moita”.

Também deve ser levado em consideração que a extensão de um termo ao ser tratado como quase-sinônimo depende, em grande parte, do assunto coberto pelo tesouro.

Em um tesouro que é especializado em uma área como Medicina Veterinária, por exemplo, é necessário que haja uma distinção de cada raça de cada animal para que as informações encontradas auxiliem o usuário a encontrar a melhor maneira de lidar com saúde daquele animal específico.

Por outro lado, em um tesouro mais geral as diferentes raças podem ser consideradas como quase-sinônimos, sem que haja a necessidade de descrição para cada característica de cada raça de animais. A decisão sobre qual abordagem se deve tomar deverá levar em conta o grau de termos distintos no momento em que for realizar a pesquisa.

Na situação “c”, quando um termo é considerado desnecessariamente específico, é representado por outro termo com escopo mais amplo. Assim, ao invés de considerar como termo preferido: “Basalto”, “granito” e “ardósia”; basta usar um termo mais geral que equivale a todos como “pedra”. Desse modo o tesouro ficaria como o exemplo:

Rocha

UP basalto

granito

ardósia

Basalto USE rocha

Granito USE rocha

Ardósia USE rocha

Quando o usuário procurar por um dos TE ele será encaminhado para o TG. No exemplo UP significa Usado Para, e isso indica que os termos pesquisados foram redirecionados para um mais amplo. À direita, a palavra USE indica ao usuário o termo preferido para aquela pesquisa.

Esse processo resulta numa diminuição de termos preferidos no tesouro, o que limita a especificidade da indexação e da busca conquistada. É benéfico quando a conteúdo a ser indexado tem muito pouca informação sobre a área de assunto em questão.

A situação “d” representa o caso de quando o termo é considerado desnecessariamente específico, assim como na situação “c”. Porém na situação “d” os termos específicos são representados por uma combinação de dois ou mais termos (conhecidos como "equivalência composta").

A norma ISO 25964 diz que “quando um termo com várias palavras é considerado inadequado como um termo preferido, mas pode ser procurado por alguns usuários, ele pode ser representado por uma combinação de dois ou mais termos preferidos”. (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011, p. 48, tradução nossa). Como por exemplo:

Mineração de Carvão

USE mineração
+ carvão

Filmes ferromagnéticos

USE materiais ferromagnéticos
+ filmes

Assim, a soma dos termos “mineração” + “carvão” conduzem o usuário para mineração de carvão, sendo que a busca pode ser feita em ambos os termos. O mesmo ocorre com filmes ferromagnéticos que são formados por dois termos equivalentes compostos: “materiais ferromagnéticos” e “filmes”.

3.2 Características das relações conceituais

As relações conceituais ou terminológicas possuem algumas características, como as relações simétricas, por exemplo, que são aquelas relações que ocorrem da

mesma maneira para os dois lados. Sendo que ao alterar a posição dos termos a relação entre os conceitos não é alterada.

Assim, quando uma relação de A para B acontece da mesma forma que a relação de B para A, essa relação é caracterizada pela simetria. Como, por exemplo na relação entre primos: João é primo de José, José é primo de João.

As relações assimétricas, em contrapartida, não acontecem da mesma maneira para os dois lados. Então, a relação de A para B não ocorre da mesma forma que a relação de B para A.

Esse tipo de relação, representa a maior parte das relações. São as relações que alteram o sentido quando invertidas as posições dos elementos. Como exemplo, pode-se citar a relação que acontece entre pai e filho: Roberto é pai de Guilherme, Guilherme é filho de Roberto. (GREEN, 2001)

De acordo com Green (2001), as relações de um-para-um acontecem quando somente um elemento pode corresponder a outros (e vice-versa). Como, por exemplo, em uma relação monogâmica entre marido e mulher: legalmente (no Brasil e alguns outros países), um marido pode ter apenas uma mulher, assim como uma mulher pode ter apenas um marido.

As relações de um-para-vários ocorrem quando um membro pode se relacionar com vários indivíduos e esses indivíduos se relacionam com esse mesmo único membro.

O exemplo pode ser visto também na relação entre pai e filho: Roberto é pai de Guilherme, Mariana e Eduardo, ou seja, Roberto pode ter vários filhos; Guilherme, Mariana e Eduardo são filhos de Roberto, um único pai. Um pai pode ter vários filhos, mas, cada filho, pode ter apenas um pai biológico (GREEN, 2001).

As relações de vários-para-vários surgem quando há relação de diversos elementos com outros diversos elementos. Como pode ser observado em uma relação parental, por exemplo, uma pessoa pode ter vários parentes e cada um destes parentes podem ter outros diversos parentes (GREEN, 2001).

As três relações acima (um-para-um, um-para-vários, e vários para vários) são, segundo Green (2001, p. 5, tradução nossa), relações binárias. Porém, a relação de vários-para-um pode acontecer de maneira lógica: “embora uma relação de vários-para-um também seja uma possibilidade lógica, é conveniente, nesse caso,

reorganizar os participantes da relação de modo a converter a relação em uma relação um-para-um.”.

Falando em relações lógicas, pode-se citar as relações de transitividade que expressam um tipo de relação, no qual uma primeira relação determina uma terceira, pois na relação de transitividade, se A está relacionado de alguma forma a B e B está relacionado de alguma forma a C, em uma relação lógica, C está relacionado da mesma maneira à A. Exemplificando: se o pai de Eduardo, é Roberto; e Roberto é pai de Mariana. Então, Eduardo e Mariana são filhos de Roberto e, por isso, são irmãos (GREEN, 2001).

Este pensamento lógico da transitividade conduz às relações de natureza lógica, descritas por Dahlberg (1978a). A autora elenca algumas características comuns às relações lógicas, sendo elas: a identidade, a implicação, a intersecção, a disjunção, e a negação.

Para compreender cada uma dessas características considere sempre dois grupos: o Grupo A e o Grupo B, no qual cada grupo representa um conceito distinto que está em relação a um outro. Ainda assim, apresentar-se-á outros exemplos para um melhor entendimento.

Na **Identidade**, Dahlberg diz que as características entre os conceitos são as mesmas, sendo que tudo o que há em um grupo A (x, x, x) corresponde exatamente a um grupo B (x, x, x). Por isso, entende-se que as características que um conceito apresenta é igual a um outro conceito.

Considere o seguinte exemplo: no Brasil, há uma planta cujo nome científico é *Manihot esculenta* mas não há um conceito comum. Porém, dependendo da região é chamada de uma maneira diferente.

A Wikipédia (2019) elenca treze palavras que são utilizadas em âmbito nacional, sendo eles: mandioca, aipi, aipim, castelinha, uaipi, macaxeira, mandioca-doce, mandioca-mansa, maniva, maniveira, pão-de-pobre, mandioca-brava e mandioca-amarga.

Cada um destes termos carrega um contexto diferente, tendo assim conceitos diferentes. No entanto, todas estas nomenclaturas se estão aproximadas conceitualmente.

Assim, é seguro afirmar que estes treze termos, e conseqüentemente, conceitos, na concepção de Dahlberg, têm uma relação lógica que é caracterizada pela Identidade

Na **Implicação**, Dahlberg diz que um conceito está contido em outro conceito, no qual um grupo A (x, x) está inserido a um grupo B (x, x, x). Isso significa dizer, portanto, que o conceito do grupo B contém uma característica a mais que o conceito do grupo B.

Considere o seguinte exemplo: o “professor” é o profissional que ensina, enquanto que “professor universitário” é o profissional que ensina em nível superior. Ambos são profissionais que estão envolvidos com o ensino, porém, o conceito de professor universitário possui nível de especificidade superior ao conceito de professor, o qual é um conceito mais amplo.

Assim, é seguro afirmar que o conceito de “banqueta” está relacionado ao conceito de “cadeira”, pois ambos conceitos se referem a um móvel bastante semelhante com a diferença do encosto na cadeira. Assim o conceito de um contém uma característica a mais que o conceito do outro.

Dessa forma, é correto declarar que os conceitos de “banqueta” e de “cadeira” apresentam uma relação lógica que é caracteriza pela Implicação designada por Dalhberg.

Na **Intersecção**, Dahlberg diz que os dois conceitos coincidem algum elemento sendo que o conceito do grupo A (x, x, o) tem algo em comum com o conceito do grupo B (x, o, o).

Considere o seguinte exemplo: O termo “estudante” refere-se a um indivíduo que está em um processo de estudos, sozinho ou em grupo e que se dedica por um período a estudar.

Enquanto isso, o termo “aluno” é o indivíduo que participa de uma instituição de ensino como alguém que estuda e recebe ensinamentos de um professor.

Ambos os termos se referem a um sujeito que busca o aprendizado. Tanto os estudantes como alunos dedicam-se ao estudo. Porém, enquanto na teoria, todo aluno é um estudante, nem todo estudante é um aluno.

Isso resulta no pensamento de que “aluno” e “estudante” são conceitos que em algum momento se coincidem, que manifestam atributos semelhantes, mas, ao mesmo tempo, não são conceitos idênticos que possam caracterizar a Identidade.

Contudo, ambos os conceitos empregam a relação lógica de Dahlberg, caracterizada como Intersecção.

Na **Disjunção**, Dahlberg diz que os conceitos se excluem mutuamente fazendo com que não haja nenhuma característica em comum. Assim, em um grupo denominado A (x, x, x) está relacionado a um grupo B denominado (o, o, o) pois não apresenta nenhuma semelhança um com o outro.

Considere o seguinte exemplo: Um “violão” é um instrumento musical que se utiliza de cordas para emitir sons. Um “espelho”, por outro lado, é um objeto que reflete a imagem que lhe é apresentada.

Os conceitos de “violão” e “espelho” são díspares, ou seja, não contém características comuns. Por essa razão, são conceitos disjuntos, não possuem similaridades.

Por serem conceitos sem atributos semelhantes, podem se relacionar de pela característica da Disjunção.

Na **Negação**, Dahlberg diz que o conceito referente ao grupo A (x, x, o), inclui uma característica cuja negação se encontra no grupo B (o, x, o). Isso quer dizer que o grupo A se opõe ao grupo B pois possui uma característica contrária.

Considere o seguinte exemplo: O conceito do termo “paz” é geralmente entendido como um estado de calma ou de tranquilidade, que ocorre quando há ausência de perturbações, de violência ou de guerra. Em contrapartida, o conceito de “guerra” é descrito como um confronto, conflito entre dois ou mais grupos, utilizando-se de armas e de violência.

Compreendendo, individualmente, o conceito de paz e o conceito de guerra, fica completamente claro que um conceito automaticamente exclui o outro, ou nega, nas palavras de Dahlberg.

Dessa forma, pode-se entender que os conceitos de paz e de guerra têm relação lógica caracterizada como Negação.

Essas características das relações lógicas podem ser aproximadas a alguns tipos de relações também expostas por Dahlberg (1978a). A relação hierárquica que a autora insere no contexto da implicação. A relação de oposição que se insere na negação e a relação funcional que se insere na intersecção. Estas associações podem ser encontradas no Quadro 4:

Quadro 4 – Síntese das ideias de Dahlberg (1978a)

Relações Lógicas	Premissa	Relações Conceituais	Premissa
Identidade	Características são as mesmas para todos os conceitos	Relação de Equivalência	Os conceitos de se assemelham a ponto de se tornar sinônimos
Implicação	Um conceito A está contido no conceito B	Relação Hierárquica	O que está na espécie precisa estar no gênero
Intersecção	Os dois conceitos coincidem algum elemento	Relação funcional (ou associativa)	Condizem processos, expressam funções. (produção, produto e produtor)
Disjunção	Nada comum entre os conceitos		
Negação	Um conceito é o contrário de outro	Relação de Oposição	Basicamente antônimos. Um se opõe ao outro

Fonte: Elaborado pelo autor

Retomando-se as discussões em relação às relações lógicas mencionadas por Dahlberg (1978a) anteriormente neste estudo, e afim de compreender o Quadro 4, serão abordadas resumidamente as relações de oposição e a relação funcional.

As relações de oposição são aquelas em que dois conceitos se opõem, como no caso dos antônimos, feliz e triste; preto e branco; bem e mal; público e privado; introvertido e extrovertido etc.

A relações funcionais condizem sobretudo a processos, expressam funções. Então, a relação entre produção, produto e produtor, por exemplo, é uma relação funcional.

Apesar de Dahlberg não as ter relacionado, as relações lógicas que ocorrem por identidade na concepção de Dahlberg, se encaixam nas relações de equivalência, pois se aproximam semanticamente.

Outro tipo de relações que podem existir são as relações bibliográficas, em uma definição mais ampla proposta por Green (2001, tradução nossa) “incluem todos os tipos de relações que estão envolvidos a catalogação descritiva de unidades bibliográficas”. Essas relações podem ser consideradas tanto, relações abstratas, como relações concretas.

Relações são associações entre duas ou mais entidades ou classes de entidades. Assim, as relações concretas são aquelas onde há associação entre os termos específicos (GREEN, 2001).

Como exemplo de uma relação abstrata, a simples relação entre pessoa e lugar se encaixa. Uma pessoa está relacionada a algum lugar, essa é uma associação genérica, pois, não se sabe qual a relação da pessoa com o lugar. Ela nasceu neste lugar? Visitou este lugar? Estudou sobre este lugar? Viveu algum tempo neste lugar? Tem parentes neste lugar? Dessa forma, esta é uma relação abstrata pois abre margem para várias interpretações.

Ao empregar atributos às relações abstratas, essas relações passam a ser consideradas relações concretas. As relações concretas são relações nas quais os elementos são específicos. Então, ao considerar que a pintora Tarsila do Amaral nasceu no Brasil, a relação: Tarsila + nascida no + Brasil é uma relação concreta, pois é uma autora específica que nasceu em um lugar específico, não abre margem para interpretações como acontece nas abstratas

Essas relações apresentadas são essenciais para a criação de sistemas de organização do conhecimento. Com elas é possível estabelecer associações entre conceitos de modo a melhor organizar um ambiente. Seja ele digital como no caso das ontologias ou físico como no caso das classificações bibliográficas.

3.3 Relações conceituais na Classificação Decimal de Dewey

Melville Louis Kossuth Dewey, ou simplesmente, Melvil Dewey procurava, em 1873, por um sistema para organizar os livros da Biblioteca do Amherst College, onde trabalhava como *student-assistant*. Nesta época, Dewey produziu uma tese que foi publicada anonimamente em 1876. Esse primeiro documento foi publicado inicialmente em forma de panfleto. Posteriormente, em 1931, Dewey publicou finalmente a Dewey Decimal Classification (DDC), isto é, a Classificação Decimal de Dewey (CDD) (ANJOS, 2008).

A princípio, a CDD tinha o objetivo de ser um instrumento com um aspecto mais prático e de aplicação simples para facilitar o trabalho dos bibliotecários e objetivava, portanto, responder “às necessidades pragmáticas de organização do conhecimento de um tipo de biblioteca que visava o livre acesso dos usuários e a rápida localização dos materiais bibliográficos nas estantes” (ANJOS, 2008).

Conforme Mitchell (2001), a CDD é a classificação mais usada no mundo, sendo utilizada em mais de 138 países para a organização das bibliotecas. Mitchell afirma também que a CDD é aplicável a todos os tipos de bibliotecas para facilitar a atividade biblioteconômica.

Como destaca Mitchell (2001), a CDD é uma ferramenta de organização do conhecimento que é estruturada em categorias que são bem definidas e possui relações hierárquicas que são bem desenvolvidas, além de dispor de uma rica rede de relações.

A CDD se utiliza de relações que se ajustam aos três modos de relações mais comuns que também estão presentes no tesouro. Sendo elas as relações associativas, relações hierárquicas e relações de equivalência (MITCHELL, 2001).

Com a classificação de Dewey, criou-se mais um instrumento que deu condições para que exemplares de um mesmo livro, que estivessem disponíveis em diferentes bibliotecas, pudessem receber um mesmo número de chamada. Este, ao menos, foi o ideal imaginado por Dewey. Assim, os usuários obtêm maior facilidade na localização dos documentos em diferentes bibliotecas, pois, dessa forma, estariam organizados de uma mesma maneira e, portanto, inseridos em um mesmo tipo de assunto. Isto resulta numa facilidade para as pessoas que procuram por um livro de um determinado assunto em uma biblioteca, porque poderiam encontrar uma mesma obra em alguma outra biblioteca que utiliza do mesmo sistema de classificação (MAI, 2003).

O sucesso de Dewey na construção da CDD, segundo Anjos (2008), foi evidenciar de que maneira a divisão decimal básica poderia possibilitar uma sequência sistemática de assuntos detalhados. Neste contexto, Anjos expõe que o conhecimento tem avançado na complexidade:

Desse modo, tornou-se evidente a necessidade de um esquema de hierarquias de termos que fosse mais do que uma imitação da relação gênero-espécie. Dewey compreendeu que esse problema existia, e providenciou tanto elementos de síntese quanto de análises em suas tabelas. Esse processo acarretou avanços no processo classificatório das classificações subsequentes (ANJOS 2008, p. 170).

Na perspectiva do conhecimento humano, a Classificação Decimal de Dewey é um sistema de organização do conhecimento que está em constante atualização e que passa por revisões para acompanhar o crescimento do conhecimento (MITCHELL; GOETZ, 2009).

A notação da CDD deve ser sempre apresentada em no mínimo três dígitos, sendo que cada algarismo do número representa uma classe, divisão ou seção diferente. Por serem sempre três dígitos para preencher as lacunas e completar três dígitos, são inseridos, quando necessário, os números zeros na notação. (MITCHELL et al., 2011).

Concomitantemente a isso, os algarismos zeros são símbolos que representam as generalidades dentro da classe, divisão ou seção. Assim, cada uma das 10 classes principais possui dez divisões e cada divisão dez seções, formando assim o sistema decimal (MITCHELL et al., 2011).

Desta maneira, as 10 Classes principais da Classificação Decimal de Dewey, em sua 23ª edição, são:

- 000 Computadores, informação e referência geral;
- 100 Filosofia e psicologia;
- 200 Religião;
- 300 Ciências Sociais;
- 400 Linguagem;
- 500 Ciência;
- 600 Tecnologia;
- 700 Artes e recreação;
- 800 Literatura;
- 900 História e geografia.

A CDD é construída em princípios que são sólidos e que fazem com que a organização do conhecimento geral seja ideal. É estruturada em notações, ou seja, em um sistema de símbolos. As notações são expressas por meio de algarismos arábicos que representam as classes da CDD. Os números dessas notações representam ao mesmo tempo cada uma das classes individualmente e suas relações entre as outras classes (MITCHELL et al., 2011).

A Classificação Decimal de Dewey é formada por dez classes básicas, desenvolvidas hierarquicamente com tópicos relacionados entre si. As classes são organizadas em disciplinas ou em campos de estudo. Cada classe principal possui dez divisões e cada divisão dez seções, formando assim o sistema decimal (MITCHELL et al., 2011).

Por separar o conhecimento em grupos de assuntos semelhantes, a CDD é organizada a partir do conceito que Green e Panzer (2010) chamam de “vizinhança”. Estando aproximadas as divisões ou seções em uma Classe elas guardam relações de proximidade conceitual entre si

Como forma de exemplificar o conceito de vizinhança observe-se a progressão de classes representada na Figura 4.

Figura 4 – Progressão das Classes

780	Música
782	Música vocal
782.2	Formas vocais não dramáticas
782.23.29	Formas vocais especificamente sacras
782.29	Formas litúrgicas
782.292	Cântico

Fonte: Elaborado pelo autor

Estas notações compõem, conforme Green e Panzer (2010), a referida vizinhança. Uma vizinhança é o conjunto que envolve todos os tópicos relacionados a uma seção específica. Assim, “Cântico” tem como vizinhos todos os itens presentes na vizinhança da Figura 4.

A partir deste exemplo, é possível constatar que o termo “cântico” da Figura 4 é o termo hierárquico mais específico desta vizinhança, pois é o termo onde a notação possui um número maior de algarismos. Assim, quanto mais dígitos uma notação tem, mais específico é o assunto; quanto menos dígitos, mais geral é o assunto (MITCHELL et al., 2011).

Como já foi exposto, as notações da CDD são construídas em no mínimo três algarismos numéricos arábicos. Em cada grupo de três algarismos, o primeiro dígito representa a classe principal; o segundo dígito refere-se à uma divisão da classe principal; o terceiro dígito indica uma seção. Observe a seguinte vizinhança:

Figura 5 – Vizinhança de Classes na CDD

<u>5</u> 00 Ciências Naturais
5 <u>9</u> 0 Animais
59 <u>8</u> Pássaros

Fonte: Elaborado pelo autor

Na vizinhança da Figura 5 a notação “**500** Ciências Naturais” é a classe principal e, dessa forma, a mais geral; “**590** Animais” é a divisão de Ciências Naturais; e “**598** Pássaros” é a seção dentro da divisão Animais e também é a mais específica do exemplo.

Além disso, os algarismos zero das notações referem-se a generalidades. Assim, na notação “**500**” o segundo e terceiro dígito indica generalidade dentro da classe principal “Ciências Naturais”, isso mostra os níveis de generalidade e especificidade nas notações da CDD, pois quanto mais dígitos zero as notações tenham mais gerais elas serão. O mesmo ocorre no oposto, quanto menos algarismos zeros mais específica é a notação de classe.

Sales (2007) expõe que as relações associativas, como o próprio nome já diz, relaciona termos por associação, ou seja, um termo que possua uma ideia semelhante a que expressa um outro termo. Em outras palavras, quando não há hierarquia, mas há relação, esta relação é associativa. Assim, o termo “professor” está associadamente relacionado ao termo “aluno”, mas não hierarquicamente. Um professor pode ser um aluno, no entanto, não necessariamente o é.

Na CDD as relações hierárquicas, conforme descritas por Mitchell (2001), são expressas por notações. Todos os tópicos da CDD, com exceção das dez classes principais, fazem parte de algum tópico maior.

A Figura 6, mostra a relação hierárquica presente na Classificação Decimal de Dewey. Analisando-a pode ser feito aquele teste de validade mencionado por Mazzocchi (2017), o teste todos-alguns.

Figura 6 – Relação hierárquica da CDD

<p>761 Processos relevográficos (block printing) 761.2 Xilografia 761.3 Linoleogravura 761.8 Gravura em relevo no metal</p>

Fonte: SILVA (2017)

Todas as “xilografias” são “processos relevográficos”. Alguns “processos relevográficos”, por sua vez, são “xilografias”, logo a relação entre “xilografia” e “processos relevográficos” é uma relação, apropriadamente, hierárquica pois se encaixa ao teste todos-alguns.

Continuando com o exemplo apresentado na Figura 6, a “xilografia” e a “linoleogravura” apresentam características que os aproximam, mas não hierarquicamente. Estes dois termos possuem uma relação associativa, pois ambos são “processos relevográficos”, ou seja, são processos que se utilizam do relevo para imprimir.

Ainda neste mesmo exemplo, apresentado na Figura 6, é possível estabelecer uma relação de equivalência que pode ser observada no conceito representado no termo entre parênteses. De modo que o conceito de *Block Printing* equivale ao conceito de Processos Relevográficos, exercendo assim, uma relação de equivalência.

A tarefa de compreender como ocorrem as relações conceituais presentes na Classificação Decimal de Dewey possibilita que este trabalho conduza a discussão para o outro sistema de organização do conhecimento a ser analisado nesta pesquisa, um instrumento menos tradicional que a CDD, isto é, a ontologia. Entender como ocorrem as relações conceituais que existem nas ontologias permite com que o trabalho de analisar a *BBC wildlife ontology* seja realizado, além, é claro, de complementar de auxiliar na comparação dos dois SOC analisados. Assim, o próximo capítulo abordará as relações conceituais que ocorrem nas ontologias.

3.4 Relações conceituais na ontologia

Assim como a Classificação Decimal de Dewey, as ontologias apareceram como mecanismos que auxiliam na necessidade de se organizar o conhecimento. Enquanto a CDD preocupou-se, e ainda se preocupa, em organizar o conhecimento registrado em ambiente físico, as ontologias surgiram para organizar o conhecimento registrado em ambiente digital.

Além disso, as ontologias surgem como uma forma de traduzir a organização para uma linguagem que possa ser reconhecida pela máquina, como por exemplo, pelo uso da linguagem *Ontology Web Language* (OWL).

A área da Ciência da Informação tem carácter multidisciplinar, isto é, dialoga com múltiplas disciplinas. Por essa razão, muitos artefatos utilizados na área são também utilizados ou provêm de outras áreas, e isso não difere quando se fala de ontologias.

O termo “Ontologia” tem origem no âmbito da Filosofia com o significado geral de estudo sobre a natureza dos seres.

Quando grafada com a primeira letra maiúscula, a Ontologia refere-se a uma disciplina da Filosofia enquanto que quando grafada a primeira letra minúscula, a ontologia refere-se ao instrumento utilizado na Ciência da Computação e na Ciência da Informação (GUARINO; GIARETTA, 1995).

As ontologias são responsáveis por criar condições lógicas para que a máquina infira automaticamente os conceitos de cada um dos termos presentes em um ambiente digital, respeitando-se, naturalmente, uma estrutura categorial previamente fornecida.

De modo similar aos tesouros, os quais guiam os usuários humanos na utilização dos termos preferidos e suas definições, as ontologias são apropriadas para, por exemplo, a navegação de agentes inteligentes.

O intuito das ontologias é, de acordo com García Marco (2007), elaborar um estoque de informações semânticas para que seja possível a inspeção de significados de um termo. Assim, segundo o autor, ontologia é um sistema de termos que descrevem e representam uma área do conhecimento, além de expressar as relações entre esses termos por meio de uma linguagem lógica legível por computadores.

Além disso, as ontologias compartilham o vocabulário e sua estrutura com os tesouros e as taxonomias utilizados na Organização da Informação. A Organização da Informação, por sua vez, usa termos relacionados para descrever um domínio ou área do conhecimento.

Dessa forma, é extremamente importante que o profissional se atente para o fato de que ao construir uma ontologia não se deve apresentar ambiguidade entre os conceitos. Um termo deve ser claro o suficiente para que se possa unificá-lo e para que, assim, a ontologia torne-se bem estruturada.

Como maneira de realizar este feito, existem instrumentos que são fundamentais para criação de ontologias, como, por exemplo, a linguagem OWL. Essa linguagem é definida por Costello e Jacobs (2003) como um conjunto que engloba elementos e atributos em *Extensible Markup Language*, ou simplesmente a linguagem XML, que possui um significado padronizado.

Ontologias em linguagem OWL contêm componentes descritivos como classes, propriedades, indivíduos que se assemelham dos componentes do *software Protégé*,

mas com terminologia distinta: classes, slots e instâncias, respectivamente (HORRIDGE, 2009).

Assim como acontece em outros sistemas de classificação mais tradicionais, as ontologias em OWL compartilham de características semelhantes. Principalmente quanto a criação de vários tipos de classes e de relações conceituais entre essas classes (ZENG; PANZER; SABALA, 2010).

Enquanto a Classificação Decimal de Dewey é um sistema que possui padronização e estrutura hierárquica bem definida, instrumentos mais formais, como as ontologias, não necessitam de adesão às padronizações presentes em outros sistemas de organização do conhecimento, como ocorre na CDD. Isso se deve ao fato de que as ontologias podem estabelecer diferentes tipos de relações conceituais, dependendo do objetivo pretendido em cada ontologia (GREEN; PANZER, 2011).

Considerando-se a perspectiva terminológica das ontologias, elas podem ser descritas como um conjunto de conceitos e termos que podem ser usados para descrever alguma área do conhecimento ou construir uma representação. O seu objetivo, como ocorre com outros SOC, está no fornecimento de sistemas de categorização que permitam ao homem organizar a realidade (MOREIRA, 2013).

Por isso, a remoção de ambiguidade entre os termos apresenta um nível elevado de complexidade, já que cada ontologia pode vislumbrar um termo a sua maneira e isso dificulta a interoperabilidade entre ontologias.

Desse modo, é possível destacar as ontologias, por serem instrumentos que podem ser lidos por máquina, possibilitam simular ou incrementar as inferências humanas.

Os detalhes na descrição de um conceito nas ontologias são importantes para que um conceito seja diferenciado de outro com precisão. Mesmo com isso em mente, é evidente que a inferência feita pela máquina, é um processo muito mais difícil de realizar, pois não consegue reproduzir em sua completude o processo de inferência que o cérebro humano pode realizar. Isso requer grande esforço na avaliação, classificação e seleção da informação de interesse (BREITMAN, 2005).

Portanto, é necessário informar às máquinas os diferentes aspectos de um termo para que este termo possa estar inserido em diferentes contextos.

Buscando fazer com que a lacuna existente entre classificações tradicionais e ontologias formais seja preenchida, Giunchiglia, Zaiherayeu e Farazi (2009)

expressam que esquemas de classificação não formalizados podem ser convertidos em linguagem OWL e esta conversão ajuda a definir a melhor organização dos esquemas de classificação.

As ontologias são instrumentos usados para capturar o conhecimento dentro de um domínio de interesse e, por isso, descrevem os relacionamentos e conceitos presentes em um domínio. Em outras palavras, para descrever os conceitos e identificar suas relações é necessária a utilização de uma linguagem ontológica. Cada tipo diferente de linguagem ontológica gera diferentes tipos de facilidades (HORRIDGE, 2009).

A OWL é uma das principais linguagens de marcação para a construção de ontologias recomendada pelo consórcio W3C e foi criada para suprir a necessidade de aplicação da web semântica (NICOLINO, 2014).

Sendo baseada em RDF (*Resource Description Framework*) e RDF Schema e utilizando-se de XML (*Extensible Markup Language*), a OWL é considerada mais adaptada e mais fácil para expressar significados e semânticas (SANTAREM SEGUNDO, 2010).

Nicolino (2014) explica que a OWL oferece três sub-linguagens, projetadas para o uso de implementadores e comunidades específicas. A *OWL Lite* que é a linguagem que dá suporte à criação de hierarquias simplificadas e com restrições simples. A *OWL DL* que é a linguagem que dá suporte aos usuários que desejam o máximo de expressividade sem perder a completude computacional. *OWL Full* que é a linguagem que possui o mesmo vocabulário da linguagem *OWL DL* e é feita para os usuários que desejam abranger o máximo de expressividade e liberdade sintática do RDF.

As ontologias em linguagem OWL puderam ser estudadas anteriormente pelo autor desta pesquisa, conforme a justificativa deste trabalho destacou, Silva (2017) criou uma ontologia com base no conjunto de classes da CDD referente à gravura e impressão representado pelas notações que vão da 760 à 769.

Desta ontologia criou-se por meio do *software* Protégé uma visualização da ontologia com suas classes e subclasses com a utilização do recurso *ontograf*, conforme pode ser observado na Figura 7.

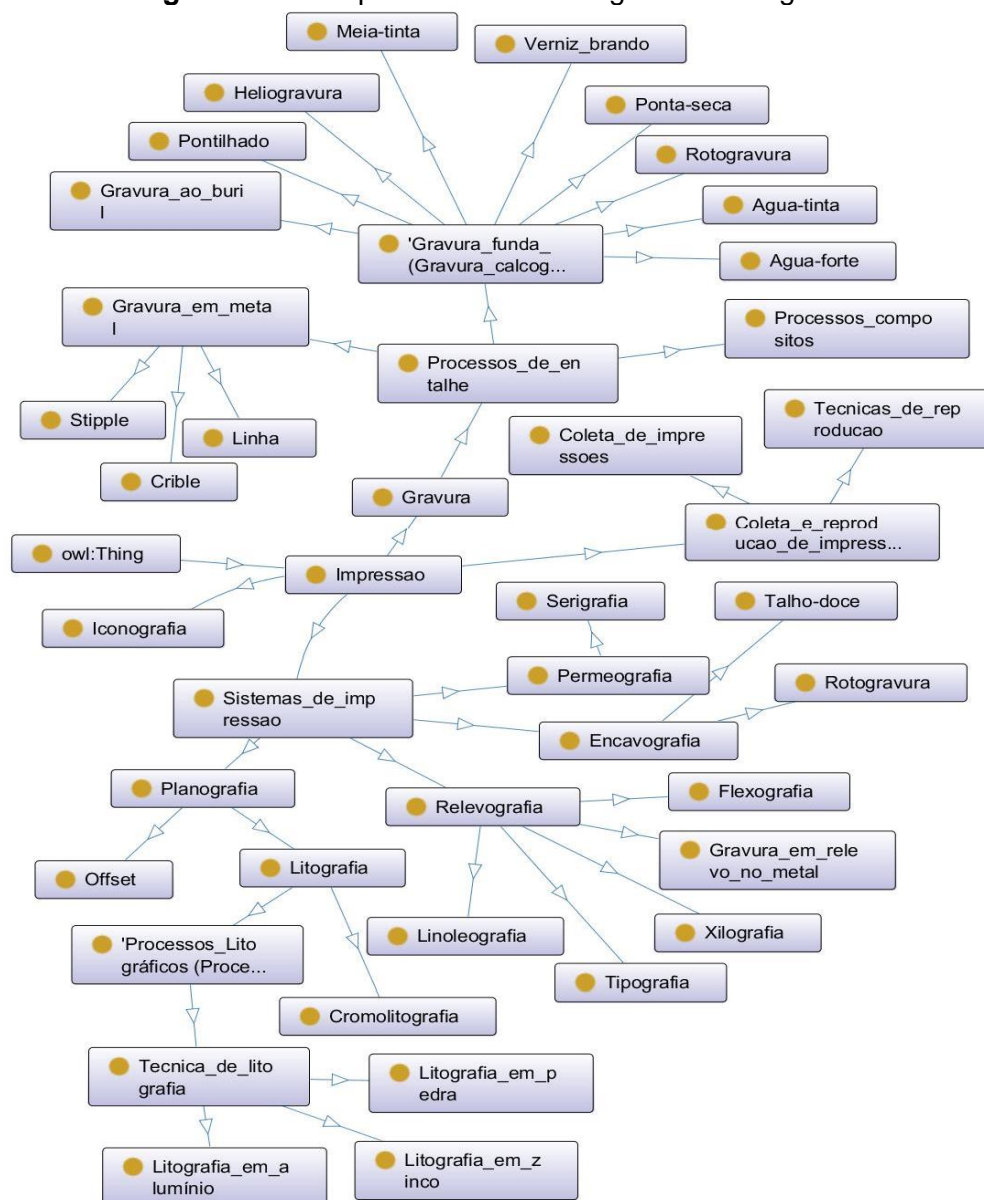
Silva (2017) percebeu que criar uma ontologia partindo-se de um sistema de classificação tradicional permite que as ontologias criadas sejam mais padronizadas

e, portanto, podem se beneficiar das classificações de outros Sistemas de organização do conhecimento, como a CDD.

Além das classes e subclasses, presentes na Figura 7, uma ontologia apresenta indivíduos, também denominados instâncias de classes, os quais são os objetos em um determinado domínio; as propriedades, são relações binárias, ou seja, relação entre dois elementos dentro dos indivíduos.

As classes são conjuntos que contêm indivíduos e são caracterizadas por usar descrições formais, isto é, descrições matemáticas (HORRIDGE, 2009). Um exemplo dessa aplicação de propriedades pode ser visto Figura 2, presente na justificativa deste trabalho.

Figura 7 – Exemplo do recurso ontograf do Protégé 5.1.0



Fonte: Silva (2017)

Conforme ressalta Carlan (2006), o que qualifica uma ontologia não é o vocabulário em si, mas os conceitos que os termos do vocabulário transmitem. Assim, ao se transferir os termos presentes em uma ontologia para outra linguagem, por exemplo, de um idioma para outro, o conceito ontológico não se modifica. Assim, no âmbito das ontologias, para se estruturar um sistema de representação do conhecimento e seu vocabulário de forma eficiente, produz-se uma análise ontológica formal do domínio.

Dessa forma, com a estipulação de termos em uma ontologia, estes termos são relacionados a outros. As relações conceituais nas ontologias acontecem, de acordo com Sales, Campos e Gomes (2008), de duas maneiras: na terminologia e na parte formal.

Segundo as autoras, na parte terminológica são usadas relações categoriais que ligariam as classes que representam conceitos, já na parte formal não basta apenas relacionar as classes, é preciso dizer a forma como acontece a relação. Esse segundo tipo de relação é chamado pelas autoras como “relações formais”. As relações formais são provindas da ciência da computação enquanto as relações categoriais são provindas da terminologia e da ciência da informação.

Assim, as relações categoriais representam aquelas relações que revelam duplas de categorias como, por exemplo: “coisa-processo”, “material-produto” etc. As relações formais, por sua vez, são aquelas que revelam o tipo de relação existente entre duplas de categorias, por exemplo: “causado_por”, “ocorre”.

Com essa divisão em mente as autoras definem então 6 tipos de relações nas ontologias como pode ser observado no Quadro 5:

Quadro 5 – Tipos de relações das ontologias

Relações Categoriais	Relações Formais
Genéricas	Genéricas
Partitivas	Partitivas
Funcionais	Funcionais

Fonte: Elaborado pelo autor com base em Sales, Campos e Gomes (2008)

Dessa forma, todas as relações explicadas abaixo são com base em Sales Campos e Gomes (2008):

- As relações categoriais genéricas e as es relações formais genéricas
- acontecem por um princípio lógico de abstração e, assim, respondem a pergunta “é um?” Logo, sempre acontecem no interior de uma mesma

categoria. Não é uma relação mutável, pois ocorre sempre da mesma maneira.

- As relações categoriais partitivas e as relações formais partitivas ocorrem quando há relação entre o todo e suas partes, como pôde ser elucidado no capítulo sobre as relações conceituais.
- As relações categoriais funcionais e as relações formais funcionais são aquelas que apresentam a relação de objeto com o mundo ou a função de um objeto em um determinado contexto.

Assim, apresentada a etapa teórica deste estudo, o próximo capítulo abordará os procedimentos metodológicos a serem usados para a obtenção dos resultados desta pesquisa a serem apresentados na defesa deste relatório de pesquisa.

4 PROCEDIMENTOS METODOLÓGICOS

Para a realização dessa pesquisa, utilizou-se materiais que foram fundamentais para a construção da bibliografia deste trabalho e também para auxiliar na resposta aos objetivos geral e específicos. Além de permitir a análise dos resultados de forma mais eficaz e coerente.

4.1 Materiais

A princípio, o trabalho preocupou-se em coletar o material bibliográfico na literatura da Ciência da Informação para a compreensão de alguns aspectos pertinentes à área como a organização e a representação do conhecimento, e alguns sistemas de organização do conhecimento, como os tesouros e, mais especificamente, a Classificação Decimal de Dewey (CDD) e as ontologias.

Juntamente a isso, para compreender melhor as ontologias foram também averiguados alguns estudos no âmbito da Ciência da Computação para o entendimento das ontologias nessa área.

Após a construção do referencial teórico, um dos importantes instrumentos utilizado neste estudo é a CDD, em sua 23ª edição, publicada no ano de 2011. Essa edição da CDD está organizada em quatro volumes. É publicada, atualmente, pela *Online Computer Library Center* (OCLC), que é responsável por manter a CDD atualizada e por deter os direitos autorais (MITCHELL, 2011).

A CDD é um instrumento utilizado em bibliotecas para a organização do acervo nas estantes, facilitando o serviço do bibliotecário e também a recuperação da informação pelos usuários da biblioteca.

Na CDD atribui-se um número de notação que auxilia na localização dos documentos. Para melhor compreensão desse SOC, o subcapítulo 3.3 deste trabalho explica como é formulada essa notação numérica, além de demonstrar a separação das classes por disciplinas.

Para a análise das relações conceituais na CDD, utilizou-se uma das classes da CDD, a classe 590 denominada “animais”, a fim de verificar se suas relações conceituais estão adequadas do ponto de vista lógico das relações conceituais.

A partir da base de dados Basel Register of Thesauri, Ontologies & Classifications (BARTOC), selecionou-se a ontologia atribuída à emissora de rádio e

televisão *British Broadcasting Corporation* (BBC). Essa ontologia, denominada *BBC Wildlife Ontology*, é referente à vida selvagem e é disponibilizada de modo aberto e gratuito.

A *BBC Wildlife Ontology*, é um:

vocabulário simples para descrever espécies biológicas e táxons relacionados. O vocabulário define termos para descrever os nomes e classificação dos táxons, além de fornecer suporte para descrever seus habitats, status de conservação e características comportamentais, etc. (BRITISH BROADCASTING CORPORATION, 2018, tradução nossa).

Para visualização da *BBC Wildlife ontology*, utilizou-se neste trabalho, o *software Protégé 5.5.0* em sua versão *desktop*. O *Protégé* é um ambiente elaborado para possibilitar a criação e edição de ontologias em linguagem OWL.

O *Protégé* é uma ferramenta desenvolvida pela Universidade de Stanford, localizada na Califórnia, nos Estados Unidos. É um *software* livre e de código aberto baseado em linguagem OWL (WIKI PROTÉGÉ, 2018). Optou-se por esse *software* por ser amplamente utilizado na criação de ontologias, por ser livre acesso. Esse *software* foi utilizado também para criar uma ontologia baseada em ambos os SOC estudados: a CDD e a *BBC Wildlife Ontology*, afim de desenvolver, a partir desses SOC um modelo que poderá servir como um exemplo de ontologia se seguir.

Para verificar a consistência das relações conceituais na classe 590 da CDD, utilizou-se a primeira parte da norma ISO 25964 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011) sobre tesouros e interoperabilidade com outros vocabulários.

Por fim, para verificar a consistência e da *BBC Wildlife Ontology* utilizou-se também a norma ISO 25964 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011). Juntamente a isso utilizou-se o *OntoClean* que é uma metodologia criada para guiar a produção de ontologias, dando mecanismos para avaliar ontologias e servindo como um guia na criação delas.

A metodologia *OntoClean* também foi usada na criação da ontologia modelo que foi feita a partir da comparação entre a classe 590 da CDD e a *BBC wildlife ontology*. Assim, o próximo subcapítulo para compreender melhor essa metodologia busca esclarecê-la.

4.2 A metodologia *OntoClean*

Para que a criação de ontologias seja melhor estruturada, a metodologia *OntoClean* surge como uma forma de tornar as ontologias mais consistentes. Vem sendo desenvolvida há alguns anos e procura fornecer um tipo de orientação baseada em noções formais altamente gerais, tão gerais a ponto de serem pertinentes a qualquer aplicação de ontologia, independente de um domínio particular.

As noções formais que devem orientar a avaliação das ontologias, segundo Guarino e Welty (2002), são denominadas: essência, rigidez, identidade e unicidade, são usadas para definir um conjunto de metapropriedades. As metapropriedades são usadas para

“caracterizar aspectos relevantes do significado pretendido das propriedades, classes e relações que formam uma ontologia. Além disso, as metapropriedades impõem várias restrições à estrutura taxonômica de uma ontologia, que ajudam a avaliar as escolhas feitas.” (GUARINO; WELTY, 2002, tradução nossa).

Assim, as noções acima comentadas serão explicadas nas próximas seções deste capítulo.

4.2.1 Essência e Rigidez

A primeira noção formal a ser discutida é a essência. Na essência, uma propriedade de uma entidade é essencial para esta entidade, se for válida para a entidade.

Para entender melhor isso considere como exemplo a propriedade “é duro”. Esta propriedade pode se referir a entidade “martelo” pois todo martelo é essencialmente duro. A entidade “esponja”, por outro lado, não é sempre essencialmente dura, mas pode ser, como ocorre com as esponjas secas, por exemplo. Assim, ser duro não é essência da entidade “esponja”, mas é da entidade “martelo”.

Ainda nesse exemplo da esponja, a fim de ficar mais claro, os autores explicam que

“Algumas esponjas (secas) são duras, e algumas esponjas específicas podem ser duras por toda a sua existência, no entanto, isso não torna “ser dura” uma propriedade essencial dessa esponja. O fato é que poderia ter sido macia em algum momento, mas ocasionou de nunca ter sido.”. (GUARINO; WELTY, 2002).

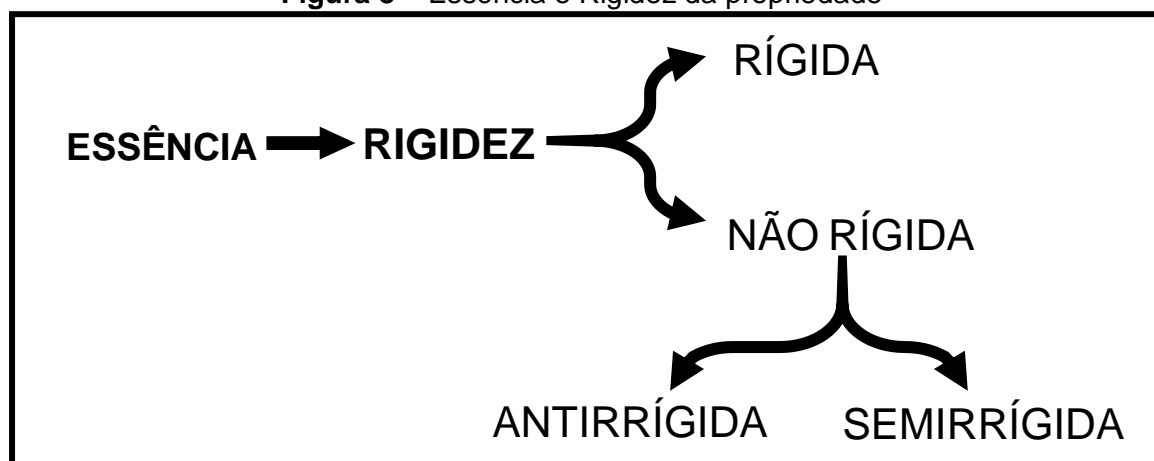
É nesse contexto de “essência” que se insere a noção de rigidez. Diz-se que uma propriedade é rígida se esta for essencial para todas instâncias. A propriedade “é uma pessoa”, por exemplo, é comumente considerada rígida, pois quando se identifica um conceito como sendo “uma pessoa” este conceito sempre será uma pessoa, não se pode deixar de ser uma pessoa, é uma essência rígida.

A propriedade “é duro” do exemplo anterior, por outro lado, não é rígida pelo fato de não ser um atributo inerente a todos os membros a todo momento como mencionado na situação da esponja que pode ser dura para sempre ou por algum período.

A rigidez acontece quando todos os membros das classes compartilham uma mesma característica, ou seja, compartilham uma mesma propriedade. Por essa razão a propriedade “é uma pessoa” delimita que todos os indivíduos (ou instâncias) que contêm essa propriedade têm a característica de ser uma pessoa. Toda pessoa é sempre uma pessoa em todas as situações.

Existem ainda propriedades que não são essenciais a todas as instâncias e, dessa forma, Guarino e Welty (2002) distinguem as propriedades como sendo: rígida ou não rígida, esta última podendo ser semirrígida ou antirrígida como pode ser visto na Figura 8.

Figura 8 – Essência e Rigidez da propriedade



Fonte: Elaborado pelo autor com base em Guarino e Welty (2002)

Uma propriedade é rígida quando se mostra essencial à todas as instâncias a todo e qualquer momento. Uma propriedade denominada “é leve” é uma propriedade rígida pois não há possibilidades de as instâncias deixarem de ser. Uma “pena” sempre será essencialmente leve.

A propriedade não rígida é aquela em que a propriedade não é essencial para todas as instâncias, podendo ser semirrígida ou antirrígida.

Quando a propriedade é essencial para algumas instâncias, mas não essencial para todas, dá-se a ela o nome de propriedade semirrígida. Uma propriedade semirrígida que pode se ter como exemplo é a propriedade “apresenta metamorfose”. As borboletas por exemplo exercem o processo da metamorfose durante sua vida se transformando de lagarta para uma borboleta. Esta é uma propriedade semirrígida porque nem toda lagarta chegará ao estágio da metamorfose, podendo perecer antes que a metamorfose aconteça.

A propriedade antirrígida é aquela propriedade que não é essencial para todas as instâncias. A propriedade “é um estudante”, por exemplo, é uma propriedade antirrígida pois a cada instância de estudante não é essencialmente um estudante, já que ele pode não estudar ou deixar de estudar a qualquer momento, não é algo rígido, nem semirrígido.

A rigidez e suas variantes, segundo Guarino e Welty (2002), são metapropriedades, e as propriedades devem ser rotuladas em uma ontologia como rígida, não rígida ou antirrígida. As metapropriedades, além de fornecer mais informações sobre o que uma propriedade pretende significar, impõem restrições à relação de inclusão, ou seja, a inclusão de alguma coisa em algo maior. Essas restrições podem ser usadas para verificar a consistência ontológica das conexões taxonômicas.

4.2.2 Identidade e unicidade

De acordo com Guarino e Welty (2002), as noções de identidade e unicidade são as noções mais importantes de serem usadas na metodologia *OntoClean*. Entendê-las é uma tarefa complexa e muitas vezes uma noção é confundida com a outra, mesmo sendo diferentes.

A identidade busca reconhecer entidades individuais como sendo iguais ou como sendo diferentes, ou seja, identificar as igualdades e diferenças. A unicidade, por outro lado, refere-se à capacidade de distinguir todas as partes que formam uma entidade individual.

Em geral, identidade refere-se ao problema de ser capaz de reconhecer entidades individuais no mundo como sendo as mesmas (ou diferentes), e unidade

refere-se a ser capaz de reconhecer todas as partes que formam uma entidade individual.

Uma das decisões mais comuns que devem ser tomadas na análise ontológica é a de se identificar se algo que era aparentemente visto como uma entidade é na verdade duas ou mais entidades.

Outra noção a ser utilizada para analisar uma ontologia é a de unicidade que que é aquela responsável por que reconhecer todas as partes em uma instância. Preocupa-se com o problema de descrever o modo como as partes de um objeto estão unidas, de forma a saber o que é parte do objeto, o que não é, e saber sob quais condições o objeto é um todo.

A classe “água”, por exemplo, não representa objetos inteiros. Uma instância dessa classe poderia ser uma quantidade de água, porém não representa um todo, já que não é reconhecível como uma entidade isolada.

Por outro lado, “oceano” pode ser uma classe que representa objetos inteiros já que uma instância como “oceano atlântico” ou “oceano pacífico”, por exemplo é reconhecível como entidades únicas.

Isso conduz também a problemas de com a inclusão. Dizer que “oceano” é uma subclasse de “água” pode fazer sentido pois todos os oceanos são água. No entanto, ao afirmar que as instâncias de “água” não são inteiras e as instâncias de “oceano” são, tem-se uma contradição. Esse tipo de problema acontece por conta da ambiguidade da linguagem natural, os oceanos não são tipos de água, eles são compostos dela.

4.3 Procedimentos de coleta de dados e de análise de dados

Para a coleta dos dados e para a construção do material de análise da pesquisa foi realizada a identificação e o mapeamento da tipologia de relações conceituais presentes na CDD, confrontando com a literatura. Como já mencionado, o estudo foi baseado na tipologia e nas definições das relações conceituais apresentadas na norma ISO 25964-1 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011) para verificar a consistência das relações conceituais na CDD.

Além disso, foi realizada a identificação e mapeamento da tipologia de relações presentes na *BBC Wildlife Ontology*, confrontando-a com a literatura; e para avaliar a

consistência das relações conceituais utilizou-se a metodologia *OntoClean* conforme pôde ser elucidado na seção anterior.

Após a coleta de dados foi, então, feita a criação da ontologia, com base nos conceitos e relações apresentados na classe 590, da CDD, com a utilização do *software* Protégé 5.5.0. O *software* foi utilizado para a compreensão dessa ontologia e para a visualização da *BBC Wildlife Ontology*, além da criação de uma ontologia a partir dos SOC.

Portanto, os resultados desta pesquisa, encontrados nos próximos capítulos, estão desdobrados em 3 capítulos.

Primeiramente, o capítulo 5 dedicado às relações conceituais presentes na Classe 590 da CDD, averiguando a consistência das relações, e identificando se a forma pelas quais estão aplicadas está em acordo com o que a norma ISO 25964 sugere.

Após isso, no capítulo 6 há uma avaliação da *BBC Wildlife Ontology* baseada na metodologia *OntoClean*. Identificando aspectos de rigidez, essência, identidade e unicidade e analisando se essas relações estão consistentes

Por fim, o capítulo 7, cumprirá o objetivo principal dessa pesquisa, isto é, a comparação entre as relações conceituais existentes na classe 590 Animais da Classificação Decimal de Dewey e na *BBC Wildlife Ontology*, identificando as semelhanças e distinções.

5 A CLASSE 590 DA CLASSIFICAÇÃO DECIMAL DE DEWEY

A Classificação Decimal de Dewey (CDD), como já exposto neste trabalho, desempenha nas bibliotecas o papel de auxiliar na organização do conhecimento nas bibliotecas.

Para tanto, Dewey separou o conhecimento em dez disciplinas principais, as quais podem ser observadas no subcapítulo 3.3 desta pesquisa. Para este atual capítulo, a análise focar-se-á na classe 590 que faz parte da classe maior, 500, referente à “ciência”. Dessa forma, a classe “590 Animais” manifesta uma relação hierárquica com a classe “500 Ciência”.

Assim, para a análise da classe 590, observe a seguinte vizinhança presente na Figura 9:

Figura 9 – Vizinhança da classe 590

500	Ciência
590	Animais
591	Tópicos específicos na história natural dos animais
592	Invertebrados
593	Invertebrados marinhos
594	Moluscos
595	Artrópodes
596	Cordados
597	Vertebrados de sangue frio
598	Aves (pássaros)
599	<i>Mammalia</i> (mamíferos)

Fonte: Adaptado da CDD

Um aspecto que salta à vista em um primeiro momento, é que as classes 592, sobre invertebrados e 593 sobre invertebrados marinhos não estão relacionados na CDD de forma hierárquica, como a norma ISO 25964 recomenda, e sim como termos associados, o que parece incoerente.

É possível perceber essa incoerência quando se aplica o teste lógico “todos-alguns”, pois ao realizar o teste de validade para identificar se a relação é hierárquica, pode-se dizer que **todos** animais invertebrados marinhos são animais invertebrados. Do mesmo modo que **alguns** animais invertebrados são marinhos.

Isso acontece porque a CDD procura dar ao profissional opções para classificar um documento. Dependendo do tipo de biblioteca, é necessário utilizar uma classe que melhor represente o tema.

Em uma biblioteca específica da Marinha, por exemplo, um documento que trata de “animais marinhos invertebrados”, encaixa-se muito mais adequadamente no contexto da classe 593, que é mais específica e, portanto, divide melhor esta área, do que na classe 592 que é mais ampla.

Por outro lado, a Classe 598, sobre aves, está apropriadamente relacionada de maneira associativa a classe 599, sobre mamíferos. Novamente, ao aplicar o teste de validade todos-alguns é perceptível que esta relação não se trata de uma relação hierárquica pois nem toda ave é um mamífero e nem todo mamífero é uma ave. Ambas são classes biológicas dos animais e, por isso, se relacionam, mas não de maneira hierárquica e sim associativa, exatamente como Dewey separou.

Considerando também a classe 599, referente a *mammalia* (mamíferos), é possível observar outra relação na classe com ela mesma. O termo representado entre parênteses expõe uma relação de equivalência entre *mammalia* e mamíferos. *Mammalia* é o termo científico, em latim, para designar os mamíferos. Sendo assim seu significado se aproxima de tal forma a se tornar um sinônimo, exercendo assim uma relação de equivalência.

Outra relação de equivalência nessa vizinhança está na relação entre aves e pássaros que, a princípio, aparenta estar correta. No entanto, apesar de estarem relacionados esses dois conceitos não se relacionam por equivalência, mas hierarquicamente.

Pássaro é um tipo de ave pequena. Então se for aplicado o teste todos-alguns nesses dois conceitos pode-se dizer que todo pássaro é uma ave e que algumas aves são pássaros.

Mais um ponto interessante a ser destacado é que a classe 594, referente a moluscos, detém uma relação com a classe 592, referente a invertebrados e também com a classe 593, referente a invertebrados marinhos, já que moluscos são animais invertebrados que podem ou não ser marinhos.

Assim sendo, as classes 592 e 593 seriam boas propriedades em uma ontologia pois agregariam características à classe dos moluscos e suas instâncias. Na CDD funcionam como relações associativas conforme sua definição na norma ISO

25964 pois são conceitos que possuem entre si relação, mas não hierárquica e nem de equivalência.

As classes “594 moluscos”, “595 artrópodes” e “596 cordados” estão relacionados na CDD corretamente de maneira associativa, apesar de também estarem relacionados associadamente de maneira incorreta, conforme os padrões definidos na norma ISO 24964, com outras classes, tais como: “592 Invertebrados”, “598 aves” e “599 Mamíferos”.

Segundo Magalhães (2018, grifo nosso), “o reino animal é dividido em diversos filos. Os principais são: poríferos, cnidários, platelmintos, nematódeos ou nematelmintos, anelídeos, equinodermos, **moluscos, artrópodes e cordados**”. Assim, as classes “594”, “595” e “596” poderiam ser inseridas dentro de uma classe denominada “filo”. Essa classe “filo” estaria dentro da classe mais ampla denominada “animais” representada pela notação “590”.

Dessa forma, as classes da CDD 594, 595 e 596 seriam classes associativas que estão hierarquicamente relacionadas à classe filo e essa classe fillo estaria relacionada hierarquicamente com a classe 590. Isso resulta em uma lacuna na classificação de Dewey que poderia estar melhor estruturada.

A classe 591 que se refere aos “tópicos específicos na história natural dos animais” está, na CDD, hierarquicamente relacionada à classe “animais 590”. Isso, porém, não representa o que uma relação hierárquica deve ser. “Tópicos específicos na história natural dos animais” não é um tipo de “animal”, essas classes estão relacionadas, mas não de maneira hierárquica. Tampouco estão relacionadas por equivalência já que um conceito não é visto como um sinônimo de outro.

Diante das considerações feitas até aqui, é seguro dizer que apesar de ser um instrumento bastante utilizado em bibliotecas para a organização de seus acervos, a Classificação Decimal de Dewey não tem uma estrutura de classificação ideal para todos os casos, como, aliás, ocorrem em qualquer sistema de classificação.

Por outro lado, Dewey construiu um sistema visando a organização do conhecimento de uma maneira mais simplificada que pudesse ser usada em todas as bibliotecas. Isso faz com que a classificação seja mais genérica e útil para acervos de maiores abrangências.

O próximo capítulo apresenta uma avaliação da *BBC wildlife ontology* com a utilização da metodologia *OntoClean* que pôde ser elucidado no subcapítulo 4.2 desta

dissertação. Nessa avaliação serão usados como recursos de análise as noções formais de essência e rigidez; e unicidade e identidade com o objetivo de examinar a consistência das relações nessa ontologia.

Por conta dessa complexidade da ontologia este trabalho se focará em identificar, com o auxílio da metodologia *OntoClean*, noções formais nas relações conceituais da *BBC wildlife ontology*.

Para analisar as relações conceituais encontradas da *BBC wildlife ontology* observe a partir dos Quadros 4 e 5 duas das classes que compõem a ontologia sobre a vida selvagem:

Quadro 6 – Classe *adaptation* da *BBC wildlife ontology*

Adaptation	
URI	http://purl.org/ontology/wo/Adaptation
Descrição¹	Uma adaptação é qualquer característica de um animal ou planta que o torna mais adequado para um habitat específico ou para uma tarefa específica. Por exemplo, ser aerodinâmico é uma adaptação para nadar rápido e ser capaz de sobreviver com pouca água é uma adaptação à vida no deserto.
Subclasses	ExtremesAdaptation, BehaviouralPattern, CommunicationAdaptation, EcosystemRole, FeedingHabit, LifeCycle, LocomotionAdaptation, MorphologyAdaptation, PredationStrategy, ReproductionStrategy, SocialBehaviour, SurvivalStrategy
Propriedades	adaptation

Fonte: Adaptado de British Broadcasting Corporation (2018)

Quadro 7 – Classe *Adapted to extremes* da *BBC wildlife ontology*

Adapted to Extremes	
URI	http://purl.org/ontology/wo/ExtremesAdaptation
Descrição	Organismos que são adaptados a extremos (conhecidos como Extremófilos) são organismos que prosperam e podem até requerer condições físicas ou geoquimicamente extremas que são prejudiciais para a maioria da vida na Terra.
Superclasses	adaptation

Fonte: Adaptado de British Broadcasting Corporation (2018)

No *software Protégé* não é possibilitada a inserção de espaços nos nomes das classes e subcasses, por isso os exemplos de ontologias presentes neste trabalho possuem classes que quando há mais de uma palavra essas palavras estão unidas. Pela falta de acentuação nas palavras inglesas, optou-se por manter os termos em inglês já que os acentos não se comportam satisfatoriamente no *software*. Nos títulos dos quadros preferiu-se manter os termos separados para melhor entendimento.

¹ Considerando que a ontologia está toda em inglês foram mantidos os nomes, classes e propriedades no original. Somente a descrição foi traduzida para melhor compreensão deste trabalho

Assim sendo, o Quadro 6 representa a classe “*adaptation*” que significa adaptação em português e dentro desta classe existem subclasses que dentre elas, a classe “*extremesadaptation*”, isto é, adaptações extremas.

Enquanto isso, o Quadro 7 apresenta uma classe denominada “Adapted to extremes” para os animais adaptados a ambientes extremos. Logo, a classe do Quadro 7 se relaciona à classe do Quadro 6.

Enquanto o Quadro 6 indica que “*extremesadaptation*” é uma subclasse de “*adaptation*”; o Quadro 7 informa que “*adaptation*” é uma superclasse de “Adapted to extremes”. Logo, pode-se dizer que “*adaptation*” está superordenada à “Adapted to extremes” e que “adaptado para extremos” está subordinado a “*adaptation*”. Isso sugere que ambas as classes possuem uma relação hierárquica.

Ao aplicar o teste de validade, confirma-se que sim, de acordo com a norma ISO 25964 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011), que se trata de uma relação hierárquica, pois, **algumas** adaptações são para extremos, **todas** adaptações extremas são adaptações. Dessa forma, a relação entre as duas classes está adequadamente hierárquica.

A classe “*adaptation*”, no entanto, possui uma falha quanto a propriedade, já que a propriedade é uma característica, parece incoerente que a propriedade de “*adaptation*” no Quadro 6, seja a própria “*adaptation*”.

Ao analisar o Quadro 8, pode-se observar como é descrita a propriedade *adaptation*:

Quadro 8 – Propriedade *adaptation* da *BBC wildlife ontology*

Adaptation	
URI	http://purl.org/ontology/wo/adaptation
Descrição	associa um rank de táxons a uma adaptação que é exibida
Domínio	Taxonrank
alcance	Adaptation

Fonte: Adaptado de British Broadcasting Corporation (2018)

A “*adaptation*” é uma propriedade que associa à uma outra classe denominada “*taxonrank*”. Assim, a propriedade “*adaptation*” se conecta com classificação taxonômica.

Isso significa que quando um usuário procurar na ontologia pela classe adaptação, a ontologia irá conectar essa classe com a propriedade “*adaptation*” que associa à taxonomia de Lineu com as classes reino, filo, classe, ordem, família,

gênero, espécie. Então, pode-se entender que a propriedade da classe “*adaptação*” é conectar com as classes taxonômicas.

Assim, dentro da noção de identidade da *OntoClean* nota-se que o termo “*adaptação*” se insere tanto na classe quanto na propriedade sendo assim duas coisas distintas que são vistas como a mesma coisa. Uma forma de evitar isso é fazer com que o domínio “*taxonrank*” seja diretamente a propriedade de *adaptação*.

Nesse mesmo exemplo é ainda válido expor as noções de essência e rigidez. Com essas noções, em específico, mostra-se que a classe “*taxonrank*” e a propriedade “*adaptation*” não estão adequadas. Isso acontece porque a *adaptação*, ou mais especificamente a taxonomia de Lineu não é essencial para existir. Em outras palavras, a classificação taxonômica não é essencial o suficiente para ser uma propriedade da classe “*adaptation*”.

No contexto da taxonomia de Lineu, a *BBC wildlife ontology* contém duas classes. A primeira pode ser observada no Quadro 9 que demonstra a classe denominada “*taxon name*”, ou em tradução livre: “nome taxonômico”. Esta classe pode ser vista a seguir:

Quadro 9 – Classe nome taxonômico da *BBC wildlife ontology*

Taxon Name	
URI	http://purl.org/ontology/wo/TaxonName
Descrição	Um nome taxonômico, descrevendo a estrutura e proveniência de um nome taxonômico.
Propriedades	<i>commonName</i> , <i>genusName</i> , <i>name</i> , <i>scientificName</i> , <i>speciesName</i> , <i>taxonomicName</i>

Fonte: Adaptado de British Broadcasting Corporation (2018)

O Quadro 9 apresenta os possíveis nomes taxonômicos como, por exemplo, os “*commonName*” ou nomes “nomes comuns”, em português, sendo aqueles nomes pelo quais são mais conhecidos ou então o “*scientificName*” referente ao nome científico. Pode-se encontrar aspectos de rigidez nessa classe ao analisar suas propriedades.

A propriedade “*scientificName*”, por exemplo, apresenta rigidez, pois todo nome científico costuma ser sempre o mesmo não havendo possibilidades de mudanças. Um cachorro, por exemplo, tem o nome científico de *Canis lupus familiaris* uma denominação em latim que é usada para definir o cachorro independente de outros idiomas.

A instância de *Canis lupus familiaris* será sempre referente ao animal cachorro e isso faz com que a propriedade “*scientificName*” seja rígida.

Já a propriedade “*commonName*” não é uma propriedade totalmente rígida, já que o nome comum de um animal pode, por exemplo, ser apresentado em termos distintos. Isso pode ser observado com sinônimos como no caso do “cão” e do “cachorro” ou até mesmo pode acontecer com termos de outros idiomas como por exemplo “cachorro”, em português, “*dog*”, em inglês, ou então “*perro*”, em espanhol. Essa propriedade é definida como semirígida pois o termo pode mudar dependendo do contexto.

A segunda classe, apresentada no Quadro 10, que pode ser visto a seguir, apresenta a outra classe dentro do contexto taxonômico.

Quadro 10 – Classe *taxonomic rank* da *BBC wildlife ontology*

Taxonomic Rank	
URI	http://purl.org/ontology/wo/TaxonRank
Descrição	Conceito genérico para uma classificação taxonômica, como um gênero ou espécie.
Subclasses	Class, Family, Genus, Kingdom, Order, Phylum, Species
Propriedades	adaptation, class, conservationStatus, family, genus, growsIn, habitat, kingdom, livesIn, name, order, phylum, species

Fonte: Adaptado de British Broadcasting Corporation (2018)

No Quadro 10 é possível encontrar uma propriedade com característica de ser antirígida, a propriedade “*conservationStatus*” ou estado de conservação, em português, por exemplo, não é uma característica essencial a todas as instâncias, por isso, é uma propriedade antirígida.

Nas subclasses do Quadro 10 pode-se também apontar aspectos de unicidade. A noção de unicidade diz que se deve reconhecer todas as partes que formam uma entidade individual.

Dessa forma, as subclasses do Quadro 10 são partes da classificação taxonômica dos animais. Então no caso já referido do cachorro, pode-se atribuir os seguintes valores a essas classes: *kingdom* (reino): animal, *phylum* (filó): cordado, *class* (classe): mamífero, *order* (ordem): carnívoro, *family* (família): canidae, *genus* (gênero): *canis*, *species* (espécie): *Canis Lupus*. Esses atributos são partes que compõem a entidade “cachorro”.

Nessa ontologia, é perceptível que os termos se repetem diversas vezes para descrever itens diferentes. Isso resulta numa confusão tanto para o entendimento humano como para a organização dos mecanismos lógicos para que produção de inferências automáticas. Somado a isso, ao analisar com as noções formais da metodologia

OntoClean nota-se que a *BBC wildlife ontology* necessita de ajustes para que ela se comporte de uma maneira melhor.

Há também que se considerar que há uma relação hierárquica apresentada no Quadro 6, entre “adaptation” e “EcosystemRole” que significa “papel do ecossistema”. No entanto, esta não parece uma relação adequada, pois ao analisar a descrição da classe, referente ao papel do ecossistema, na própria ontologia, constata-se que o papel do ecossistema é a classe que diz respeito “a parte que um animal ou planta desempenha na manutenção ou sustento do habitat ao seu redor” (BRITISH BROADCASTING CORPORATION, 2018, tradução nossa).

No Quadro 7 é possível estabelecer, mesmo que não esteja explicitamente relacionado na ontologia, uma relação de equivalência entre a classe “adaptado para extremos” com o termo “extremófilos” que é descrito como sendo a mesma coisa que o título da classe. Definir essa relação por equivalência informa à ontologia que ambos os termos se referem ao mesmo conceito (ou o conceito mais aproximado).

Entendendo essas relações conceituais presentes na classe 590 da CDD e as relações conceituais presentes na *BBC Wildlife Ontology*, possibilitam com que seja feita a comparação no próximo capítulo.

7 COMPARANDO A CLASSE 590 COM A *BBC WILDLIFE ONTOLOGY*

O primeiro ponto que deve ser destacado é que apesar da ontologia da BCC ter guiado esta pesquisa à escolha da classe 590 na CDD, os dois SOC não se referem a exatamente mesma coisa. Enquanto na CDD os animais são abordados de forma geral, na *BBC Wildlife Ontology* aborda-se apenas a vida selvagem, ou seja, é especializada na vida selvagem animal.

Ambos os SOC são utilizados para organizar uma área conhecimento. Para a CDD apresentam-se os “animais” como objeto de estudo, dentro da classe Ciência. A classe 590 uma classe mais simples que a ontologia e pode ser adaptada dependendo do contexto do acervo da biblioteca.

A *BBC Wildlife Ontology* possui um nível de detalhamento maior para que possa ser utilizada por computadores e ainda emprega características a cada uma das classes. Isso acontece pois na ontologia não há a interpretação humana que ocorre na CDD.

Na CDD o profissional pode, por exemplo, compreender que “ganso” é um tipo de “ave” que é um tipo de animal e encaixá-lo na classe 598. Na ontologia seria preciso dar mecanismos para a máquina associar que é um ganso, o que é um animal, e o que é uma ave, além de informar que há relação entre todos esses termos.

Isso indica que a CDD por já possuir uma relação hierárquica definida, permite fazer com que o trabalho de criação das relações conceituais na ontologia, não necessite ser feito totalmente no escuro, pois pode-se observar a estrutura de relações em instrumentos já consolidados na Ciência da Informação.

Além disso, com a metodologia *OntoClean* para avaliar a ontologia, foi possível constatar que utilizando-a fica mais fácil construir uma classificação coerente para estabelecer relações e empregar as propriedades.

Tanto a CDD como a ontologia da BBC apresentam inconsistências nas relações conceituais em ambos os SOC e na definição das propriedades, no caso das ontologias.

A CDD, apesar de tradicional, sendo bastante usada em bibliotecas para organização dos acervos e mesmo com suas constantes atualizações, mostrou que organizar o conhecimento não é algo simples e de fácil classificação. Pelo contrário, mostrou-se que definir classes suficientemente representativas é uma tarefa de alto nível de complexidade.

As ontologias, que surgiram mais recentemente, já se mostram definitivamente mais complexas do que as estruturas mais tradicionais como a da CDD. O alto grau de detalhes que exige uma ontologia demonstra o quão complicado é entender como o conhecimento pode ser dividido.

Da mesma forma, que se pode analisar as relações conceituais de uma ontologia a partir da norma ISO 25964, é possível também inserir a classe 590 da CDD na metodologia *OntoClean*. Como exemplo a isso, pode-se definir a identidade na classe 598 a indica aves e pássaros num mesmo nível de relação. Aves e pássaros são conceitos diferentes e devem estar em lugares diferentes de acordo com a metodologia *OntoClean*.

Porém, é certo que a CDD não necessita da inserção de propriedades como no caso das ontologias. Então seria errôneo analisar a CDD com a metodologia *OntoClean*, já que esta analisa as propriedades das ontologias. Enfim, são recursos diferentes para análises diferentes.

Pode-se estabelecer uma comparação entre esses dois sistemas de organização do conhecimento. Ambos apresentam elementos da classificação taxonômica de Lineu mas de formas diferentes. Enquanto na CDD são apresentados “moluscos, artrópodes e cordados”, a *BBC wildlife ontology* apresenta a classe biológica dos animais filo.

As três classes da CDD apresentadas inserem-se na classe filo da ontologia, sendo então subclasses de filo. Assim sendo, há relação hierárquica na qual “filo” está superordenado em relação a moluscos, artrópodes e cordados. Enquanto isso, moluscos, artrópodes e cordados são conceitos coordenados e estão subordinados à classe Filo.

Dentro da Classificação taxonômica de Lineu também pode ser inserida a classe 599 que agrega os mamíferos. Mamíferos, ou *mammalia*, constituem a classe taxonômica “classe”.

Apesar das dificuldades encontradas para comparar os dois SOC esta pesquisa possibilitou perceber que há algumas semelhanças e possivelmente a ontologia da BBC pode ser beneficiada pela estrutura da CDD que mesmo não sendo perfeita, possui uma lógica e juntamente a isso, emprega padrões de organização.

Uma ontologia quando criada sem qualquer base ou padrão acaba por não contribuir com a interoperabilidade entre ontologias, já que para que haja

interoperabilidade é necessário algum padrão, pois assim, poder-se-ia dar mecanismos para a máquina associar dois elementos iguais e auxiliar na organização em diversos ambientes diferentes. Fazendo assim com que os sistemas possam “conversar” e aprender de maneira lógica uns com os outros.

A partir da análise dos dois SOC, produziu-se uma ontologia como recurso metodológico para a análise das relações conceituais. Neste trabalho refere-se a esta ontologia como “ontologia modelo”.

Dentre os resultados obtidos na comparação realizada nesta pesquisa, há um ponto que se mostrou relevante para dar início à execução deste capítulo. A partir das classes referentes à moluscos, invertebrados e invertebrados marinhos foi possível iniciar a criação da estrutura da ontologia modelo.

Como pôde ser observado no capítulo anterior, a Classe 590 da Classificação Decimal de Dewey sobre animais e a ontologia da BBC sobre vida selvagem têm classes que se aproximam do sistema de classificação biológica de Lineu.

Assim, após a classe maior da ontologia denominada *Thing*, ou coisa, que é uma característica do Protégé que significa que todas as coisas a partir desta classes estão dentro de um determinado domínio

Esta ontologia modelo criada não é uma ontologia completa pois servirá apenas como um referente para a criação de ontologias. Assim, optou-se por não definir classes que pudessem separar as plantas.

Da mesma forma, a classe “animais” também não está classificada por completo, caso contrário a ontologia modelo seria demasiadamente complexa para ser entendida como algo referencial.

Por não ser o foco desta pesquisa entender os seres vivos e seus papéis biológicos, os dados biológicos obtidos nesta pesquisa foram todos retirados da Wikipédia (2019).

Para prosseguir com a criação da ontologia utilizou-se da metodologia de Noy e McGuinness (2001) para a criação de ontologias e juntamente a isso a Metodologia Ontoclean que servirá como instrumento de avaliação para a consistência das relações conceituais. Definindo bem as propriedades, permite que as relações conceituais criadas na ontologia sejam coerentes.

Na metodologia para criação de ontologias de Noy e McGuinness (2001) é recomendada a execução de sete passos para realizar o desenvolvimento de uma ontologia. Observe o Quadro 11.

Quadro 11 – Sete passos para criar uma ontologia

Passo 1	Determinar o domínio e escopo;
Passo 2	Considerar o reuso de ontologias existentes;
Passo 3	Enumerar os termos importantes
Passo 4	Definir as classes e sua hierarquia;
Passo 5	Definir as propriedades das classes;
Passo 6	Definir as facetadas das propriedades;
Passo 7	Criar instâncias.

Fonte: Elaborado pelo autor com base em Noy e McGuinness (2001)

Ao se considerar os passos enumerados acima, para cumprir o passo 1 esta pesquisa optou por determinar o domínio aproveitando-se da análise da *BBC wildlife ontology* e da Classe 590 da Classificação Decimal de Dewey. Assim, a área de abrangência da ontologia modelo são os moluscos, ou seja, um tipo de Filo da taxonomia de Lineu.

Para não se tornar uma análise extensa e de difícil compreensão utilizou-se como objeto de análise a instância Polvo Paul, conforme poderá ser observado no decorrer do capítulo. Decidiu-se pela criação de uma ontologia mais simples tendo em vista os objetivos da análise.

O passo 2 estabelece que se deve considerar o reuso de ontologias já existentes. Nesse passo, utilizou-se da ontologia *BBC wildlife ontology* para auxiliar na definição de categorias dentro da ontologia, mas não para sua efetiva criação.

A classe “adaptacao”, por exemplo, reutiliza a estrutura da classe “adaptation” na *BBC wildlife ontology*. Utilizando porém, outras propriedades. Além disso, a classe “adaptacao” não apresenta acentos para facilitar no reconhecimento do *software*. Assim, o reuso da ontologia não foi feito de forma integral, mas utilizou-se de sua estrutura para sua composição.

O passo 3 da metodologia de Noy e McGuinness (2001) sugere que seja feita a enumeração de termos importantes para a ontologia. Assim, os termos que representam papel relevante na ontologia modelo são: moluscos: polvos; classificação taxonômica: reino, filo, classe, ordem; habitat: marinhos; adaptação: estratégia de sobrevivência; vertebrados e invertebrados.

Após a delimitação de termos importantes para a ontologia definidos no passo 3 o passo 4 inicia a elaboração das relações conceituais entre os termos escolhidos.

Conforme poderá ser observado neste capítulo, a ontologia modelo se enfocará em demonstrar o exemplo do polvo.

Assim, foram sendo delimitadas as primeiras classes e suas relações hierárquicas e juntamente a elas percebeu-se que poderiam ser criadas também algumas propriedades. Isso pode ser visualizado por meio do Quadro 12:

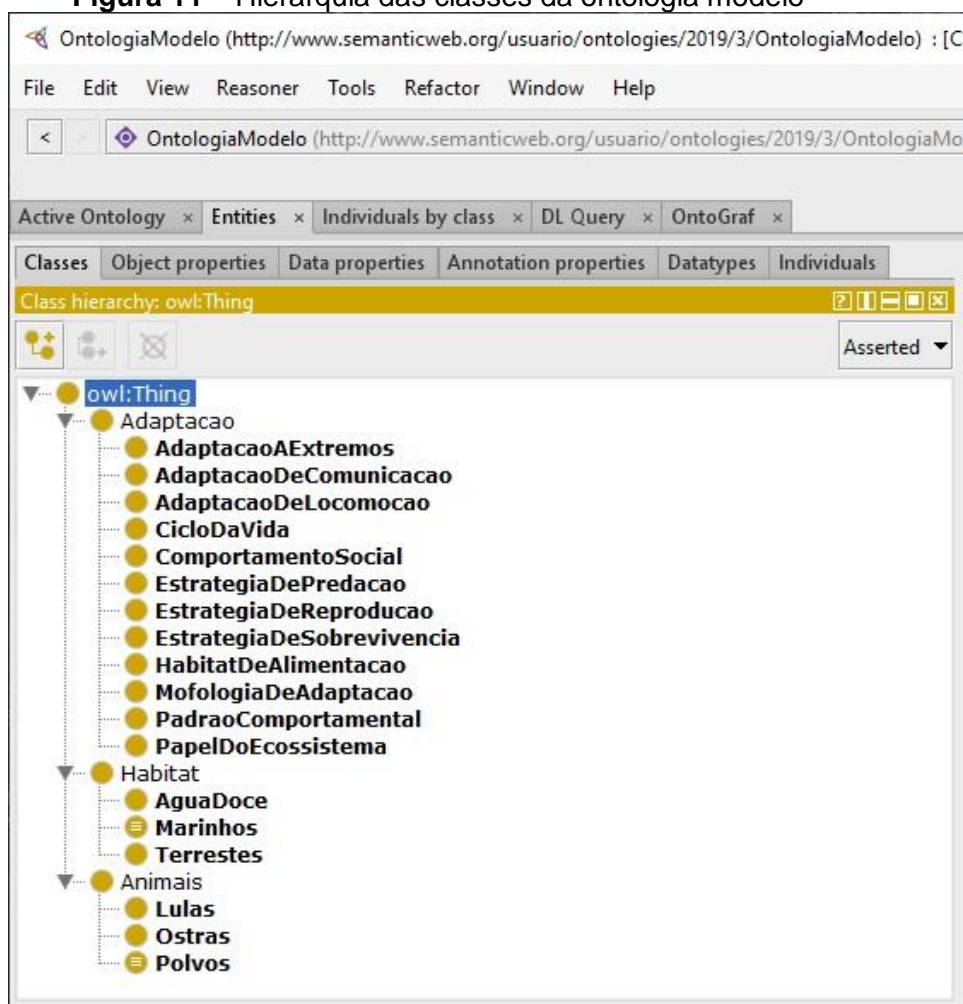
Quadro 12 – Classes, entes abstratos e propriedades da ontologia modelo

Classes, entes abstratos e propriedades	Descrição
Classificação taxonômica	envolve todas as classes que organizam os seres vivos iniciando-se na mais geral (o reino, no caso desta ontologia modelo) até o mais específico (a ordem, no caso desta ontologia. Definindo esses termos notou-se que a ontologia se comportam melhor quando são inseridos como propriedades, isto é características dos seres vivos, conforme pode ser observado na Figura 11.
Habitat	envolve três classes que definem o ambiente que o animal vive. Sendo elas: terrestres, marinhos e água-doce. Para o exemplo do polvo a ser demonstrado neste capítulo a classe “marinhos” é a mais relevante. Nesta classe Habitat está superordenado às outras classes que estão coordenadas. As propriedades da ontologia conduzem à uma subclasse da classe Habitat.
Adaptação	envolve as mesmas classes da <i>BBC wildlife ontology</i> porém, a mais relevante a ser comentada dentro do exemplo é a classe “EstrategiasDeSobrevivencia”. Essa classe será associada ao polvo que pode ser visto no quadro central da Figura 11.
Vertebras	envolve características dos animais sendo eles vertebrados ou invertebrados. Na CDD continha as classes invertebrados e invertebrados marinhos. Apesar de envolver as características dos animais foram definidas como não como propriedades mas como entes abstratos. Já o termo “marinhos” for desconsiderado nesses entes abstratos pois está inserido à classe Habitat.

Fonte: elaborado pelo autor

Como pode ser analisado no Quadro 12, juntos esses termos importantes definidos no passo 3 formaram as classes e as propriedades da ontologia modelo. Além dessas primeiras classes foram definidas também alguns entes abstratos como o Quadro 12 mencionou.

Figura 11 – Hierarquia das classes da ontologia modelo



Fonte: Elaborado pelo autor

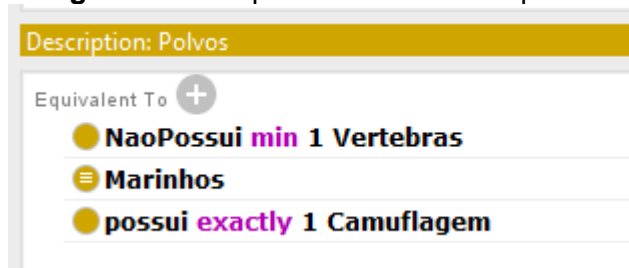
Como foi explicado no Quadro 12, as classes taxonômicas de Lineu foram incluídas na ontologia modelo como propriedades. Isso aconteceu para que pudessem ser incluída uma classe denominada “animais” na ontologia modelo de modo a associá-los às classes taxonômicas.

A partir da classe retirada da *BBC wildlife ontology* denominada “adaptacao” foi possível informar que os polvos são moluscos que se utilizam da camuflagem e de sua capacidade de largar tinta como estratégia de defesa. Para relacioná-las fez o semelhante à classe habitat na qual é estabelecida uma relação conceitual associativa entre as classes para identificar meios de adaptação e habitat.

Os polvos por exemplo possuem estratégias de defesa que faz com que eles se adaptem ao ambiente modificando de cor para se camuflar. Junto a isso os polvos são animais que tem como habitat o ambiente marinho, ou seja, podem viver em mares, oceanos ou em aquários.

Considere a Figura 12:

Figura 12 – Propriedades da classe polvos

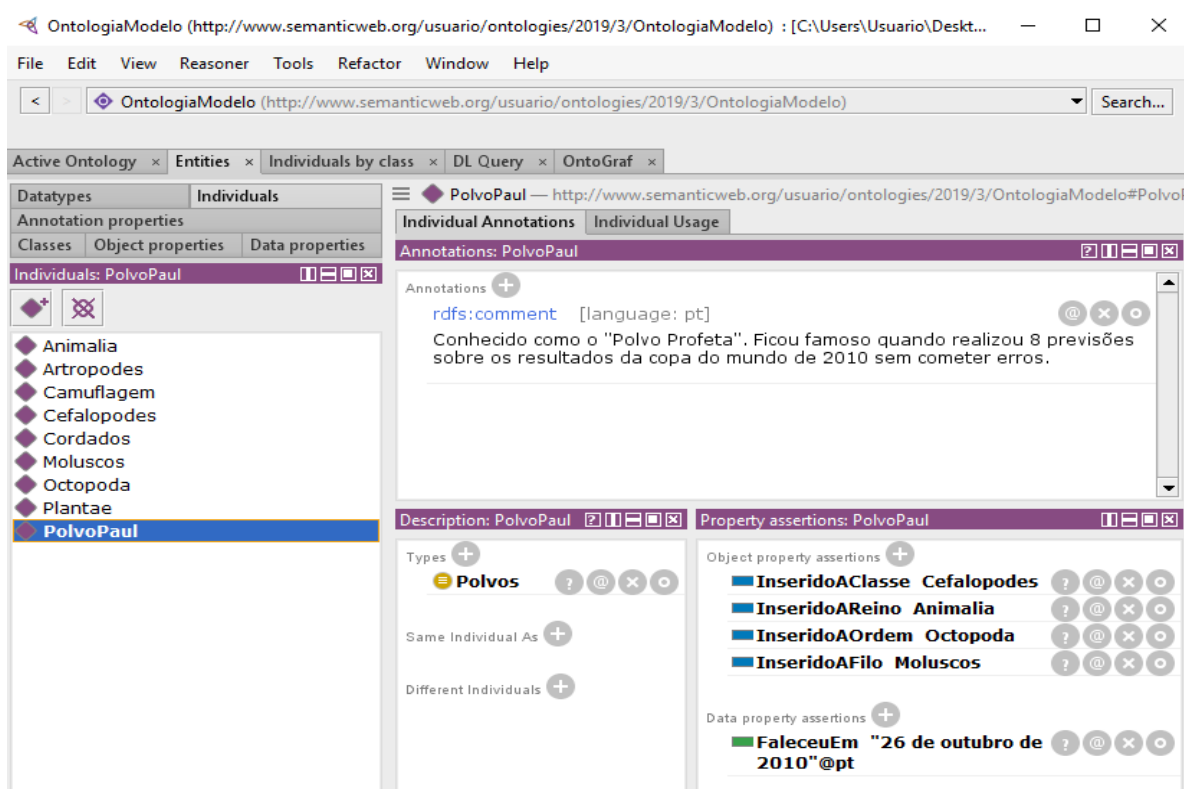


Fonte: elaborado pelo autor

A Figura 12 apresenta as propriedades da classe polvo. Na figura é mostrado que os polvos não possuem o mínimo de uma vertebra, estando assim inseridos no grupo dos invertebrados. Além disso mostra que são animais marinhos, isto é, invertebrados marinhos. Somado a isso, a terceira propriedade informa que os polvos possuem uma estratégia de sobrevivência que se denomina camuflagem.

Para cumprir o passo 5 as separações dentro das classes taxonômicas foram inseridas como propriedades. Assim, um polvo está em uma classe “animal” que tem como propriedade “inseridoAOrdem”, ou “inseridoAClasse”. Isso pode ser observado na Figura 13:

Figura 13 – Visão geral das classes e propriedades da ontologia modelo



Fonte: Elaborado pelo autor

Assim, à esquerda da Figura 13 podem ser observadas as instâncias criadas para a ontologia modelo, sendo que a maioria é referente ao animal polvo. A última que está selecionada em azul sob o nome de “PolvoPaul” representa a instância da qual serão analisadas as propriedades neste capítulo.

A instância “PolvoPaul” refere-se à um polvo em específico. Paul foi o nome dado a um polvo que em 2010 ficou famoso por sempre acertar os resultados dos jogos da copa do mundo naquele ano. Mantinham Paul em um aquário e colocavam dois alimentos iguais lado a lado no aquário. Cada alimento tinha uma bandeira dos times que estavam disputando a partida. A previsão consistia em observar qual dos dois alimentos seria escolhido primeiro pelo polvo como sua refeição. O número de acertos consecutivos tornou o polvo famoso. Esta explicação está resumida no quadro superior direito da Figura 13.

Ao centro dessa mesma figura pode ser observada a classe à qual o polvo Paul está inserido “polvos” que é uma classe que está localizada dentro da classe “Animais” na Figura 11.

No canto inferior direito da Figura 13 pode-se encontrar um dado a respeito do Paul especificando a data de seu falecimento que foi no mesmo ano em que se tornou famoso em 2010 sendo esta uma propriedade.

Baseando-se na metodologia OntoClean, é possível inserir esta propriedade no conceito de semirrigidez. Um falecimento sempre será essencialmente o mesmo pois, não se pode morrer mais que uma vez. Porém, a data não necessariamente é a mesma para os diferentes países do planeta terra.

Apesar dos humanos estarem vivendo ao mesmo tempo, não estão compartilhando todos do mesmo horário ou dia.

Sendo assim, a propriedade “FaleceuEm” não é uma propriedade essencial em todas as situações e contextos, por isso, esta é uma propriedade que tem a característica de ser antirrigida.

O passo 6 e 7 foram criados juntamente ao passo 5 pois todos são a partir das propriedades. Sendo assim a Figura 14 demonstra as propriedades juntamente com as instâncias.

Figura 14 – Propriedades da instância “PolvoPaul”



Fonte: Elaborado pelo autor

A primeira propriedade a ser analisada na Figura 14 é a “InseridoAClasse”, essa propriedade significa que o polvo Paul está designado a uma Classe taxonômica que engloba os Cefalópodes. É também uma propriedade rígida pois todos os polvos são essencialmente pertencentes à classe dos cefalópodes.

O mesmo acontece com as outras propriedades da ontologia como pode ser visto na segunda propriedade designada “InseridoAREino” que apesar de não fazer sentido gramatical foi grafada dessa forma para manter um padrão. Na classe taxonômica “reino” são divididas as plantas dos animais. Então esta propriedade informa que polvo é um tipo de animal, ou seja, do reino “animália”.

Por terem sido criadas com base na classificação taxonômica, as propriedades da ontologia modelo se tornaram essencialmente rígidas pois, a princípio, não existem alterações quanto está classificação.

Utilizar a metodologia Ontoclean permitiu com que a criação da ontologia modelo estabelecesse melhores relações conceituais pois já foram pensadas respeitando as noções de essência e rigidez, identidade e unicidade. As noções de identidade, por exemplo permitem com que sejam diferenciados os artrópodes dos moluscos. E além disso entender que a instância “PolvoPaul” representa uma instância única pois o polvo paul foi algo integral, completo como o conceito de “água” e “oceano” explicados anteriormente.

Criar a ontologia modelo permitiu perceber que com a sua criação, baseando-se em metodologias que auxiliam na definição dos termos e relações conceituais a tarefa de criar ontologias seja mais respaldada a alguma base do que apenas ser uma criação empírica.

Diante de todos os pressupostos teóricos e de todos resultados possibilitados pela análise da Classe 590 da Classificação Decimal de Dewey, da análise da *BBC*

Wildlife Ontology, da comparação dos dois SOC e da elaboração de uma ontologia modelo, esta pesquisa apresenta, na sequência, o último capítulo desta dissertação, as considerações finais.

8 CONSIDERAÇÕES FINAIS

Este trabalho teve o principal intuito o de comparar as relações conceituais presentes em dois sistemas de organização do conhecimento (SOC): A Classificação Decimal de Dewey (CDD) e uma Ontologia.

Para realizar este feito, foi realizada a análise de uma classe específica da CDD, a 590 sobre “animais”. Essa classe foi selecionada por indicação da ontologia analisada. Essa ontologia foi a *BBC wildlife ontology*, a qual foi localizada na base de dados BARTOC. A ontologia foi escolhida por ser de uma instituição de renome e julgou-se mais confiável do que uma ontologia de autoria anônima.

Para que a comparação fosse realizada de forma segura, foi necessário compreender os tipos de relações conceituais e suas características. Desse modo, elencou-se as três principais relações conceituais também encontradas nos tesouros, as associativas, hierárquicas e de equivalência. Então, de modo a dar base à futura análise, utilizou-se da norma ISO 25964 sobre tesouros.

Utilizou-se então a classe 590 referente a animais da CDD como forma de identificar as coerências e incoerências presentes nessa classe. Foi destacado que a CDD organiza o conhecimento de forma mais ampla para que seja adaptável a todas as bibliotecas.

Apesar de todos os méritos da CDD em organizar bibliotecas sua classificação não é perfeita, e foram encontradas classes que não estão relacionadas de maneira eficiente na CDD.

Como exemplo, notou-se que pássaros e aves são vistos como conceitos equivalentes, mesmo estes sendo coisas distintas. Além disso, constatou-se que algumas classes poderiam estabelecer relações hierárquicas ao invés de associativas. Isso pôde se percebido nas classes “invertebrados” e “invertebrados marinhos”, as quais estão relacionadas de maneira associativa quando na verdade são hierárquicas.

Para investigar a consistência das relações conceituais na *BBC wildlife Ontology* utilizou-se a metodologia para avaliação de ontologias denominada *OntoClean*. Nessa metodologia, conforme pôde ser elucidado nos procedimentos metodológicos, utilizou-se as noções formais propostas por seus criadores Guarino e Welty (2002), sendo elas: essência e rigidez, identidade e unicidade.

A princípio houve dificuldade em compreender cada um desses tópicos para que se pudesse fazer a análise. Com o uso deles foi possível perceber a inconsistência da *BBC wildlife ontology*.

Identificar a identidade e unicidade na ontologia demonstrou que as relações conceituais não estavam bem estabelecidas. Isso porque os termos por se repetirem constantemente acabaram por não serem tão diferenciados como deveriam.

Com a essência e rigidez também se notou que não foram feitas propriedades suficientemente rígidas para dar suporte às classes.

Por meio da análise de ambos os sistemas de organização do conhecimento de maneira separada foi possível, então compará-los.

A comparação foi realizada a princípio observando as razões pelas quais as ontologias necessitam de mais detalhamentos do que SOC mais tradicionais como a CDD.

Também se apontou para o fato de que apesar da ontologia ter conduzido à classe 590 da CDD, ambos os SOC não se referem exatamente às mesmas coisas. Nem mesmo contêm classes iguais para a comparação.

Dessa forma iniciou-se a comparação analisando que tanto um SOC quanto o outro possui incoerências em alguns aspectos e que isso se deve à complexidade em organizar o conhecimento.

Ainda que não existam expressivas semelhanças entre os SOC analisados, foi possível mostrar que havia pontos de concordância entre ambos. Na classificação taxonômica notou-se que a classe Filo da ontologia e as classes moluscos, artrópodes e cordados da CDD estavam diretamente relacionadas entre si.

Com este trabalho espera-se que haja aumento na quantidade de trabalhos se dediquem aos estudos voltados para as ontologias dentro do universo da ciência da informação. Uma das razões para este desejo é a de que as ontologias possam ser melhores compreendidas e, por conta disso, serem inseridas na realidade dos profissionais que destinam seus esforços ao trabalho com disciplina denominada ciência da informação.

REFERÊNCIAS

ALVARENGA, Lídia. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. **Encontros Bibli**, n. 15, jan./jun. 2003.

ANJOS, Liane dos. **Sistemas de classificação do conhecimento na filosofia e na biblioteconomia**: uma visão histórico-conceitual crítica com enfoque nos conceitos de classe, de categoria e de faceta. Tese (Doutorado em Ciência da Informação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2008.

ARAUJO, Vera Maria Araujo Pigozzi de. **Documentação, terminologia e lingüística**: uma interface produtiva. 2006. 163 f. Dissertação (Mestrado em Estudos da Linguagem) - Universidade Federal do Rio Grande do Sul. Porto Alegre, 2006

AULETE, Francisco Júlio de Caldas; VALENTE, Antonio Lopes dos Santos. **iDicionário digital Aulete**. Rio de Janeiro: Lexikon, 2018. Disponível em: <http://www.aulete.com.br>. Acesso em: 3 jul. 2018.

BARITÉ ROQUETA, Mario Guido. Sistemas de organización del conocimiento: una tipología actualizada. **Informação & Informação**, v. 16, n. esp., p. 122-139, jan./jun. 2011.

BARITÉ, Mario. Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación. In: CARRARA, K. (Org.). **Educação, Universidade e Pesquisa**. Marília: Unesp-Marília-Publicações; São Paulo: FAPESP, 2001. p. 35-60.

BARITÉ, Mario. et al. **Diccionario de organización del conocimiento**: clasificación, indización, terminología. Montevideo: PRODIC, 2013.

BEAN, Carol A.; GREEN, Rebecca (Ed.). **Relationships in the Organization of Knowledge**. Springer Science & Business Media, 2001

BRÄSCHER, Marisa; CARLAN, Eliana. Sistemas de organização do conhecimento: Antigas e novas linguagens. In: ROBREDO, Jaime; BRÄSCHER, Marisa (Orgs.). **Passeios no Bosque da Informação**: Estudos sobre Representação e Organização da Informação e do Conhecimento. Brasília DF: IBICT, 2010, 335, p. 147-176.

BREITMAN, Karin. **Web semântica**: a internet do futuro. Rio de Janeiro: LTC, 2005.

BRITISH BROADCASTING CORPORATION. **Wildlife Ontology**, London: BBC, 2018. Disponível em: <https://www.bbc.co.uk/ontologies/wo>. Acesso em: 04 out. 2018.

CAFÉ, Lígia Maria Arruda; SALES, Rodrigo de. Organização da informação: conceitos básicos e breve fundamentação teórica. In: ROBREDO, Jaime; BRÄSCHER, Marisa (Orgs.). **Passeios no bosque da informação**: estudos sobre representação e organização da Informação e do conhecimento. Brasília: IBICT, 2010. p. 115-129.

CAMPOS, Maria Luiza Almeida; GOMES, Hagar Espanha. Taxonomia e classificação: a categorização como princípio. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 8., 2007, Salvador. **Anais [...]**. Salvador: UFBA, 2007.

CARLAN, Eliana. **Ontologia e web semântica**. 2006. 60 f. Monografia (Graduação em Biblioteconomia)-Faculdade de Economia, Administração, Contabilidade e Ciência da Documentação e Informação, Universidade de Brasília, Brasília, 2006.

CINTRA, Anna Maria Marques; TÁLAMO, Maria de Fátima Golçalves Moreira; LARA, Marilda Lopes Gimenez de; KOBASHI, Nair YUMIKO. **Para entender as linguagens documentárias**. 2. ed. rev. ampl. São Paulo: Polis/APB, 1994.

CLARKE, Stella G. Deste. Thesaural relationships, In: BEAN, Carol A.; GREEN, Rebecca (Ed.). **Relationships in the Organization of Knowledge**. Springer Science & Business Media, 2001, p.37-52.

COSTELLO, Roger. L.; JACOBS, David B. **A quick introduction to OWL web ontology language**. 2003. 34 Slides.

DAHLBERG, Ingetraut. Teoria do conceito. **Ciência da informação**. Tradução: Astério Tavares Campos, v. 7, n.2, p.101-107, 1978a.

DAHLBERG, Ingetraut. Uma teoria para o interconcept: teoria analítica do conceito voltada para o referente. **International Classification**. Tradução: Vânia Teixeira Gonçalves, v. 5, n. 3, p. 142-151, 1978b.

DODEBEI, Vera Lucia Doyle Louzada de Mattos. **Tesouro**: linguagem de representação da memória documentária. Rio de Janeiro: Interciência, 2002.

GARCÍA JIMÉNEZ, Antonio. Instrumentos de representación del conocimiento: tesauros versus ontologías. **Anales de Documentación**, n. 7, p. 79-95, 2004.

GARCÍA MARCO, Francisco Javier. Avances en organización del conocimiento en España: los 11 encuentros sobre organización del conocimiento en sistemas de información y documentación. In: GARCÍA MARCO, Francisco Javier (Ed.). **Organización del conocimiento en sistemas de información y documentación**: actas del II Encuentro de ISKO-España, Getafe, 16 y 17 de noviembre de 1995. Zaragoza: Universidad de Zaragoza, 1995.

GARCÍA MARCO, Francisco Javier. Ontologías y organización del conocimiento: retos y oportunidades para el profesional de la información. **El Profesional de la Información**, v. 16, n. 6, p. 541-550, nov./dic. 2007.

GARCÍA MARCO, Francisco Javier. Paradigmas científicos en representación y recuperación de la información. In: GARCÍA-MARCO, Francisco Javier (Ed.). **Organización del conocimiento en sistemas de información y documentación**: actas del I Encuentro de ISKO-España, Madri, 4 y 5 nov. 1993. Zaragoza: Universidad de Zaragoza, 1993.

GIUNCHIGLIA, Fausto; ZAIHERAYEU, Ilya; FARAZI, Feroz. Converting classifications into OWL ontologies. In: ARTIFICIAL INTELLIGENCE AND SIMULATION OF BEHAVIOUR CONVENTION, 2009, Edinburgh. **Proceedings...** Edinburgh: [s.n.], 2009.

GNOLI, Claudio. Ontological foundations in knowledge organization: the theory of integrative levels applied in citation order. **Scire**, v. 17, n. 1, p. 29-34, ene./jun. 2011.

GREEN, Rebecca. Relationships in Organization of knowledge: an overview. In: BEAN, A.; GREEN, Rebecca (Ed.). **Relationships in the Organization of Knowledge**. Springer Science & Business Media, 2001.

GREEN, Rebecca; PANZER, Michael. Relations in the notational hierarchy of the Dewey Decimal Classification. In: SLAVIC, Aida; CIVALLERO, Edgardo (Ed.). **Classification and ontology: formal approaches and access to knowledge: proceedings of the International UDC Seminar, 19-20 September 2011, The Hague, The Netherlands**. Würzburg: Ergon Verlag, 2011. p. 161-176.

GREEN, Rebecca; PANZER, Michael. The ontological character of classes in the Dewey Decimal Classification. In: GNOLI, Claudio; MAZZOCCHI, Fulvio. (Ed.). **Paradigms and conceptual systems in knowledge organization: proceedings of the 11th International ISKO Conference, 23-26 feb. 2010, Rome, Italy**. Würzburg: Ergon Verlag, 2010. p. 171-179.

GUARINO, Nicola; GIARETTA, Pierdaniele. Ontologies and knowledge bases: towards a terminological clarification. In: MARS, Nicolaas J. I. (Ed.). **Towards very large knowledge bases: knowledge building and knowledge sharing**. Amsterdam: IOS, 1995. p. 25-32.

GUARINO, Nicola; WELTY, Christopher Evaluating Ontological Decisions with OntoClean. **Communications of the ACM**, v. 45 n.2 p. 61–65 2002.

HJØRLAND, Birger. Concept theory. **J. Am. Soc. Inf. Sci.**, n.60, p.1519–1536, 2009.

HJØRLAND, Birger. Knowledge organization = information organization? In: NEELAMEGHAN, A.; RAGHAVAN, K. S. (Eds.). **Categories, contexts and relations in knowledge organization: proceedings of the 12th International ISKO Conference, 6-9 aug. 2012, Mysore, India**. Würzburg: Ergon-Verlag, 2012. p. 8-14.

HJØRLAND, Birger. Theories of knowledge organization: theories of knowledge. **Knowledge Organization**, v. 40, n. 3, p. 169-181, 2013.

HODGE, Gail. **Systems of knowledge organization for digital libraries: beyond traditional authority files**. Washington: Council on Library and Information Resources, 2000.

HORRIDGE, Matthew. **A practical guide to building OWL ontologies using Protégé 4 and CO-ODE Tools**. [S.I.]: University of Manchester, 2009.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 25964-1:2011**: Information and documentation – thesauri and interoperability with other vocabularies – part 1: Thesauri for information retrieval. Genebra, 2011.

JACOB, Elin K. Classification and categorization: a difference that makes a difference. **Library Trends**, v. 52, n. 3, p. 515-540, 2004.

LARA, Marilda Lopes Ginez de. Propostas de tipologias de KOS: uma análise das referências de formas dominantes na organização do conhecimento. **Encontros Bibli**, v. 20, n. esp. 1, p. 89-107, fev. 2015.

LIMA, José Leonardo Oliveira; ALVARES, Lillian. Organização e representação da informação e do conhecimento. In: ALVARES, Lillian (Org.). **Organização da informação e do conhecimento**: conceitos, subsídios interdisciplinares e aplicações. São Paulo: B4 Editores, 2012. p. 21-48.

MAGALHÃES, Lana. **Reino animal**. [S.l.]: Toda Matéria, 2018. Disponível em: <https://www.todamateria.com.br/reino-animal>. Acesso em: 23 mar. 2019.

MAI, Jens-Erik. The future of general classification. **Cataloging & Classification Quarterly**, v. 37, n. 1-2, p. 3-12, 2003.

MAZZOCCHI, Fulvio. Relations in KOS: is it possible to couple a common nature with diferente roles? **Journal of Documentation**, v. 73, n. 2, p. 368-383, 2017.

MIRANDA, Marcos Luiz Cavalcanti de. Organização e representação do conhecimento: fundamentos teórico-metodológicos na busca e recuperação da informação em ambientes virtuais. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 7., 2005, Marília. **Anais [...]**. Marília: UNESP/FFC, 2005.

MITCHELL, Joan S. et al. (Ed). **Dewey Decimal Classification and relative index**: edition 23. Dublin: OCLC, 2011.

MITCHELL, Joan S. Relationships in the Dewey Decimal Classification System. In: BEAN, Carol A.; GREEN, Rebecca (Ed.). **Relationships in the organization of knowledge**. Dordrecht: Kluwer Academic, 2001, p. 211-226.

MITCHELL, Joan S.; GOETZ, Diane Vizine. The Dewey Decimal Classification. In: BATES, Marcia J.; MAACK, Mary Niles (Ed.). **Encyclopedia of Library and Information Science**. 3. ed. Boca Raton: CRC Press, 2009.

MOREIRA, Alexandra. **Tesauros e ontologias**: estudo de definições presentes na literatura das áreas das ciências da computação e da informação, utilizando-se o método analítico-sintético. 2003. 150 f. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2003.

MOREIRA, Walter. Teoria da classificação e ontologias: em busca do diálogo necessário. **Scire**, v. 19, n. 2, p. 63-70, jul./dic. 2013.

MUNK, Timme B.; MORK, Kristian. Folksonomy, the power law & the significance of the least effort. **Knowledge Organization**, v. 34, n. 1, p. 16-33, 2007.

NICOLINO, Maria Elisa Valentim Pickler. **Diretrizes para a utilização de ontologias na indexação automática**. 2014. 101 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.

NOY, Natalya F.; MCGUINNESS, Deborah L. **Ontology development 101**: a guide to creating your first ontology. Stanford: Stanford University, 2001.

SALES, Luana Farias.; CAMPOS, Maria Luiza Almeida; GOMES, Hagar E. Ontologias de domínio: um estudo das relações conceituais. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.13, n. 2, p. 62-76, maio/ago., 2008.

SALES, Luana Farias. Relações conceituais para instrumentos de padronização terminológica: um novo modelo para uso em ontologias. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 8., 2007, Salvador. **Anais [...]**. Salvador: UFBA, 2007.

SALES, Rodrigo de; CAFÉ, Lígia Maria Arruda. Diferenças entre tesouros e ontologias. **Perspectivas em Ciência da Informação**, v. 14, n. 1, p. 99-116, jan./abr. 2009.

SANTAREM SEGUNDO, José Eduardo. **Representação Iterativa**: um modelo para repositórios digitais. 2010. 224 f. Tese (Doutorado em Ciência da Informação)- Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

SANTOS, Cibele Araújo Camargo Marques dos; MAZINI, Elizabeth Sardelli. Organização do conhecimento: das classificações e vocabulários controlados às taxonomias e ontologias na web. In: VALLS, Valéria Martin; VERGUEIRO, Waldomiro (Org.). **Tendências contemporâneas na gestão da informação**. São Paulo: Sociologia e Política, 2011. p. 125-142.

SETZER, Valdemar W.; Dado, Informação, Conhecimento e Competência. **DataGramZero**, 1999.

SILVA, André Felipe Pereira da. **Conversão de sistemas de classificação em ontologias**: análise de viabilidade utilizando a CDD e o Protégé. 2017. 72f. Trabalho de Conclusão de Curso (Bacharel em Biblioteconomia) – Universidade Estadual Paulista Marília, 2017.

SMIT, Johanna Wilhelmina. **O que é documentação**. 2.ed. São Paulo: Brasiliense, 1987.

SOERGEL, Dagobert. The rise of ontologies or the reinvention of classification. **Journal of the American Society for Information Science**, v. 50, n. 12, p. 1119-1120, oct., 1999.

SOLER MONREAL, Concha; GIL LEIVA, Isidoro. Posibilidades y límites de los tesauros frente a otros sistemas de organización del conocimiento: folksonomías, taxonomías y ontologías. **Revista Interamericana de Bibliotecología**, v. 33, n. 2, p. 361-377, jul./dic. 2010.

SOUZA, Renato Rocha; TUDHOPE, Douglas; ALMEIDA, Maurício Barcellos. Towards a taxonomy of KOS: dimensions for classifying knowledge organization systems. **Knowledge Organization**, v. 39, n. 3, p. 179-192, 2012.

VIGNOLI, Richele Grengé; ALMEIDA, Patrícia Ofélia Pereira de; CATARINO, Maria Elisabete. Folksonomias como ferramenta da organização e representação da informação. RDBCI: **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, SP, v. 12, n. 2, p. 120-135, mai. 2014. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/1606>. Acesso em: 21 ago. 2018.

WIKIPEDIA. **Wikipédia**: a enciclopédia livre. 2019. Disponível em: <https://pt.wikipedia.org/>. Acesso em: 3 abr. 2019.

WIKI PROTÉGÉ. **Welcome to the Protégé wiki!**. 2018. Disponível em: https://protegewiki.stanford.edu/wiki/Main_Page. Acesso em: 3 out. 2018.

ZENG, Marcia Lei; PANZER, Michael; SALABA, Athena. Expressing classification schemes with OWL 2. In: GNOLI, Claudio; MAZZOCCHI, Fulvio. (Ed.). **Paradigms and conceptual systems in knowledge organization**: proceedings of the 11th International ISKO Conference, 23-26 feb. 2010, Rome, Italy. Würzburg: Ergon-Verlag, 2010. p. 356-362.