

3 CICLO DE ESTUDOS
DOUTORAMENTO EM INFORMAÇÃO E COMUNICAÇÃO EM PLATAFORMAS DIGITAIS

Análise das relações semânticas
implícitas em folksonomias:
contributos para uma ontologia sobre
diabetes

Jacqueline A. Souza

D

2018



Jacqueline A. Souza

**Análise das relações semânticas implícitas em
folksonomias: contributos para uma ontologia sobre
diabetes**

Tese realizada no âmbito do Doutoramento em Informação e Comunicação em Plataformas Digitais, orientada pela Professora Doutora Olívia Pestana, Professora Auxiliar da Faculdade de Letras da Universidade do Porto.

Faculdade de Letras da Universidade do Porto

Janeiro de 2018

FICHA CATALOGRÁFICA

J729a Souza, Jacqueline Aparecida de.
Análise das relações semânticas implícitas em folksonomias:
contributos para uma ontologia sobre diabetes/ Jacqueline Aparecida de
Souza. – Porto, 2018.
243p. il.
Tese (Doutoramento em Informação e Comunicação em Plataformas
Digitais) – Universidade do Porto/Universidade de Aveiro.
1. Organização da Informação. 2. Indexação Social. 3. Indexação
colaborativa. 4. Terminologia. 5. Ontologia

CDU 025.4

O Júri

**Análise das relações semânticas implícitas em
folksonomias: contributos para uma ontologia sobre
diabetes**

Jacqueline A. Souza

Tese realizada no âmbito do Doutoramento em Informação e Comunicação em Plataformas Digitais, orientada pela Professora Doutora Olívia Manuela Marques Pestana.

Membros do Júri

Professora Doutora Maria da Graça Lisboa Castro Pinto - Presidente
Faculdade de Letras da Universidade do Porto

Professora Doutora Maria da Graça Melo Simões - Vogal
Faculdade de Letras da Universidade de Coimbra

Professora Doutora Helena Topa Valentim - Vogal
Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa

Professora Doutora Olívia Manuela Marques Pestana - Vogal
Faculdade de Letras da Universidade do Porto

Professor Doutor Óscar Mealha - Vogal
Departamento de Comunicação e Arte da Universidade de Aveiro

Professora Doutora Maria Elisa Cerveira - Vogal
Faculdade de Letras da Universidade do Porto

À memória de Elenice Braga de Souza

Sumário

Declaração de honra	9
Agradecimentos	11
Resumo.....	13
Abstract	14
Lista de figuras	15
Lista de quadros	17
Introdução	19
Contextualização e motivação	19
Questão de investigação.....	28
Hipóteses	29
Objetivos	29
Estrutura da tese	29
Capítulo 1 - Contributos teórico-metodológicos para estruturação semântica	33
1.1 A classificação como estrutura do pensamento: breves apontamentos	33
1.2 Classificação facetada: estrutura dinâmica e multidimensional	39
1.2.1 Princípios ranganathianos e os elementos da estrutura classificatória ...	46
1.3 Terminologia: abordagens e princípios	52
1.3.1 Especificidades terminológicas em saúde	58
1.4 Teoria do conceito: elementos constitutivos	64
1.5 Tesouros: em foco a estruturação conceitual	67
1.5.1 Estrutura e componentes fundamentais.....	73
Capítulo 2 - Esquemas para organização da informação em contexto digital.....	82
2.1 Folksonomias: fonte para estudos semânticos	85
2.1.1 Conceitos associados	94
2.1.2 Sistemas Baseados em Folksonomias.....	96
2.2 Ontologias: inter-relações conceituais complexas	101

2.3 Ontologia e folksonomia no âmbito da saúde	106
2.4 Integração entre folksonomia e ontologia.....	111

Capítulo 3 - Relações semânticas na perspectiva da Ciência da

Informação	125
3.1 Semântica: concepções lacônicas.....	125
3.2 A conexão entre semântica e organização da informação	132
3.3 Relações semânticas: conceito e propriedades.....	137
3.4 Tipologia das relações semânticas.....	143
3.4.1 Relações de equivalência	150
3.4.2 Relações hierárquicas.....	153
3.4.3 Relações associativas.....	154
3.5 Relações semânticas no domínio da saúde.....	164

Capítulo 4 - Percurso investigativo..... 166

4.1 Abordagens metodológicas	167
4.1.1 Análise de domínio.....	167
4.1.2 Análise conceitual	169
4.2 Instrumentos e procedimentos.....	171
4.3 Descrição das etapas.....	173

Capítulo 5 - O domínio da Diabetes: caracterização do seu léxico a partir de folksonomia

176	176
5.1 Apontamentos e especificidades sobre diabetes.....	176
5.2 Qualificadores do decs e as tags.....	192
5.3 Categorização a partir das tags e do contexto.....	195
5.4 Relações semânticas no domínio da diabetes.....	199

Considerações finais

213	218
Referências	218

Apêndices.....	232
Apêndice 1 - Indexações livres para o termo diabetes	233
Apêndice 2 - Protocolo para validação preliminar dos conceitos	245
Apêndice 3 - Protocolo parcial – retorno do especialista	246

DECLARAÇÃO DE HONRA

Declaro que a presente tese “**Análise das relações semânticas implícitas em folksonomias: contributos para uma ontologia sobre diabetes**” é de minha autoria e não foi utilizada previamente noutro curso ou unidade curricular, desta ou de outra instituição. As referências a outros autores (afirmações, ideias, pensamentos) respeitam escrupulosamente as regras da atribuição, e encontram-se devidamente indicadas no texto e nas referências bibliográficas, de acordo com as normas de referência. Tenho consciência de que a prática de plágio e autoplágio constitui um ilícito académico.

Porto, 18 de junho de 2018

Jacqueline Aparecida de Souza

AGRADECIMENTOS

À Professora Doutora Olívia Pestana, pela oportunidade, confiança e profissionalismo.

À Faculdade de Letras da Universidade do Porto e a todos os professores. Um agradecimento especial ao Professor Doutor Armando Malheiro da Silva, pelo acolhimento e pelas orientações.

À Universidade Federal do Rio Grande do Norte e aos professores do Departamento de Ciência da Informação, por viabilizarem este aprimoramento.

À João Castro pelas orientações e sugestões.

À Gracinéia Oliveira pelas recomendações.

À minha família pelo amor, confiança e, em especial, minha irmã Daiane.

À Mário Gonzaga pelo afeto e carinho.

Aos amigos do doutoramento, Monica, Thais, Laerte, Dolores e Fernando, por todos os incentivos, conversas descontraídas e partilhas enriquecedoras.

À Ana Roberta pela amizade e companheirismo.

Aos amigos da residência Campo Alegre III, Aline Alvarenga, Fernanda, Aurineide, Cidália, Dinângela.

Aos amigos queridos que residem no Brasil e foram essenciais para conclusão dessa jornada: Francisco de Assis N. G. de Araújo, Madalice A. Jorge, Aline Brito, Manuela Oliveira e a prima Danusa Andrade. Muito obrigada pelo apoio e amizade dedicados a mim.

RESUMO

Esta tese aborda um fenômeno característico do paradigma tecnológico no escopo da Organização da Informação, a folksonomia. Enquanto sistema de etiquetagem livre, que apesar de apresentar limitações decorrentes do uso da língua natural e a falta de controle do vocabulário, detém um conhecimento útil para aprimorar o tratamento da informação em contexto digital. Como uma das vantagens, por meio de conceitos e relacionamentos semânticos implícitos, oferece contributos para a estruturação semântica. Considerando o desafio de aliar os contributos da folksonomia à ontologia no contexto da saúde, o objetivo é perceber como as **relações semânticas** implícitas nas **folksonomias** podem ser incorporadas na **estruturação** de uma **ontologia sobre diabetes**. Especificamente, pretende sistematizar um núcleo essencial de relações semânticas, perceber aspectos da representação da informação sobre diabetes em contexto digital, a partir dos sistemas baseados em folksonomias, reconhecer quais são os qualificadores sobre o conceito diabetes, identificar os padrões de relações semânticas sobre o tema e verificar os contributos para estruturação de uma ontologia a partir de um núcleo de relações semânticas extraído da folksonomia. Para consecução da investigação de caráter exploratório, adotamos como abordagem metodológica a análise de domínio, e a análise conceitual. Foram coletadas de 500 indexações livres no CiteULike, com 184 tags. O estudo revelou que a diabetes é debatida expressivamente sob a perspectiva da Epidemiologia, assim como coocorre com outras doenças, transtornos e disfunções, referindo-se ao risco de desenvolvê-las. Neste sentido, discute-se o conceito de Risco como uma possibilidade de relação estruturante. Adicionalmente, verifica-se a predominância das relações Ação, Processo e Causa/efeito. Conclui-se que, em certa medida e no domínio estudado, a folksonomia fornece subsídios para a estruturação de uma ontologia.

Palavras-chave: Organização da Informação. Indexação Social. Indexação colaborativa. Terminologia. Ontologia.

ABSTRACT

This thesis deals with a phenomenon that is a characteristic of the technological paradigm present in Information Organization, folksonomy. As a free-form tagging system, it presents limitations due to use of natural languages and lack of controlled vocabulary. However, it provides useful knowledge that enables information treatment in the digital context. One of its advantages is that it contributes to semantic structuring. This is done by means of the use implicit semantic concepts and relationships. Considering the challenge of allying the contribution of folksonomy to ontology in the context of health, the objective of this work is to understand how implicit semantic relations in folksonomies can be incorporated in the structuring of an ontology regarding diabetes. Specifically, it aims to systematize an essential nucleus of semantic relations in this realm, perceive the aspects present in the information representation in the digital context considering Folksonomy-based systems. It also aims to recognize qualifiers present in the concept of diabetes as well as in the semantic relations patterns and verify the contributions for structuring an ontology from a nucleus of semantic relations extracted from folksonomy. In order to achieve these objectives, the research of exploratory nature, relied on a domain analysis approach, a case study and conceptual analysis. Thus, 500 free-form tagging was collected through in CiteULike, with 184 tags. The study revealed that diabetes is debated expressively from the perspective of Epidemiology, as well as co-occurring with other diseases, disorders and dysfunctions, referring to the risk of developing them. In this sense, the concept of Risk is discussed as a possibility of structuring relationship. In addition to this, there is predominance of the relations Action, Process and Cause/effect is verified. It is possible to conclude that, to a certain extent and in the studied field folksonomy provides subsidies for the structuring of an ontology.

Keywords: Organization of information. Social indexing. Collaborative indexing. Terminology. Ontology.

LISTA DE FIGURAS

Figura 1- Árvore de Porfírio.....	38
Figura 2 - Poli-hierarquia no DeCS.....	63
Figura 3 - Triângulo conceitual.....	65
Figura 4 - Termos e a composição dos tesauros	74
Figura 5 - Exemplo de relações hierárquicas no DeCS	76
Figura 6 - Exemplo de relação de equivalência	77
Figura 7 - Exemplo de termo relacionado	79
Figura 8 - Relação associativa: diabetes mellitus e diabetes insípido	80
Figura 9 - Conjunto de princípios e práticas da web 2.0.....	83
Figura 10 - Triângulo da folksonomia	85
Figura 11 - Formas de visualização das tags.....	86
Figura 12 - Folksonomia: processos colaborativos ou coletivo	92
Figura 13 - Folksonomia aberta x folksonomia restrita	93
Figura 14 - Compartilhamento de tags entre os utilizadores	94
Figura 15 - Processamento de linguagem natural e folksonomia.....	100
Figura 16 - Ontologia organizada com base na complexidade semântica.....	104
Figura 17 - Integração da folksonomia e ontologia para aquisição e organização ..	116
Figura 18 - Ontologia baseada em folksonomia.....	118
Figura 19 - Tagging e modelo ontológico	119
Figura 20 - Ontologia estruturada por tags.....	120
Figura 21 - Folskonomia assistida	123
Figura 22 - Ontologia folksonomizada	124
Figura 23 - Semântica lexical - componentes	131
Figura 25 - Diferença entre relação sintagmática e paradigmática	147
Figura 26 - Relações semânticas e distintas denominações.....	148
Figura 27 - CiteULike - coleta das tags	172
Figura 28 - Busca por conceitos no DeCS	172
Figura 29 - Áreas limítrofes	191
Figura 30 - Edição preliminar da ontologia.....	210
Figura 31 - Ontologia: Classes e subclasses	210
Figura 32 - Ontologia: indivíduos	211

Figura 33 - Ontologia: propriedades	211
Figura 34 - Triplos semânticos.....	212

LISTA DE QUADROS

Quadro 1 - Categorias de Aristóteles	36
Quadro 2 - Tabelas principais e auxiliares da CDD	42
Quadro 3 - Tabelas da classificação da biblioteca do congresso.....	42
Quadro 4 - Tabelas principais e auxiliares da CDU	44
Quadro 5 - Símbolos utilizados na CDU.....	44
Quadro 6 - Cânones do plano ideacional	48
Quadro 7 - Categorias por Ranganathan x CRG	51
Quadro 8 - Características dos conceitos	67
Quadro 9 - Relações associativas segundo a norma Z39.19	80
Quadro 10 - Aspectos do vocabulário controlado e das folksonomias.....	88
Quadro 11- Principais sistemas baseados em folksonomia	97
Quadro 12 - Enfoques e funcionalidades do LibraryThing e CiteULike.....	98
Quadro 13 - Métodos para construção de ontologias	105
Quadro 14 - Ontologias na área da saúde	107
Quadro 15 - Comparação entre ontologia e folkonomia	114
Quadro 16 - Abordagens de estudos semânticos em CI	134
Quadro 17 - Propriedade gerais das relações semânticas.....	139
Quadro 18 - Propriedades lógicas	140
Quadro 19 - Propriedades lógicas nas relações paradigmáticas	141
Quadro 20 - Relações semânticas e situações reais.....	142
Quadro 21 - Ocorrências significativas nas normas ISO 25964 e ANSI/NISO Z39.19-2005	145
Quadro 22 - Classes de sinônimos na norma ISO 25964	151
Quadro 23 - Tipos e exemplos de variantes lexicais	152
Quadro 24 - Relações associativas	160
Quadro 25 - Relações semânticas em saúde	164
Quadro 26 - Tipos de análise conceitual	170
Quadro 27 - Exemplificação das indexações livres	176
Quadro 28 - Lista de tags.....	178
Quadro 29 - Qualificadores no DeCS e as tags	193
Quadro 30 - Categorias e manifestação.....	196

Quadro 32 - Relações associativa no DeCS para o termo diabetes	200
Quadro 33 - Exemplos de relações semânticas sobre diabetes	203
Quadro 34 - Tag diabetes, risco e fator de risco	204
Quadro 35 - Síntese das relações semânticas sobre diabetes	207

INTRODUÇÃO

A tese intitulada “**Análise das relações semânticas implícitas em folksonomias: contributos para uma ontologia sobre diabetes**” desenvolve-se no âmbito do programa doutoral **Informação e Comunicação em Plataformas Digitais**, com o intuito de investigar as relações semânticas implícitas nas folksonomias, com vista a estruturar uma ontologia sobre Diabetes.

Perceber as relações semânticas¹ entre as tags² que compõem as folksonomias, embasadas em uma abordagem epistemológica da Organização da Informação (OI) e em consonância com os novos paradigmas infocomunicacionais, permitirá refletir com aprofundamento sobre a temática, assim como apontar métodos e técnicas para melhorar a recuperação da informação em contextos digitais. Pressupõe-se que, nas representações do utilizador, via folksonomia, há um conhecimento semântico estabelecido por meio das relações implícitas entre os conceitos, que, ao tornarem-se explícitas, contribuirão para delinear um sistema de conceitos e relações, como em uma ontologia.

CONTEXTUALIZAÇÃO E MOTIVAÇÃO

Estimulada pelas Tecnologias de Informação e da Comunicação (TIC), a sociedade contemporânea caracteriza-se pela alta interatividade e pela conexão entre muitos mundos, com destaque para a redução de intermediários nos processos de produção, armazenamento, recuperação, transmissão e acesso à informação. A lógica de rede, enquanto estrutura social básica, constantemente impõe desafios em todos os contextos, os quais se refletem

¹ Khoo e Na (2006) consideram relações semânticas como relações significativas entre dois ou mais conceitos, entidades ou conjunto de entidades. Comumente, as relações semânticas abarcam as relações de equivalência (sinonímia; antonímia), hierarquia (gênero e espécie ou todo e parte) e as associativas (que não são de equivalência, nem apresentam subordinação).

² Nesta tese, ao referir-se a atribuição de rótulos a recursos informacionais de forma aberta e compartilhada pelo utilizador, em ambiente digital, tratar-se-á como tags com vistas à precisão terminológica.

nos relacionamentos humanos, na comunicação e na própria composição da rede. O conjunto de mudanças paradigmáticas apresenta conflitos e desafios, os quais são importantes reconhecer e discutir, a fim de conhecer suas fronteiras, limitações e buscar por novas metodologias. Vive-se num mundo que se tornou digital e o que é designado como revolução tecnológica (Castells, 2003) não é a centralidade do conhecimento e da informação, mas a aplicação destes na produção de dispositivos de processamento/comunicação da informação, num ciclo de realimentação cumulativo entre inovação e uso.

Neste cenário, configurado pela alta interatividade, no qual cruzam-se conceitos-chave - *web 2.0* ou *web social*, ciberespaço e inteligência coletiva (Levy, 2009), infoesfera (Floridi, 2010), cultura de convergência (Jenkins, 2009) e apoiado pela mediação tecnológica, é que a Ciência da Comunicação e da Informação (CCI) se articula para desenvolver pesquisas e consolidar seu arcabouço teórico-conceitual.

A Organização da Informação, por ser a subárea que se dedica à proposição de teorias e metodologias para análise e condensação de conteúdo informacional e desenvolver produtos documentários³, busca lidar com um novo elemento do paradigma tecnológico, demarcado pela autonomia do utilizador na produção e consumo de informação e o enfoque nas linguagens naturais que, tradicionalmente, tinham seus fenômenos atenuados pela aplicabilidade das linguagens de indexação⁴.

A linguagem natural, também compreendida como texto livre, sinônimo de discurso comum, utilizada habitualmente na fala e na escrita, sem padronização ou controle de vocabulário, apresenta como fenômenos naturais a ambiguidade, polissemia e sinonímia. Em face a sua natureza plurissêmica, as atividades de tratamento da informação em ambientes tradicionais pretendiam fixar significados. Neste seguimento, Lancaster (2004) reconhece

³ Entende-se por produtos documentários os resumos, lista de cabeçalhos, tesouros, ontologias.

⁴ Linguagens de indexação trata-se de uma lista de termos autorizados, em forma de uma estrutura semântica, cuja finalidade é controlar sinônimos, diferenciar homógrafos, ligar termos cujos significados apresentem alguma relação (Lancaster, 2004). São instrumentos que profissional auxiliam o profissional a representar o conteúdo informacional, para aprimorar e mediar o acesso ao conteúdo ou recurso informacional.

que a representação via linguagem natural será mais específica e atual e também destaca que é dinâmica e funciona bem em muitos contextos para distintos objetivos.

Caracterizada pela autonomia do utilizador e com enfoque na linguagem livre, a folksonomia apresenta-se como uma nova forma de classificar o conteúdo na *web*. É resultante da atribuição de tags a recursos disponíveis em ambientes digitais, as quais podem ser compreendidas como uma indexação livre e pessoal, em linguagem natural, sem regras e controle do vocabulário. Os Sistemas Baseados em Folksonomia (SBF) como Citeulike⁵, Delicious⁶ e Flickr⁷, tornaram-se populares por serem considerados simples: basta ter o recurso digital (sítios da *web*, foto, trabalho acadêmico, vídeo) e atribuir uma etiqueta (*tag*), a qual deve ser compartilhada e aberta a outros utilizadores, promovendo o acesso e conferindo visibilidade ao recurso informacional.

As folksonomias exibem as tags agrupadas (em formato de nuvens ou listas), sem impor uma estrutura fechada, pré-estabelecida, no qual a relação semântica entre os conceitos não é explícita. Ademais, promove alta contextualização semântica, decorrente da relação associativa entre as tags. Apesar de suas limitações, como a falta de controle terminológico, erros ortográficos e ambiguidade, são abordadas sob diversas perspectivas, a saber: para conhecimento de determinada comunidade, para observação de novas práticas sociais, como uma nova metodologia e estratégia para indexação.

Nesse ínterim, a Organização da Informação tem como desafio tratar os registros de conhecimento disponíveis em ambiente digital, onde as representações são efetuadas via conjunto de conceitos, no qual as relações semânticas (implícitas ou explícitas) são um elemento-chave comum entre as folksonomias e as ontologias. Associada a noção de *web semântica*, as ontologias são uma “especificação explícita de uma conceitualização (Gruber, 1993)”, cujo propósito é garantir inteligibilidade nas comunicações entre humanos e máquinas.

⁵CiteULike – os recursos informacionais etiquetados são trabalhos acadêmicos: <http://www.citeulike.org/>

⁶ Delicious – os recursos informacionais etiquetados são websites: <https://delicious.com/>

⁷ Flickr : os recursos informacionais etiquetados são fotografias: <https://www.flickr.com/>

Consoante a isso, Basso, Ferreira e Silva (2009) reforçam que uma forma de melhorar significativamente a organização da informação, e sua posterior recuperação, é pelo uso de ontologias, as quais representam conhecimento estruturado por meio de conceitos, instâncias, atributos e relações que são modelados na forma de um grafo ou rede. As ontologias permitem a definição de vários tipos de relacionamentos tais como: o pai-filho, o todo-parte, o causa-efeito, o de associação, entre outros. Estas relações semânticas mais estruturadas podem tornar o processo de recuperação da informação menos custoso para o utilizador, pois ele terá outras formas de acessar a mesma informação. O autor ainda expõe que o problema de utilizar ontologias para classificação da informação é que esta é uma tarefa que também tem um alto custo cognitivo para o utilizador. Isso ocorre, principalmente, porque as ontologias, assim como as taxonomias, são difíceis de construir e manter, pois a coerência da ontologia deve ser mantida após a classificação de uma nova informação.

Os estudos dos aspectos semânticos em Organização da Informação devem pautar-se em teorias semânticas que forneçam respaldo teórico e metodológico capaz de orientar a análise conceitual e os desafios do processamento automático de textos impostos pelas demandas de recuperação inteligente da informação, que levam a modelos de representação cada vez mais complexos (Café & Medeiros, 2011). Neste sentido, reforça-se o necessário diálogo entre distintas áreas do conhecimento, especificamente Língua, Computação, Comunicação e Ciência da Informação, com vistas a estruturação semântica.

Acerca das relações semânticas, Maculan e Lima (2014, p.185), expõem que o termo “relações” ou “relacionamentos” possui um sentido amplo que abarca a pluralidade de conexões existentes entre objetos. Estas relações, comumente são de equivalência, hierarquia ou associativas e cada uma contém noções semânticas lexicais e cognitivas. Os diferentes sistemas de Organização da Informação ou as linguagens de indexação possuem diferentes propósitos e níveis de controle da terminologia. A exemplo disso, o tesauro como linguagem artificial e normalizada, com fins de controle

terminológico de uma área de especialidade, estrutura elementos linguísticos que o compõem e apresenta relacionamentos existentes entre eles.

Entretanto, investigar sobre as relações semânticas que compõem parte de um universo do conhecimento, especificamente na área de saúde, permite apontar a redução das diferenças entre as representações e configurações conceituais apresentadas por sistemas de informação e as representações mentais do conhecimento especializado de seus utilizadores. A perspectiva de ampliação da dimensão semântica de sistemas artificiais pode ainda contribuir para uma maior aproximação entre atores humanos e artificiais, otimizando a interação e o diálogo entre ambos. As relações semânticas estabelecidas entre as unidades de conhecimento, os conceitos, auxiliam nos processos de raciocínio e na tomada de decisão, e os sistemas de informação têm suas potencialidades ampliadas (Biolchini, 2001).

Evangelista, Simões e Guimarães (2016) apoiados em Kim (2006) afirmam que as relações semânticas entre termos-conceitos são identificadas de forma mais precisa a partir de meios intelectuais. Embora a indexação automática seja uma realidade, o estabelecimento de relações semânticas entre termos e conceitos é uma operação mais consistente quando é feita por humanos, o que destaca ainda mais o papel da especificidade no definir dessas relações.

Implícitas nas folksonomias, as relações associativas, segundo Simões (2008, p.123), são mais enriquecedoras e permitem ao utilizador uma navegação horizontal, ou seja, possibilita pesquisar na mesma categoria conceitual, assim como pesquisar em diferentes categorias conceituais. Estas relações, complementares às hierárquicas, assumem a função de evitar a poli-hierarquia, pois tornam possível substituir a dependência de um termo específico sob dois ou mais termos genéricos, por um termo relacionado, contribuindo para que a estrutura seja simples, de fácil compreensão e constitua um valor acrescentado aos utilizadores de áreas distintas, multidisciplinares.

No tocante às relações semânticas em folksonomia e ontologia, Peters e Weller (2008) expõem três observações:

First of all, the current efforts to establish a Semantic Web demand for more sophisticated depiction of knowledge relationships. Ontologies, the core technique for indexing documents with semantic representations, can make use of a wider range of relations than classical KOS⁸. A closer look at the expressiveness of logic-based ontology languages and an examination of current ontologies will help to redefine a classification of relationship types. The second and less obvious stimulus lies in upcoming folksonomies and the relations implicitly existent between user-created tags in social tagging systems. Both new methods of knowledge representation may be a valuable resource for reconsidering existing relations and establishing new generalizable ones. Finally, we point out several aspects in establishing knowledge-relationships, which are of importance for information structuring and retrieval and should become subjects of future discussions (Peters e Weller, 2008, p.100).

Complementarmente, para derivar ontologias deve-se levar em conta múltiplas dimensões, uma vez que as folksonomias consistem em uma estrutura tríade: as **tags** são associadas por **pessoas** aos **recursos informacionais**. Este aspecto deve ser considerado ao pretender extrair relações de equivalência, hierarquia e associativa para auxiliar na composição de ontologias. Ademais, em decorrência da versatilidade e capacidade de integrar novos vocabulários, as folksonomias são uma boa oportunidade para preencher ontologias ou sugerir novos conceitos (Limpens, Gandon e Buffa, 2009, p.6).

Relativamente a função das relações semânticas para recuperação da informação, Hjørland (2007) afirma que

In information retrieval the basic functions for semantic relations may be conceived as contributing to the increase of recall and precision. The inclusion of synonyms and broader terms in a query may, for example, contribute to increased recall. The differentiation of homonyms and the specification of terms may increase precision. In this way the widely use of standard semantic relations used in thesauri may be explained functionally (Hjørland, 2007, p.20)

Ainda, acerca da folksonomia e ontologia, Basso, Ferreira e Silva (2009) ressaltam que, se por um lado a tarefa de organizar a informação na forma de uma ontologia tem um alto custo cognitivo, organizar a informação utilizando tags é uma tarefa bem simples. Vários sistemas hoje em dia permitem que os próprios utilizadores façam uso de tags escolhidas por eles mesmos para

⁸ Knowledge Organization System

organizar algum tipo de informação. O valor do processo de atribuição de tags é derivado do fato de que as pessoas usam seu próprio vocabulário, adicionando, assim, significado explícito ao recurso, o qual pode vir do entendimento delas sobre a informação ou sobre o objeto que está sendo categorizado, tornando a classificação da informação uma tarefa de baixo custo.

Assim, relativamente a integração entre ontologia e folksonomia, Sharif (2009) ressalta e complementa que

The insufficient involvement of users in the construction of ontologies is a significant cause for the current shortage of satisfying coverage found in domain ontologies (Damme, Hepp and Siorpaes, 2007: 57) For this reason, collaborative environments as provided in social tagging may inspire new ways of efficient ontology engineering (Weller, 2007:112). Gruber, in his article “ontology of folksonomy: a mash-up of apples and oranges” presents the different roles of ontologies and folksonomies in semantic web. He shows that not only both techniques are not completely opposed to each other but it is possible to get them complement each other (quoted in Echarte, Astrain, Córdoba, and Villadangos, 2007). Presently, there are some built ontologies available on the web. Semantic web tends to reuse ontologies whenever they are available so we can use the benefits of both ontologies and folksonomies to complete the process of knowledge representation on web. In this approach it is possible to have the flexibility of use and implementation of folksonomies and the structured model of knowledge in ontologies at the same time (Sharif, 2009, p.3).

Diante do exposto, para elaborar sistemas de conceitos, sejam linguagens documentais ou ontologias, faz-se necessário selecionar uma área, subárea ou domínio do conhecimento, o qual apresente alguma demanda social, cultural ou econômica, com vistas a aprimorar o tratamento informacional. Nesta perspectiva, considerando a complexidade da informação para a saúde, sobretudo no que tange a representação da informação em intersecção às vivências e experiências dos indivíduos interagindo nos contextos digitais, o domínio selecionado para o desenvolvimento desta tese foi diabetes, doença incurável que constitui um dos maiores desafios da saúde pública.

No Egito, em 1500 a.C., foi constatado o primeiro caso de diabetes como uma doença desconhecida e essa denominação foi usada pela primeira vez por Apolonio e Memphis em 250 a.C. Diabetes em grego quer dizer sifão (tubo

para aspirar a água) devido a sintomatologia da doença que provoca sede intensa e grande quantidade de urina. A diabetes só adquiriu a terminologia mellitus no século I d.C. e Mellitus, que em latim significa mel, logo a patologia passa a ser chamada de urina doce (Marcelino & Carvalho, 2005).

Essa doença crônica caracteriza-se pela elevação da glicose (açúcar) no sangue acima da taxa normal (hiperglicemia). A taxa normal é de aproximadamente 60 a 110 mg%. Ele é causado por fatores genéticos (herdados) e ambientais, isto é: a pessoa quando nasce já traz consigo a possibilidade de ficar diabética. Quando, aliado a isso, se traz fatores como obesidade, infecções bacterianas e viróticas, traumas emocionais, gravidez etc., a doença pode surgir mais cedo (Zagury, Zagury & Guidacci, 2000, p. 16).

Há dois tipos de diabetes: o tipo 1 e tipo 2, os quais apresentam diferenças quanto a sintomatologia e tratamento, assim como a população que atingem. Comumente, a diabetes tipo 1 atinge crianças e adolescente e o tipo 2 adultos.

Também, considerada uma das principais epidemias do século XXI, tornando-se um problema global independente das condições de desenvolvimento social e econômico de cada país. Segundo a OMS⁹ (2016), diabetes é uma “doença crônica, de fundo metabólico na qual existe, por parte do organismo, uma incapacidade total ou parcial de retirar a glicose (além de outras substâncias) do sangue” e que, apesar de não apresentar mortalidade elevada, quando comparada a outras, constitui um importante fator de risco pela comorbidade e disfunções como cegueira, insuficiência renal e Acidente Vascular Cerebral.

Ainda, conforme a OMS (2016), 422 milhões de adultos no mundo foram diagnosticados com a doença em 2014, enquanto em 1980 eram 108 milhões. Dados publicados em 2015 no Relatório Anual do Observatório Nacional da Diabetes (Portugal) revelam que, em 2014, a prevalência estimada da diabetes na população portuguesa com idades entre os 20 e 79 anos (7,7 milhões de indivíduos) foi de 13,1%, ou seja, mais de 1 milhão de portugueses nesta faixa

⁹ Organização Mundial de Saúde

etária tem diabetes. Também, em termos de composição da taxa de prevalência da diabetes, em 56% dos indivíduos havia sido diagnosticada e em 44% ainda não. Acrescenta-se ainda que, mais de um quarto das pessoas entre os 60-79 anos tem Diabetes.

Pelo exposto, verifica-se que a diabetes confere um impacto na qualidade de vida dos indivíduos, assim como apresenta desdobramentos de ordem econômica, sendo um domínio fértil para desenvolver pesquisas no âmbito da OI. Sobretudo, considerando o cenário marcado pela cultura do compartilhamento de informações.

Antes de apresentar a questão de investigação, objetivos e estrutura da tese, é conveniente mencionar uma interface desta investigação, que em certa medida, aproxima-se de atividades desenvolvidas no âmbito do Processamento de Linguagem Natural ou Processamento de Língua Natural (PLN). Para Gonzalez e Lima (2003) o PLN trata computacionalmente os diversos aspectos da comunicação humana, como sons, palavras, sentenças e discursos, considerando formatos e referências, estruturas e significados, contextos e usos, visando fazer o computador se comunicar em linguagem humana, nem sempre necessariamente em todos os níveis de entendimento e/ou geração de sons, palavras, sentenças e discursos.

Estes níveis são: **fonético e fonológico** (relacionamento das palavras com os sons que produzem), **morfológico** (relativo a construção das palavras a partir unidades de significado primitivas e de como classificá-las em categorias morfológicas), **sintático** (relativo ao relacionamento das palavras entre si, cada uma assumindo seu papel estrutural nas frases e de como as frases podem ser partes de outras, constituindo sentenças), **semântico** (relativo ao relacionamento das palavras com seus significados e de como eles são combinados para formar os significados das sentenças) e **pragmático** (relativo ao uso de frases e sentenças em diferentes contextos, afetando o significado) (Gonzalez e Lima, 2003).

Relativamente ao processamento semântico, enquanto a sintaxe corresponde ao estudo de como as palavras agrupam-se para formar

estruturas em nível de sentença, a semântica está relacionada ao significado, não apenas de cada palavra, mas ao conjunto resultante delas. Neste sentido, o processamento semântico é considerado um dos maiores desafios do PLN, pois vincula-se com a morfologia e a estrutura sintática, assim como com informações pragmáticas. Por exemplo, o significado de uma frase, pode originar-se do significado de cada palavra. Este princípio revela a importância das relações semânticas que ocorrem entre os itens lexicais.

Em consonância à recuperação da informação, Gonzalez e Lima (2003) reforçam que há perdas significativas para a RI, quando as abordagens são apenas estatísticas e não há nenhum esforço em nível de processamento semântico. Por fim, os autores, apoiados em Voorhees (1999), realçam que o processamento de linguagem natural aplicado à recuperação da informação deve procurar aproximar da prática a teoria linguística, e esta aproximação deve ser feita da forma mais perfeita possível, para trazer vantagens. Sem abandonar os métodos estatísticos, há concretamente dois motivos para que os pesquisadores busquem aplicar PLN com conhecimento linguístico à RI: (a) as abordagens estatísticas ainda não resolveram satisfatoriamente o problema da RI; e (b) a RI trata inerentemente de processamento de textos, e esta característica faz com que o conhecimento linguístico não possa ser descartado. Mesmo que a RI ainda deva esperar muito do PLN motivado por conhecimento linguístico, a linguística computacional, em contrapartida, tem encontrado na RI um campo experimental que tem possibilitado evoluções significativas. Isto decorre da exigência de grande precisão nos procedimentos aplicados, incentivando aprimoramentos.

QUESTÃO DE INVESTIGAÇÃO

Face ao exposto, verifica-se que a convergência entre os aspectos-chave desta tese, nomeadamente relações semânticas, folksonomia, ontologia e diabetes subsidiam uma variedade de questões de investigação, a saber:

- Há um padrão de relações semânticas nas folksonomias?
- Na área da saúde, quais os contributos da folksonomia para o desenvolvimento de ontologia?

Não obstante a estas indagações, neste âmbito, a questão de investigação desta tese é: **em que medida as relações semânticas implícitas nas folksonomias permitem estruturar uma ontologia sobre diabetes?**

HIPÓTESES

Diante do que foi exposto, apresentam-se as seguintes hipóteses:

- Há um núcleo essencial de relações semânticas implícitas nas folksonomias que são basilares para estruturar ontologias.
- Embora a folksonomia não seja uma única e abrangente fonte para compilação da variedade de conceitos que devem compor um tesouro ou ontologia, elas fornecem subsídios para estruturação e configuram-se como uma estratégia de garantia do uso do termo.

OBJETIVOS

O objetivo desta tese doutoral é perceber como as **relações semânticas** implícitas nas **folksonomias** podem ser incorporadas na **estruturação** de uma **ontologia sobre diabetes**. Especificamente, tem a pretensão de:

- Sistematizar um núcleo essencial de relações semânticas na perspectiva da Ciência da Informação.
- Perceber aspectos da representação da informação sobre diabetes em contexto digital, a partir dos sistemas baseados em folksonomias.
- Reconhecer quais são os qualificadores sobre o conceito diabetes.
- Identificar os padrões de relações semânticas sobre diabetes.
- Verificar os contributos para estruturação de uma ontologia a partir de um núcleo de relações semânticas extraído em folksonomia.

ESTRUTURA DA TESE

Com o intuito de contemplar os eixos temáticos que compõem esta investigação, a tese será composta por cinco capítulos, além desta introdução e da conclusão.

O capítulo 1, “**Contributos teórico-metodológico para estruturação semântica**”, tem como objetivo apresentar e contextualizar o diálogo existente entre teorias que abordam o conceito e suas relações como elemento fundamental na estruturação de esquemas de Organização da Informação. Nomeadamente, a Teoria da classificação, especificamente os contributos de Ranganthan, com ênfase na proposta do método analítico-sintético e a noção das facetas. A Teoria da Terminologia, nas perspectivas de Eugen Wuster, sobretudo o método onomasiológico e Maria Teresa Cabré que acrescentou o aspecto comunicativo do termo, considerando-o também integrante da linguagem natural. Complementarmente, seguiu-se com a Teoria do Conceito, advogada por Dalhberg, cuja contribuição versa sobre a composição do conceito, suas características e a análise conceitual. Por fim, apresentam-se algumas definições sobre Tesouro, com ênfase na sua estrutura e composição, nomeadamente os termos e relações semânticas, considerando a linguagem de indexação tradicional mais flexível e adaptável ao contexto digital. Em decorrência da extensão de cada um dos tópicos mencionados, não houve a pretensão de aprofundar cada uma das teorias, mas apresentar um breve panorama sob a perspectiva da estruturação via sistema de conceitos e suas relações.

O capítulo 2, “**Novos esquemas para Organização da Informação em contexto digital**”, teve como objetivo conceituar a folksonomia, apresentando algumas variações terminológicas, a saber: indexação social, indexação democrática, classificação social, etiquetagem colaborativa. Também, apontar suas vantagens e desvantagens, bem como reconhecer os desafios impostos e o papel que as linguagens naturais assumem neste contexto. Em seguida, faz-se uma breve discussão sobre ontologia, que nos últimos anos, tem sido abordada como uma relevante alternativa para representação da informação em ambiente digital, considerando-a como um sistema de conceitos explícitos e formal¹⁰. Também, foi apresentada sua integração com a folksonomia e, a título de exemplificação, reuniu-se alguns exemplos de como ambas se impõe no domínio da saúde.

¹⁰ Formal refere-se a ser legível por computadores.

Na sequência, o capítulo 3, “**Relações semânticas na perspectiva da Ciência da Informação**”, objetivou sistematizar as relações semânticas, por vezes, com distintas denominações e tipificações, a título de exemplo: paradigmáticas, sintagmáticas. Conjuntamente, abordam-se as normas ISO 25964 - *Information and documentation - Thesauri and interoperability with other vocabularies*, parte 1 - *Thesauri for information retrieval* (2009) e parte 2 - *Interoperability with other vocabularies* (2013), assim como a *BS 8723-2:2005 - Structured vocabularies for information retrieval: guide* e a *ANSI/NISO Z39.19-2005* (r. 2010). Também, foram apresentados o conceito de semântica, sua conexão com a Organização da Informação, especificamente para a recuperação da informação, as propriedades lógicas das relações semânticas e um breve exemplo de relações semânticas no domínio da saúde.

Sucessivamente, o capítulo 4, “**Percurso investigativo**”, propõe apresentar as abordagens metodológicas adotadas, como análise de domínio, uma vez que, a compreensão de domínios do conhecimento reflete o compartilhamento de conceitos entre diferentes membros de uma comunidade. Também, a análise conceitual, por integrar a identificação, determinação, estruturação, validação dos conceitos e relações semânticas apropriadas para a representação da informação. Ademais, descrevem-se os procedimentos aplicados, como consulta ao DeCS e demais dicionários na área da saúde, o sistema baseado em folksonomia utilizado para coletar as indexações livres e as tags, especificamente o CiteULike, assim como o Protege utilizados para edição de ontologias e o protocolo para comunicação com os especialistas sobre o assunto.

O capítulo 5, “**O domínio do Diabetes: caracterização do seu léxico a partir de folksonomia**” reúne os resultados obtidos. Primeiramente, segue a lista de candidatos a descritores do domínio da diabetes extraído da folksonomia, seguida das áreas do conhecimento que compartilham conceitos e apresentam uma relação estreita no que tange o assunto diabetes. Ademais, segue a verificação de qualificadores apresentados no DeCS e sua ocorrência na folksonomia, assim como categorias estabelecidas, com exemplos de manifestação, um núcleo de relações semânticas com base nos conceitos mais

expressivos e os contributos para a estruturação de uma ontologia. Adiciona-se também uma breve discussão acerca do conceito risco e fator de risco como uma possibilidade de relacionamento entre os conceitos. Neste capítulo são descritas e relatadas as especificidades e implicações na representação da informação na área da saúde, que envolve os processos cognitivos empreendidos.

Por fim, apresentam-se as considerações finais, as quais discorre-se sobre a questão de investigação, os objetivos da tese, limitações da investigação e trabalhos futuros.

CAPÍTULO 1 - CONTRIBUTOS TEÓRICO-METODOLÓGICOS PARA ESTRUTURAÇÃO SEMÂNTICA

Esta tese tem como elemento nuclear as relações semânticas implícitas nas folksonomias como possibilidade para estruturação de uma ontologia por meio dos conceitos inter-relacionados. Considerando as teorias e os processos metodológicos necessários para a definição de um sistema de conceitos e suas relações, faz-se necessário contextualizar e apresentar contributos teórico-metodológicos que alicerçam e promovem o desenvolvimento de sistemas de organização da informação emergentes em ambientes digitais, os quais caracterizam-se atualmente pelo papel que a linguagem natural assume, considerando os seus fenômenos, e a autonomia do utilizador na execução das atividades.

Para tanto, não há a pretensão de aprofundar, esgotar ou contrapor teorias, mas abordá-las e apresentá-las sob a perspectiva do conceito e suas relações enquanto elemento fundamental para estruturação semântica. No âmbito da Ciência da Informação, abordou-se a Teoria da Classificação, especificadamente a Teoria da Classificação Facetada, Teoria do Conceito e Tesouros. Em Ciências da Linguagem discorre - se acerca dos contributos da Teoria da Terminologia.

1.1 A CLASSIFICAÇÃO COMO ESTRUTURA DO PENSAMENTO: BREVES APONTAMENTOS

A classificação enquanto ato e processo de classificar, elemento auxiliar na organização da informação, é um fenômeno inerente à natureza humana. De forma consciente ou inconsciente, permeia o dia a dia dos indivíduos, estabelecendo ordem intelectual, física, sendo dinâmica e adaptável às demandas da sociedade. Quando é objeto de estudo e análise, permite ser observada sob distintas perspectivas, revelando sempre sua arbitrariedade.

Enquanto atividade provocada pela necessidade que o homem tem de possuir modelos mentais de referência, a classificação proporciona orientação em determinado universo, sendo manipulável. Dessa maneira, evidenciam-se aspectos relacionados à organização física e abstrata, a composição de quadros mentais, os quais revelam visões de mundo (Simões 2011).

Acerca da naturalidade das classificações, Pombo (1998, p.1) salienta que

[...] nada nos parece mais "natural", óbvio e indiscutível que as classificações dos entes, dos factos e dos acontecimentos que constituem os quadros mentais em que estamos inseridos. Elas constituem os pontos estáveis que nos impedem de rodopiar sem solo, perdidos no desconforto do inominável, da ausência de "idades" ou "geografias". Só elas nos permitem orientar-nos no mundo à nossa volta, estabelecer hábitos, semelhanças e diferenças, reconhecer os lugares, os espaços, os seres, os acontecimentos; ordená-los, agrupá-los, aproximá-los uns dos outros, mantê-los em conjunto ou afastá-los irremediavelmente (Pombo, 1998, p.1).

Este excerto permite identificar aspectos referentes à ordem cognitiva, ao processo mental, a formação metódica e sistemática de agrupamentos, e ainda, ao estabelecimento de critérios, conscientes ou não, para aproximações ou distinções. Na esfera geral, Pombo (1998, p.4) pontua cinco características de toda classificação real:

- 1) Por trás de cada classificação há um determinado mecanismo classificador que executa, melhor ou pior, as operações necessárias à classificação.
- 2) Cada classificação persegue uma mais ou menos sistemática multiplicidade de fins que, em última análise, vão determinar a sua estrutura.
- 3) Cada classificação exerce-se sobre um domínio da realidade cujas estruturas internas tornam mais ou menos fácil as operações necessárias à classificação.

4) Cada classificação constrói-se no contexto das classificações precedentes da mesma área, ou seja, há uma inexorável historicidade das classificações ao longo da qual as áreas classificadas podem ser modificadas, as divisões podem ser completadas, novos critérios de classificação podem ser acrescentados.

5) Para cada classificação existe um produto externo da atividade classificadora que se apresenta como uma árvore genealógica mais ou menos regular. Toda classificação supõe uma dupla operação: o estabelecimento de equivalências entre classes do espaço classificatório global; o estabelecimento de hierarquias entre subclasses no interior das classes previamente estabelecidas.

Estas características, embora gerais, ressaltam a finalidade das classificações, a delimitação a partir do universo a que se referem, bem como sua estrutura firmada nas classes, subclasses e relações estabelecidas.

Referente às aplicações e finalidades das classificações, mesmo em um contexto mais específico, Langridge (1977, p.11) nota que a “classificação transforma impressões sensoriais isoladas e incoerentes em objetos reconhecíveis e padrões recorríveis”. Ao distinguir classificação natural e classificação artificial, ele reforça e exemplifica que o Homem busca por classificações que são mais fundamentais e, neste sentido, os cientistas servem-se de classes nas quais os membros possuem maior número possível de características em comum, como, por exemplo, os zoólogos que classificam os animais de acordo com as semelhanças estruturais. Nesta lógica, o autor elucida que baleias pertenceriam à classe dos “mamíferos” como cavalos, vacas, ratos. Entretanto, aponta que o leigo poderia pensar nas baleias como pertencentes à mesma classe dos peixes, uma vez que vivem no ambiente aquático, apesar de ser uma característica isolada, pois a baleias e os peixes não têm muito em comum. Dessa forma, questionou e reforçou as distintas aplicações e finalidades das classificações associadas ao interesse e as necessidades que elas devem atender, considerando sua flexibilidade e arbitrariedade.

Concernente às questões de classificação, por muito tempo considerou-se a lógica, enquanto procedimento de raciocínio, a regra básica. Neste sentido, sem a pretensão de analisar profundamente a lógica clássica e seguir uma abordagem filosófica, em face a sua influência em esquemas de classificação tradicionais para fins de tratamento e recuperação da informação, é oportuno explicar brevemente o legado de Aristóteles.

Filósofo, aluno de Platão, Aristóteles (384 a.C – 322 a.C) nasceu em Estagira (Grécia) e seus estudos abrangeram diversos temas, como: a lógica, física, astronomia, retórica, metafísica, psicologia e ética. Pertinente à classificação do saber, é considerado um dos precursores e seu sistema surgiu como a necessidade de abarcar e expressar a realidade de modo organizado e preciso. Assim, buscou identificar e estabelecer conceitos básicos que orientassem os agrupamentos e as partições do real operado pelo pensamento humano, ou seja, por meio de categorias primordiais, conforme expõe Aranalde (2009). Este autor acrescenta que, na concepção aristotélica

as categorias são identificadas como conceitos elementares, isto é, como princípios que permitem identificar as notas essenciais que caracterizam um objeto do conhecimento. A partir desta operação mental de identificação, é possível formular conceitos empíricos, ou seja, buscar uma equivalência entre como o objeto se apresenta e a representação mental que se faz dele e de suas relações com outros objetos. As categorias são concebidas como metaconceitos que permitem a efetiva conceitualização de objetos passíveis de serem conhecidos, organizados e classificados. Portanto, elas são elementos intermediários entre os conceitos e a realidade cognoscível (Aranalde, 2009, p. 87).

Isso revela as categorias enquanto unidades básicas de sentido, que norteiam e subsidiam análises e interpretações em um universo de conhecimento. Nesta perspectiva, foram estabelecidas as dez categorias de Aristóteles:

Quadro 1 - Categorias de Aristóteles

Categorias	Exemplo
Substância/ matéria	Homem, cavalo, computador.
Quantidade/ extensão	Comprido, grande, metro, quilos.

Qualidade	Azul, virtuoso, filósofo, professor.
Relação	Pesado, barulhento.
Lugar	Europa, Brasil, fazenda, escola, hospital.
Tempo	Ontem, Século XX, primavera.
Estado	Em pé, sentado.
Hábito	Queimado, processado.
Ação	Correr, processar, falar.
Paixão	Cortado, processado, curtido.

Fonte: Araújo (2008).

Integrantes dos processos classificatórios e imprescindíveis para estruturação dos sistemas de classificação, estas categorias tiveram como alicerce os cinco predicados de Aristóteles, segundo Araújo (2008):

- a) Gênero: classe ou grupo de seres ou objetos que possuem um determinado número de características em comum;
- b) Espécie: ser ou coisa que possui uma diferença específica que a distingue de seu gênero próximo; a espécie é obtida do gênero pelo acréscimo de uma diferença;
- c) Diferença: é a característica que serve para gerar uma espécie; cada acréscimo de diferença gera uma nova espécie;
- d) Propriedade: algo próprio de cada elemento de uma classe mas que não é imprescindível à definição da classe;
- e) Acidente: qualidade não obrigatória a todos os elementos de uma classe, isto é, que pode ou não estar presente em um conceito (Araújo, 2008, p.122).

Complementarmente, outro contributo de Aristóteles é a árvore de Porfírio, que, atualmente, é abordado em discussões acerca da elaboração de mapas conceituais, taxonomias e ontologias. A seguir encontra-se a representação gráfica da estrutura arbórea citada.

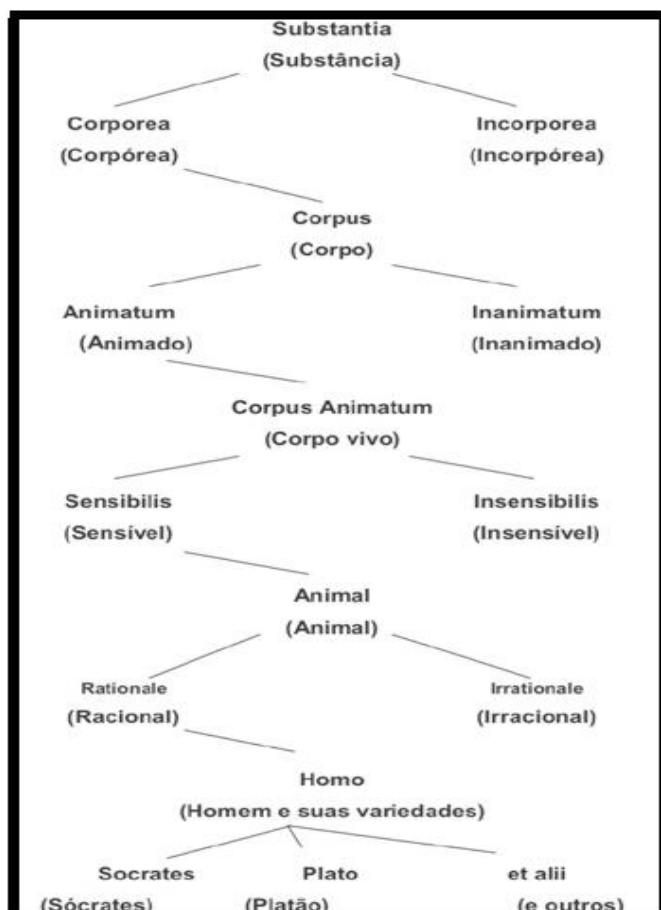


Figura 1- Árvore de Porfírio
 Fonte: adaptada de Pombo (1998).

A estrutura exposta constitui-se como um conjunto hierárquico finito de gêneros e espécies, identificados por dicotomias sucessivas, as quais procedem do geral para o particular, da extensão à maior compreensão, do gênero às espécies ínfimas (Pombo, 1998; Moreira & Lara, 2011). Por ela, é possível perceber a relação entre os conceitos como elemento estruturante dos esquemas de organização tradicionais e emergentes.

Entretanto, mesmo desempenhando um papel fundamental para o desenvolvimento de outras classificações, também revelam limitações a despeito da relatividade da sua construção e da seleção dos atributos. Neste sentido, Pombo (1998) reitera que a escolha de uma ou outra propriedade tem como resultado diferentes arranjos das realidades a classificar e, conseqüentemente, a constituição de diferentes classificações.

Schreiner (1976) complementa que a classificação baseada no princípio de relação gênero-espécie, conforme propôs Aristóteles, implica na escolha arbitrária de uma característica como diferença, sendo todas as demais abstraídas, não importando sua relevância para outras formas de abordagem do conhecimento e resultando em uma classificação rigidamente monohierárquica, que não corresponde à complexidade da natureza. Entretanto, é um princípio básico utilizado na constituição de esquemas de classificação bibliográfica, com destaque às classificações facetadas.

Pelo exposto, pretendeu-se de forma breve e genérica, contextualizar aspectos da classificação, os quais foram a base de outros esquemas. A classificação não se aplica mais na mesma perspectiva esboçada por Aristóteles, uma vez que se faz necessário dar conta da complexidade do intelecto humano, correspondente ao complexo padrão operacional da mente. Ademais, dissertar-se-á sobre a classificação facetada, considerando-a a mais importante para a Teoria da Classificação nos últimos cem anos (Broughton, 2004), pois introduziu o conceito de facetas, propondo uma perspectiva multidimensional, capaz de alicerçar uma estruturação dinâmica e mais complexa.

1.2 CLASSIFICAÇÃO FACETADA: ESTRUTURA DINÂMICA E MULTIDIMENSIONAL

A necessidade de organizar recursos informacionais em ambiente digital tem impulsionado e requerido soluções em diversas áreas do conhecimento. A Ciência da Informação fornece princípios básicos de classificação, em uma perspectiva teórica e metodológica, com vistas a organizar a informação “dispersa” na internet por meio da elaboração e aplicação de novos esquemas de organização da informação, como por exemplo, as ontologias.

Neste âmbito, a compreensão sobre classificação mais recorrente na literatura da área refere-se a uma divisão em grupos ou classes, segundo diferenças ou semelhanças, de forma sistemática. Assim, Buchanan (1979) reitera:

Classification is the act of grouping like things together. All the members of a group - or class - produced by classification share at least one characteristic which members of other classes do not possess. The "things" which are classified may be concrete entities, the ideas of such entities, or abstractions. The recognition of likenesses may be performed by machines or by human beings (Buchanan, 1979, p.9).

Morfologicamente, oriunda da palavra latina *classificare*, ou seja, *ficare* em *classis*, classificação é um substantivo derivado do verbo classificar e, inserida no uso geral da língua, apresenta quatro acepções distintas: 1) colocar em determinadas classes; 2) determinar a classe a qual pertence uma entidade; 3) definir classes e, 3) alcançar determinável nível. Estas acepções indicam os processos de rotulação¹¹, seguido do processo de busca de uma classe para a entidade, de pesquisa para qual classe o conceito está associado e verificação e validação de um conceito perante uma estratificação (Schiessl & Shintaku, 2012). Percebe-se, então, que a organização e estruturação de um domínio, para além de classificar, separar por semelhanças e diferenças, estabelece-se relações de identidade, similaridade e de associações, expressas por meio de classes de conceitos.

Diferenças terminológicas associadas à classificação no âmbito da CI são inexpressivas, segundo Lancaster (2004). Contudo, por permear todas as atividades pertinentes ao armazenamento e recuperação da informação, o autor aponta que há alguma confusão com a expressão indexação¹² e que se deve à incapacidade de distinção entre as etapas de análise conceitual e de tradução da indexação. O autor considera que a análise é uma atividade intelectual, no qual o especialista coloca um item em uma classe conceitual e que a tradução envolve a representação da análise mediante um termo ou termos. Assim, a classificação refere-se à formação de classes de itens com base no conteúdo temático e o processo intelectual é o mesmo. Esta observação é relevante para esta tese uma vez que as folksonomias podem

¹¹ Atribuição de um ou mais descritores ou tags.

¹² Segundo Lancaster (2003) a indexação implica a preparação de uma representação do conteúdo temático. O indexador descreve o conteúdo ao empregar um ou vários termos de indexação, comumente selecionados de algum tipo de vocabulário controlado.

ser analisadas sob a perspectiva da indexação, no sentido de atribuição de *tags*, ou da classificação social, que é convergente a este estudo.

Na década de 1930 a discussão sobre classificações bibliográficas foi ampliada, tendo Shiyali Ranganathan como o principal disseminador e precursor da classificação por facetas. Antes de dissertar sobre o conceito e as características deste tipo de classificação considera-se relevante apresentar classificações precedentes que motivaram e influenciaram sua proposta. A saber, têm as classificações enumerativas (com destaque a *Library of Congress Classification* e Classificação Decimal de Dewey) e as mistas (Classificação Decimal Universal), as quais tinham a função de enumerar os assuntos, mas eram esquemas descritivos e ofereciam pouca flexibilidade para inserção de novos assuntos.

A Classificação Decimal de Dewey (CDD) foi a primeira classificação bibliográfica popularmente conhecida e prenunciou muitas ideias. Concebida em 1876 por Melvil Dewey (1851-1931), foi publicada anonimamente sob o título: *A classification and subject index for cataloging and arranging the books and pamphlets of a library* e herdou características da classificação de Harris¹³. Simões (2011) expõe que Dewey apresentou como novidade a notação decimal, situação que permite representar os conceitos de uma forma mais específica. A notação consiste em uma sequência de números que são tratados como se fossem precedidos de zero, vírgula. Outra característica é que o sistema é construído *a priori*, ou seja, sua construção é independente dos documentos. Trata-se de um sistema mono-hierárquico, no qual as notações são subordinadas umas as outras e de uma linguagem pré-coordenada¹⁴. Também, é composto por tabelas principais e tabelas auxiliares. A primeira compreende que o assunto está dividido em dez classes e, sucessivamente, subdividido. As segundas devem especificar o assunto principal por diferentes perspectivas.

¹³ Simões (2011, p.59) William Torrey Harris concebeu um sistema de classificação para St. Louis Public School Library sob influência indireta de Francis Bacon. Ele dividiu o conhecimento em dez classes principais, sendo cada uma constituída por um número variável de divisões.

¹⁴ Linguagem pré-coordenada é quando os descritores são combinados no momento da indexação, na entrada ao sistema.

Quadro 2 - Tabelas principais e auxiliares da CDD

Tabelas principais		Tabelas auxiliares
000	Informática, informação e generalidades	Tabela 1: Subdivisões comuns. Tabela 2: Regiões geográficas, períodos históricos e pessoas. Tabela 3: Subdivisões para as artes, a literatura e gêneros literários. Tabela 4: Subdivisões de língua e de famílias de língua Tabela 5: Raças, etnias e nacionalidades. Tabelas 6: Línguas. Tabela 7: Pessoas.
100	Filosofia e Psicologia	
200	Religião	
300	Ciências Sociais	
400	Línguas	
500	Ciência	
600	Tecnologia	
700	Arte e lazer	
800	Literatura	
900	Geografia e História	

Fonte: adaptado de Simões (2011)

Outra classificação de base enumerativa é a Classificação da Biblioteca do Congresso dos Estados Unidos¹⁵. Criada no final do século XIX, com a finalidade de ordenar os documentos considerados no seu aspecto físico e não em termos de conteúdo, como na CDD, deu lugar a interesses utilitários, específicos do congresso. Conseqüentemente, o sistema de classificação esteve confinado a uma biblioteca particular, especificamente a coleção de Thomas Jefferson¹⁶, e suas classes figuraram apenas as matérias que faziam parte do acervo. Ao contrário da CDD, caracteriza-se por um sistema a posteriori, ou seja, sua construção depende dos documentos que existem na biblioteca. Também é um sistema mono-hierárquico, pré-coordenado. Não é um sistema decimal e as subdivisões encontram-se organizadas dentro de cada classe ou subclasse na seguinte ordem: primeiro aparecem as notações que representam o assunto geral, seguida das notações sobre aspectos específicos (Simões, 2011).

Relativamente às tabelas principais e auxiliares, as primeiras contemplam 21 classes, independente uma da outra e identificada por uma letra maiúscula e as segundas contemplam as tabelas que permitem maior especificidade, conforme apresentam-se sumarizado no quadro abaixo:

Quadro 3 - Tabelas da classificação da biblioteca do congresso

Tabelas principais		Tabelas auxiliares
A	Obras gerais	<ul style="list-style-type: none"> • Tabelas de forma • Tabelas geográficas • Tabelas cronológicas • Tabelas de subdivisão de
B	Filosofia, Psicologia e Religião	
C	História e Ciências auxiliares	
D	História universal	
E	História da América	

¹⁵ A Biblioteca do Congresso foi inaugurada em 1800, quando o presidente John Adams transferiu a sede do governo da Filadélfia para a nova capital, Washington.

¹⁶ Terceiro presidente dos Estados Unidos (1801-1809).

F	História da América	assunto • Tabelas mistas
G	Geografia, Antropologia e Folclore	
H	Ciências sociais	
J	Ciência política	
K	Direito	
L	Educação	
M	Música	
N	Belas artes	
P	Língua e Literatura	
Q	Ciência	
R	Medicina	
S	Agricultura	
T	Tecnologia	
U	Ciência militar	
V	Ciência naval	
Z	Bibliografia. Biblioteconomia	

Fonte: adaptado de Simões (2011)

A iniciativa pelo desenvolvimento deste sistema de classificação deve-se a Herbert Putnam, bibliotecário da instituição. Posto isto, Simões (2011) sumariza as seguintes características para as classificações enumerativas:

- A estrutura mono-hierárquica, ou seja, as classes obedecem ao princípio de dependência lógica, nas quais os conteúdos mais específicos se encontram compreendidos em conteúdos gerais.
- As matérias apresentam-se em classes e subclasses, geralmente, em números consecutivos de dez subdivisões.
- As notações são numéricas ou alfanuméricas e, na maioria dos casos, hierárquicas.
- Unidimensionais: um assunto apenas pode ser perspectivado sob um ponto de vista e não representa os vários aspectos pelo qual pode ser abordado.
- São linguagens consideradas analíticas, mas não funcionam para representar matérias que num documento se encontram perspectivadas em vários aspectos.

As classificações mistas apresentam características mencionadas, porém acrescenta-se o uso de tabelas auxiliares compostas por sinais (por exemplo: adição (+), extensão (/), relação (:), entre outros) e subdivisões que permitem analisar assuntos em diferentes aspectos, possibilitando também a construção

de assuntos compostos e representações mais específicas. Nesta circunstância é que se insere a Classificação Decimal Universal (CDU).

A CDU teve origem a partir da 5ª. Edição da Classificação Decimal de Dewey e nasceu do objetivo de ordenar a compilação de um repertório universal de bibliografia e organizar toda a documentação. Para tanto, Paul Otlet e Henry la Fontaine consideraram os seguintes fatores: o conhecimento humano organizado em taxonomias, a notação constituída apenas por algarismos árabes, com vistas a obter uma linguagem universal, e o princípio decimal subjacente à sua estrutura, cuja ideia era intercalar conceitos de forma infinita. Dessa forma, Gil Urdiciain (1998) salienta que a CDU é uma classificação sistemática que apresenta uma sucessão de conceitos organizados por relações essenciais e que se trata de um sistema misto, o qual enumera o conhecimento em 10 grandes grupos chamados tabelas principais, podendo combinar todos os temas reconhecidos neste conjunto hierárquico com uma série de subdivisões auxiliares comuns e especiais, que possibilitam a formação dos conceitos. Segue o princípio de subordinação lógica e cada grupo de conceitos que compõem as tabelas principais se organizam em outras 10 subdivisões. Posto isto, seguem as tabelas principais.

Quadro 4 - Tabelas principais e auxiliares da CDU

Tabelas principais	
0	Generalidades. Ciência e conhecimento. Organização. Informação. Documentação. Biblioteconomia. Instituições. Publicações.
1	Filosofia. Psicologia
2	Religião. Teologia
3	Ciências sociais. Estatística. Política. Economia. Comércio. Direito. Administração pública. Forças armadas. Assistência Social. Seguros. Educação. Etnologia
4	Classe não utilizada
5	Matemática e Ciências naturais
6	Ciências aplicadas. Medicina. Tecnologia
7	Arte. Recreação. Entretenimento. Esporte
8	Língua. Linguística. Literatura
9	Geografia. Biografia. História
Tabelas auxiliares	
	Auxiliar comum de forma, lugar, raça, tempo, ponto de vista, Materiais, pessoas.
	Auxiliares especiais

Fonte: adaptado de UDC Consortium

Quadro 5 - Símbolos utilizados na CDU

Símbolos	
+	Adição
*	Asterisco

/	Barra inclinada	A/Z	A barra z – extensão alfabética
:	Dois pontos	.00	Ponto zero zero
::	Dois pontos duplos	-03 e -05	Hífen zero três e hífen zero cinco
[]	Colchetes	-1/-9	Hífen um a hífen nove
=	Igual	.0	Ponto zero
(0...)	Parênteses zero	'	Apóstrofo
(1/9)	Parênteses um barra nove	...1/...9	Reticências um barra reticências nove
(=...)	Parênteses igual	"..."	Aspas

Fonte: Adaptado de UDC Consortium

Além de incorporar aspectos da multidimensionalidade conferida pela aplicação das tabelas auxiliares e símbolos, ressaltam-se as seguintes características (Simões, 2011):

- Linguagem específica e exaustiva, com base na extensão de conceitos que podem ser apresentados.
- Sistema analítico-sintético: o qual permite decompor a análise de um determinado assunto nas suas várias perspectivas, ao mesmo tempo em que proporciona uma representação de forma sintética, resultado dessa análise.

Apesar de estas classificações adotarem os princípios aristotélicos, uma vez que o conhecimento é dividido em categorias, demonstravam-se pouco flexíveis em sua estrutura, pois não permitiam representar coordenando as classes. Nesse sentido, Broughton (2007) reitera que

The difference between the traditional kinds of classification systems, the enumerated kinds and the facet-analytic schemes is that in the enumerative kinds all classes are listed (for example all countries and all periods). In the faceted classification, by contrast, classes are constructed when needed by combinations of building blocks, much like building objects out of Meccano parts¹⁷ (Brought, 2007, s.n.).

Ranganathan introduziu alterações importantes, concebendo em 1933 a *Colon Classification* ou Classificação dos Dois Pontos, no qual a utilização do sinal de dois pontos (:) promove a conexão de ideias diferentes e assim os assuntos passariam a ser analisados em facetas antes que o número de classes pudesse ser atribuído (Duarte, 2010).

¹⁷ Ranganathan, na elaboração da Colon Classification, teve como inspiração a Meccano, uma marca comercial de miniaturas de máquinas para montar, utilizando como analogia o conjunto Meccano de classificação (Brought, 2007)

Campos (2001, p.30) ressalta que o esquema de classificação proposto por ele evidencia dois espaços de ação diferenciados: o espaço do documento e o espaço do conhecimento. Respectivamente, um evidencia a problemática da manifestação do conhecimento que se dá através dos assuntos dos documentos (quais são os modos de formação desses assuntos, para que a tabela de classificação, através de sua notação possa representá-los), ou seja, analisa os assuntos para identificar seus elementos constitutivos. Estes, mas não mais os assuntos, passam a ser unidade da estrutura classificatória, capaz de ser dinâmica o bastante para abarcar todos os assuntos advindos do desenvolvimento científico, com suas peculiaridades.

A seguir, apresentam-se sucintamente os princípios ranganathianos e os elementos da estrutura classificatória, com o propósito de clarificar e sumarizar seus contributos para a estruturação de novos esquemas de organização da informação, aspecto que será debatido na sessão subsequente.

1.2.1 PRINCÍPIOS RANGANATHIANOS E OS ELEMENTOS DA ESTRUTURA CLASSIFICATÓRIA

Dalhberg (1976) aponta que o sistema ranganathiano diferia dos demais pelo fato de que não utilizava classes pré-estabelecidas e prontas, mas criava classes no momento da análise, segundo os elementos conceituais do assunto, e sintetizado segundo as regras das fórmulas de facetas ligadas às disciplinas.

Como o principal disseminador do método analítico-sintético, Ranganathan introduziu o conceito de facetas definindo-as como um termo genérico utilizado para denotar algum componente de um assunto composto. Apresentou a concepção de faceta básica como primeiro elemento do contexto especificado, agrupando assuntos básicos de determinada área do conhecimento, assim como a concepção de faceta isolada tendo a função de formar renques, termos e números (Campos, 2001). Assim, entende-se que a faceta é uma coleção de termos, os quais apresentam relacionamentos e refletem a aplicação de um princípio de divisão (Silva & Lima, 2011).

Os elementos que constituem a estrutura classificatória proposta por Ranganathan são características, renques e cadeia. As características são compreendidas como um atributo (propriedade, qualidade, medida quantitativa de uma entidade) e utilizadas para comparar os elementos classificatórios com o objetivo de formar classes. Os renques referem-se às classes derivadas de um universo em uma única característica em um espaço de divisão para estabelecer o arranjo completo na sequência preferida, ou seja, são classes formadas a partir de uma única característica de divisão. E a cadeia é uma sequência formada de classes e seu universo de deslocamento, especificamente, são séries verticais de conceitos em que cada conceito tem uma característica a mais ou a menos, conforme a cadeia descendente ou ascendente (Campos, 2001).

A estruturação está ancorada em três planos de trabalho: o plano ideacional, verbal e notacional, os quais possuem princípios normativos próprios. Relativamente ao plano ideacional, este é considerado superior aos demais, pois é onde se originam as ideias, se analisam conceitos que compõem um sistema de classes e precede qualquer manifestação, capacidade de expressão, linguagem articulada e tradução para símbolos. Refere-se ao processo de pensar e situa-se no campo da abstração. É nesta perspectiva que Ranganathan (1967) expõe que um esquema de classificação implica antes o conceito de esquemas de classes, o qual envolve cinco outros conceitos: características, sucessão de características, renque, cadeias e sequência de filiação.

Ranganathan elaborou um conjunto de princípios normativos para organização e estruturação de conceitos. A seguir, apresentam-se sumarizados nas tabelas os cânones que compõem o plano das ideias.

Quadro 6 - Cânones do plano ideacional

PLANO IDEACIONAL	CÂNONES	CANON	EXPLANAÇÃO (Gomes, Motta & Campos, 2006)
	CARACTERÍSTICAS	Diferenciação	Uma característica usada como base para a classificação de uma área deve diferenciar algumas de suas entidades, dando origem à ao menos duas classes.
		Relevância	Uma característica deve ser relevante para o propósito da classificação, considerando o público, tradição, necessidades etc.
		Verificabilidade	A característica deve ser definitiva e verificável. Assim, somente as características verificáveis devem ser escolhidas para a divisão do esquema de entidades.
		Permanência	Uma característica deve permanecer enquanto não houver mudança no propósito da classificação.
	SUCESSÃO DE CARACTERÍSTICAS	Concomitância	Em um esquema de classificação não se deve usar características que produzam o mesmo renque.
		Sucessão relevante	A sucessão destas características deve ser relevante para os propósitos da classificação.
		Sucessão consistente	Este princípio garantirá consistência quando da introdução de uma nova característica. Para evitar futura inconsistência a sucessão de características deve ser estabelecida no momento do planejamento de um esquema de classificação.
	RENQUE	Exaustividade	Mapear todo o universo de conhecimento de forma exaustiva, ou seja, o maior número de conceitos de forma a prever futuras necessidades e garantir a inserção de novos tópicos para compor a estrutura.
		Exclusividade	Nenhum componente da estrutura pode pertencer a mais de uma classe no renque.
Sequencia útil		Ordenar os elementos de um renque e suas classes e é dependente do propósito da classificação	
Sequencia consistente		Estabelecer que classes semelhantes em diferentes renques devem apresentar uma sequência paralela em todos aqueles renques.	
CADEIAS	Extensão decrescente	A classe mais abrangente deve sempre preceder a mais específica em que se divide. A extensão tem por medida o número de entidades compreendidas na classe ou no isolado. Extensão de um conceito compreende a totalidade ou o número de conceitos que abarca, ou seja, uma medida quantitativa de uma classe ou de um isolado ordenado.	
	Modulação	Uma cadeia de classes deve compreender uma classe em cada passo da subdivisão. Neste sentido, é depende do uso correto das características relevantes e da sequência de emprego destas características.	

SEQUÊNCIA DE FILIAÇÃO	Classe subordinada	Em uma sucessão de classes cada uma deve seguir imediatamente a outra, sem que seja interrompida pela inclusão de uma ou mais classes de natureza distinta.
	Classe coordenada	As classes originadas de uma mesma classe superordenada têm uma sequência de filiação horizontal, formando um renque de classes.

Fonte: adaptado de Ranganathan (1967).

Conforme a tabela exposta, no princípio referente as Características o foco é o conteúdo de um assunto, ou de uma ideia, sem a preocupação com sua relação com outros assuntos. Enquanto no princípio de Sucessão de Características, a característica é vista em seu potencial de organização e visa à sucessão das grandes classes num universo de assunto. No conjunto dos princípios relativos ao Renque, o foco está em sua formação e ordem dos elementos em seu interior. Quanto ao conjunto de princípios relativos à Cadeia, o enfoque está em sua formação e ordenação dos elementos em seu interior e em Sequência de filiação estão os princípios para a construção de uma hierarquia rígida e consistente (Gomes, Motta & Campos, 2006).

Coexistente ao plano ideacional, o plano verbal refere-se à manifestação das ideias, a capacidade de desenvolver uma linguagem articulada. É relativo à representação por meio da linguagem articulada, no qual surgem os fenômenos naturais da língua, como os homônimos e sinônimos (Campos, 2001). Neste sentido, é que são evidenciados o caráter mediador da linguagem e seus ruídos no processo de comunicação das ideias.

The language of our conversation is notoriously vague. The language of even carefully prepared documents lends itself to several interpretations; and it hides or confuses the original intention to such an extent that society is forced to maintain the costly profession of advocates. it has to divert to this profession - by the lure of disproportionate emoluments - some of the best brains which should be used in substantial and creative work more beneficial to society [...] such words are vague, because their meaning shades off imperceptibly into the meanings of other words (Ranganathan, 1967, p. 201).

Portanto, refere-se ao plano das palavras. O outro plano é o notacional, o qual se refere ao plano dos números para representar os conceitos. É relativo à manifestação do que foi articulado no plano ideacional e verbal, permitindo a

representação e inserção nos esquemas de classificação e a manipulação de arranjos.

Campos (2001) enfatiza que este plano garante que novos conceitos criados sejam inseridos nas tabelas, por meio do princípio da hospitalidade¹⁸. Este princípio torna possível a introdução de diversos mecanismos para que o esquema de classificação acompanhe a dinâmica do conhecimento, a saber, pela ampliação da base notacional, através de uma notação alfanumérica, pela ampliação dos renques e a organização da estrutura classificatória em categorias fundamentais PMEST (Personalidade, Matéria, Energia, Espaço e Tempo) e a adoção do método analítico-sintético.

Considerando o universo de conhecimento como uma espiral em movimento contínuo agregando novos conceitos, o postulado das categorias representa cinco ideias fundamentais que permitem recortar um universo de assunto¹⁹ em classes bastante abrangentes, sendo o primeiro corte classificatório, o qual permite o entendimento global da área. Neste seguimento, apresentam-se breves esclarecimentos acerca de cada categoria (Campos, 2001):

- Personalidade: esta categoria é considerada de difícil identificação ou indefinível. Costuma ser aplicada após exclusão das demais categorias.
- Matéria: considera-se uma categoria complexa, no qual manifesta-se via material ou propriedade, que podem constituir todas as espécies. Também compreende métodos, operações.
- Energia: compreendida como a ação, movimento, técnica, tratamento, procedimentos e operações.
- Espaço: tem como função localizar o assunto num espaço geográfico, como continentes, países, estados.
- Tempo: exemplifica-se com ideias isoladas no tempo comum, tem a função de localizar o assunto num espaço cronológico; apresentando também correspondência com tabelas cronológicas.

¹⁸ Princípio utilizado inicialmente por Cutter como Expansividade, depois por Mevil Dewey.

¹⁹ Campos (2001, p.54) corpo de conhecimento organizado e sistematizado.

Dois aspectos devem ser ressaltados: inicialmente, Ranganathan traz para a Documentação a preocupação com o universo de conhecimento, no qual as unidades que constituem uma estrutura não são mais apenas os assuntos, mas os conceitos, e estes, reunidos por uma combinação, permitem formar assuntos. Concomitantemente, elaborou uma série de princípios os quais permitem que estes conceitos sejam estruturados sistematicamente. Os conceitos se organizam em renques e cadeias, estruturadas por classes, facetadas, e dentro das categorias (Campos, 2001, p.57).

Sobre análise facetada Aitchison e Clarke (2004) esclarecem que

The technique analyses complex subjects into constituent categories of the same inherent type. These fundamental categories include actions, comprising processes and operations; entities, such as natural objects, products, materials; agents, including personnel and equipment; and time and place. In the field of education, for instance, there would be at eaching methods facet, arising from the operations category; an educational personnel facet from the personnel category; a teaching aids facet from the equipment category, and so on. The facets are mutually exclusive, and the terms within each facet share a common characteristic. A “facet indicator”, or “node label”, is often inserted to name that common characteristic. The terms so grouped tend to be short and simple, and may then be used in combination with other simple terms to express compound subjects, either post coordinately when indexing using a thesaurus, or as pre coordinated class marks in the context of a faceted classification scheme (Aitchison & Clarke, 2004, p.11).

Entretanto, entusiasmados com a proposta de Ranganathan, em 1952, no Reino Unido, formou-se o *Classification Research Group (CRG)*, inicialmente composto por Derek Austin, Eric Coates, Jason Farradane, Robert Fairthorne, Douglas Foskett, Barbara Kyle, Derek Langridge, Jack Mills, Bernard Palmer, Jack Wells e Brian Vickery. Embasados na ideia de Ranganathan, ampliaram as categorias fundamentais, conforme apresentado no quadro:

Quadro 7 - Categorias por Ranganathan x CRG

Categorias	
Ranganathan	Classification Research Group
Personality	Thing Kind Part Property

Matter	Material
Energy	Processo Operation
	Patient Product By product Agent
Space	Space
Time	Time

Fonte: Broughton (2006).

Para o CRG as categorias são derivadas na natureza dos assuntos e que nem todos os assuntos se encaixam nas categorias fundamentais. Neste sentido, sugere identificar as categorias fundamentais pelo contexto do próprio assunto, concluindo que nenhuma lista de categorias fundamentais deve ser exaustiva ou imposta mecanicamente aos assuntos. Comparadas ao PMEST, as categorias do CRG oferecem mais flexibilidade e os classificadores podem moldar assuntos com maior especificidade, além de formar categorias mais distintas e definidas (Lima, 2004).

A teoria desenvolvida por Ranganathan tem influenciado os demais sistemas de organização da informação, desde os tesouros aos sistemas emergentes no contexto digital, em decorrência do método analítico-sintético, assim como a dinamicidade e a perspectiva multidimensional atribuída pela ideia das facetas.

1.3 TERMINOLOGIA: ABORDAGENS E PRINCÍPIOS

O desenvolvimento, estabelecimento e o progresso de todas as áreas do conhecimento são dependentes do processo de comunicação da informação. Neste contexto, compreendeu-se que os elementos de pensamento (conceitos) nem sempre assumem o mesmo significado para os participantes deste processo, seja em nível nacional, internacional e em um mesmo domínio de especialidade. Neste contexto, Barros (2006) ressalta que a transmissão do saber ocorre por meio de textos que possuem características peculiares, em nível semiótico, pragmático, sintático, semântico e, sobretudo, lexical, uma vez que é principalmente por meio de uma terminologia própria que esse tipo de texto veicula os conhecimentos especializados.

Nesta lógica, a terminologia constitui-se enquanto disciplina que estuda os termos das áreas de especialidade, como a denominação de novos conceitos e sua harmonização. Assim, apesar de correntes divergentes, a terminologia como teoria é um conjunto de premissas, argumentos e conclusões necessário para explicar o relacionamento entre conceitos e termos especializados. Como prática é um conjunto de métodos e atividades voltado para coleta, descrição, processamento e apresentação de termos e, como produto, é um conjunto de termos, ou vocabulário, de uma determinada especialidade (Dias, 2000).

As motivações que originaram a prática terminológica associam-se a necessidade de relacionar denominações aos conceitos científicos. Entretanto, no século XIX tal prática foi estimulada pelo progresso da área científica, no qual cientistas descreveram a regra de formação dos termos para cada domínio de especialidade, a saber:

- Lavoisier e Berthold em Química;
- Linneu em Botânica e Zoologia.

Inicialmente, o interesse era a fixação de denominações dos conceitos científicos direcionados aos especialistas e com a internacionalização progressiva da ciência, estes manifestaram de forma sistemática a necessidade de dispor de regras de formação de termos para cada área do conhecimento. Assim, botânicos (1867), zoólogos (1889), químicos (1892) se expressaram em seus respectivos colóquios internacionais. Para os especialistas, a terminologia é o reflexo formal da organização conceitual de uma especialidade, e um meio inevitável de expressão e de comunicação profissional (Cabré, 1993, p.37).

O objetivo teórico da metodologia é

[...] el de describir formal, semántica y funcionalmente las unidades que pueden adquirir valor terminológico, dar cuenta de cómo lo activan y explicar sus relaciones con otros tipos de signos del mismo o distinto sistema, para hacer progresar el conocimiento sobre la comunicación especializada y las unidades que se usan en ella. El objetivo de la terminología aplicada es el de recopilar las unidades de valor

terminológico en un tema y situación determinados y establecer sus características de acuerdo con esta situación. Dentro de sus características puede figurar su condición de unidad normalizada (Cabré, 1993, p.37)

A autora reitera que a terminologia como todas as outras matérias científicas interdisciplinares, se define em relação a outras matérias, e toma emprestado um conjunto específico de conceitos. Neste seguimento, a relação entre terminologia e documentação é estreita. Acerca disso, Cabré (1993, p.112) ressalta que a terminologia constitui uma das bases dos tesouros e classificações temáticas, uma vez que a expressão de qualquer relação de conteúdo entre os conceitos se efetuam através do termo. Assim, nestes instrumentos documentais os termos representam relações e a utilização da uma terminologia padronizada para indexar os documentos confere a documentação uma sistematicidade e uma univocidade que o uso de palavras da língua comum nunca poderiam lograr. Por tudo isso, a terminologia padronizada é também um elemento de controle para o documentalista. Nesta perspectiva, uma linguagem de indexação comporta três tipos de operações de caráter terminológico: estabelecer uma classificação prévia de noções; efetuar uma seleção das noções e estabelecer relações entre elas, revisando passo a passo as séries verticais de noções; efetuar uma seleção de termos eliminando sinônimos e, se possível, os homônimos. Posto isto, verifica-se que os documentos técnicos são a base de um trabalho terminológico, pois neles aparecem os termos que efetivamente utilizam os especialistas:

Las denominaciones que el terminólogo debe recoger se encuentran en los documentos, y a través de ellos el terminólogo informa cada término en relación con su semántica, función, forma y situación, para proceder después a una propuesta denominativa adecuada [...] Efectivamente, el terminólogo se ve obligado a servirse de documentación para conocer la materia y su estructuración conceptual, para encontrar los términos que utilizan en una materia, para confirmar la calidad de los primeros datos que ha recogido, para informar los datos desde los distintos puntos de vista, para ordenar y presentar esos datos en forma de recopilación impresa o automatizada, o para preparar las propuestas alternativas de denominación con vista a una acción de normalización (Cabré, 1993, p.113).

Posto isso, todo o trabalho terminológico se dá com base nos dados veiculados por textos das áreas de especialidade, que são identificados, analisados e organizados pela documentação. O produto final do trabalho do terminólogo, os dicionários especializados, também é catalogado pelos documentalistas.

O engenheiro austríaco Eugen Wuster (1898-1977) é considerado o propulsor da terminologia moderna. Em 1931 ele apresentou sua tese doutoral intitulada *A normalização Internacional da Terminologia Técnica*. Seu legado está fixado na Teoria Geral da Terminologia (TGT), cuja intenção era eliminar das linguagens especializadas a imprecisão, a diversificação e a polissemia, considerando a terminologia um instrumento para desambiguação da comunicação científica e técnica, preocupando-se com os aspectos metodológicos, normativos. Assim, a TGT estabelece como objeto de análise as unidades unívocas normalizadas próprias dos domínios técnico-científicos e como atividade principal, a compilação de conceitos e termos para sua normalização. Para tanto, também estabeleceu as seguintes características (Barros, 2004):

- Terminologia como disciplina autônoma;
- Objeto de estudo são os termos técnico-científicos concebidos como unidades específicas de um âmbito especializado. Ou seja, não considera suas variáveis e o uso está circunscrito a um domínio;
- O conceito precede a denominação, o que caracteriza o método onomasiológico de análise;
- Os conceitos mantêm entre si relações de diferentes tipos. O conjunto destas relações entre os conceitos constitui a estrutura conceitual de uma área de conhecimento;
- O valor do termo se estabelece pelo lugar que ele ocupa na estrutura conceitual;
- O objeto de estudo dos termos é a normalização conceitual e denominativa;
- A finalidade aplicada da normalização terminológica é garantir a precisão e univocidade da comunicação profissional

Dentre estas características da TGT, convém destacar o método onomasiológico, processo no qual parte-se do conceito para a denominação. Primeiramente, é estabelecida uma lista de conceitos que formam parte da estrutura nocional mais ou menos formalizada. Os conceitos pertencentes a esta estrutura se encontram inter-relacionados lógicamente e ontologicamente e seu conjunto constitui o sistema conceitual de uma disciplina ou atividade especializada. Em cada elemento conceitual é atribuído uma denominação, que corresponde à forma que utilizam efetivamente os especialistas quando se referem a um conceito da estrutura nocional (Cabré, 1993, p. 92).

Segundo Barcellos Almeida (2003) as limitações acerca do legado de Wuster referem-se à idealização da realidade, marcada pela homogeneidade e universalidade do conhecimento especializado e o desejo de unificação das formas de expressão e a univocidade da comunicação. Tal perspectiva tem como suposição que o conceito pré-existe à expressão, o conhecimento técnico e científico é universalmente uniforme, a estruturação de uma área é única em todos os grupos e contextos. Considera que o conhecimento científico é neutro, sem interferências culturais, sociais ou ideológicas e o termo normalizado representaria as características pertinentes para todos os grupos especializados. Para além do idealismo, também há o reducionismo decorrente do aspecto normalizador, uma vez que a unidade terminológica limita-se a sua denominação, sem considerar os aspectos sintáticos, negando a variação e outros aspectos comunicativos dos termos. A partir destas duas limitações, formulam-se os seguintes apontamentos (Barcellos Almeida, 2009):

- A realidade só pode ser conceitualizada cientificamente pela ciência que, por meio da lógica, ordena o conhecimento científico. Os conceitos estruturam lógicamente e ontologicamente de maneira hierárquica.
- O conhecimento científico, concebido como universalmente homogêneo, é o modelo que deve ser seguido para organizar os conceitos e todas as áreas profissionais, sem considerar as suas diferenças, os contextos socioculturais, as áreas geográficas, as realidades socioeconômicas e as línguas.

- A uniformização se alcança por meio da normalização. Se os conceitos podem unificar-se, as denominações podem normalizar-se. Com esse processo de uniformização, silencia-se a diversidade denominativa e conceitual da realidade. Os conceitos são estáticos e os termos não tem valor pragmático nem apresentam variação semântica, não é concebido como unidade da língua natural.

Dessa forma, Cabré Castellví (2003) reitera que a proposta de Wuster inserida em um marco comunicativo plural, não dá conta das unidades terminológicas. Em contraste as bases da TGT, a autora ressalta a relevância do legado de Wuster e propõe uma abordagem de base comunicativa, a Teoria Comunicativa da Terminologia (TCT), no qual o ponto de partida é considerar que o termo pode ser observado a partir de diferentes facetas, ou seja, de forma multidimensional. Alguns princípios da TCT são (Cabré Castellví, 2003):

- Explicar as concomitâncias e diferenças entre o conhecimento geral e o especializado sem dissociá-los da competência do falante-especialista, mas conservando a idiosincrasia de cada um. Assim, deve-se assumir que existem traços diferenciadores do conhecimento especializado, mas que este conhecimento não está interiorizado de forma independente na mente dos falantes.
- Explicar a interdisciplinaridade dos termos e dar conta da diversidade de visões que os especialistas possuem sobre os termos, o que constitui sua poliedricidade²⁰. Ou seja, os termos podem ser definidos por distintas facetas, o que explica a diversificação de acepções de um termo segundo o tipo de especialista ou especialidade.
- Dar conta de como um conceito pode formar parte da estrutura conceitual de distintas disciplinas, conservando, trocando ou matizando suas características, explicando se trata ou não de um mesmo conceito e como se produz esta circulação conceitual.
- Admitir a sinonímia como um fenômeno real dentro da comunicação especializada natural e apontar critérios para estabelecer o distinto valor

²⁰ Analogia utilizada na Teoria Comunicativa da Terminologia, no qual os termos ou as unidades terminológicas são poliédricas.

das unidades, se este for o caso. Para TCT, a sinonímia na comunicação especializada é um fato real, quantitativamente dependente do nível de especialização de um discurso. Quanto mais especializado é o texto, maior é sua sistematicidade e menor seu grau de variação denominativa. A TCT supõe que os termos, em relação de sinonímia, podem ter um valor similar ou muito diferente, conforme os tipos de traços pragmáticos associados a cada variante.

- Os termos ocorrem de forma natural no discurso e, conseqüentemente, tem uma projeção sintática mais além de seus limites denominativos e variam em função do discurso.

Também, tem como fundamento, o termo como objeto de estudo e integrado, não dissociado da linguagem natural. Assim como os conceitos admitem relações de tipos diferentes entre si e o conjunto destas relações constitui a estrutura conceitual de uma área de especialidade. Neste sentido, ressalta-se que o valor de um termo decorre do lugar que ocupa nesta estruturação conceitual. Cada domínio de conhecimento pode ser estruturado a partir de diferentes perspectivas e em diferentes concepções, assim como cada temática pode ser abordada a partir de outras. Um conceito pode participar em mais de uma estrutura com o mesmo ou diferente valor, de forma que os termos não pertencem a um domínio, mas são usados em um domínio com valor singularmente específico (Cabré, 1993).

Pelo exposto, os aspectos-chave e sobressalentes que convergem com esta investigação centralizam-se no termo enquanto elemento fundamental da estrutura conceitual de uma área do conhecimento, evidente nas linguagens de indexação e o termo como parte da língua natural, ou seja, que contempla os mesmos fenômenos, como a sinonímia e a ambigüidade. Antes de dissertar acerca das contribuições da Teoria do conceito, é relevante destacar algumas especificidades da terminologia na área da saúde.

1.3.1 ESPECIFICIDADES TERMINOLÓGICAS EM SAÚDE

Os utilizadores diretos da terminologia são os especialistas de cada campo científico, no qual a terminologia é um instrumento de comunicação

necessário e elemento importante de conceitualização da própria disciplina. Isso explica o interesse da terminologia pela normalização como processo de estabelecimento dos conceitos e da fixação por denominações correspondentes. Suas necessidades comunicativas partem do conhecimento do conceito e de necessidade de transmiti-lo na comunicação, valendo-se assim de formas linguísticas. Além disso, são intermediários da terminologia os profissionais que precisam da terminologia para desenvolver sua profissão como facilitadores da comunicação, como tradutores, redatores, intérpretes, (Cabré,1993, p. 38), documentalistas e classificacionistas.

Silva, Castellví e Aymerich (2008), em relação aos conceitos especializados, expõem as seguintes premissas:

- a) Os conceitos especializados são elementos do pensamento que surgem em um contexto cultural, socioprofissional e temporal concreto;
- b) São resultado das mesmas operações de categorização que ocorrem no conhecimento geral e sua estruturação está vinculada a fatores situacionais e funcionais;
- c) Os conceitos especializados possuem uma estrutura interna ou intra-categorial e uma estrutura externa ou Inter-categorial que é determinada por suas relações com os demais conceitos em um determinado âmbito;
- d) Os conceitos especializados possuem uma estrutura multidimensional que admite diferentes configurações em função de variáveis cognitivas e comunicativas;
- e) Os conceitos se definem, delimitam e atualizam nos textos produzidos por especialistas e seu conteúdo se ajusta aos usos discursivos.

Tais premissas revelam que as linguagens de especialidades são dinâmicas, variáveis e moldam-se conforme o uso, assim como a sua estruturação. Nesta perspectiva, com base na literatura, discorre-se acerca de algumas características gerais da terminologia médica, cuja finalidade assenta na simplificação da linguagem, para permitir expressar em poucas palavras fatos e conceitos que, de outro modo, demandariam locuções e frases extensas. Assim, tal como ocorre em outras áreas do conhecimento, cada termo médico caracteriza um objeto, indica uma ação ou representa a síntese

de uma ideia ou de um fenômeno, a definição de um processo, que contem em si, muitas vezes, verdadeira holofrase, cujo sentido está implícito na própria palavra (Rezende, 2004, p. 31).

Para aprimorar o processo de comunicação, convencionou-se que os termos médicos seriam construídos com formantes (radicais, sufixos e prefixos) gregos e latinos, como por exemplo: hematúria, cardiopatia, hipotermia, adenocarcinoma, arterite, poliangeíte microscópica. Outra característica é o uso de epônimos (do grego *epónimo*), adjetivo que dá ou empresta o nome, como por exemplo: Mal de Parkinson, Alzheimer, Doença de Chagas. Além disso, os termos são multivocabulares, ou seja, recorre-se a mais de um vocábulo para caracterizar um fenômeno ou um processo corresponde a uma espécie de descrição do que está em foco, por exemplo: ducto hepático, ducto hepático comum; vasculite de vasos de pequeno calibre, vasculite de vasos de médio calibre (Krieger, 2010).

Para além das características mencionadas, na comunicação entre médico e paciente, evidencia-se a ocorrência de termos populares, porém de cunho terminológico. Neste sentido, considerando que a linguagem médica não está fechada em si mesma, Santiago e Krieger (2009) salientam que nestas condições revelam-se focos temáticos de uma área, como por exemplo:

- Dor de cabeça/ Cefaleia;
- Toxoplasmose/ Doença do gato;
- Hepatite/ Inflamação do fígado/ Amarelão;
- Acidente Vascular Cerebral/ AVC/ Derrame cerebral;

Segundo Rezende (2004), os termos que possuem origem grega e figuram na terminologia médica podem ser de dois tipos, a saber:

a) termos do grego clássico, que passaram pelo latim antes de chegar às línguas modernas e, por isso, sofreram alterações prosódicas e morfológicas, mantidas até hoje;

b) termos formados de elementos gregos e latinos de forma direta, após o declínio do latim enquanto língua das ciências, o que ocorreu a partir do século XVII.

Outro aspecto marcante refere-se à comunicação internacional com vistas a melhorar a precisão, evitando ambiguidades e, apesar de distintos objetivos, propõem a integração de informações sobre a área médica, a obtenção de dados estatísticos e a padronização terminológica para indexação. Consequentemente, desenvolvem-se glossários, dicionários, sistemas de classificação, os quais sistematizam, ordenam e contemplam o vocabulário médico. Alguns estão sumarizados abaixo.

- **SNOP e SNOMED:** o *Systematized Nomenclature of Pathology* (SNOP), publicado em 1965, consistiu na organização de termos e códigos utilizados especificamente para descrever casos de patologia, sendo o precursor do *Systematized Nomenclature of Medicine* (SNOMED), publicado em 1974. Este último acompanhou os avanços da informática médica e nos anos 2000 foi desenvolvido o *SNOMED Referents Terms* (SNOMED RT), que consistiu em uma terminologia de referência baseada em conceito, com ampla cobertura, de modo a fornecer infraestrutura para o registro eletrônico em saúde. Em 2002 apresentou-se o *SNOMED Clinical Terms* (SNOMED CT), considerada a mais abrangente terminologia de referência clínica internacional e multilíngue, que fornece conteúdo clínico para registro e documentação. É mantido pela *International Health Terminology Standards Development Organisation* (IHTSDO), associação sem fins lucrativos, situada na Dinamarca (Brogan, 2009).
- **ICD:** a *International Classification of Diseases* tem a finalidade de monitorar, mapear, fornecer dados estatísticos sobre incidências de doenças e demais problemas relacionados à saúde, fornecendo um panorama situacional dos países e populações. Consiste em um instrumento que viabiliza a padronização de diagnósticos para epidemiologias, gestão da saúde e registros clínicos, os quais fornecem bases para extração de estatísticas, a saber: estatísticas nacionais de

mortalidade e morbidade pelos Estados-Membros da OMS. AICD 10 foi aprovada pela OMS em 1990, mas entrou em vigor a partir de 1994 (OMS, 2016).

- **MESH e DECS:** o *Medical Subject Headings* (MESH), publicado em 1954 pela *U. S. National Library of Medicine* (NLM), consiste em um conjunto de termos denominados descritores dispostos em uma estrutura alfabética e hierárquica que permite realizar pesquisa em vários níveis de especificidade. É aplicado no processo de indexação e catalogação de assunto da literatura sobre informação biomédica e, atualmente, é composto por 27.883 descritores, com mais de 87.000 termos de entrada (National Library of Medicine, 2016). Derivado do MESH o DECS (Descritores em Ciências da Saúde), criado pela BIREME²¹ e com a primeira edição datada em 1987, é um vocabulário estruturado e trilingue (inglês, português e espanhol), participa no projeto de desenvolvimento de terminologia única e rede semântica em saúde, a UMLS (*Unified Medical Language System*) da NLM com a responsabilidade de atualizar e enviar termos em português e espanhol. Além dos termos médicos extraído do MESH, foram desenvolvidas áreas específicas, a saber: Saúde Pública (3.492 termos), Homeopatia (1.945 termos), Ciência e Saúde (218 termos), Vigilância Sanitária (827 termos). Os conceitos são organizados em uma estrutura hierárquica, totalizando 31.865 descritores, sendo destes 27.232 do MeSH e 4.633 exclusivamente do DeCS. Existem 2.084 códigos hierárquicos de categorias DeCS em 1.516 descritores MeSH (Biblioteca Virtual em Saúde, 2016).

Enquanto vocabulário controlado, convém apresentar algumas de suas características, nomeadamente, a poli-hierarquia e a pré e pós-coordenação:

- **Poli-hierarquia:** considerando a natureza multidisciplinar dos conceitos, caracteriza-se a poli-hierarquia quando um descritor pode

²¹ Bireme: O Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde é um centro especializado da OPAS (BIREME/OPAS/OMS) com a missão de contribuir para o desenvolvimento da saúde nos países da AL&C por meio da democratização do acesso, publicação e uso de informação, conhecimento e evidência científica.

pertencer a mais de uma área ou categoria ou a vários ramos da categoria, conforme apresentado na Figura 2.

The image shows a screenshot of the DeCS (Descritores em Ciências da Saúde) interface. The main content area displays the following information for 'Diabetes Insipidus':

- Descritor Inglês: **Diabetes Insipidus**
- Descritor Espanhol: **Diabetes Insípida**
- Descritor Português: **Diabetes Insípido**
- Categoria: [C12.777.419.135](#), [C13.351.968.419.135](#), [C19.700.159](#)
- Definição Portuguesa: Doença caracterizada por [micção](#) frequente, excreção de grandes quantidades de URINA diluída e SEDE excessiva. A etiologia do diabetes insipidus compreende deficiência do hormônio antidiurético (também conhecido como HAD ou VASOPRESSINA) secretado pela [NEURO-HIPÓFISE](#), resposta do RIM ao HAD e regulação hipotalâmica da sede prejudicadas.
- Relacionados Português: [Diabetes Mellitus](#)

On the right side, a box titled 'CATEGORIAS' lists three hierarchical paths:

- 1 - Doenças urogenitais masculina > Doenças urológicas > Nefropatias > Nefropatias associadas a AIDS > DIABETES.
- 2 - Doenças urogenitais e complicações na gravidez > Nefropatias > Nefropatias associadas a AIDS > DIABETES.
- 3 - Doenças do sistema endócrino > Doenças da hipófise > DIABETES.

A large green arrow points from the 'Categoria' field to the 'CATEGORIAS' box, illustrating the multi-hierarchical nature of the classification.

Figura 2 - Poli-hierarquia no DeCS
Fonte: Adaptada de DeCS (2016)

- **Pós-coordenação e pré-coordenação:** esta característica refere-se ao ato de coordenar (combinar) dois ou mais conceitos no momento da busca (pós-coordenação) com a finalidade de restringir o escopo do assunto ou no momento de indexar a literatura (pré-coordenação).

Para além das distinções entre a Teoria Geral da Terminologia e a Teoria Comunicativa da Terminologia, verificam-se seus contributos para estruturação de esquemas de representação da informação, com destaque à preocupação em obter a precisão terminológica no processo de comunicação, aos termos em linguagem de especialidade também admitirem os fenômenos da língua natural, constituindo-se como parte dela. Também, evidencia-se a relação entre pensamento, linguagem e realidade (Campos, 2001) e compreende-se que a terminologia na área da saúde, como outras linguagens de especialidade, é um produto convencional, histórico, constituído por fatores linguísticos, sociais, culturais.

A seguir, abordar-se-á a Teoria do Conceito, que na Ciência da Informação teve como propulsora Ingrid Dahlberg. Essa teoria esclarece que o termo denota e é representação do conceito, além de abordar que a análise conceitual ocorre por meio do referente, àquilo a que se refere.

1.4 TEORIA DO CONCEITO: ELEMENTOS CONSTITUTIVOS

A teoria do conceito ou teoria analítica do conceito, no âmbito da Ciência da Informação, ocupa-se da natureza, dos elementos constitutivos e inter-relações dos conceitos, assim como a natureza da análise conceitual e suas implicações.

Implica em que cada conceito tem um referente (seja este um conjunto de objetos, um único objeto, uma atividade, um fato, um tópico, etc.), sobre o qual afirmações verificáveis podem ser feitas. Todas essas afirmações podem ser sumarizadas e/ou sintetizadas por um termo que, então representará um conceito em qualquer processo de comunicação (Dahlberg, 1978, cit. por Motta, 1987, p.31).

Esta teoria contribui para a elaboração de linguagens de indexação, especificamente na elaboração de tesouros. Gonçalves e Souza (2013) afirmam que o entendimento de conceito por Dahlberg em suas formas (conceitos individuais, gerais e especiais) associado a outras teorias pode auxiliar na distinção das características que “unificam” nosso entendimento sobre um dado conceito em uma rede maior de conceitos, como por exemplo, na formação de uma área do conhecimento.

Com o apoio da linguagem, o homem foi capaz de elaborar enunciados sobre coisas, fenômenos, processos, acontecimentos etc. expressos por conceitos individuais, aqueles delimitados por tempo e espaço, diferentemente dos objetos gerais, os quais prescindem das formas do tempo e do espaço. Dessa maneira, formulam-se os conceitos individuais (por exemplo: Universidade do Porto/ Faculdade de Letras da Universidade do Porto (FLUP)) e conceitos gerais (as universidades). Estes enunciados permitem elaborar os conceitos relativos a objetos diversos, de modo que cada enunciado faz referência a algum dos elementos do conceito, conforme exemplificado a seguir:

- **Objeto individual:** Faculdade de Letras da Universidade do Porto

- **Enunciados:** é uma instituição de ensino e investigação; situa-se no Porto; possui mais de 3.000 alunos; oferece licenciatura, mestrado e doutoramento.

A soma destes enunciados sobre a FLUP fornece o conceito sobre o mesmo. Se tomar o conceito geral instituição de ensino e investigação, também é permitido formular outros conceitos. Assim, cada enunciado faz referência a algum dos elementos do conceito.

Nesta perspectiva, o conceito é constituído por elementos que se articulam numa unidade estruturada e que os elementos contidos nos conceitos gerais encontram-se nos conceitos individuais, sendo possível reduzir os conceitos individuais aos gerais e ordená-los de acordo com os conceitos gerais. Nesta sequência, o conceito é definido como a compilação de enunciados verdadeiros sobre determinado objeto, fixada por um símbolo linguístico (Dalhberg, 1978).

Os elementos dos conceitos são chamados de características dos conceitos. “As características dos conceitos são seus elementos constituintes e sua soma total representa os próprios conceitos, ou unidades de conhecimento”. Dalhberg apresentou o seguinte modelo:



Figura 3 - Triângulo conceitual
 Fonte: Motta (1987)

Este triângulo conceitual revela que o conceito não é um todo composto de partes distintas, mas que está contido, a um só tempo, no referente (fatos,

coisas, fenômenos, processos etc.), nas predicções a ele feitas (características) e na forma verbal adotada para designá-lo (Motta, 1987).

Deste modo, o conceito passa de um artefato mental, abstrato, para instrumentar pesquisa e manipulação para a construção de vocabulários e tesouros – principalmente devido à univocidade do conceito: para cada conceito, um conjunto particular de enunciados.

Em síntese, reconhece-se que (Dalhberg, 1978):

1. As características dos conceitos são obtidas por meio dos predicados (enunciados).
2. Os conceitos possuem elementos que são as respectivas características.
3. Um conjunto de características determina um conceito.
4. Os conceitos são unidades de conhecimento constituídas pelas características dos objetos associadas a elementos linguísticos. O aspecto teorico-quantitativo dos conceitos até aqui exposto auxilia o esclarecimento da natureza das relações entre os mesmos conceitos.

Além destes apontamentos, Dalhberg esclarece que para a construção de linguagens de indexação, o que interessa é a relação entre o referente e o termo, pois ela expressa a relação existente entre conceitos e sua manifestação concreta, o conteúdo de registros bibliográficos. Esta teoria torna possível definir o conceito através de suas características e o nomear adequadamente. Para tanto, a identificação de características (análise conceitual) de um conceito facilita o entendimento geral, explica a existência de relações entre conceitos de um sistema e possibilita estabelecer o relacionamento entre eles e avaliar a natureza desse relacionamento (Motta, 1987)

No tocante à análise do conceito (análise conceitual), refere-se ao processo no qual ocorre à decomposição do conceito, coletando os enunciados

verdadeiros sobre determinado objeto, assim, os elementos do conceito são obtidos pelo método analítico-sintético, no qual cada enunciado apresenta um atributo predicável do objeto que, no nível de conceito, se chama característica. Assim, 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 atingir uma característica tão geral que possa ser considerada uma categoria (Dalhberg, 1978).

Diante disso, tendo por base categorias aristotélicas, Dahlberg sintetizou uma possibilidade de características:

Quadro 8 - Características dos conceitos

ESPÉCIES DE CARACTERÍSTICAS	EXEMPLOS
Matéria (substância)	De madeira, de metal, de couro, de vidro, etc.
Qualidade	Possuir determinada estrutura, determinada forma, ser redondo, denso, colorido etc.
Quantidade (extensão)	Possuir comprimento, largura, peso, etc.
Relação	Ser o dobro, ser mais largo, ser causa de, ser condição de, etc.
Processo (atividade)	Começar, continuar, terminar, realizar algo, etc.
Modo de ser	Estar em pé, sentado, voando, etc.
Passividade	Ser cortado, pressionado, etc.
Posição	Estar em cima, em baixo
Localização (lugar)	Estar em Brasília, no Rio de Janeiro, etc.
Tempo	Em fevereiro de 1978, etc

Fonte: Dalhberg (1978)

Pelo exposto, o conhecimento das características dos conceitos facilita a determinação do número de funções: ordenação classificatória dos conceitos e respectivos índices; definição dos conceitos e formação dos nomes dos conceitos. Também, quando dois conceitos diferentes possuem uma ou mais característica em comum, verifica-se que possuem relações e Dalhberg também faz importantes contribuições acerca deste tema. Entretanto, dedicar-se-á o capítulo 3 para abordar as relações semânticas.

1.5 TESAUROS: EM FOCO A ESTRUTURAÇÃO CONCEITUAL

Esta seção tem por objetivo fornecer uma visão geral acerca dos aspectos históricos, definição, tipologia e componentes dos tesouros, uma vez que é a linguagem de indexação que tem influenciado e fornecido embasamento teórico e prático para propor e solucionar questões relacionadas à recuperação da informação em ambientes digitais, sendo considerada adaptável às possibilidades informáticas. A aplicabilidade do tesouro em ambiente digital teve sua pesquisa ampliada a partir da década de 1970 com o uso da internet e de sua principal aplicação, a *World Wide Web*, a partir da década de 1990, conforme Maculan e Lima (2014). As autoras citam que os tesouros fazem o controle terminológico em um domínio de especialidade, como um sistema conceitual que relaciona descritores entre si, estabelecendo uma estrutura no qual os relacionamentos apresentam-se de forma explícita, compondo uma rede semântica que facilita o acesso e fluxo de informação.

A palavra tesouro é oriunda do grego *Thesaurós*, com o significado de coleção, repertório e em latim adotou-se o vocábulo no sentido de tesouro. Enquanto expressão, *Thesaurus* foi utilizada pela primeira vez por Bruneto Latini (1220-1294) na enciclopédia "*Livre dou Trésor*" (entre 1262 e 1268) e, tanto na Idade Média quanto no Renascimento sua denominação restringia-se aos dicionários monolíngues. Assim, verifica-se que o tesouro teve como acepção inicial a ideia de repertório, compêndio de termos ou vocabulário, que se originou com vistas a apoiar a descrição da linguagem, auxiliando no seu uso (Gil Urdiciain, 1998).

Simões (2008, p. 51) esclarece que o emprego deste termo ocorreu por volta de 1531, por Robert Estienne, com o título *Thesaurus Linguae Latinae*, no qual buscou contextualizar as palavras no seu campo semântico recorrendo a exemplos de aplicação. Entretanto, o precursor do que é conhecido como tesouros em Ciência da Informação foi publicado em 1852, o *Thesaurus of English words and phrases classified and arranged so as to facilitate the expression Of ideas and assist in literary composition*, por Peter Mark Roget²². Esta obra caracterizou-se pela sua formatação, no qual formou uma cadeia de

²² Secretário da Royal Society, instituição responsável pela promoção do conhecimento científico, fundada em 1660.

sinônimos, que forma categorias conceituais, apresentando as palavras de acordo com as ideias que expressam, de acordo com o seu significado.

Especificamente, o tesouro de Roget recolheu todas as possíveis acepções de uma mesma palavra e agrupou vocábulos por categorias semânticas, organizando também em função de sua morfologia. Dessa maneira, Roget transformou ideia em palavras, possibilitou expressar diferentes aspectos de um conceito determinado e demonstrou que é o princípio que determina o estabelecimento de relações semânticas entre elementos componentes de tesouros documentais (Gil Urdiciain, 1998). Neste tesouro²³ parte-se de um significado, de uma ideia, para se chegar às palavras que melhor a representam, que trata-se de um esquema de classificação com um índice alfabético remissivo (Campos 2001).

Entretanto, foi a partir da década de 50 que no contexto da recuperação e armazenamento da informação foram publicados os primeiros tesouros documentais. A primeira menção foi em 1957 por Helen Brownson na *Dorking Conference on Classification*, no qual propôs um projeto de um novo instrumento de linguagem que consistia na transformação de conceitos e das suas relações, como se expressavam nos documentos, em uma linguagem normalizada, com sinônimos controlados e as estruturas sintáticas simplificadas (Urdiciain, 1998; Simões, 2008). Isto divergia da acepção inicial de tesouro linguístico, apenas similar e confundido com um dicionário, cuja estrutura é baseada nas relações entre os conceitos mais significativos de uma língua, com a função de facilitar a localização de termos, enquanto nos tesouros documentais possuem uma estrutura alicerçada nas relações entre os conceitos dos conteúdos dos documentos, cuja finalidade é contribuir para a coerência da análise, normalizar a representação de conceitos e a pertinência de uma pesquisa (Simões 2008).

Coexistindo com outros instrumentos de organização da informação, como as classificações enumerativas, os tesouros documentais surgem para

²³ Gil Urdiciain (1998) esclarece que primeiramente se realizava busca no índice por uma palavra que descrevia uma ideia, se elegia a entrada cujo significado se aproximava mais do conceito desejado. Em seguida, mediante um código que acompanha a entrada, dispunha-se o conceito acompanhado de todos os seus possíveis sinônimos e ideias associadas.

cumprir diferentes objetivos: facilitar a manipulação apresentando conseqüentemente uma nova estrutura, enquanto linguagem de indexação. Entretanto, as motivações que estiveram presentes na sua origem estão associadas ao que Simões (2008) ressaltou como razões contextuais genéricas e específicas. A primeira refere-se às questões socioeconômicas e políticas, com destaque ao papel que a informação adquiriu após a segunda guerra mundial e a necessidade de construir linguagens de indexação com vistas a promover o acesso à informação em menor espaço de tempo e obter resultados de pesquisas pertinentes e precisos, solução que perpassa pelo controle de vocabulário e a qualidade da representação do conteúdo dos documentos. A autora ainda complementa que neste cenário bélico, foram elaborados os primeiros tesouros: em 1959 o The Engineering Information Center of E.I. Dupont Nemours, em 1960, a *Armed Services Technical Information Agency (ASTIA)*, pelo Departamento de Defesa dos Estados Unidos, em 1961 o *American Institute of Chemical Engineers (AIChE)* publicou o *Chemical Engineering Thesaurus* e em 1994 o *Engineering Joint Council* publicou o *Thesaurus of Engineering Terms*. Quanto as razões contextuais específicas, estas se referem a explosão da informação que ocorreu desde meados do século XX e aumentou após a segunda guerra mundial, assim como o desenvolvimento da ciência e a conseqüente produção de informação técnico-científica. Além disso, verificou-se a necessidade de mecanizar os processos técnicos.

As influências para desenvolvimento do tesouro estão relacionadas à simbiose entre linguagem controlada e não-controlada, assim como pré-coordenada e pós-coordenada. Assim, uma das vertentes considera o Sistema Uniterm (1951), de Mortimer Taube, também reconhecido como um prenúncio do tesouro e sua principal característica é que a representação do assunto é por palavras únicas (Uniterm) extraídas do texto de um documento sem qualquer forma de controle. Isso possibilitava a composição do assunto no momento da recuperação da informação (pós-coordenada) e se opunha aos cabeçalhos de assuntos. Outra vertente advoga a influencia da Teoria da Classificação Facetada (Lancaster, 2003; Campos, 2001; Simões, 2008; Foskett, 1973), caracterizada por sugerir que a organização dos conceitos de

uma área se dá a partir da categorização do mesmo em facetas. Dessa forma, Mendes, Reis e Maculan (2015) esclarecem que cada termo é apresentado tanto na classificação, determinando a faceta e hierarquia, quanto no tesauro, indicando hierarquias alternativas e relações semânticas que estão incluídas no arranjo da classificação, cumprindo também a função de controle de vocabulário. Nesta perspectiva, o tesauro é uma linguagem pré e pós-coordenada.

O tesauro documental é entendido como uma “linguagem documentária dinâmica que contém termos relacionados semântica e logicamente, cobrindo de modo compreensivo um domínio do conhecimento” (Gomes, 1990, p.16). Tal definição recobre, para além da sua composição e estrutura, o caráter flexível dos tesauros, devido à possibilidade de inserção e supressão de termos, assim como a sua especificidade e delimitação ao se referir a um domínio do conhecimento.

Currás (1995) considera uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos linguísticos que a compõem – termos, simples ou compostos – encontram-se relacionados entre si sintática e semanticamente.

Lancaster (2004) especifica em sua definição que nos tesauros o arranjo explícito dos termos é alfabético, mas existe uma estrutura hierárquica implícita, incorporada à lista alfabética por meio de remissivas. A tradicional lista de cabeçalhos de assuntos é similar por ser de base alfabética e por não distinguir claramente as relações hierárquicas das associativas. Os outros tipos de vocabulário controlam sinônimos, distinguem homógrafos e agrupam termos afins, mas empregam métodos um tanto diferentes para alcançar objetivos.

A ANSI/NISO Z39.19-2005 (2010, p.09) define tesauros como “*a controlled vocabulary arranged in a known order and structured so that the various relationships among terms are displayed clearly and identified by standardized relationship indicators. Relationship indicators should be employed reciprocally*”.

Entretanto, reforçando a aplicabilidade nos ambientes tradicional e digital, Aitchison, Clarke e Gilchrist (2005) definem tesouro como “[...] *a vocabulary of controlled indexing language, formal organized so that a priori relationships between concepts are made explicit, to be used in information retrieval systems, ranging from the card catalogue to the internet* (Aitchison, Clarke & Gilchrist, 2005).

Maldonado-Martínez (2010, p. 178) ressalta que o tesouro é utilizado para controlar o vocabulário, filtrando os termos precedentes da linguagem natural pré-selecionada por indexadores e utilizadores como termos de indexação pelos primeiros e como termo de busca pelos segundos, o que inseri esta definição na perspectiva do tesouro facetado. Referente à diversidade de tipos de tesouro, estes seguem os seguintes critérios (Simões, 2008):

- **Abrangência territorial:** refere-se a tesouros nacionais ou internacionais.
- **Entidade produtora:** trata-se de quando o tesouro é produzido por uma organização, seja ela governamental ou não-governamental.
- **Ponto de vista linguístico:** são os tesouros monolíngues, bilíngues e multilíngues.
- **Dimensão:** refere-se a megatesouro, cujo conteúdo é muito amplo, assim como os multidisciplinares, semelhantes às classificações. Há também o macrotresouro, que representa as grandes áreas do conhecimento e o microtesouro, que é mais especializado e cobre temas mais delimitados e específicos.

Esta breve retrospectiva e conceituação revelam que o tesouro, tem o seu aprimoramento com base em distintas vertentes, mas constantemente teve seu principal elemento a estrutura classificatória dada pelo conjunto de conceitos e suas relações. Sua finalidade, apesar de moldar-se e adaptar-se as questões culturais e sociais, sempre versaram a questão da sobre os elementos da significação.

1.5.1 ESTRUTURA E COMPONENTES FUNDAMENTAIS

Uma vez que os tesouros são sistemas estruturados, tem como elementos fundamentais os termos e as relações semânticas estabelecidas, com o propósito de auxiliar a representação do conteúdo de recursos informacionais, potencializando a recuperação da informação. Dentro de uma estrutura conceitual, cada termo tem sua posição e estão organizados de modo a estabelecer ligação com ao menos um outro termo do sistema, refletindo a ordenação e estrutura de alguma área do conhecimento (Naumis Peña, 2007, p. 129).

Por conseguinte, um tesouro compreende os atributos usuais da linguagem: o léxico²⁴, considerando o vocabulário em si próprio e uma estrutura, uma vez que envolve um sistema de regras pelo os termos no léxico são arranjados, através do inter-relacionamento dos termos. Assim, aponta dois aspectos importantes: a seleção de termos de um determinado domínio e o estabelecimento de relações entre estes termos, as quais geram a estrutura.

Em princípio, segundo Curras (1995), seguindo a vertente apoiada pelo sistema Uniterm, o enfoque era a palavra-chave, isolada, como representativa do conteúdo. Entretanto, passou-se a considerar um conjunto de palavras, as quais incluíam autores, entidades, locais geográficos e que estas palavras expressavam conceitos. Desta forma, já incorporando traços da Teoria da Terminologia e Teoria do conceito, caracteriza-se o termo como um dos elementos fundamentais na composição dos tesouros. Este argumento é demonstrado na figura 4, a qual contempla outras denominações.

²⁴ O termo “léxico” significa uma relação de palavras com suas categorias gramaticais e seus significados. Em relação a uma determinada língua, um léxico é o universo de todos os seus itens lexicais, que seus falantes utilizam, já utilizaram ou poderão vir a utilizar [Scapini, 1995, citado por Gonzalez e Lima, 2003).

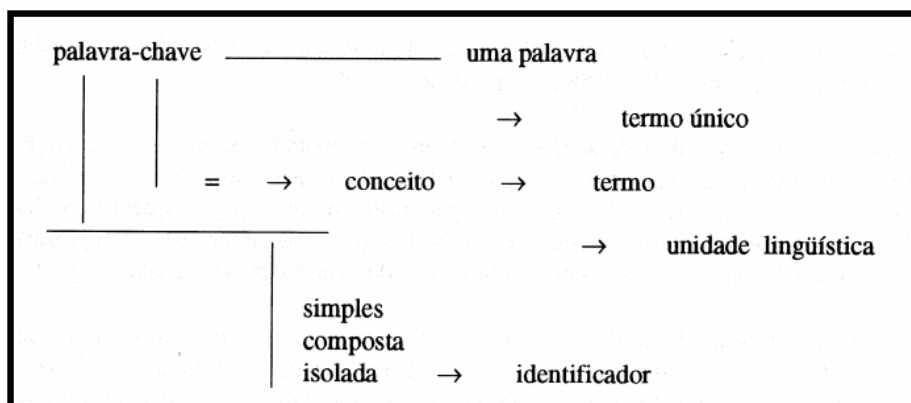


Figura 4 - Termos e a composição dos tesauros
 Fonte: Currás (1995)

Pela figura exposta, compreende a composição, os termos que podem ser simples e compostos, assim como identificadores, aqueles que auxiliam na identificação temática, mas não são determinantes e não necessitam ser incorporadas aos tesauros, e os indicadores de função são aqueles que auxiliam a classificação.

Termos não-descritores ou não-preferidos são sinônimos ou quase sinônimos ou ainda denominados como equivalentes e são contemplados nos tesauros para indicar que não são autorizados para representação do assunto. Segundo Currás (1995) os termos serão descritores (preferido)²⁵ e não-descritores (preterido)²⁶

Os termos são derivados do léxico de uma área do conhecimento especializada e, conforme Simões (2008) ressalta, a circunstância de se basear numa linguagem pré-existente, faz considerá-la uma metalinguagem, no qual constitui-se uma linguagem usada para representar um assunto, e o próprio léxico usado no documento, o léxico do utilizador também.

Um tesouro é caracterizado pela sua forte estrutura semântica, no qual os termos estão delimitados em uma estrutura conceitual, ou seja, que cada termo tenha um significado fixo e distinto em relação ao significado dos outros

²⁵ Termo preferido: aquele que foi eleito para padronização.

²⁶ Não descritores: não preferido, aqueles que são sinônimos ou quase-sinônimos, chamados de equivalentes.

termos, de forma que esta pluralidade conceitual se caracterize pela precisão terminológica. Assim, sua estrutura é o resultado da teia de relações semânticas que se estabelecem entre as unidades léxicas que o constituem, por um lado entre os não-descritores e descritores, que se designam por relações de equivalência; por outro lado relações associativas (Aitchison & Gilchrist, 1972; Simões, 2008 p. 93)

Diferentemente dos cabeçalhos de assunto, no qual as listas alfabéticas são deficientes e as relações entre os termos não são intrínsecas e recíprocas, conseqüentemente com uma estrutura deficitária, a riqueza dos tesouros deve-se ao número de relações semânticas e a quantidade de termos que o compõem (Schiessl & Shintaku, 2012).

Relativo às relações semânticas entre os termos, Foskett²⁷ (cit. por Simões, 2008, p.109) assinala que as relações entre os termos se efetuam em dois níveis:

- As que exprimem assuntos relacionados: caracterizam-se por possuir um caráter permanente.
- Aquelas cujos assuntos não se encontram relacionados e que representam assuntos compostos: caracterizam-se por ser temporário, fruto de associação pontual, baseada nos assuntos versados nos documentos, os quais o autor entende por relações sintáticas.

Para Currás (1995) as relações semânticas são estabelecidas a partir do significado de cada termo, os quais podem ser organizados em grupos que apresentam afinidades semânticas. Tais grupos figuram termos superiores (termo genérico - TG) e termos de menor conteúdo de significado, os termos subordinados (termos específicos - TE), evidenciando o estabelecimento de relações hierárquicas, conforme exemplo esboçado na figura 5.

²⁷ Foskett, A.C. *The subject approach to information*, p. 72-73.

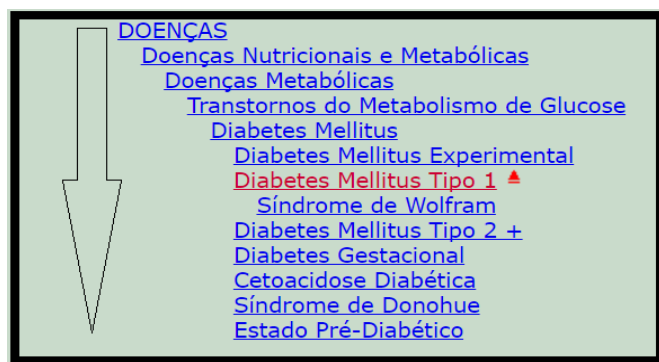


Figura 5 - Exemplo de relações hierárquicas no DeCS

Conforme este exemplo, o termo Diabetes Mellitus Tipo 1 é subordinado a Diabetes Mellitus, assim como tem o termo subordinado Síndrome de Wolfram, evidenciando que as relações hierárquicas são estabelecidas na relação de um termo com o outro. Este agrupamento ocorre em nível vertical e quando há uma associação de ideias em um nível horizontal, sem a possibilidade de estabelecer uma relação hierárquica, evidenciam-se relações associativas.

Com a finalidade de controlar a sinonímia, a ANSI/NISO Z 39.19 (2005, p.159) define que a relação de equivalência é “a relationship between or among terms in a controlled vocabulary that leads to one or more terms that are to be used instead of the term from which the cross-reference is made; begins with the word SEE or USE”. Simões (2008) ainda complementa que ocorre quando há dois ou mais termos que expressam o mesmo conceito e que são traduzidos por um mesmo descritor e este tipo de sinonímia aparece em termos que apresentam o mesmo significado, a saber:

- Entre um termo antigo e um atual.
- Entre um termo popular e um científico.
- Entre termos diferentes de distintas etimologias.
- Entre siglas, acrônimos e seus desenvolvimentos.
- Entre termos provenientes de diferentes culturas dentro da mesma língua.

A autora também destaca o fenômeno da sinonímia documental, que é quando termos que tem significado diferente, mas que por redução de

vocabulário são considerados como sinônimos (ou quase-sinônimos) e realça as seguintes categorias:

- Termos com significado demasiado específico para constituírem um tesouro deverão ser substituídos por termos genéricos. Por exemplo: frigorífico, máquina de lavar louça/ electrodoméstico.
- Termos variantes de um mesmo conceito. Por exemplo: jornal/ periódico.
- Termos antônimos. A exemplo: seca/ humidade.

Os operadores utilizados para indicar esta relação são:

- *USE*: para indicar um termo preterido (não descritor) a partir de um descritor.
- *UP (Used for)*: para indicar o uso do termo preferido (descritor) em relação a outro sinónimo ou quase-sinónimo (não descritor).

A seguir, apresenta-se um exemplo com o descritor Diabetes Mellitus Experimental e seus sinônimos.

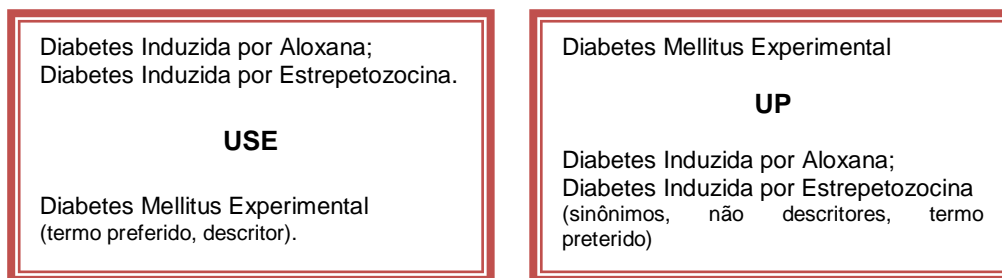


Figura 6 - Exemplo de relação de equivalência

Fonte: DeCS (2016)

Concernente a relação hierárquica a ANSI/NISO Z 39.19 (2005, p.160) define como “a relationship between or among terms in a controlled vocabulary that depicts broader (generic) to narrower (specific) or whole-part relationships; begins with the words broader term (BT), or narrower term (NT)”. As relações hierárquicas expressam a subordinação entre dois termos em função de sua significação. Ou seja, quando um termo genérico engloba conceitualmente o

outro, no qual a significação do termo genérico sempre engloba o termo específico. Há também as relações partitivas, as quais se ocorrem entre o todo e suas partes. Comumente, refere-se a um objeto ou organização social, política, administrativa.

Todos os termos que constituem um tesouro são passíveis de serem relacionados, desde que haja uma associação de ideias entre eles. Os termos que pertencem a categorias diferentes são os que pretendendo a tipos conceituais diferentes, estão semanticamente implicados com outros (Simões, 2008, p.122).

A ANSI/NISO Z 39.19-2005 (2010, p.3) define a relação associativa como “a relationship between or among terms in a controlled vocabulary that leads from one term to other terms that are related to or associates with it”. Ainda enfatiza que

[...] this relationship covers associations between terms that are neither equivalent nor hierarchical; yet the terms are semantically or conceptually associated to such an extent that the link between them should be made explicit in the controlled vocabulary, on the grounds that it may suggest additional terms for use in indexing or retrieval. The most common associative relationship used in thesauri is symmetrical and is generally indicated by the abbreviation RT (related term), but more semantically refined associations may also be developed to capture both symmetric and asymmetric associations (ANSI/NISO Z39.19-2005, 2010, p.51)

Assim, as relações associativas se estabelecem entre os termos próximos entre si, os quais representam conceitos que podem ocorrer mutuamente por associação de ideias, uma vez que não são equivalentes, nem apresentam entre eles uma relação hierárquica. Neste sentido, esta é uma relação que pode ter uma natureza muito diversificada ao estar embasada na associação de ideias e, conforme Maldonado-Martínez (2010, p. 185) “*Se puede asociar entes com processos, processos com objetos, objetos com propriedades, etc. La asociativa es um tipo de relacion muy aberta*”. Para o autor, a maioria dos casos tem se dedicado maiores esforços a construção de estrutura hierárquica

e ao encaixe de cada um dos termos nesta estrutura do que no estabelecimento de relações associativas. Ainda enfatiza que

um término cuantos más términos relacionados tenga mejor, siempre y cuando sean relaciones consistentes. No se debe de ser restrictivo a la hora de establecer relaciones asociativas. En la elaboración de un tesouro se deve de hacer um esfuerzo em el establecimiento de relaciones de asociación, pues ayudan a ampliar el número de relaciones por término, lo que enriquece la información sobre el mismo (Maldonado-Martínez, 2010, p. 186).

Nesta mesma perspectiva, Simões (2008, p.123) reforça que a estrutura destas relações permite ao utilizador uma navegação horizontal, ou seja, possibilita pesquisar na mesma categoria conceitual, assim como pesquisar em diferentes categorias conceituais. Estas relações, complementares as hierárquicas, assumem a função de evitar à poli-hierarquia, pois tornam possível substituir a dependência de um termo específico sob dois ou mais termos genéricos, por um termo relacionado, contribuindo para que a estrutura do tesouro seja simples, de fácil compreensão e constitui um valor acrescentado aos utilizadores de áreas multidisciplinares.

A seguir, apresenta-se um exemplo de relação associativa com os termos Diabetes Mellitus e Diabetes Insípido, utilizando o operador RT:



Figura 7 - Exemplo de termo relacionado
Fonte: Adaptada do DeCS (2016)

Este exemplo foi extraído do DECS (2017) e apresenta as seguintes definições para cada termo.

Diabetes Mellitus: Grupo de transtornos heterogêneos caracterizados por hiperglicemia e intolerância à glicose.

Diabetes Insípido: Doença caracterizada por micção frequente, excreção de grandes quantidades de urina diluída e sede excessiva. A etiologia do diabetes *insipidus* compreende deficiência do hormônio antidiurético (também conhecido como *had* ou vasopressina) secretado pela neuro-hipófise, resposta do rim ao *had* e regulação hipotalâmica da sede prejudicadas

Figura 8 - Relação associativa: diabetes mellitus e diabetes insípido

Fonte: Adaptada do DeCS (2016)

Com base na tabela, verifica-se que ambas as denominações referem-se à doença, entretanto não apresentam uma subordinação conceitual, nem equivalência, mas os termos ocorrem no mesmo campo associativo.

De acordo com a ANSI/NISO Z 39.19-2005 (2010, p.53) as relações associativas são “related term references are often made between etymologically related term [...] terms that contain the same root, but which do not represente the same kind of thing”. Para exemplificação, segue a tabela abaixo.

Quadro 9 - Relações associativas segundo a norma Z39.19

Relações Associativas	
Z39.19	Exemplos
Process/ Agent	Controle de temperatura RT Termostato
Process/ counterange	Flame RT Flame reatardants
Action/ Property	Votação RT Opinião pública Limpeza ambiental RT poluição
Action/ Product	Tecelagem RT Tecido
Action/ Target	Colheita RT Cultivo
Cause/ Effect	Patógeno RT Infecção
Concept or object/ property	Veneno RT Tóxico
Concept or object/ Origins	Método socrático RT Civilização grega
Concept or object/ Units or Mechanisms of measurement	Corrente elétrica RT Ampere
Raw material/ product	Trigo RT Farinha
Discipline or field/ object or practitioner	Neurologia RT Sistema Nervoso; Matemática RT Matemático
Noun is not a true broader term	Pato RT Pato de borracha
Antonyms as associative	Baixo RT Alto

Fonte: NORMA Z39.19

Simões (2008, p. 125), com base em Van Slype, Lancaster, Chaumier, Aitchison e Gilchrist, Foskett e na ISO 2788:1986, apresenta um quadro comparativo no qual aponta especificidades e características das relações associativas. A autora verificou há uma pluralidade de categorias, mas uma uniformidade na tipologia dessas categorias em relação ao número e à nomenclatura.

O tesauro, ligado a operações documentais, foi o primeiro instrumento a explicitar as relações semânticas e, até os dias atuais, por refletir peculiaridades de uma área do conhecimento e pela estrutura classificatória, é importante instrumento de organização da informação, adaptável em sistemas computacionais. Neste ínterim, contemplando a taxonomia em sua estrutura, ao ser incorporado em ambiente web auxilia a recuperação da informação e a navegação, reforçando a importância da teoria da classificação e dos instrumentos de indexação em contextos digitais.

Ao final, reconhece-se que os pontos convergentes entre os tópicos apresentados neste capítulo, para além da estrutura classificatória por meio dos conceitos e suas relações, reforça o método analítico-sintético e a multidimensionalidade.

CAPÍTULO 2 - Esquemas para organização da informação em contexto digital

A organização da informação compreende a organização de um conjunto de objetos informacionais para arranjá-los sistematicamente em coleções, seja em bibliotecas, museus, arquivos, tanto nos tradicionais quanto nos eletrônicos (Café & Brascher, 2008). Atualmente, sistemas com interface gráfica promovidos pela internet, sobretudo na arquitetura de sítios da web, incorporam elementos para disponibilizar e acessar informação nomeados tradicionalmente como catalogação, classificação, indexação.

Conforme apresentado no capítulo 1, a classificação tradicional enumerativa monodimensional hierárquica, quando aplicada a uma arquitetura de sítios da web, trata o conhecimento como um todo integrado que é dividido e subdividido em grupos específicos, como uma estrutura de árvore. Em uma estrutura hierárquica, um objeto é tipicamente alocado em uma categoria, não sendo possível expressar as propriedades e relacionamentos multidimensionais dos objetos digitais. Por essa razão, noções tradicionais de simples hierarquias de classificações e precisam ser aumentadas com métodos mais eficientes (Brandt, 2009, p.38).

Inicialmente, acreditou-se que o conteúdo de sítios da web apenas analisados por facetas seria o mais adequado para a recuperação da informação. Entretanto, com a web 2.0 emergiram outras formas de compartilhar, acessar e organizar informação. Algumas perspectivas versam a web 2.0. Para O'Reilly (2005) foi uma mudança cujo objetivo foi desenvolver aplicativos que aproveitem a inteligência coletiva em rede. Em outra perspectiva, Davis (2005) afirma que se refere mais a uma atitude e não uma tecnologia, pois trata da possibilidade e o incentivo a participação através de aplicações socialmente abertas. Moura (2009) reforça que a web 2.0 possibilitou a personalização tecnológica, abrindo novas possibilidades para o desenvolvimento de produtos e serviços agregados e ampliaram a presença do utilizador comum na coordenação de ações de produção, organização e

difusão de informações voltadas a públicos segmentados pela natureza das redes sociais aos quais pertencem.

Para O'Reilly (2005) muitos conceitos importantes oriundos da web 2.0 não possuem um limite rígido, mas sim, um núcleo gravitacional, conforme a figura 8²⁸.

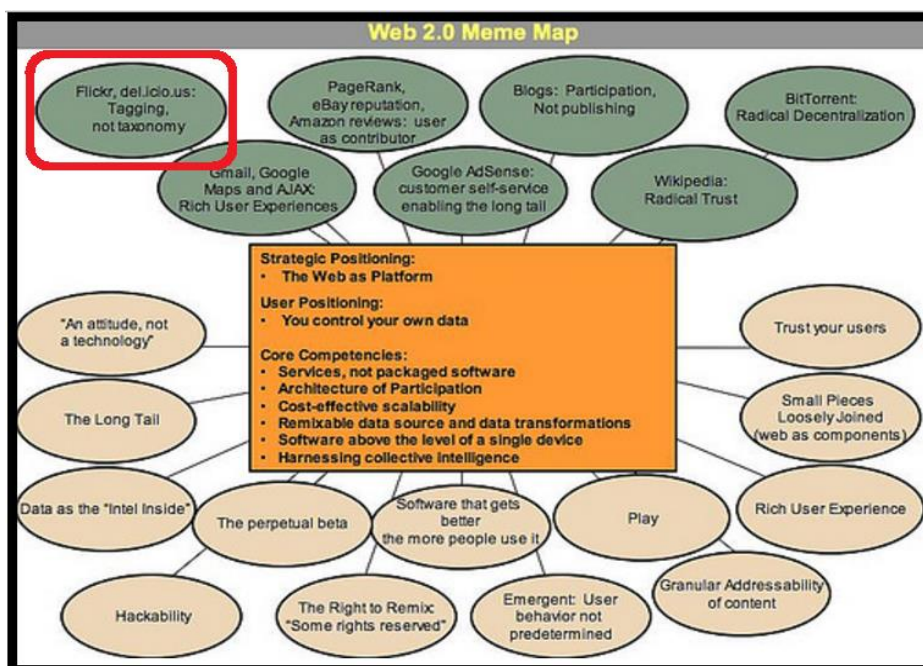


Figura 9 - Conjunto de princípios e práticas da web 2.0
 Fonte: O'Reilly (2005).

Esta figura permite visualizar um conjunto de princípios e práticas que unem um verdadeiro sistema solar de sites que demonstram alguns ou todos esses princípios, revelando as muitas ideias que irradiam do núcleo Web 2.0, a destacar na figura a menção aos SBF: Flickr, Delicious, Tagging e Folksonomy.

Aitichison e Clarke (2004, p.14) pontuaram alguns desafios impostos pela internet e a web 2.0, sobretudo ao que se refere ao controle do vocabulário a partir de linguagens de indexação. Os autores apontam que:

²⁸ A Figura foi desenvolvida em uma sessão de brainstorming durante a FOO Camp, uma conferência na O'Reilly Media (O'Reilly, 2005). Fonte: <http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html#mememap>

- O acesso à informação prossegue através de uma variedade de número de diferentes portais, entradas e motores de busca, voltados para públicos específicos. Ou seja, não há um tesouro universal, mas uma multiplicidade de diferentes vocabulários para diferentes aplicações.
- Um item publicado pode ser reutilizado e não é possível prever em quais sistemas ou redes um determinado documento pode eventualmente aparecer. Assim, os profissionais devem se empenhar mais para prever todas as necessidades que possam surgir para um determinado documento.
- Com a tarefa de entrada de dados / indexação distribuído a partir de um grande número de autores, webmasters, administradores de sistemas, etc., o controle de qualidade não pode ser imposta através das fronteiras organizacionais.
- Os utilizadores não querem lidar com qualquer coisa complicada. Um tesouro é percebido como muito complicado e estas estruturas sistemáticas bem apresentadas, cuidadosamente projetadas para selecionar o termo correto e preferido (s) para cada conceito necessário, são muitas vezes rejeitadas, consideradas uma barreira entre o utilizador e seu objetivo.

Neste cenário, caracterizado pela interação e pela imaterialidade do objeto informacional, cuja representação é de forma livre, a folksonomia enquanto atribuição de tags a recursos informacionais em contexto digital, de forma compartilhada e aberta promove flexibilidade na organização da informação. Assis e Moura (2013, p.102) reiteram que “a utilização do potencial da linguagem contextualizada que emerge das folksonomias implica em desafios que envolvem a formalização e a interoperabilidade entre os diferentes sistemas de representação da informação e do conhecimento que constituem a web e a investigação de novas metodologias e modelos conceituais no âmbito da organização da informação”.

Face ao exposto, argumentar-se-á sobre as folksonomias como fonte para estudos semânticos, levando em conta o conceito, suas características, vantagens e desvantagens, denominações associadas, os principais sistemas

atualmente em uso e suas estruturas. Ao final, faz-se breves apontamentos da aplicabilidade das folksonomias na área da saúde e, posteriormente, apresenta-se as ontologias, como estrutura formal e explícita.

2.1 FOLKSONOMIAS: FONTE PARA ESTUDOS SEMÂNTICOS

Em 2004, ao ter conhecimento de serviços online que permitiam ao utilizador adicionar e pesquisar *bookmarks* sobre qualquer assunto, que o arquiteto da informação Thomas Vander Wal formulou o termo *Folksonomy*, junção dos termos *folk* (povo, pessoas) com *taxonomy* (classificação), propondo a seguinte definição:

the result of personal free tagging of information and objects (anything with a URL) for one's own retrieval. The tagging is done in a social environment (usually shared and open to others). Folksonomy is created from the act of tagging by the person consuming the information (Vander Wal, 2007).

Entende-se que é de forma livre, pois pode ser utilizado ou alterado por outro utilizador e, o fato da escolha pela tag não estar submetida a nenhum controle ou regra de vocabulário, neste ambiente aberto, promovendo a construção coletiva e possibilitando uma abordagem colaborativa. Pelo excerto, as folksonomias tem como alicerce três elementos: as tags, os recursos informacionais (blogs, fotos, sítios da web, vídeos, etc) e as pessoas, conforme a ilustração abaixo:



Figura 10 - Triângulo da folksonomia

Fonte: Scot Project (2010) cit. por Guedes, Moura, Dias (2001)

A folksonomia pode aparecer sob diferentes perspectivas: como produto, torna-se resultado da etiquetagem de recursos, um vocabulário, uma lista ou conjunto de termos, um sistema de classificação socialmente construído, uma coleção de conceitos, conjunto informal e orgânico de terminologia relacionada, ou conjunto de etiquetas; como abordagem, metodologia, sistema, sendo um conceito que representa processo (Catarino & Batista, 2007). Nesta perspectiva, a ilustração abaixo esclarece as formas de visualização das tags no SBF CiteULike:

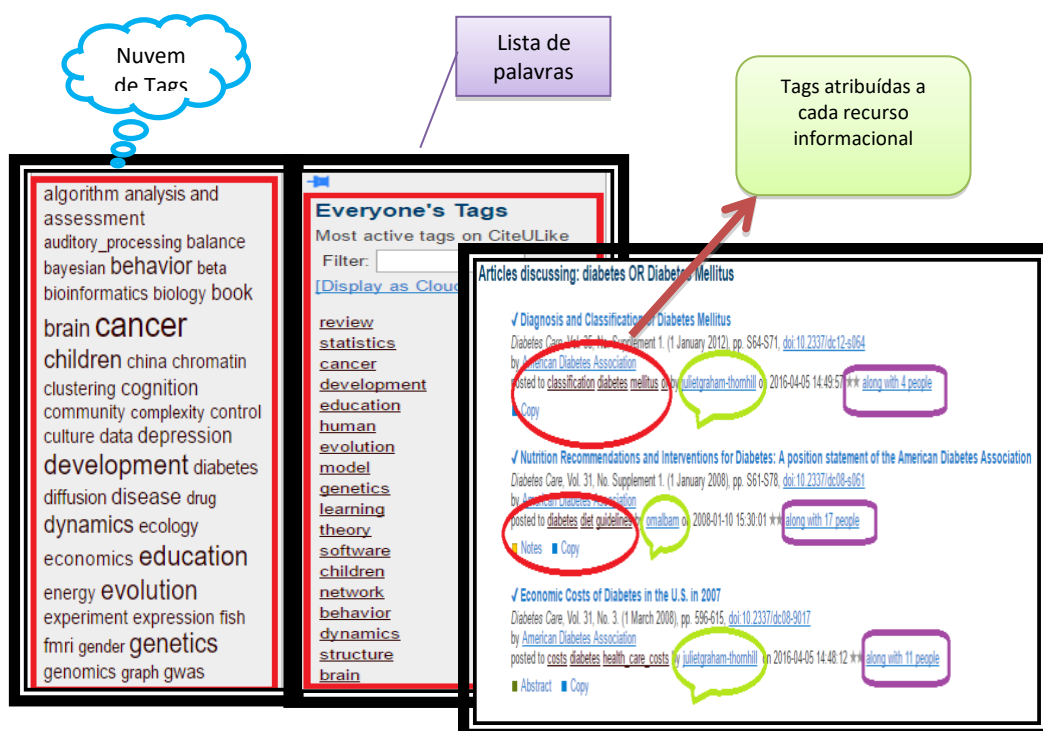


Figura 11 - Formas de visualização das tags

Fonte: Catarino e Batista (2007)

As características fundamentais (Yedid, 2013) estão centralizadas nos seguintes aspectos:

- São baseadas em linguagem natural, sem padronização e controle de vocabulário.
- As etiquetas são atribuídas aos recursos pelos utilizadores sem qualquer interferência e mediação.
- Sempre ocorre em ambientes digitais e de acesso a *web*.
- Implicam no compartilhamento e na colaboração entre os utilizadores.

- Não se refere apenas à atribuição, mas também à criação de etiquetas.

As vantagens que elas apresentam são a abordagem colaborativa e o compartilhamento, permitindo a visualização e a análise de como um objeto informacional é representado por diferentes utilizadores; a contextualização semântica esboçada pelas tags atribuídas; a informação distribuída; a criação automática de comunidades e grupos que partilham interesses comuns; a estrutura plana; a fácil visualização; a serendipidade; e o baixo custo do processo.

No entanto, suas limitações e desvantagens referem-se à baixa precisão no momento da busca, uma vez que uma mesma tag pode assumir diferentes significados, termos imprecisos, diferenças linguísticas, presença de erros ortográficos, manifestação dos fenômenos da língua natural (polissemia, ambiguidade, sinonímia). Assim, é oportuno mencionar que, apesar destes fatores, a língua natural compõe as taxonomias, as linguagens de indexação, ontologias e que, sob a ótica da representação e recuperação da informação, se apresenta como um desafio.

Peters (2009) expõem a diferente abordagem das folksonomias em relação à classificação e estrutura de recursos informacionais em ambientes digitais:

Instead of choosing a classification criterion and filling it with resources, it is now the resources that are allocated the criteria. Folksonomies turn the classification system from a criteria-centric into a resource-centric approach. This means that multiple storage no longer refers to the resources but to the multiple allocations of the "folders", "drawers" or "shelves" that are the tags of the folksonomy. Pinned to the information resources are as many tags as are necessary to adequately describe and retrieve them. Thus tags enable the most diverse criteria to be allocated to the resources and in this way guarantee a much broader access to them, which, due to the collaborative construction of the folksonomy, is also independent of the guardian. In the digital world, however, this approach always requires an indexing and retrieval system to render the folksonomy-based classification system manageable. The user may have created numerous access paths to the information resources, but a system will be needed to aggregate the tags and so provide links to the desired resources. So in order to structure and classify resources, folksonomies sidestep onto a meta-level, which represents the resource via (a whole lot of) tags (Peters, 2009, p. 3).

À vista disto, a autora ressalta que a abordagem focada no termo, no uso de vocabulário controlado e realizada por profissionais treinados, conseqüentemente a recuperação da informação independe de sua representação via linguagem.

A seguir, encontra-se o quadro 10, que descreve e estabelece diferentes aspectos do vocabulário controlado e das folksonomias.

Quadro 10 - Aspectos do vocabulário controlado e das folksonomias

ASPECTOS	VOCABULÁRIO CONTROLADO	FOLKSONOMIAS
Objetivo	Representar e buscar informação	Etiquetar informação
Origem	Década de 1950	Década de 2000
Cobertura	Restrita a um campo do saber	Geral
Entorno	Analógico e digital	Digital
Entidades	Estável	Instável
Fontes	Autorizadas	Subjetivas
Produtores	Especialistas, com base em normas	Internautas (utilizador comum)
Autoria/propriedade	Instituição ou pessoa que elaborou	Inexistente
Linguagem	Normalizada e controlada	Natural e dinâmico
Custo do processamento	Elevado	Baixo
Atualização	Periódica	Imediata
Valor adicionado	Organização conceitual de um campo do saber	Estudo do vocabulário usado por uma comunidade e sua evolução
Tipos de relação	Hierárquicas, associativas, equivalência	Associativas (co-ocorrência de palavras)
Utilizadores	Profissionais da informação, requer formação acadêmica	Utilizador de internet, não requer formação acadêmica
Exaustividade/precisão	Menor exaustividade, mas maior precisão	Maior exaustividade, mas menos precisão.
Estrutura	<i>Top-down</i>	<i>Bottom up</i>

FONTE: adaptado de YEDID (2013)

Relativamente ao quadro 10, convêm destacar o item ‘estrutura’, onde, no vocabulário controlado, os elementos são criados na abordagem *top-down*, ou seja, de cima para baixo, sob regras rígidas, estabelecendo relações hierárquicas; e, nas folksonomias não há estrutura, pois, os elementos são criados de baixo para cima, sem estarem submetidos a regras ou classes hierárquicas.

Assim, diferente dos sistemas tradicionais, incluindo bibliotecas digitais, na abordagem *bottom-up* as tags são livremente escolhidas pelos utilizadores sem uma taxonomia pré-definida facilitam a navegação de um sítio da web. Acerca disso e dos aspectos semânticos, Zhang, Wu & Yu (2006, p.171) reiteram que, apesar do aparente caos ocasionado pelo uso das tags, os sistemas baseados em folksonomias atraem muitos utilizadores, pois fornecem um mecanismo viável e eficaz para organizar os recursos na web, em decorrência das seguintes razões: baixas barreira na entrada dos dados, comunicação assimétrica e *feedback*, e aspectos individuais e comunitários. Uma vez que não é necessário um conhecimento sofisticado sobre taxonomias ou ontologias para fazer a anotação, marcação semântica ou categorização, as barreiras para entrada de dados são reduzidas. As etiquetas abertas e compartilhadas fornecem um retorno imediato (*feedback*) a um recurso informacional disponibilizado, pois outros utilizadores visualizam o recurso e etiquetam, utilizando a mesma tag ou não. Os utilizadores se adaptam às normas “implícitas” da comunidade reutilizando as mesmas tags ou podem influenciar o grupo fazendo outras sugestões. Assim, os utilizadores estão “negociando” significados de termos em uma comunicação assimétrica, proporcionada também favorecida pela estrutura up-bottom. Nesta perspectiva, acerca dos aspectos semânticos e vocabulários compartilhados entre os utilizadores, Merholz (2004, cit por Zhang; Wu; & Yu, 2006) fez uma interessante analogia denominada “*Desire line*”:

Desire lines are the foot-worn paths that sometimes appear in a landscape over time. The emergent semantics is like the desire lines. It emerges from the actual use of the tags and web resources and directly reflects the user’s vocabulary and can be used back immediately to serve the users that created them. In the following of the paper, we quantitatively analyze the folksonomy and show that emergent semantics indeed can be inferred statistically from it (Merholz, 2004, cit. por, Zhang; Wu; & Yu, 2006).

As tags (etiquetas) são à base dos espaços sociais semânticos e podem ser estudadas em função das palavras (linguística), dos símbolos programáticos (Ciência da Computação), das significações (Semiótica e Psicologia), da lógica simbólica (Filosofia), dos metadados (Ciência da

Informação) (Moura, 2009). Considera-se como importante contributo o fato das folksonomias surgirem como um meio de autoexpressão de um grupo e, em um contexto geral, sugerirem possibilidades de agregação, de análise e de funcionamento de comunidades que podem acelerar o compartilhamento e a especificação de linguagens de referência (Quintarelli, 2005).

Os estudos que versam sobre folksonomia envolvem a linguagem, a colaboração e as redes sociais, e estas últimas incorporam três dimensões (Marteleto, 2005):

- **sócio-comunicacional:** envolve elos, motivações e interações entre os atores sociais;
- **linguístico-discursiva:** incorpora aspectos cognitivos e informacionais envolvidos no compartilhamento social;
- **produção de sentidos:** refere-se ao fluxo e à dinâmica da ação colaborativa partilhada.

Assim, tanto o processo quanto o produto gerado pela folksonomia podem ser considerados esquemas de representação. São instrumentos que fazem a tradução dos conteúdos informacionais, sejam eles documentos originais e completos. Para um esquema estruturado, que representa esse conteúdo, a finalidade principal é organizar a informação e o conhecimento e, conseqüentemente, facilitar a recuperação das informações (Dahlberg, 2006 cit. por Carlan & Medeiros, 2011).

As representações são construídas socialmente por uma comunidade ou grupo de sujeitos. Assim, para representar algo é necessário que os fenômenos observados e suas representações estejam assentados na consciência do grupo. O ato de representar trata-se de um processo cognitivo, no qual tem como resultado a expressão dos pensamentos, observações e metodologias aplicadas pelo autor da representação. Para este processo é necessário que o autor utilize uma linguagem apropriada, condizente com o meio social (Almeida 2005 cit. por Boccato 2009).

Com o advento da cibercultura tornou-se possível a efetivação de ações de natureza coletiva, com participação de sujeitos geograficamente dispersos. A chave que permite o encontro entre tais subjetividades deixou de ser a empatia fundada nas relações interpessoais para efetivar-se pela convergência aparente de perfis numéricos e pelas evidências da similaridade de interesses, demarcadas pelos vestígios eletrônicos. O ciberespaço promove e intensifica a efetivação de uma singularidade ordinária e avessa às tentativas de centralização, pois nele é possível escolher um personagem, uma vida: “[...] máscaras de agir como um nômade nas diversas subjetividades *prêt-à-porter* da sociedade contemporânea” (Lemos, 2000 cit. por Moura, 2009).

Moura (2009) ressalta que as folksonomias contribuíram enormemente para a popularização de novas perspectivas de classificação de documentos digitais e ampliaram as possibilidades de compartilhamento de novas significações para termos e conceitos socialmente estabelecidos e debatidos em ambientes virtuais.

É neste panorama que Vander Wal esclarece o funcionamento de um sistema baseado em folksonomia e o uso indiscriminado dos conceitos ‘colaboração’ e ‘coletivo’. Tal equívoco deve-se ao fato de ferramentas serem apresentadas como colaborativas, entretanto são apenas baseadas e promotoras da interação social. A figura 12 torna compreensível o ponto de vista do autor.

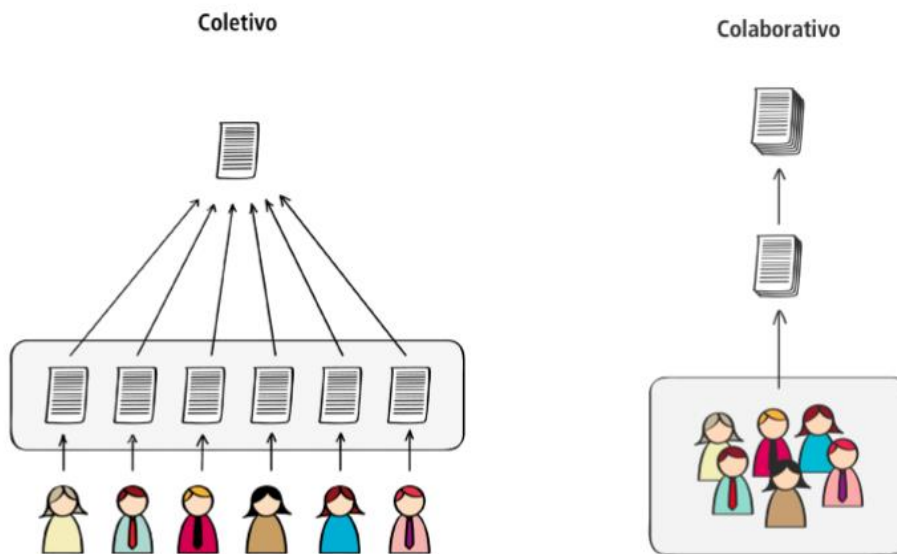


Figura 12 - Folksonomia: processos colaborativos ou coletivo
 Fonte: Brandt (2009)

A figura exposta esclarece que um único objeto informacional é compartilhado, visualizado por vários indivíduos e etiquetado por várias tags, enquanto no processo colaborativo vários indivíduos visualizam um único objeto e colaborativamente constroem uma única tag. Entretanto, com base no conceito de folksonomia formulado por Wal (2008) a contribuição de cada indivíduo é realizada isoladamente e pode ser selecionada e agregada para promover uma compreensão mais completa sobre o objeto, assim como a etiquetagem é feita por e para o indivíduo ler, visualizar, recuperar e compartilhar o objeto. O aspecto colaborativo existe, mas é vinculado e dependente de uma necessidade, pois é destinado para determinadas situações e isso não representa a folksonomia em si (Vander Wal, 2008).

Concomitantemente, há duas tipologias de folksonomia: a aberta, mais utilizada pelos *social bookmarks*, como o *Delicious*, consideradas abertas a qualquer utilizador; e a fechada, que é restrita ao utilizador que detém o objeto ou a outro utilizador permitidos pelo detentor. Esta última é mais utilizada para fotos e vídeos, como o YouTube e *Flickr*. A seguir, na figura 13, estão representadas as tipologias mencionadas.

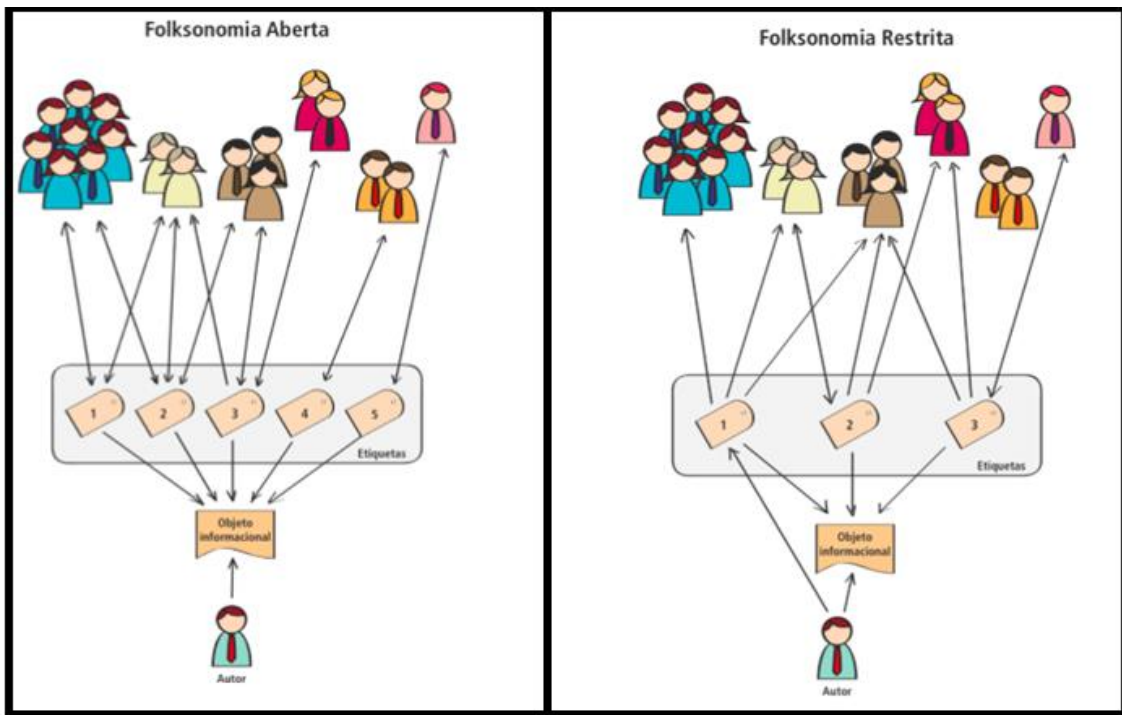


Figura 13 - Folksonomia aberta x folksonomia restrita

Fonte: Brandt (2009)

Com base na figura 13, verifica-se que na folksonomia aberta vários utilizadores colocam etiquetas no mesmo objeto e cada um pode etiquetar como preferir, utilizando seu próprio vocabulário. Dessa forma, um mesmo objeto pode receber *tags* de diversas pessoas. Entretanto, na folksonomia fechada, um grupo menor de pessoas, às vezes apenas uma, faz a etiquetagem do objeto e esta pessoa permite ou não que algumas pessoas visualizem e etiquetem o objeto, resultando em um número menor de etiquetas.

Neste íterim, os sistemas baseados em folksonomia aberta permitem identificar duas características: a lei da potência e a cauda longa (*power law* e *long tail*) (Brandt, 2009).

A lei da potência significa dizer que muitas etiquetas se repetem, pois muitos usuários utilizam o mesmo termo para descrever determinado objeto. Ao mesmo tempo, há muitas etiquetas usadas poucas vezes, formando a chamada cauda longa: “é onde se encontra uma pequena minoria de pessoas que denominam o objeto por um termo (Brandt, 2009, p.44).

Essas características são importantes, pois a primeira permite a identificação de tendências, destacando as etiquetas mais populares (ver Figura 2), e a segunda possibilita que o objeto seja recuperado por outros

utilizadores que façam uso do mesmo vocabulário das minorias que o classificaram, com etiquetas menos populares (Brandt, 2009). A seguir, apresenta-se a figura 14, que reproduz o compartilhamento de *tags* pelos utilizadores.

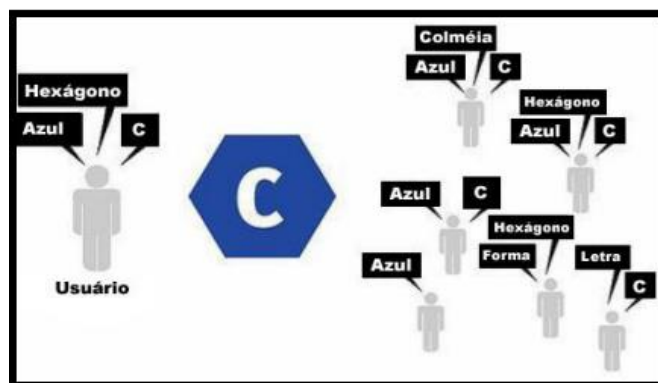


Figura 14 - Compartilhamento de tags entre os utilizadores

Fonte: Brandt (2009)

De acordo com a figura exposta, verifica-se que a folksonomia permite visualizar e acompanhar as preferências do utilizador, com o intuito de definir a relevância dos conteúdos. O conjunto de opiniões esboçado pela nuvem de *tags*, ou por alguma *toplist* (*top tags*), refletirá sobre a avaliação do conteúdo, sem encerrar as possibilidades de novas alterações (Nascimento, 2009). A figura representa ainda uma nova forma de organizar o site, e no âmbito da OI, pode ser aplicada enquanto garantia de uso do termo, atendendo necessidades de um contexto no qual a informação é distribuída e descentralizada.

2.1.1 CONCEITOS ASSOCIADOS

Conforme já mencionado, há distintas maneiras de abordar os estudos sobre folksonomias. Alguns destacam o papel do utilizador no processo de representação da informação, outros a questão do comportamento e produção de significados. Neste cenário, referente às múltiplas perspectivas, Trant (2009) ressalta:

We can think of tagging as a process (with a focus on user choice of terminology); of folksonomy as the resulting collective vocabulary (with a focus on knowledge organization); and of social tagging as a socio-technical context within which tagging takes place (with a focus on social computing and networks) (Trant, 2009, p. 4).

Consequentemente, uma diversidade de conceitos associados encontra-se disponível na literatura, seja ao fazer a intersecção com conceitos tradicionais oriundos da Ciência da Informação, ou relacionados à nova configuração perfilada pela web 2.0, a saber: indexação colaborativa ou democrática (Lancaster, 2004; Hilderley e Rafferty, 2007), indexação social, classificação social, etiquetagem social²⁹ (Trant, 2009), etnoclassificação e catalogação social.

A indexação democrática refere-se aos métodos colaborativos, no qual utilizadores compartilham um objeto informacional indexado, podem alterar os termos atribuídos, de modo a refletir uma “visão particular” acerca do objeto. Segundo Rafferty e Hilderley (2007), este processo considera quantos utilizadores associaram determinado termo a determinado objeto, defendendo que um objeto informacional pode ser lido sob múltiplas perspectivas. Também ratificam que os métodos colaborativos são viáveis em ambientes digitais, principalmente ao tratar imagens, enfoque de pesquisas desenvolvidas por estes autores.

The principle of Democratic Indexing is that individuals will have their own, potentially different, interpretation(s) of an image: the differences may be manifested as a different focus on parts of the image and different terms to describe the image. Democratic Indexing has incorporated a number of novel features: the information recorded for each information item includes descriptive cataloguing and subject indexing based on user perceptions of the item; the collection of user-generated indexes will be used to compile a “public” index through a process called “reconciliation”; and the ability of individual users to record their private indexes, offering a democratic approach to indexing (Rafferty & Hilderley, 2007, p.406).

²⁹ Social Tagging

Nesta concepção, os aspectos-chave são a interpretação e visão do utilizador acerca do objeto informacional, diferentemente da concepção da indexação tradicional, mediada por especialistas. Leva-se em conta o papel ativo dos utilizadores na construção e determinação de significados, via conjunto de tags atribuídas.

Na perspectiva do utilizador como único agente no processo de atribuição de tags a conteúdos na web, ocorre também o conceito de etiquetagem colaborativa (*Collaborative tagging*), conceituada por Golder e Huberman (2006) como um processo em que muitos utilizadores adicionam metadados para partilhar conteúdos. Os autores esclarecem que a etiquetagem ou organização de conteúdo eletrónico não é algo novo, mas a forma colaborativa sim. Para complementar, ao comparar este formato com os tradicionais, destacam a não hierarquia entre as tags e a possibilidade de múltiplas formas de organizar o conteúdo.

Outra noção que comumente associada à folksonomia é a etnoclassificação (Merholz, 2004), um outro neologismo, no qual “*etno*” refere-se a um grupo de determinada origem e costumes comuns, acrescido do termo classificação. Siqueira (2012) distingue folksonomia e etnoclassificação, considerando que o segundo apresenta em sua carga etimológica a noção de cultura e identidade.

Diante do exposto, observa-se que os conceitos associados as folksonomias refletem a perspectiva na qual o processo de etiquetagem e as tags serão abordados, mesmo que contemplem o termo colaborativo, o qual o Wal considera equivocado. Consequência das múltiplas formas com que pode ser abordada, pode-se verificar que todas as ocorrências mantêm as características das folksonomias, como a dinamicidade, flexibilidade, não hierarquização, a possibilidade de cooperação entre os indivíduos, sobretudo que estruturas sociais são compartilhadas.

2.1.2 SISTEMAS BASEADOS EM FOLKSONOMIAS

Há diversificadas maneiras de se referir aos sistemas que incorporam a folksonomia: Sistemas de Etiquetagem Social, Sistemas de Etiquetagem Colaborativa, Sistemas de Classificação Social e Sistemas de Classificação Distribuída, além de *Social bookmarking*, que refere-se a prática de etiquetar um recurso, semelhante aos “ favoritos”, ferramentas para o armazenamento de *bookmarks* em serviços online, ou os *tagging systems*. Para Braly e Froh (2006, cit por Peters, 2009, p.24) *social bookmarking systems* são a “*class of collaborative applications that allow users to save, access, share and describe shortcuts to web resources*”. Neste sentido, optou-se por utilizar a denominação Sistema Baseado em Folksonomia.

Em função do aumento e popularidade de elementos, como a interação social, provenientes da web 2.0 os SBF dependem de estruturas sociais compartilhadas, como as estruturas conceituais e linguísticas relacionadas a uma comunidade de utilizadores. Paralelamente, Peters (2009) insere os SBF na perspectiva de Serviços de Informação Colaborativa (*Service Collaborative Information*). Segundo a autora, a estrutura destes serviços revela que o acesso e a recuperação dos recursos informacionais ocorre por meio de interfaces que são representadas por nuvens tags, social bookmarks, pelas recomendações, ranking das tags mais populares.

O primeiro SBF desenvolvido foi o Delicious ao final de 2003, por Joshua Schachter. Ao longo destes anos, outros sistemas incorporaram as folksonomias, entretanto, alguns foram desativados. A seguir, encontram-se relacionados alguns SBF.

Quadro 11- Principais sistemas baseados em folksonomia

SBF	Recursos
Bibsonomy	Trabalhos científicos
CiteULike	Trabalhos acadêmicos
Delicious	Links favoritos
Diggo	Blogs
Flickr	Fotos
Frassle	Blogs
LastFM	Músicas
LibraryThing	Livros
LiveJournal	Blogs
Mendeley	Referências bibliográficas
Steve	Obras de arte

Technorati	Blogs
Tumblr	Blogs
Twitter	Microblog
Youtube	Vídeos

Fonte: Autoria própria.

Conforme exposto na tabela, verifica-se que os SBF diferem-se em seus objetivos e tipos recursos informacionais. Em geral, para além de suas especificidades isoladas, estes sistemas caracterizam-se por:

- Armazenar, gerencial e compartilhar recursos informacionais;
- Descobrir e pesquisar diferentes temas na mesma plataforma;
- Formar comunidades, promovendo redes de interação;
- Promover a colaboração e cooperação entre os utilizadores;
- Viabilizar o ranqueamento de temas mais populares, revelando pontos de vistas;
- Visualizar por nuvem de tags ou lista de palavras;
- Personalizar a representação dos recursos de informação armazenados.

Na sequência, no escopo desta tese, apresenta-se os enfoques e algumas funcionalidades específicas dos SBF LibrayThing e CiteULike

Quadro 12 - Enfoques e funcionalidades do LibraryThing e CiteULike

Enfoques e funcionalidades	LibraryThing	CiteULike
Cooperação	Sim	Sim
Formação de comunidades/ redes de	Sim	Sim, indica grupos de interesse
Formas de visualização	Lista	Lista de palavra e nuvem de tags
Gestão de tags	Adicionar, remover e renomear tags	Adicionar, remover e renomear tags
Recomendação de tags utilizadas por outros utilizadores	Sim	Sim
Convergência midiática ³⁰	Sim	Sim
<i>Trending</i> ³¹	Sim	Sim

Fonte: autoria própria.

³⁰ Entende-se por convergência midiática o fluxo de conteúdos através de múltiplas plataformas de mídia, à cooperação entre múltiplos mercados midiáticos e ao comportamento migratório dos públicos dos meios de comunicação, que vão a quase qualquer parte em busca das experiências de entretenimento que desejam (Jenkins, 2008).

³¹ Lista de temas mais pesquisados.

O quadro expõe as especificidades de cada sistema, com pontos em comum, mas revela principalmente que constituem ferramentas que alicerçam e viabilizam estudos semânticos a partir das folksonomias, disponibilizando uma base linguística. No tocante à convergência midiática, considerando a etiquetagem como parte do processo comunicacional espontâneo, facilitadora da interação entre utilizadores, as ferramentas convergem com outras, como, por exemplo, Twitter e Facebook, seja para o utilizador efetuar o login ou para compartilhar o recurso informacional.

Em comparação a esquemas de organização da informação, como tesouros, a folksonomia não apresenta estruturação semântica explícita, mas com base nas características e vantagens já mencionadas, viabiliza e expõe novas e importantes práticas sociais. Conforme os objetivos da tese, a seguir, apresenta-se breve argumentação acerca das ontologias, considerando-a um esquema de organização da informação associado ao contexto digital, cuja estrutura semântica é explícita e formal.

Antes de avançar sobre o tema ontologias, convém circunscrever brevemente a folksonomia no âmbito do Processamento de Linguagem Natural. Segundo Peter (2009), o PLN

Is a subarea of information linguistics that serves the disambiguation of natural language, in particular written language, and aims towards an automatic processing of language in computerized environments. It mainly draws upon the findings of linguistics and computer linguistics Peter (2009, p.139).

Portanto, a título de ilustração, segue a figura 15, de Peter (2009, p.141) que esboça um algoritmo típico do PLN relativo a folksonomia.

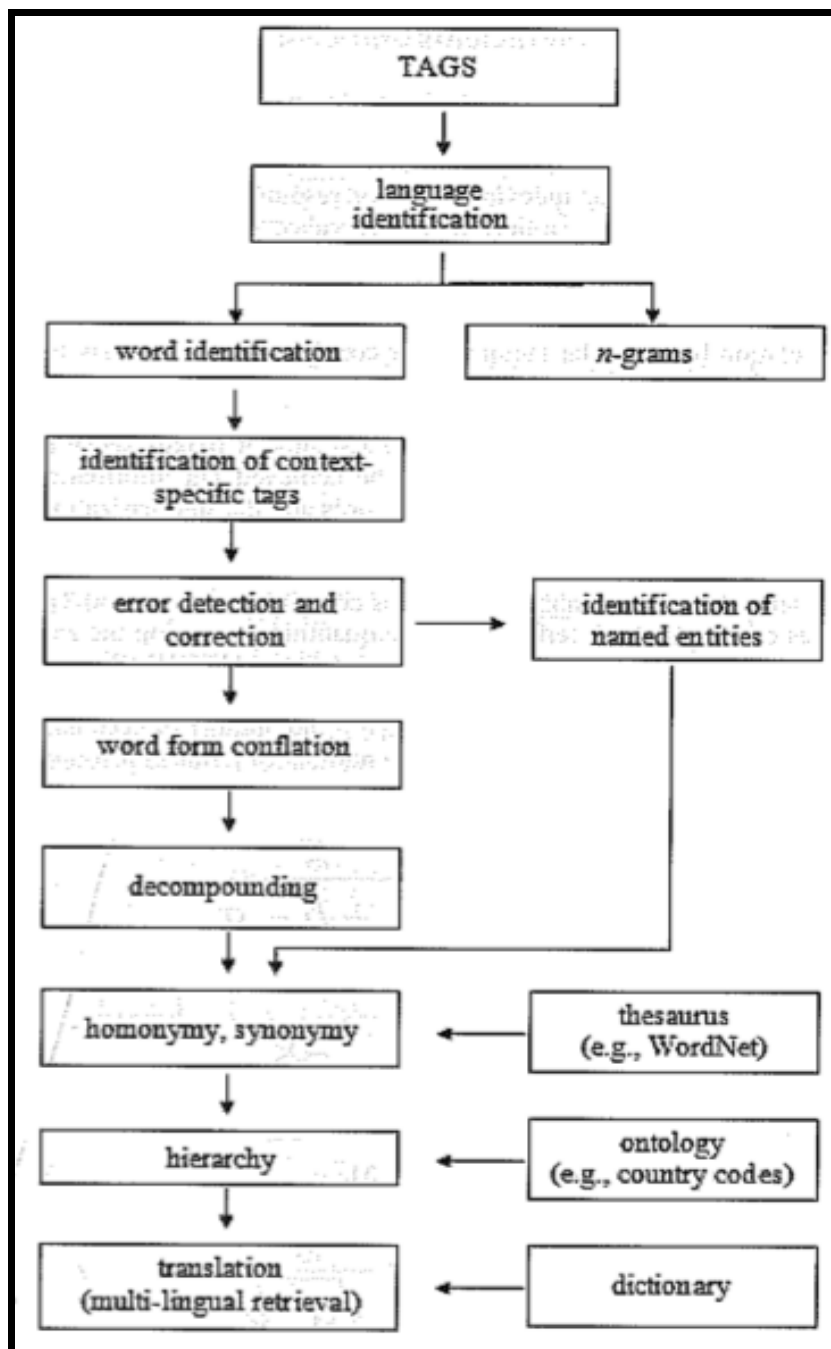


Figura 15 - Processamento de linguagem natural e folksonomia
 Fonte: Peter (2009).

Por meio da figura apresentada, compreende-se que o uso das folksonomias em PLN contempla várias etapas: inicialmente faz comparações com padrões de linguagem, distribuição de palavras ou n-gramas³² que revelam a linguagem dos termos ou tags. Sequencialmente, as tags são analisadas por palavras ou n-gramas, depois é feita a detecção e recuperação

³² Sequência de palavras.

de erros, bem como a descoberta de entidades, através do alinhamento da lista de palavras e nomes. Mesmo os sinônimos simples, como as siglas e a ortografia completa, são processados nesta etapa. As variantes ortográficas e morfológicas, como letras maiúsculas e minúsculas ou formas singular e plural são adicionadas no resumo das formas de palavras e unificadas via *stemming*³³. Já a lematização, mais recorrente ao conhecimento linguístico, trata-se de um processo de deflexionar uma palavra para determinar o seu lema (as flexões chamam-se lexemas). Ambos os processos unificam os tipos de variações existentes. O uso das tags com os processos típicos do PLN e a técnicas de alinhamento com tesouros, ontologias e dicionários aprimoram a recuperação da informação (Peter, 2009).

2.2 ONTOLOGIAS: INTER-RELAÇÕES CONCEITUAIS COMPLEXAS

Questões relacionadas à recuperação da informação aproximam a Ciência da Informação à Ciência da Computação, pois, imersas ao que é digital, em busca de aprimoramentos e soluções para organizar e compartilhar conteúdos na *web*, elas propõem teorias e metodologias acerca das ontologias.

Uma ontologia se refere a um artefato constituído por um vocabulário, usado para descrever uma certa realidade, além de um conjunto de fatos explícitos e aceitos, que dizem respeito ao sentido pretendido para as palavras do vocabulário (Almeida & Bax, 2003). Ou seja, configura-se como uma estrutura de organização da informação, semelhante às tradicionais linguagens de indexação, como tesouros, mas com inovações que permitem inferências automáticas.

Gruber (1993) expõe que uma ontologia é:

an explicit specification of a conceptualization. The term is borrowed from philosophy where an Ontology is a systematic account of existence.

³³ Stemming ou stemização: trata-se da redução de um termo ao seu radical, removendo as desinências, afixos e vogais temáticas. Com sua utilização, os termos derivados de um mesmo radical são contabilizados como um único termo.

For AI (Artificial Intelligence) systems, what “exists” is that which can be represented. When the knowledge of domain is represented in a declarative formalism, the set of objects that can be represented is called the universe of discourse. This set objects, and the describable relationships among them, are reflected in the representational vocabulary with which a knowledge-based program represents knowledge. Thus, in the context of AI, we can describe the ontology of a program by defining a set of representational terms. In such an ontology, definitions associate the names of entities in the universe of discourse (e.g. classes, relations, functions, or other objects) with human-readable text describing what the names mean, and formal axioms that constrain the interpretation and well-formed use of these terms. Formally, an ontology is the statement of a logical theory (Gruber, 1993, p. 908).

Com base nesta asserção, há dois elementos proeminentes, a saber: conceituação (conjunto de relações entre os conceitos) explícita e formalismo. A conceituação deve ser explícita no momento em que traduz do implícito para o explícito, descrevendo os elementos essenciais e particulares de um objeto, com vista a eliminar as ambiguidades. O formalismo refere-se ao fato da ontologia ser sistematizada, estruturada em modelos que permitam a união do explícito e formal, ser processada por máquinas e entendida por humanos (Shiessl & Shintaku, 2012). Já a conceituação deve ser consensuada, ou seja, aceita e compartilhada pela comunidade.

Dessa maneira, as ontologias são formadas por classes, relações, axiomas e instâncias. Por classes, entende-se como sendo as representações dos conceitos de um domínio estruturado em taxonomia, ou seja, com o estabelecimento de relações hierárquicas. Já as relações são a representação do tipo de interação entre os conceitos que constituem o domínio. Os axiomas são as estruturas, os atributos de restrição e interpretação dos conceitos, e, por fim, as instâncias são a representação dos objetos específicos, os próprios dados (Gruber, 1993).

Neste seguimento, Mika (2007) reforça que as ontologias, enquanto esquemas de representação de um domínio, apresentam duas características que carregam a noção de compartilhamento de significados ou semântica:

1. Ontologies are expressed in formal languages with a well-defined semantics.

2. Ontologies build upon a shared understanding within a community. This understanding represents an agreement among members of the community over the concepts and relationships that are present in a domain and their usage (Mika, 2007, p.67).

Referente à definição e as relações semânticas, Aitchson e Clarke (2004) esclarecem que

‘Ontology’ is another term used quite loosely indifferent circles, but in the artificial intelligence (AI) community, it is often defined as ‘a formal, explicit specification of a shared conceptualization.’ Implicit in this somewhat impenetrable definition is the requirement that concepts in the ontology, and the constraints on them, should be carefully defined. Relationships between the concepts should be specified much more precisely than is customary in thesauri. For example, where as most thesauri use the abbreviations BT and NT to cover all hierarchical relationships, whether generic, instantial, or partitive, an ontology used for AI applications would need to discriminate between these three. Similarly, the associative (RT) relationship would need to be separated out into cause/effect, material/property, etc. As Green points out, “The intuitive understanding that humans bring to relationships is not shared by computational devices”. Provided the terms in the system are defined and the nature of their inter-relationships is specified, a computer can make inferences when communicating with another computer (Aitchison & Clarke, 2004, p. 16).

Pelo exposto, nas ontologias as relações são mais detalhadas, dependentes de especialistas de determinado domínio e em termos de nível semântico, apresenta maior complexidade que outros esquemas de organização e representação da informação, conforme fica explícito na figura abaixo.

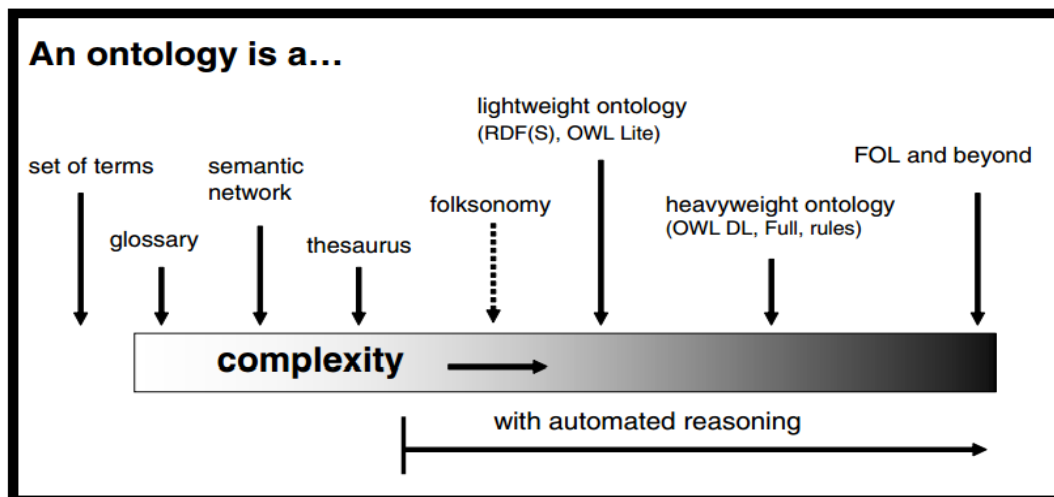


Figura 16 - Ontologia organizada com base na complexidade semântica
 Fonte: Mika (2007).

Shiessl e Shintaku (2012) apresentam quatro classificações para ontologias:

- **Genéricas:** apresentam conceitos gerais, independente de domínio. São embasadas em conhecimento comum e permitem aplicações mais amplas;
- **Domínio:** representam um universo do conhecimento, restringem sua abrangência por área e permitem ser utilizadas em várias aplicações ou sistemas;
- **Aplicação:** descreve métodos de solução geral de problemas. Podem ser aplicadas de diferentes maneiras e contemplam a terminologia necessária para que o método possa ser compartilhado e internalizado pelos utilizadores;
- **Tarefa:** tipo de ontologia mais específica, construída para uma aplicação peculiar, costuma descrever atividades, atores, funções e outros itens de um domínio bastante restrito, que impede a sua reutilização.

Diante disso, uma questão importante e preliminar sobre ontologia é sua finalidade. Uma vez que a intenção é partilhar o conhecimento e promover a interoperabilidade, Gruber (1993) propõe um conjunto de critérios para desenvolvê-la: clareza, coerência, extensibilidade, compromisso ontológico.

Nesse contexto, a especificação da ontologia cria compromissos ontológicos, os quais significam um acordo sobre a adoção de um vocabulário consistente, embora incompleto, com a teoria especificada por uma ontologia.

Alguns métodos para construção de ontologias são: TOVE (*Toronto Virtual Enterprise*), Enterprise, Método 101, *Methontology*, entre outros que estão reunidos no quadro.

Quadro 13 - Métodos para construção de ontologias

Métodos	Descrição
TOVE	Criada por Gruninger e Fox (1995), esta metodologia criada com base no desenvolvimento do projeto Toronto Virtual Enterprise (TOVE), cujo objetivo era o de criar um modelo de senso comum ou conhecimento compartilhado sobre empresas. Esta metodologia serviu de base para o projeto e avaliação de ontologias integradas em domínios corporativos, incluindo propostas de construção de ontologias e extensões de ontologias já existentes.
Enterprise	Método de Ushold e King (1995) foi desenvolvido com base na prática da construção da ontologia de alto nível Enterprise Ontology. Tem como propósito principal descrever o conhecimento sobre domínios corporativos ou de negócios.
Método 101	Prevê a construção de uma hierarquia de classes sobre o domínio estudado, bem como a definição de propriedades e instâncias para estas classes
Methontology	Metodologia desenvolvida no Laboratório de Inteligência Artificial da Universidade Politécnica de Madri entre 1996 e 1997, que possibilita a construção de uma ontologia por reengenharia sobre outra ontologia, utilizando-se do conhecimento do domínio tratado. Tal metodologia pode ser usada, segundo seus autores, em quaisquer domínios do conhecimento, embora seu uso mais conhecido e citado ocorreu na criação de uma ontologia no domínio da química.
Kactus	Método recursivo derivado do projeto Kactus que permitiu a reutilização de conhecimento em sistemas de complexidade técnica, tal como o domínio de redes elétricas, e a construção de ontologias nesse domínio como suporte a tais sistemas
Sensus	Método derivado da ontologia Sensus, a qual foi desenvolvida pelo grupo Information Sciences Institute (ISI) com o propósito de ser usada para fins de processamento de linguagem natural. O método Sensus propõe alguns processos para estabelecer as ligações entre os termos específicos e os termos da ontologia de alto nível, que corresponde à ontologia Sensus. Na prática, o método Sensus foi aplicado no desenvolvimento de uma ontologia no domínio de planejamento de uma operação militar aérea.
CYC	Método usado na construção da ontologia CYC, que considera o conhecimento consensual do mundo e é indicada pelos autores na criação de ontologias para fundamentar sistemas inteligentes
On-to-Knowledge Methodology (OTKM)	Metodologia desenvolvida para a construção de ontologias para aplicações de gestão do conhecimento, com o foco em Processo de Conhecimento (Knowledge Process) e em Conhecimento do Processo Meta (Knowledge Meta Process). Na prática, tal metodologia foi aplicada em um estudo de caso em gestão de competências de uma empresa internacional localizada na Suíça – a Swiss Life
Metodologia UP for ONtology	Metodologia de construção de ontologias derivada e baseada no padrão de engenharia de software conhecido como Processo Unificado – do inglês Unified Software Development Process ou Unified Process (UP) – do qual

(UPON)	foram derivados metodologias de software como o Rational Unified Process (RUP) e o PROcesso para Aplicativos eXTensíveis e Interativos (PRAXIS). Apesar da UP for ONtology poder ser usada em diferentes domínios do conhecimento, segundo seus autores, seu uso mais comum é no domínio do e-bussines
Metodologia NeON	Metodologia para construção de redes ontológicas baseado em um desenvolvimento colaborativo e argumentativo de ontologias. Tal metodologia foi desenvolvida em uma abordagem híbrida que combina o trabalho metodológico da área de Engenharia de Software e algumas metodologias para construção de ontologias, especificamente, a Methontology, a On-To-Knowledge, a DILIGENT e outros métodos ontológicos, como o de Grüninger e Fox. A NeOn Methontology inclui: (i) o Glossário NeOn de Processos e Atividades (ao todos tem-se 59 processos e atividades definidos), o qual identifica e define os processos e atividades potencialmente envolvidos quando redes ontológicas são construídas colaborativamente; e (ii) duas redes ontológicas sobre os modelos de ciclo de vida.
Metodologia MFPFO	Metodologia de construção de ontologia multi-facetada, anotada semanticamente, para a modelagem de uma família de produtos. Tal metodologia é capaz de sugerir, automaticamente, anotações semanticamente relacionadas, baseadas no design e no repositório de construção.
Ciclo de Vida de Schiessl e Bräscher (2011)	Embora não seja propriamente uma metodologia de construção de ontologias, o ciclo de vida ontológico, descrito por Schiessl e Bräscher (2011), inclui todas as etapas necessárias no processo de construção de ontologias, destacando o papel de cada etapa e as tarefas contidas em cada uma delas.

Fonte: adaptado de Mendonça (2015).

Embora cada método apresente suas particularidades e diretrizes, perpassam pelo ciclo de vida, cuja sequência é: especificação, conceituação, formalização, implementação e manutenção.

A especificação refere-se à definição do domínio, objetivo da ontologia, necessidades que ela deverá suprir e profissionais que estarão envolvidos. A conceituação envolve os requisitos especificados na etapa anterior, descreve um modelo conceitual de acordo com especificações, selecionando os elementos que comporão a ontologia e, a partir daí, constrói os relacionamentos. Posteriormente, a etapa de formalização estrutura e normaliza a ontologia, transformando o modelo conceitual em um modelo formal. Para tanto, a conceituação é representada em linguagem formal, a fim de restringir ambiguidades e delimitar as interpretações. A etapa de aplicação trata da implementação em um sistema de informação, seguida da manutenção (Shiessl & Shintaku, 2012).

2.3 ONTOLOGIA E FOLKSONOMIA NO ÂMBITO DA SAÚDE

As ontologias, atendendo a interesses de diversas áreas, são utilizadas em projetos relacionados ao desenvolvimento de sistemas para executar funções voltadas a recuperação da informação, ao apoio ao ensino e aprendizagem em áreas específicas, ao gerenciamento de dados e demais atividades que envolvem pessoas e máquinas, reforçados pelo fato de uma ontologia expressar um modelo, uma visão de mundo, a representação de um universo. Concernente à esta investigação, com a finalidade de esclarecer e exemplificar, encontram-se relacionadas no quadro 14, adaptada de Mendonça (2015) algumas ontologias desenvolvidas no âmbito da saúde.

Quadro 14 - Ontologias na área da saúde

Ontologia	Descrição
Ontologia de referência para o domínio da cardiologia	Ontologia de aplicação, define os conceitos, subclasses e instâncias em cada uma das subteorias do domínio da cardiologia, anatomia do coração, diagnóstico, terapia, avaliação clínica e exames complementares. (Adaptado de Mendonça, 2015)
Ontologia para sistemas de emergência: SAMU (Serviço de Atendimento Móvel de Urgência) e Corpo de Bombeiros	Ontologia de aplicação para facilitar o reuso e provável integração futura entre os sistemas de emergência em hospitais, clínicas e aplicações de prontuário eletrônico. Pode ser usada no desenvolvimento de sistemas baseados em conhecimento (Mendonça, 2015).
Ontologia de doenças sexualmente transmissíveis (DST)	Ontologia de aplicação para gerenciar o conhecimento referente a Acquired Immunodeficiency Syndrome (AIDS), Human Papiloma Virus (HPV), sífilis, sífilis congênita, gonorreia, herpes genital, linfogranuloma venéreo e tricomoníase (Mendonça, 2015).
Ontologia SODont	Ontologia de aplicação, define um vocabulário comum para médicos e um conjunto de dados com suas estruturas para softwares que precisam compartilhar informações de segunda opinião diagnóstica (Mendonça, 2015).
Ontologia para o atendimento emergencial de pacientes	Ontologia de aplicação, objetiva trocas de informações entre sistemas de informação hospitalar em atendimento de urgência a pacientes, englobando um vocabulário comum entre sistemas heterogêneos para o compartilhamento de informações relevantes sobre emergências (Mendonça, 2015).
Ontologia do cartão nacional de saúde	Representa conceitos e relacionamentos definidos pelo padrão de troca de mensagens do Sistema do Cartão Nacional de saúde. Engloba dados como os do atendimento do paciente, tipos de alta médica, entre outros (Nardon e Moura Jr., 2003).
Ontologia no domínio da doença neurológica	Desenvolvido na Universidade de Aveiro (Portugal) o projeto tem por objetivo melhorar pesquisas na web sobre saúde, atribuindo mais informações semânticas. Associado a plataforma FAST, que inclui

	navegação de ontologia nas classes e relações da entidade para orientar consultas; Uso das classes de domínio mais frequentes nos documentos relacionados a uma busca para fornecer formas adicionais de navegação (Mendonça, Rosa, Oliveira & Teixeira, 2013).
BioCaster – Ontologia de doenças infecciosas	O BioCaster é um sistema de mineração de texto baseado em ontologia para detectar e rastrear a distribuição de surtos de doenças infecciosas a partir de sinais linguísticos na Web (Collier, Doan, Kawazoe, Goodwin, Conway, Tateano & Shigematsu, 2008).
Ontologia – Gestão integrada de doenças crônicas	Effective use of routine data to support integrated chronic disease management (CDM) and population health is dependent on underlying data quality (DQ) and, for cross system use of data, semantic interoperability. An ontological approach to DQ is a potential solution but research in this area is limited and fragmented.
Ontologia sobre mortalidade	O objetivo foi utilizar as possíveis contribuições de uma ontologia para representar o conhecimento de um domínio da área de saúde constituído pelas informações sobre mortalidade no Brasil, diante de sua relevância para os estudos epidemiológicos e demográficos. (Mendonca, Cardoso & Drumond, 2009)
OntoCancro³⁴ – Inflamação crônica	Desenvolvido pela Universidade Federal de Santa Maria (Brasil) o projeto Ontocancro tem como objetivo integrar dados, de bases de dados públicas, de genes referentes ao câncer.
Food Ontology	Iniciativa do departamento de Ciência da Computação da Univerisdade de Liverpool, da Universidad Politecnica de Valencia e do Intituto Scientifico Universitario San Raffaele (Milão), trata-se de uma ontologia de alimentos orientada para o domínio nutricional e de saúde que é usado para compartilhar conhecimentos entre as diferentes partes interessadas envolvidas no projeto PIPS (<i>Personalised Information Platform for Health and Life Services</i>)
Ontologia - Doenças Tropicais Negligenciáveis – NTDO	Objetiva representar um conjunto de conhecimento complexo sobre a transmissão de Doenças Tropicais Negligenciáveis e os possíveis processos que ocorrem a partir do desenvolvimento destas, como o falecimento de indivíduos, em uma ontologia: a NTDO (Neglected Tropical Disease Ontology) (Silva, 2012).
OntoS – Ontologia Saúde – vigilância epidemiológica	A OntS (Ontologia Saúde), a princípio representando os conceitos relacionados a vigilância epidemiológica e os agravos de notificação compulsória, conforme a Portaria GM/MS Nº 5 de 21 de fevereiro de 2006 [9], e o desenvolvimento de uma ferramenta para auxiliar aos utilizadores na busca e recuperação de informações com relação aos agravos (Costa Medeiros, Oliveira & Sousa, 20??).
Ontologia de imagens do domínio da nefrologia	No âmbito do Projeto/Universal, com objetivo básico de construir uma ontologia de imagem desse

³⁴ Disponível em: <http://ontocancro.inf.ufsm.br/>

	domínio de conhecimento, dando ênfase às nefropatias, visando ao mapeamento, a recuperação e a gestão de imagem em ambiente eletrônico (Pinto, Campos, Ferreira & Freitas, 2013).
Ontologias para cuidados de saúde mental (Mental Health Care – MHC)	Descreve diferentes elementos do domínio MHC e o uso potencial das ontologias no desenvolvimento de sistemas interativos que podem apoiar intervenções de saúde mental, assim como são importantes para o gerenciamento de riscos (Coyle & Doherty, 2008).
Ontologia – Diagnóstico e tratamento da cefaleia	O objetivo foi buscar um meio de facilitar a geração de redes bayesianas, baseadas em diretrizes clínicas, a partir das ontologias criadas no domínio da cefaleia (Zanatta, Rodrigues, Cazella, Flores & Bez (2012).
OncoTerm	OncoTerm é um projeto de pesquisa interdisciplinar sobre a terminologia médica cujo objetivo principal é o desenvolvimento de um sistema de informação sobre subdomínio biomédica da oncologia em que os conceitos estão ligados a uma ontologia. Estrutura conceitual que subjaz a todos os outros componentes e fornece o link entre os termos em línguas diferentes (Inglês, Espanhol e Alemão). O sistema é projetado tanto para profissionais de comunicação e saúde e pacientes, usando tanto informações de dicionários e corpus especializados, tal como previsto por especialistas textos. Na gestão do conhecimento terminológica, identificando o campo base de conhecimentos e representações conceituais multidimensionais de organização conceitual é muito importante (López Rodríguez, Sánchez & Faber Benítez (2006).
DDO: diabetes mellitus diagnosis ontology	Este projeto objetivou desenvolver uma ontologia de diagnóstico de diabetes mellitus. O software Protégé 5 foi utilizado para a construção da ontologia. O DDO é desenvolvido dentro do quadro da ontologia formal básica e da ontologia para a ciência médica geral para representar entidades no domínio da diabetes. O DDO contém 6444 conceitos, 48 propriedades, 13.551 anotações e 27.127 axiomas. O DDO pode servir como base de conhecimento sobre diabetes e apóia o raciocínio automático. Isso representa um passo importante para o desenvolvimento de uma nova geração de ferramentas de suporte à decisão centradas no paciente.

Fonte: autoria própria.

O quadro acima apresenta o uso potencial das ontologias no âmbito da saúde. Certamente, não expõe os inúmeros projetos existentes, mas exemplifica distintas finalidades. Ademais, relativamente a diabetes, domínio selecionado para esta investigação, convém destacar o projeto Food Ontology, uma vez que visa atender pacientes diabéticos, que tem necessidades distintas, utilizando uma plataforma personalizada, com informações nutricionais, incluindo o tipo e quantidade de nutrientes e a ingestão diária

recomendada, exemplificando a utilidade da ontologia no contexto social e de melhoria da qualidade de vida de pacientes com doenças crônicas.

Relativamente à folksonomia, ela vem fomentado muitas questões acerca dos aspectos sobre colaboratividade, compartilhamento, inteligência coletiva, sua utilidade e comparação em relação às técnicas tradicionais de representação da informação, sobre a qualidade, tipo de linguagem, termos ou descritores que são utilizados, seu nível de especificidade. Não obstante a essas discussões, a título de elucidação, também convêm apresentar brevemente como a folksonomia tem sido aplicada na área da saúde. Assim, seguem descritos alguns casos:

- O PatientsLikeMe é uma comunidade de redes sociais online e seu público são pacientes com 3 diagnósticos: esclerose lateral amiotrófica, esclerose múltipla e doença de Parkinson. Os membros da comunidade podem descrever seus sintomas online em linguagem natural, resultando em tags populares disponíveis para análise clínica e para procurar por outros utilizadores e pacientes. A iniciativa comprovou os desafios da comunicação online dos pacientes, não apenas dos profissionais. Também, reconheceu o potencial da folksonomia para reduzir barreiras entre consumidores, pacientes e especialistas (Frost & Massagli, 2008).
- O artigo produzido por Linehan, Doughty, Lawson, Kirman, Olivier & Moynihan (2010) descreve o projeto e a avaliação inicial do Tag-iatelle, um aplicativo de marcação colaborativa para incentivar a alimentação saudável. Os utilizadores fotografaram suas próprias refeições e carregam essas fotos em um sítio, onde outros utilizadores fazem a etiquetagem do conteúdo anonimamente. Os resultados sugeriram que a marcação do conteúdo referente a alimentação é uma atividade muito popular e que há possibilidade de extrair informações nutricionais por meio das tags geradas.
- A Agência de Pesquisa e Qualidade da Saúde dos Estados Unidos (*U.S. Agency for Healthcare Research and Quality*) criou um site público para divulgar informações sobre suas iniciativas em tecnologia da informação em saúde. No último projeto de melhoria de qualidade

contínua, o NRC (*National Resource Center*) usou os *logs*³⁵ de busca do site para extrair frases de pesquisa geradas pelo utilizador. As frases foram então comparadas ao vocabulário controlado do site em relação à linguagem, gramática e precisão de busca. Os resultados da comparação demonstram que os dados do registro de pesquisa podem ser uma maneira econômica de melhorar os vocabulários controlados, bem como a recuperação da informação. As frases de pesquisa inseridas pelo utilizador também encontraram muitas semelhanças com tags (Dixon & McGowan (2010).

- Kristine Adland e Lykke (2012) relatam a pesquisa que o objetivo foi explorar como a marcação social poderia ser útil em um site de informações para pacientes com câncer e seus familiares. Verificou-se o potencial das tags para descrever e fornecer acesso ao conteúdo do site, assim como elas podem melhorar a interação e a comunicação entre utilizadores leigos e especialistas de domínio em um site de informações dentro dos cuidados de saúde.

Pelo exposto, foram reforçadas e exemplificadas as diferentes aplicabilidades das ontologias e das folksonomias na área da saúde.

2.4 INTEGRAÇÃO ENTRE FOLKSONOMIA E ONTOLOGIA

Com base no que foi exposto anteriormente, observa-se que, tanto a folksonomia quanto as ontologias associam-se às técnicas tradicionais de organização da informação, porém de maneiras distintas. Enquanto as folksonomias incorporam as dimensões sociais oriundas da atribuição das tags, de forma livre, as ontologias possibilitam a estruturação de um vocabulário formal, e assim ambas as abordagens reforçam discussões sobre metadados na web (Weller, 2007).

Quanto as folksonomias, Weller (2007) menciona suas desvantagens, já discutidas neste texto, e destaca a relação custo-benefício, viabilidade para indexação de grandes coleções de dados, o fácil manuseio, assim como o fato

³⁵ Log de dados é uma expressão utilizada para descrever o processo de registro de eventos relevantes em um sistema computacional.

da marcação social tornar grandes comunidades conscientes do uso de metadados de conteúdo descritivo, enfatizando também o aspecto comportamental dos utilizadores na atribuição das tags, no sentido de que podem aprender, compartilhar e usufruir as tags de outros utilizadores .

Outro ponto relevante refere-se ao fato de nem sempre o processo de atribuição das tags ter objetivos sociais e coletivos para auxiliar a recuperação de recursos de uma comunidade e grupo, mas assume objetivos estritamente pessoais. Neste sentido, Weller (2007) cita os estudos de Al-khalifa e Davis (2007), os quais discutem a natureza das tags, agrupando-as em três categorias: pessoal, subjetiva e factual.

Segundo Al-khalifa e Davis (2007) na categoria pessoal (*personal tags*) as tags tem um público-alvo³⁶ e são frequentemente usadas para organizar recursos do utilizador, podendo ser classificadas em tags de auto-referência, tags de gerenciamento de tarefas e tempo. A de auto-referência inclui qualquer tag que se aproxime do interesse do utilizador, como datas, nomes próprios. As tags de gestão de tarefa e tempo funcionam como lembretes e listas de tarefas para gerenciar atividades futuras do utilizador. Ainda nesta categoria, ressalta-se o uso de preposições, frases inteiras e abreviaturas.

Na categoria subjetiva (*subjective tags*) as tags expressam as opiniões sobre o recurso e podem ser classificadas como compostas ou informais. As compostas consistem em um qualificador subjetivo como “Beautifulsite”, ‘goodfor” e as informais incluem palavras produzidas pelo próprio vocabulário do utilizador, como “kool”, “Kickass”. Por fim, na categoria factual, as tags identificam fatos sobre o recurso, como autoria, local, principais conceitos, título, idioma, sinônimos, acrônimos, variação ortográfica ou linguagem específica de uma área de aplicação ou uso. Esta categoria fornece metadados usuais.

Complementarmente, Weller (2007) também apresenta as contribuições de Golder e Huberman (2006) acerca das funções das tags:

³⁶ Audience

1. Identificar sobre o que (ou quem) é um recurso: as tags identificam os tópicos de itens marcados. Esses itens incluem substantivos comuns de muitos níveis de especificidade, bem como muitos substantivos próprios, no caso de conteúdo que discute pessoas ou organizações.

2. Identificar o que é o próprio recurso: Tags podem identificar que tipo de coisa um item marcado é, além do que é sobre. Por exemplo, artigo, blog e livro.

3. Identificar quem é o proprietário do recurso: algumas etiquetas estão associadas a quem possui ou cria o conteúdo marcado.

4. Refinar categorias: algumas tags não ocorrem isoladas e, ao invés de estabelecer categorias próprias, refina ou qualifica categorias existentes.

5. Identificar qualidades ou características: geralmente são adjetivos que expressam a opinião do utilizador acerca do recurso etiquetado. Por exemplo: engraçado, assustador.

6. Auto-referenciar: as tags geralmente se iniciam com o pronome *meu*, de forma a identificar a relação do recurso e do conteúdo com o utilizador.

7. Organizar de tarefa: tags relacionadas à execução de tarefas que podem auxiliar o utilizador, ou um grupo.

Portanto, a noção sobre a natureza e tipos de função assumidos pelas tags esclarecem os metadados que são produzidos no processo de etiquetagem e como os mesmos podem ser utilizados. Por exemplo, no âmbito do processamento de linguagem natural, conhecer os erros ortográficos permite auxiliar na recomendação ou correção de textos.

Concernente à ontologia, elas são um componente relevante no desenvolvimento de pesquisas com abordagem semântica. De forma contrária às folksonomias, elas são projetadas por especialistas e utilizadas para tornar explícito e inequívoco o significado dos recursos, não apenas para comunicação interpessoal, mas também na perspectiva da interação homem-computador. As conceitualizações são expressas em sistemas de conceitos

(classes), instâncias (indivíduos) e as relações entre eles, incluindo possibilidades de definição de novos tipos de relações entre os conceitos, para além dos propostos nos sistemas de organização tradicionais, e permitem a adição de regras e axiomas. Sua criação é laboriosa e requer consideração meticulosa sobre como representar adequadamente um domínio de conhecimento, conforme reitera Weller (2007). Ainda, para este autor, as ontologias, no espectro dos sistemas de organização da informação ampliam as possibilidades dos métodos tradicionais, sendo um sistema semanticamente mais rico que os demais, contendo informações explícitas sobre os objetos e o contexto. Isso é realizado pelo uso de atributos ou restrições adicionado às relações de conhecimento, pela inclusão de instâncias.

Relativamente à comparação das ontologias com a folksonomias e como convergem, Binzabiah e Wade (2012) enfatizam o problema das tags não se ligarem entre si e de forma explícita, entretanto, ressaltam que elas podem ser utilizadas para fornecer recursos, tais como determinar o significado, o nível de aproximação e a relação entre os termos. No quadro 15, os autores pontuam as distinções.

Quadro 15 - Comparação entre ontologia e folkonomia

Formal taxonomies or ontologies	Folksonomies and free tagging
Categorization or model is seen as something static that can be created in advance.	Something that is created and updated as a part of ongoing activity.
Ontologies often are based on hierarchical structures.	Folksonomy creates an entirely flat namespace.
Hierarchical structures provide much more expressiveness and support for reasoning of various kinds.	Less expressiveness and support for reasoning of various kinds.
Hierarchical structures they are also more sensitive to changes.	Less sensitivity for changes.
The namespace in a Ontology is normally entirely closed. Users are free to choose whatever tags they want to describe entity.	The namespace in a folksonomy is normally entirely open. Users are free to choose whatever tags they want to describe an entity.
Provide a framework to handle structured information and to extract conclusions from such structured information.	Does not provide such a framework.
Ontologies are difficult to maintain.	Easy to maintain.
On the spectrum of knowledge representation systems, the most expensive in creation and maintenance is an ontology.	Easier to create, edit, use and reuse.
Requires consensual agreement on its contents from community members.	Does not require such consensual agreement.
Metadata is generated only by experts.	Metadata is generated not only by experts but also by creators and consumers of the content.
Usually, controlled vocabulary is used.	Usually, freely chosen keyword are used

	instead of a controlled vocabulary.
	Folksonomic tagging is intended to make information increasingly easy to search, discover, and navigate over time.
	The number of websites that support tagging has rapidly increased since 2004.
	Multidimensional: users can assign a large number of tags to express a concept and can combine them.
	Uncontrolled tagging can result in a mixture of types of things, names of things, genres and formats.

Fonte: adaptado de Binzabiah e Wade (2012)

Nesta mesma perspectiva, com ênfase na funcionalidade, Tjondronegoro e Spink (2008) fazem os seguintes esclarecimentos:

A well-developed folksonomy (community-based taxonomy) can act as a shared vocabulary that is both originated-by and familiar-to its primary users to make information increasingly easier to search, discover, and navigate over time. [...] As folksonomies are developed in social Web-based environments, users can generally discover who created a particular folksonomy tag, and see the other tags that this person created. This allows a more intuitive search as a user can learn from another user who tends to interpret and tag content in a way that makes sense to them, thus increasing their capacity to find related contents. For a narrower-collection of multimedia database, users would start with navigation, instead of searching as they would not know what is available for them. Most of general search engines emphasize on search, while most of specialized engines provide semantic-based browsing. One of the key reasons is that, collection-based search engines can provide users with narrow-enough topics and quantity of documents to be browsed. Hence, multi-document summaries such as topic-based clustering of related videos will be very useful to assist users in navigating the multimedia collections.

Complementarmente, sobre ontologia os autores esclarecem:

Hierarchical ontology should be used to classify and visualize keywords, topics, and other metadata that users and applications generate. A well-defined annotation dictionary [...] is desired as it allows the standardization of various multimedia contents descriptions. For search formulation, ontology-based classification can help users in redesigning their query if it is too specific. For example, instead of looking for “aloe vera”, users can be suggested to search on “green plants”. Moreover, a unified indexing on keywords and semantic summaries will enable search engines to support users in finding related topics.

Ambos os excertos enfocam a funcionalidade, sobretudo, acerca do processo de busca da informação, navegação e descoberta de recursos, podendo ser considerado um ponto convergente entre folksonomia e ontologia. Por conseguinte, no âmbito da abordagem integrada nos sistemas de representação do conhecimento, Sharif (2007) apresenta dois modelos de um sistema integrado, no qual cada uma destas tecnologias tem reduzido seus limites atuais, e exploram o potencial de cada uma. Neste seguimento, o autor apresentou uma ontologia da folksonomia para mostrar diferentes elementos em um espaço dinâmico e como as relações semânticas implícitas emergem de redes complexas implícitas nas folksonomias.

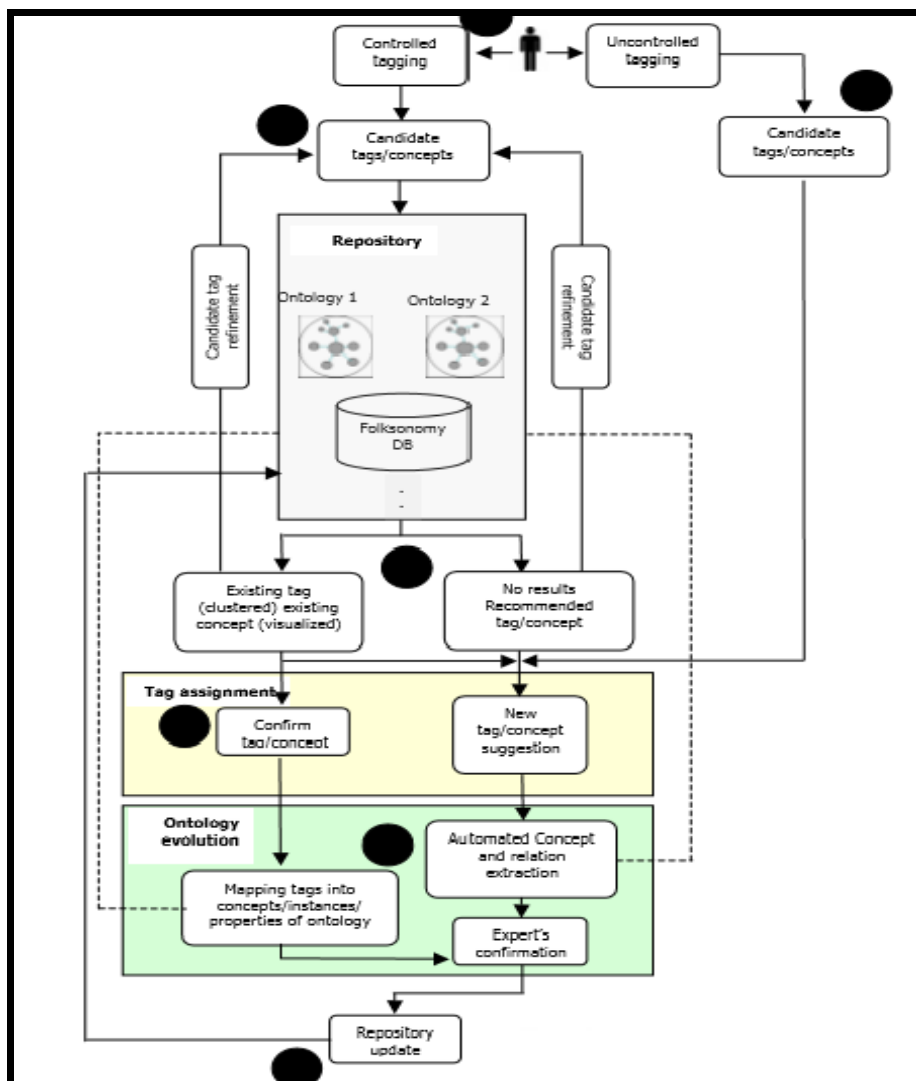


Figura 17 - Integração da folksonomia e ontologia para aquisição e organização
 Fonte: Sharif (2006)

Este primeiro modelo, exposto na Figura 17 (Sharif, 2006, p.9), apresenta um processo de marcação baseado em ontologia. O autor considera que a maioria dos utilizadores apesar de serem os provedores de informação não está familiarizada com as ontologias. Neste sentido, a inserção de novos conteúdos (conceito, relação, instância) trata-se de um processo difícil para os utilizadores e questiona como explorar os benefícios da ontologia sem incomodá-los e impor grandes desafios. Na sua proposta, ele reforça que a comunidade molda colaborativamente as definições formais de um domínio através do compartilhamento, sem a necessidade de estabelecer ou impor acordos acerca do significado de uma tag, pois seu significado gradualmente emerge com o uso do próprio sistema. Novas tags podem ser introduzidas quando necessário e a construção e manutenção estará ligada ao uso real do sistema. Neste sentido, as folksonomias reagem, refletem às mudanças e respondem às novas necessidades dos utilizadores.

Outro ponto exposto na figura é que diferentes níveis de colaboração são possíveis: a comunidade pode trabalhar com uma ontologia existente e sugerir novos conceitos ou instâncias, o que contribui para manutenção, expansão e atualização da ontologia. Também os utilizadores podem atribuir tags a um objeto sem que a tag seja sugerida ou já tenha sido utilizada, sem controle. Com base no modelo exposto, há um repositório que auxilia na recomendação de tags apropriadas e auxilia o utilizador.

Relativamente aos sistemas baseados em folksonomia, considera-se a relevância da navegação basear-se em nuvem de palavras, apesar de não se apresentar como a forma mais sofisticada, mas que pode reforçar também as ontologias. Conforme apresentado no modelo exposto na figura 18, a lista alfabética de tags, conceitos, atores e grupos podem ser desenvolvidos com base em um repositório de ontologias e folksonomias existentes. Também é possível ter visualização de conceitos, bem como nuvens tag e ator. Estes ajudam o utilizador a navegar inteligentemente. Também, caso o utilizador decida por consultar as ontologias incorporadas no sistema, isso levaria a melhor precisão e revocação. É patente que uma pesquisa de folksonomia baseada em ontologia reduz a ambiguidade. Na abordagem proposta (figura

18) observa-se o enriquecimento semântico por meio das tags explorando ontologias.

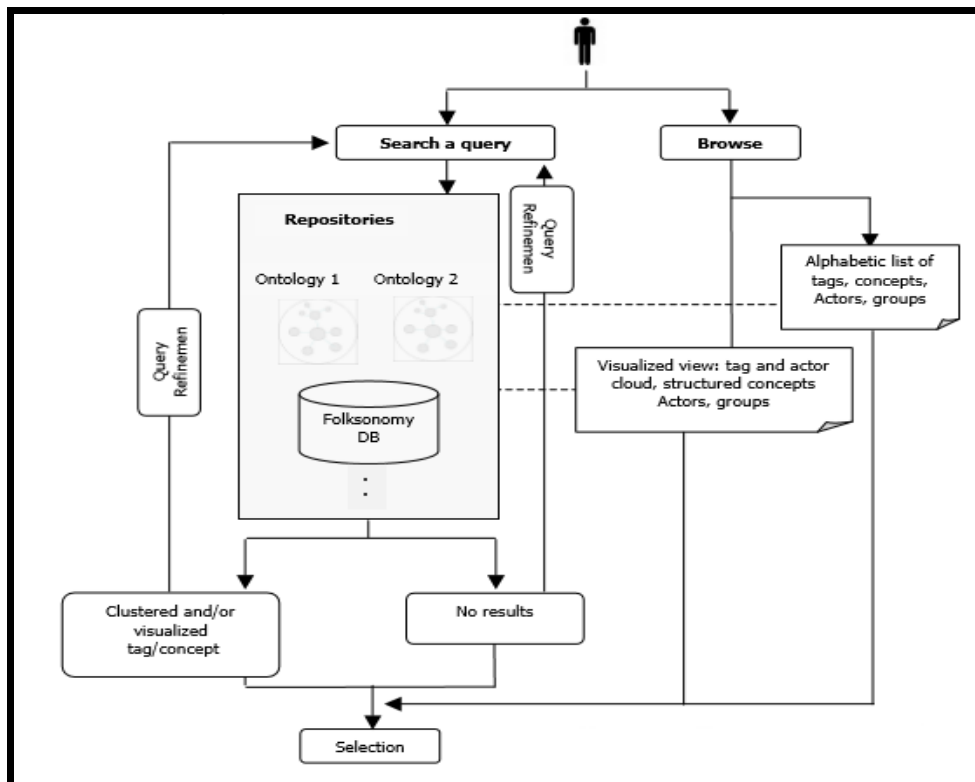


Figura 18 - Ontologia baseada em folksonomia
Fonte: Sharif (2006)

Após o exposto, tanto as diferenças, como convergências entre folksonomias e ontologias são utilizadas para potencializar os ideais da web semântica. Neste sentido, o autor, baseando-se em outras pesquisas, reitera

[...] that intelligent techniques may well be inside the system, but should be hidden from the user. So users in such integrated systems can exploit the benefits of ontology enhanced folksonomy search without the knowledge of the existence of neither the ontology infrastructure nor the related ontologies. They also help ontology evolution thanks to one more step [...] compared to other tagging systems. With this approach to integrating folksonomies and the semantic web, as Specia and Motta (2007) have mentioned, we intend to show ultimately both (1) that the ontologies provided by the semantic web can be used to structure folksonomies semantically and (2) that the dynamic knowledge provided by folksonomies can be used as a resource for bottom up knowledge acquisition to support ontology evolution (Sharif, 2006, p.11).

Para complementar a convergência entre estas tecnologias, Basso, Ferreira e Silva (2009) desenvolveram um processo para gerar ontologia com diversas relações (hierárquicas e não hierárquicas) que facilite a tarefa da recuperação da informação a partir das tags de um utilizador. Além de continuarem a ressaltar os problemas ocasionados pela folksonomia, principalmente ambiguidade, uso de acrônimos, homônimos e sinônimos, também mencionam que com o aumento no número de categorizações, devido aos SBF utilizarem listas ou nuvem de tags para iniciar a recuperação da informação, o que exige bastante esforço cognitivo por parte do utilizador. Enfatiza que a única relação entre as tags é a de co-ocorrência, não existindo muitas outras opções de visualização e acesso à informação, por ser semanticamente fraca, esta relação gera apenas uma estrutura plana entre as tags. Neste sentido, os autores apresentam na figura 19 um modelo ontológico³⁷ que representa o processo de atribuição de tags (*tagging*) estendido a partir do modelo de Knerr com relacionamentos semânticos entre as tags.

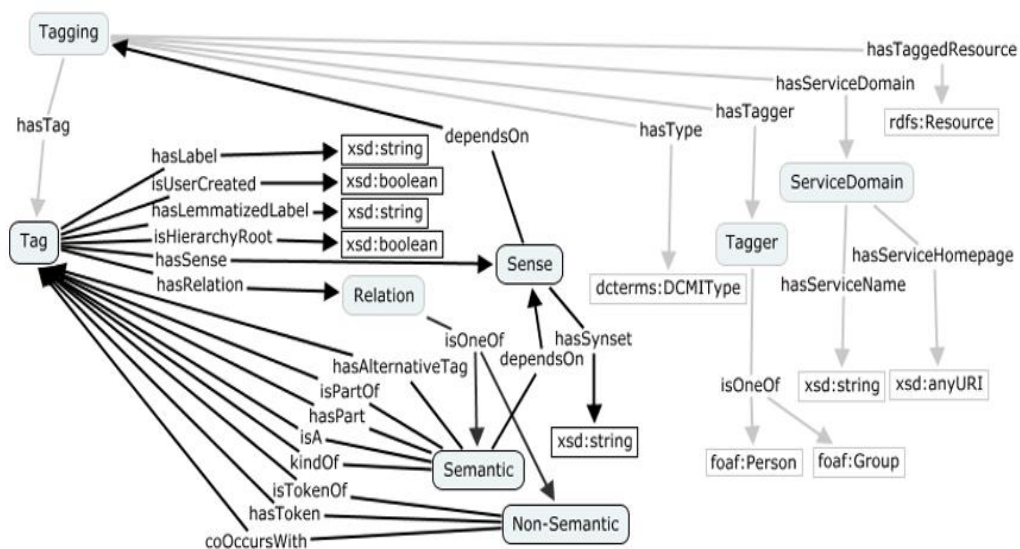


Figura 19 - Tagging e modelo ontológico
 Fonte: Basso, Ferreira e Silva (2009)

Relativamente à figura 19, destaca-se que a relação *hasAlternative* é ideal para resolver problemas associados a sinonímia, uma vez que ela relaciona

³⁷ Modelo de dados que representa conceitos.

uma tag a formas alternativas de escrita. As relações *hasPart* e *isPartOf* agrupam partes a um todo e vice-versa. Quanto às relações *isA* e *kindOf*, elas associam semanticamente termos mais abrangentes com tags cujo significado é mais específico e vice-versa. Também, podem formar estruturas hierárquicas se forem obtidas recursivamente até um nó raiz de uma ontologia. Deste modo, segundo os autores, esta taxonomia pode substituir a nuvem ou lista de tags e permitir ao utilizador especializar uma busca até obter o resultado esperado. Ademais, permite fazer busca mais abrangente, retornar a todos os objetos categorizados com tags mais específicas (gênero/espécie, por exemplo). Sequencialmente, apresenta-se a figura 20, na qual consta um conjunto de tags (programmingLanguage, Java e Prolog), cuja única relação é a de co-ocorrência, evoluindo para uma ontologia com diversas relações semânticas entre as tags (Basso, Ferreira & Silva, 2009).

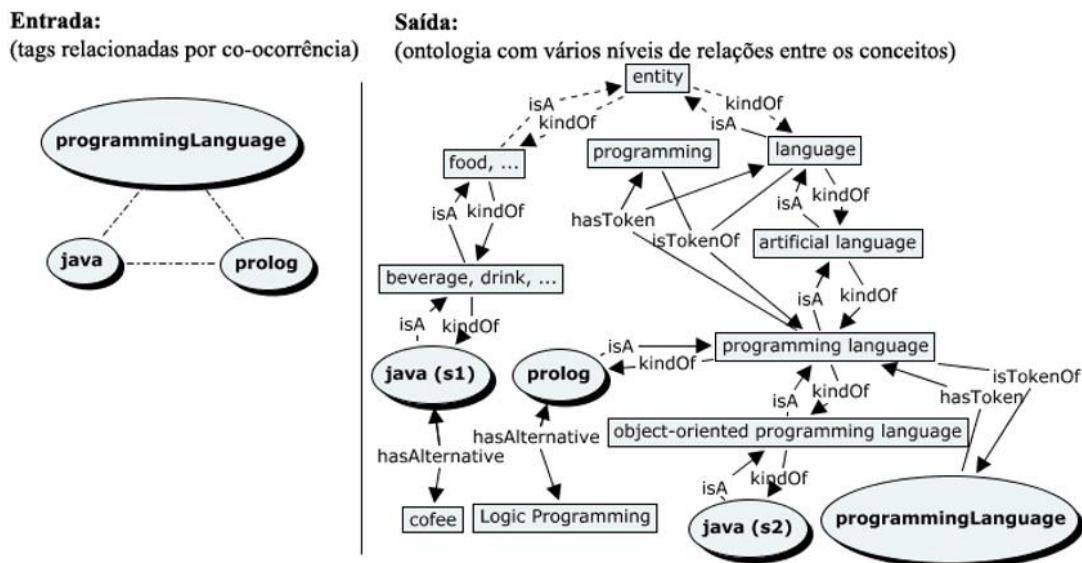


Figura 20 - Ontologia estruturada por tags
Fonte: Basso, Ferreira e Silva (2009)

Conforme a figura 20, observa-se que mais de um sentido foi atribuído a tag “java”, assim como foram atribuídas outras tags para mostrar o conjunto de tags em forma hierárquica (Basso, Ferreira & Silva, 2009). Os autores seguiram uma abordagem bastante automática e apresentaram como as ontologias podem ser utilizadas nos sistemas de recuperação e para recomendação de conteúdo, uma vez que elas possuem informações sobre o interesse dos utilizadores.

Também, Specia e Motta (2007) reiteram como as ontologias podem ser utilizadas para aprimorar tarefas nos sistemas baseados em folksonomia, pontuando 3 utilidades:

1. Desambiguação de consulta (tag): nas pesquisas de tags, a extensão de consulta para todas as tags relacionadas (ou um subconjunto deles) pode ser oferecida ao utilizador. As heurísticas simples, com base nos tipos de tags, podem ser usadas para estender a pesquisa a um subconjunto das tags relacionadas. Além disso, se a tag pesquisada for ambígua, o utilizador pode receber as tags relacionadas em cada cluster que aparece para escolher o sentido. A pesquisa pode então ser restrita a esse sentido adicionando outra tag (do cluster) à consulta.

2. Visualização: os clusters de tags relacionadas (e as relações entre eles) podem ser apresentados graficamente para fornecer uma melhor compreensão sobre o uso da tag pesquisada no sistema.

3. Sugestão de tag: ao marcar um recurso, o utilizador pode ser oferecido sugestões de tags "adequadas", com base em outras tags usadas por outras pessoas para esse recurso ou tags relacionadas que são altamente frequentes em um determinado cluster.

Diante do que foi exposto, verifica-se que a integração da folksonomia com as ontologias ocorre por duas vertentes: a primeira como fonte para popular as ontologias, sobretudo por representar a garantia do utilizador no que tange a nomear ou indexar recursos e a segunda como apoio ao utilizador ao realizar uma pesquisa nos sistemas, no próprio processo etiquetagem, de forma que ambas oferecem benefícios. Entende-se que algumas perspectivas utilizam análise estatística, por frequência e co-ocorrência das tags, outras enfocam a importância das relações semânticas serem explícitas, ou seja, como emergir uma estrutura com base nas tags. Também, percebe-se a intenção do enriquecimento semântico e da Folksonomia Assistida.

Baseando-se em Santarem Segundo (2010) entende-se que a Folksonomia Assistida, é um processo de apoio ao utilizador, no momento de definir os termos mais adequados para as tags que farão referência ao seu recurso informacional. O processo é composto por duas partes principais:

- **Primeira:** para a implementação da Folksonomia Assistida, deverá ser alterada a interface de comunicação do utilizador com o sistema informático, ou então desenvolvida uma nova interface, para a inserção de informações no campo palavra-chave, utilizado como referência para a inserção de conteúdo para as tags. Ou seja, deve-se apresentar ao utilizador, no momento da digitação da tag, de forma sistemática, um conjunto de informações previamente inseridas no sistema como sugestão de tags. A busca de informação para a sugestão é baseada em busca sintática.
- **Segunda:** o sistema deverá receber os termos enumerados pelo utilizador e proceder à pesquisa de relacionamento da informação dada pelo utilizador em relação ao conjunto de informações internas que a ferramenta dispõe. O processo de relacionamento em questão é justamente uma busca de relações dentro de uma estrutura de representação do conhecimento das áreas de especialidades, visto que esta pode ser caracterizada por um tesouro ou ainda por uma ontologia, que permitirão uma busca hierárquica horizontal, mas, principalmente, uma busca hierárquica vertical de relacionamento de termos.

O processo mencionado é exposto na figura 21 e expressa a principal característica da folksonomia assistida, que é oferecer ao utilizador um conjunto de termos que já estão sendo empregados no sistema, de forma que ele possa utilizar a base de conhecimento de um repositório específico para qualificar a descrição de seu recurso.

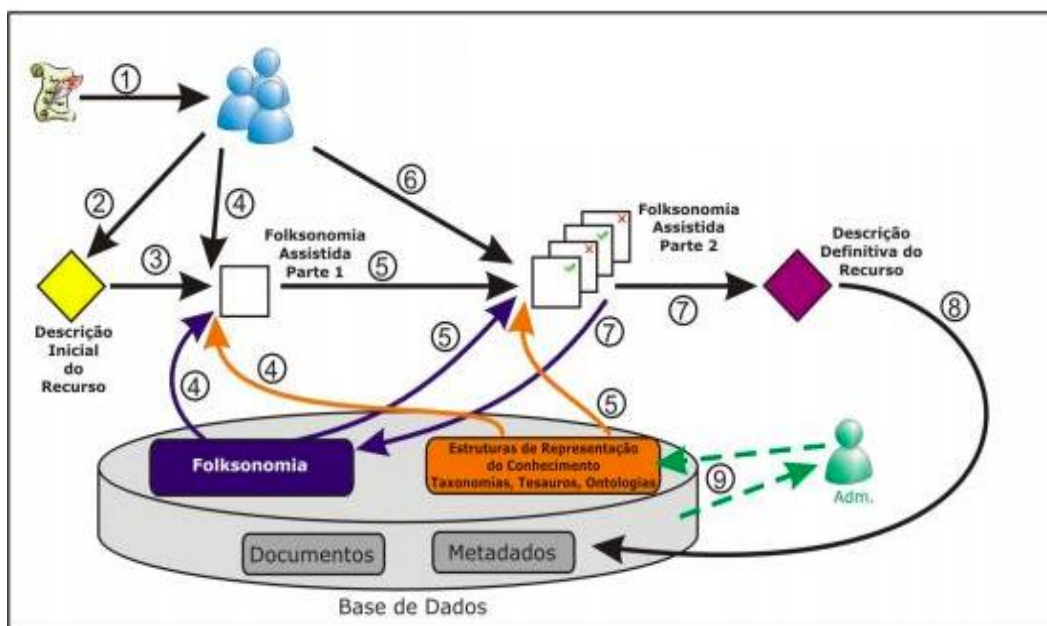


Figura 21. Folksonomia assistida
Fonte: Saitara e Segundo (2017)

O processo apresentado na figura 21 converge com a necessidade de fazer com que as tags tenham maior grau de significado em relação ao objeto inserido, principalmente dentro do contexto em que está sendo utilizada.

Por fim, em consonância ao que já foi exposto, outro conceito recorrente na literatura é Ontologia Folksonomizada (*folksonomized ontology* - FO).

We define a folksonomized ontology as an ontology aligned with terms of a folksonomy and enriched with their contextual data. By contextual data we mean data which emerges from a statistical analysis of a folksonomy, e.g. tag frequency, co-occurrence and information content. [...] The choice of the domain covered by the ontology and the folksonomy have direct impacts in the results. The results will be as good as the overlap between their domains (Wang, Wang, Zhuang & Fei, 2014, p. 179).

A técnica proposta aborda por um lado a FO que está alinhada com tags, e que conduz uma combinação de classificação baseada em semântica, categorização e tag mais rica. Por outro lado, os dados contextuais são usados para rever e melhorar a ontologia. A figura 22, de Wang, Wang, Zhuang & Fei (2014), esquematiza os papéis desempenhados por uma ontologia e uma folksonomia na construção de uma ontologia folksonomizada.

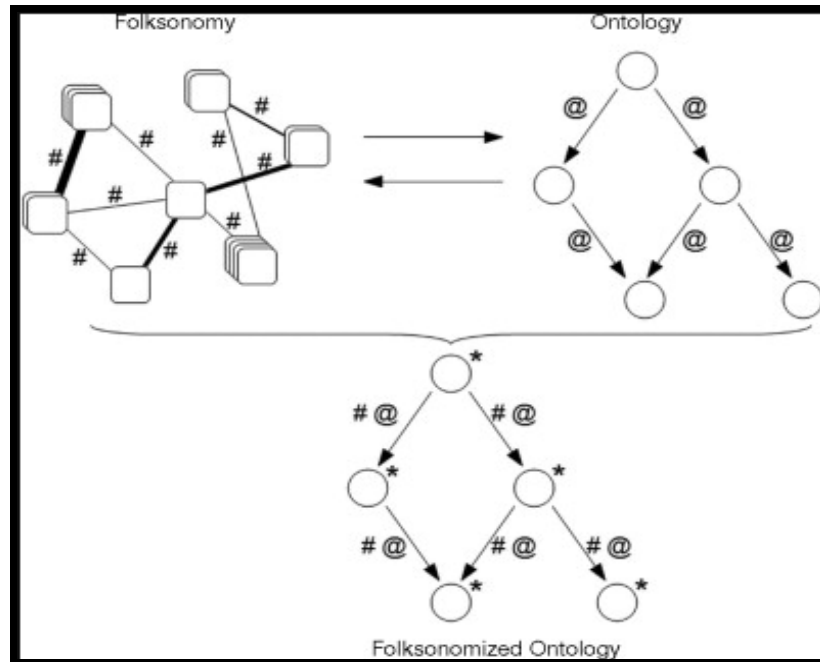


Figura 22 - Ontologia folksonomizada
 Fonte: Wang, Wang, Zhuang & Fei (2014)

Após o exposto, a FO é apresentada como uma entidade híbrida que funde folksonomias e ontologias. Trata-se de uma combinação simbiótica, aproveitando ambas as organizações semânticas. As ontologias fornecem uma base semântica formal, que é contextualizada por dados locais, melhorando as operações sobre tags com base em ontologias. Por outro lado, as OP foram utilizadas como ferramentas para analisar a ontologia e apoiar o processo de evolução da ontologia, mostrando discrepâncias entre o conhecimento emergente de uma comunidade e a representação formal desse conhecimento na ontologia.

CAPÍTULO 3 - RELAÇÕES SEMÂNTICAS NA PERSPECTIVA DA CIÊNCIA DA INFORMAÇÃO

“Toda relação é exterior a seus termos” (David Hume)

O objetivo deste capítulo é apontar como as relações semânticas são contempladas na perspectiva da Ciência da Informação, especificamente no âmbito da organização e da representação da Informação com vistas à recuperação da informação. Evidentemente, decorrente da característica interdisciplinar da Ciência da Informação, visitam-se outras áreas e subáreas do conhecimento, a saber: Linguística, Semântica e Lógica. A relevância deste capítulo assenta em esclarecer conceitos e reforçar abordagens teórico-metodológicas desta investigação, sobretudo ao que tange a relações semânticas na perspectiva da semântica lexical³⁸. Neste sentido, não se apresenta uma discussão exaustiva sobre Semântica, dada a dispersão e aos diversos aspectos em torno de sua definição, suas diversificadas linhas de pensamento e sistematização, mas algumas breves noções relevantes para esta tese. Em seguida, discutem-se acerca das relações semânticas, nomeadamente conceito, propriedades lógicas e tipologia.

3.1 SEMÂNTICA: CONCEPÇÕES LACÔNICAS

Semântica refere-se ao estudo do significado. Sua definição popularmente difundida é como a ciência que estuda a significação (Ilari & Geraldi, 1985), o estudo dos significados (Lyons, 1981). Rector e Yunes (1980) ressaltam que a Semântica se ocupa do significado dos signos em geral e é paralela à Semiótica e à Linguística, na medida em que se preocupa com o significado do signo linguístico ou não linguístico. Enfatizam que a análise semântica de uma dada língua deve explicar como as frases desta língua são compreendidas, interpretadas e relacionadas aos estados, processos e objetos do universo.

³⁸ Segundo Cançado (2013), a Semântica Lexical, vista como uma ampla área de investigação, trata do significado cognitivo que envolve a relação entre a língua e os construtos mentais que de alguma maneira representam ou estão codificados no conhecimento semântico do falante. Esta teoria faz parte da semântica estruturalista que, assim como Saussure, preocupa-se com a linguagem e não com as coisas (mundo real).

Antes de discorrer acerca das abordagens e vertentes da Semântica, convém apresentar sumariamente em que medida diferem-se da Semiótica. Rector e Yunes (1980) esclarecem que por Semiótica entende-se uma teoria geral de todos os tipos de signos. Entretanto, as autoras observam que as fronteiras Semióticas são móveis e movimentam-se em diferentes direções. A Semiótica interroga a natureza do signo, estuda um código com relação a outro e o produto desta relação. Elas enfatizam que na perspectiva do teórico Greimas a Semiótica não constitui a teoria dos signos, mas a da significação: o signo está na estrutura superficial e somente por meio de sua desestruturação chegamos à significação que já é subjacente. Entretanto, define-se a importância da Semântica, que se ocupa do significado dos signos em geral e pode-se dizer que é “paralela” à Semiótica e à Linguística, na medida em que se preocupa com o significado do signo, linguístico ou não (Rector e Yunes, 1980, p.5). Ainda, segundo as autoras, Semiótica³⁹, a grosso modo, é o estudo do signo e este é “alguma coisa que está no lugar de algo”. O componente do signo que tem um impacto físico ou perceptível no intérprete é o significante. O componente, na “mente” do intérprete, chamado geralmente de ideia ou de conceito, e que pode, eventualmente, revelar-se uma entidade neuroquímica, é conhecido pelos semioticistas como significado. Um signo é entendido não somente como um significante físico, mas ambos: um significante e um significado conceitual que, juntos, estão no lugar de algo. Os dois componentes do signo, significado e significante, são de certo modo ativados no intérprete que os aceita como um substituto para o objeto (evento). A Semiótica, portanto, não é meramente o estudo dos signos, mas é o estudo de como os intérpretes efetivam, realizam um número infinito de relações semióticas que existem potencialmente no universo.

Já os estudos semânticos manifestam-se de diversas formas e abordagens. Rector e Yunes (1980) apresentam a Semântica subdividida em filosófica, geral e linguística:

³⁹ Alguns semioticistas são: Roman Jakobson, Charles Morris, Algirdas Greimas, Claude Lévi-Strauss, Décio Pignatari, Jacques Derrida, Lucia Santaella, Yuri Lotman, Roland Barthes, Vilém Flusser.

- Semântica filosófica: também conhecida como lógica do signo, estuda a linguagem como instrumento do conhecimento, cujo objetivo da disciplina é livrar o pensamento do condicionamento social e construir sistemas de signos originais, o que facilita precisar a estrutura e o mecanismo de significação e mostrar a originalidade da linguagem e sua verdadeira função.
- Semântica geral: conhecida também como psico-sociologia do signo, ela estuda os princípios gerais que regem a comunicação humana, cujo objetivo é ser uma aplicação de métodos, hábitos e pontos de vista da ciência aos problemas vivenciais diários.
- Semântica linguística: também conhecida como teoria da informação, na perspectiva apresentada por Rector e Yunes (1980), na qual se estuda a função e a natureza das “palavras” no processo da comunicação linguística, atendendo ao sistema linguístico em si mesmo.

Complementarmente, Muller, Viotti e Foltran (2003) declaram que há uma variedade de semântica: textual, cognitiva, discursiva, argumentativa, lexical e assim por diante. Consequentemente, reforça a dificuldade em sistematizar os conceitos, acepções e abordagens. Neste sentido, observa-se que os estudos semânticos podem manifestar-se nas seguintes abordagens:

- Semântica Formal: considera como uma propriedade central das línguas humanas o ser sobre algo, o fato de que as línguas naturais são utilizadas para falar sobre objetos, indivíduos, fatos, eventos, propriedades, descritos como externos à própria língua. Assim, a referencialidade é tomada como uma de suas propriedades fundamentais. Por esta razão, na Semântica Formal, o significado é entendido como uma relação entre a linguagem por um lado e, por outro, aquilo sobre o qual a linguagem fala. Este 'mundo' sobre o qual falamos quando usamos a linguagem pode ser tomado como o mundo real, parte dele, ou mesmo outros mundos ficcionais ou hipotéticos (Muller, Viotti e Foltran, 2003);
- Semântica estrutural: oriunda dos estudos desenvolvidos por Ferdinand Sausurre, que compreende que o signo é composto pelo significado

(conceito, plano das ideias) e pelo significante (imagem acústica, plano de expressão) (Saussure, 2004);

- Semântica da enunciação: também conhecida como Semântica Argumentativa, estuda o sentido construído pelo linguístico. Pretende mostrar que é possível explicar o significado do signo sem fazer intervir o extralinguístico, ou seja, sem fazer intervir um domínio diferente do domínio da língua. Não se trata de estudar a relação entre o sentido de uma palavra e a informação que a palavra dá sobre a realidade extralinguística, nem sobre uma realidade psicológica, nem sobre condições de verdade. Mas seu objeto de estudo é o sentido linguístico que se produz, não na língua, mas no discurso, ou seja, no emprego da língua. Em síntese, é o modo de explicar o sentido essencialmente pela noção de relação que decorrem do objeto de estudo e das características da teoria: a de ser uma semântica, porque vai em busca da explicação do sentido; a de ser uma semântica linguística, porque explica o sentido construído pela relação entre palavras, enunciados, discursos; a de ser uma semântica linguística do discurso, isto é, do emprego da língua, não da palavra ou da frase isoladas; a de ser uma teoria explicativa do sentido do discurso, sempre olhando a linguagem a partir das bases epistemológicas que a sustentam (Barbisan, 2013).

Ainda, relativamente às abordagens sobre Semântica, apoiados em Cançado (2005), Almeida, Coelho, Souza e Teixeira (2010, p. 387) expõem três:

- a abordagem referencial⁴⁰, que advoga a relação entre o referente no mundo e as palavras que o representam. A abordagem referencial fundamenta-se no processo de comunicação, nas formas através das quais as regras e os princípios gramaticais combinam palavras e expressões significativas, baseando-se em condições de verdade. Cançado (2005, p. 79) apresenta como exemplo dessa abordagem a sentença - Noam Chomsky é um famoso linguista - explicando que o nome Noam Chomsky refere-se a um único indivíduo no mundo, que tem

⁴⁰ Segundo Cançado, também compreendida como Semântica Formal

como características famoso e linguista. Exemplos dessa vertente são a semântica formal de Lyons e de Kempson, bem como a semântica lexical de Cruse e de Evens;

- a abordagem mentalista, na qual o sentido está em um nível intermediário entre o mundo e as palavras, denominado nível de representação mental. Fundamenta-se na construção do significado através de representações mentais criadas pelas pessoas a respeito do que dizem ou observam. Essas representações são esquemas mentais que podem ser imagéticos ou conceituais. Exemplos de abordagens dessa linha de pensamento são a semântica gerativista de Chomsky, a semântica representacional de Jackendoff e a semântica cognitiva de Lakoff;

- a abordagem pragmática, a qual estuda os usos situados da língua, ou seja, os usos que correspondem ao modo como as palavras são pronunciadas em certas circunstâncias. Na abordagem pragmática, a linguagem “é mais do que a descrição dos estados das coisas, como na linguagem referencial, e mais do que estabelecer a relação entre o mundo e os conceitos mentais, como tratado na linguagem mentalista” (Cançado, 2005, p. 127). O significado é, nesse caso, atribuído à prática e às convenções sociais. Como exemplo, a sentença - O cachorro está no jardim - tem um significado que, ao interagir com a capacidade humana de contextualização, permite sugerir atos distintos como: Tire o cachorro do jardim! Ou Socorro! Exemplos de abordagens relevantes dessa linha de pensamento são a teoria dos atos da fala de Austin e a semântica argumentativa de Ducrot.

Após o exposto, a perspectiva adotada nesta investigação é a da semântica lexical, uma vertente relativa aos estudos semânticos. Segundo Cançado (2013), essa teoria faz parte da semântica estruturalista que, assim como o estruturalismo linguístico, proposto por Saussure, se preocupa com a linguagem e não com as coisas (mundo real). Na semântica lexical, as palavras são definidas umas em relação às outras. A autora complementa reiterando que a Semântica Lexical, vista como uma ampla área de investigação, trata do significado cognitivo que envolve a relação entre a língua e os construtos

mentais que de alguma maneira representam ou estão codificados no conhecimento semântico do falante. Teorias que tratam do significado cognitivo olham para dentro do aparato linguístico do falante e não estão preocupadas com o “mundo público”, que envolve a comunicação linguística”. Em uma breve apresentação, segue-se a trajetória da semântica lexical:

O primeiro estágio da história da Semântica Lexical pode ser datado de 1830 a 1930 e é conhecido como Semântica Histórico-Filológica. Com uma orientação histórica, a preocupação principal concerne às mudanças dos sentidos das palavras: a identificação, classificação e explicação das mudanças semânticas. Por volta de 1930 a 1960, temos a Semântica Estruturalista, evidentemente influenciada e orientada pelo trabalho de Saussure. Em meio a uma variedade de posições teóricas e métodos descritivos que surgem da concepção estruturalista, podemos apontar três pontos que distinguem essa corrente teórica: o aparecimento do conceito “campo lexical”, a análise componencial (traços semânticos) e a semântica relacional (sinonímias, hiponímias, antonímias e meronímias). [...] Portanto, grosso modo, não podemos falar de uma teoria classificada como Semântica Lexical, mas de várias “semânticas lexicais”, e o que teria de comum entre esses estudos é ter como principal objeto a relação entre a língua e a sua representação mental (Cançado, 2013, p. 127).

A seguir, encontra-se a figura 23 que representa a semântica lexical e seus componentes:

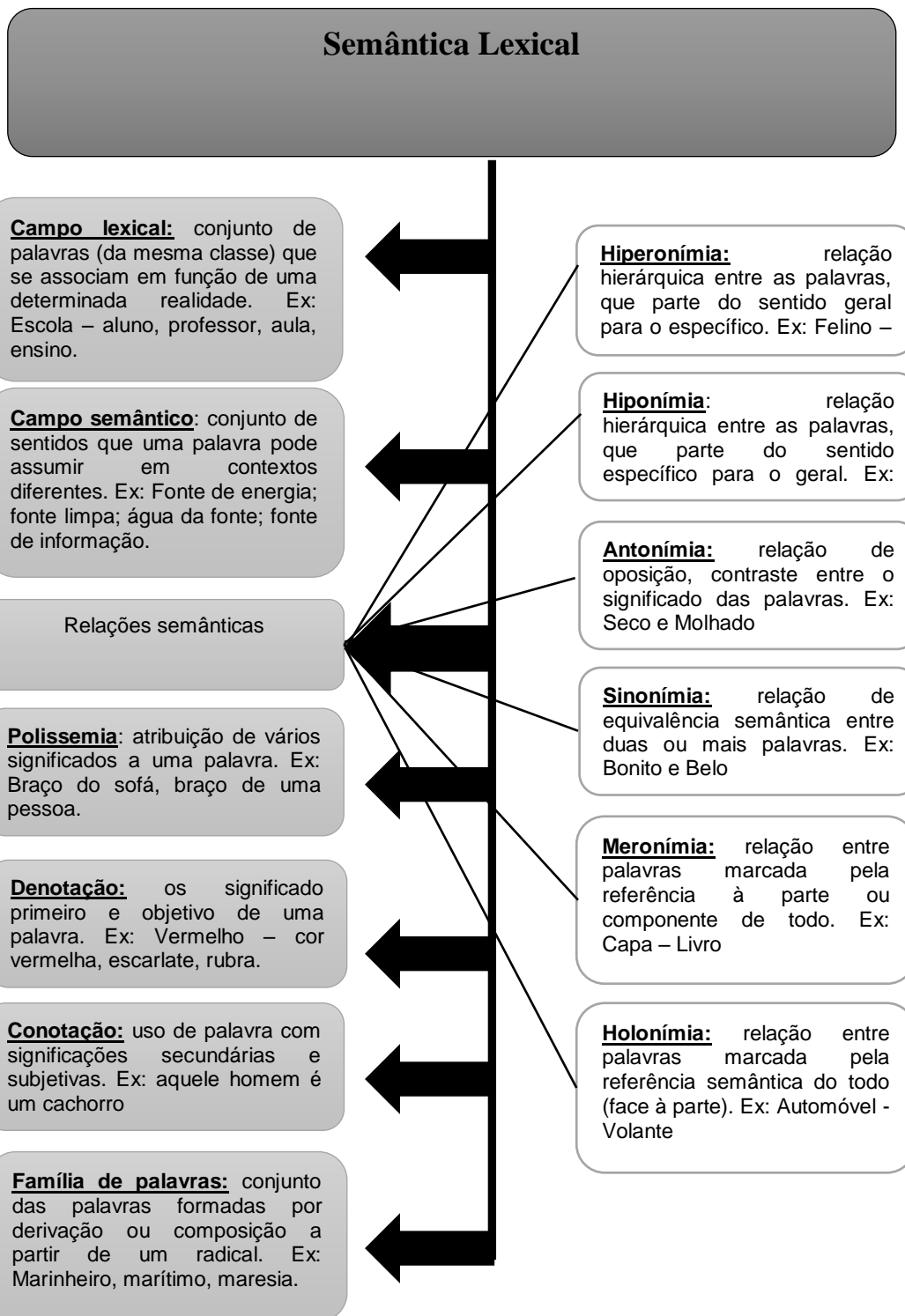


Figura 23 - Semântica lexical - componentes

Fonte: adaptada de Moreira e Pimenta (2012).

Diante disso, algumas teorias oriundas da linguística podem contribuir para a estruturação de referencial teórico e metodológico que sirva de base para a análise e representação de conceitos e seus relacionamentos.

3.2 A CONEXÃO ENTRE SEMÂNTICA E ORGANIZAÇÃO DA INFORMAÇÃO

No âmbito da Ciência da Informação, a representação está relacionada com as formas de simbolizar a informação e o conhecimento. Por vezes, muitos conceitos são utilizados indiscriminadamente e de forma imprecisa. A ocorrência dos termos organização e representação da informação e do conhecimento associam-se muito aos processos de organização, utilizando-se de representações das informações e dos conhecimentos cujos elementos, para serem compreendidos, necessitam da compreensão de outros conceitos, como a do que é a Ciência da Informação, o conceito de informação e conhecimento (Lima & Alvares, 2012). Entretanto, não obstante a importante discussão acerca destes conceitos, nesta seção discorre-se sobre organização da informação para elucidar sua conexão com a semântica via recuperação da informação, considerando que, em um sistema de informação, a qualidade obtida na recuperação da informação depende dos procedimentos e instrumentos utilizados na organização desta, o que requer que os padrões de organização sejam definidos desde a concepção do sistema, para permitir que a informação seja recuperada (Lima & Alvarez, 2012).

Conforme Brascher (2008) aponta, o objetivo do processo de organização da informação é possibilitar o acesso ao conhecimento contido nela. Amparada em Svenonius (2000), a autora detalha os objetivos da Organização da Informação:

1. Localizar entidades⁴¹ em arquivo ou base de dados como resultado de uma busca por meio de atributos e relacionamentos entre as entidades.

⁴¹ Para a autora, as entidades são aquelas relacionadas aos objetos da descrição bibliográfica que incluem, por exemplo, a obra, o documento, a superobra, a edição, o autor e o assunto.

2. Identificar uma entidade, ou seja, confirmar que a entidade descrita em um registro corresponde à entidade desejada ou distinguir entre duas ou mais entidades com características similares.
3. Selecionar uma entidade que é apropriada às necessidades dos utilizadores.
4. Adquirir ou obter acesso à entidade descrita.
5. Navegar numa base de dados, isto é, encontrar obras relacionadas a determinada obra por meio de generalização, associação, agregação.
6. Encontrar atributos relacionados por equivalência, associação e hierarquia.

Neste sentido, percebe-se a OI como um processo que compreende a descrição física e de conteúdo dos objetos informacionais, e o produto desses processos descritivos é a representação dos objetos informacionais, entendida como um conjunto de elementos descritivos que representam os atributos de um objeto informacional específico (Brascher, 2012).

O problema da recuperação da informação, em contexto digital ou não, versa sobre recuperar itens úteis para satisfazer a necessidade de informação do utilizador e evitar itens inúteis. A estas características deu-se o nome de coeficiente de precisão (*precision*) e coeficiente de revocação (*recall*), no qual uma melhoria na revocação em geral implica perda da precisão e vice-versa. Nomeadamente, a revocação designa a capacidade de recuperar documentos úteis e a precisão a capacidade de evitar documentos inúteis (Lancaster, 1993).

A recuperação da informação tem como pontos centrais a questão da relevância e pertinência, assim como recorre-se a mais dois princípios básicos: a exaustividade e a especificidade. O primeiro refere-se à quantidade de conceitos que caracterizam o conteúdo íntegro do recurso informacional e, quanto mais descritores são atribuídos, maior a exaustividade. A especificidade refere-se à qualidade na seleção dos termos de indexação, a precisão e a profundidade dos termos que representam o conteúdo, os conceitos especificados (Lancaster, 1993; Gil Leiva, 1999).

A semântica, independente da perspectiva adotada, busca contribuir para a melhoria da qualidade da informação. Hjørland (2007) discutiu acerca dessa questão da semântica para a organização do conhecimento⁴² e, pautando-se em seus estudos, a seguir encontram-se relacionados no quadro alguns autores e as distintas perspectivas que versam sobre a conexão entre as abordagens semânticas e a CI.

Quadro 16 - Abordagens de estudos semânticos em CI

Autores	Abordagens
Bean e Green (2001)	Relações semânticas na organização do conhecimento.
Beghtol (1986)	Validade e garantia semântica; Garantia cultural, filosófica, educacional; Sistemas de classificação – sintaxe.
Blair (2003)	Recuperação da informação e Filosofia da linguagem.
Bonnevie (2001)	Teoria da informação semântica de Dretske; Teoria da comunicação matemática.
Brooks (1995, 1998)	Recuperação semântica; Distância semântica; Avaliação da relevância.
Budd (2004);	Semântica; Informação, significado e realidade/verdade.
Café & Brascher (2011)	Organização da informação; Linguística; Sistemas de organização da informação; Teoria da Valência, Gramática de Casos, Gráficos Conceituais e a Teoria da Gramática Funcional; Representação conceitual.
Chaffin & Herrmann (1984)	Similaridade semântica; Relações semânticas.
Dahlberg (1978, 1995)	Teoria do conceito.
Doerr (2001)	Tesouros; Terminologia; Interoperabilidade; Diferenças semânticas.
Evens (2009)	Léxico; Semântica relacional; Modelo relacional do léxico, Linguística computacional; Recuperação da informação; Terminologia; Tesouros.
Farradane (1980)	Relações semânticas; Teoria do conceito; Indexação relacional
Frohmann (1983)	Classificação; Bases semânticas.
Hammwohner e Kuhlen (1994)	Hipertexto.
Hedlund, Pirkola, and Kalervo (2001)	Recuperação da Informação; Linguística computacional; Processamento de linguagem natural; Morfologia; ambiguidade lexical.
Hjørland (2007)	Semântica, Pragmática; Organização da informação e do conhecimento.
Khoo & Myaeng (2002) e Khoo, Christopher S. G. Na, Jin-Cheon	Recuperação da informação; Relação semântica; Relações conceituais; Processamento de linguagem natural.
Lara (2015)	Linguagens documentárias; Terminologia; Teorias semânticas, Normalização.
Moreiro Gonzáles (2011)	Ontologia; Linguística, Vocabulário controlado, Documentação.
Simões (2011, 2008)	Tesouros, Relação semântica, Resumo, Indexação.
Song e Galardi (2001)	Relações semânticas.
Sowa (1983)	Relações conceituais, Linguagem de representação; Inteligência artificial; Linguística; Lógica; Filosofia; Processamento de

⁴² O autor adota o termo Organização do Conhecimento.

	linguagem natural.
Stokolova (1977)	Recuperação da informação; Vocabulário e sintaxe; Grau de relevância.
Storey (1993)	Modelagem semântica; Relações semânticas; Processamento de linguagem natural.
Sun, Chen, Zeng, Lu, Shi, & Ma (2004)	Recuperação da informação; Indexação Semântica Latente (LSI); Semântica latente; Motores de busca.
Svenonious (2000)	Linguagens documentárias – aspectos semânticos, sintáticos e pragmáticos.
Vickery (1997)	Organização da informação; Tesouros; Inteligência Artificial; Linguística; Estruturação semântica, Classificação facetada.

Fonte: autoria própria.

Com base no quadro, verifica-se que distintas abordagens aproximam a semântica e a organização da informação. Os contributos são de difícil sistematização, mas pode-se inferir que dialogam com a filosofia e com a psicologia cognitiva. Também verificam-se os aspectos técnico das pesquisas, assim como novas tecnologias e ferramentas semânticas, evidenciando-se a natureza interdisciplinar da CI, da Organização da Informação, associada à tecnologia da informação. No contexto mais tecnicista e metodológico, Frohmann (1983 *apud* Hjørland, 2007) demonstra que os problemas de classificação devem ser vistos como problemas relacionados a teorias semânticas. Conceitos como "cão", "gato", "baleia", "pique" e "coruja" podem ser agrupados ou classificados de diferentes maneiras. Por exemplo, um princípio de divisão separa o conjunto de acordo com as características noturnas e diurnas. Nesse caso, "gato" e "coruja" pertencem à primeira categoria, e os outros termos na segunda. Outro princípio de divisão separa mamíferos de não-mamíferos. Nesse caso, "cão", "gato" e "baleia" pertencem à primeira categoria, enquanto "pique" e "coruja" pertencem ao segundo.

A recuperação da informação pode ser melhorada com a intensificação da exploração das informações semânticas intrínsecas aos recursos informacionais, de forma a expandir a compreensão e os padrões de significado em textos, imagens e outras mídias. Também pelo desenvolvimento de novas possibilidades de marcação semântica utilizando metalinguagens, criando registro de metadados que possam ser facilmente manipulados e identificados por computadores e outros dispositivos pelos próprios utilizadores, com maior contextualização, de forma que as relações entre os conceitos e os contextos linguísticos subjacentes sejam evidentes e permitam ao utilizador

refinar os resultados por meio da definição das conexões pertinentes e exclusão das associações erradas (Souza, 2006).

Referente às relações semânticas e recuperação da informação, Hjørland (2007) enfatiza que

In information retrieval the basic functions for semantic relations may be conceived as contributing to the increase of recall and precision. The inclusion of synonyms and broader terms in a query may, for example, contribute to increased recall. The differentiation of homonyms and the specification of terms may increase precision. In this way the widely use of standard semantic relations used in thesauri may be explained functionally (Hjørland, 2007, p. 20).

Nesta mesma perspectiva, Campos (2009) explica que o papel das relações semânticas retoma os princípios da indexação e expõe a seguinte situação: dois conceitos, A e B, são os descritores selecionados através de uma ferramenta (vocabulário controlado, por exemplo) para representar o conteúdo de um recurso informacional. Sem especificar a relação semântica entre os dois conceitos, falsas associações entre os conceitos poderão ser interpretadas (pelo utilizador no momento da busca e pelo sistema ao processar os resultados da busca) e resultarem em prejuízo para a revocação e precisão. A probabilidade de ocorrência de falsas associações desse tipo aumenta consideravelmente com a exaustividade da indexação. No entanto, mais que evitar falsas associações, a especificação das relações entre os conceitos representativos de um assunto está no fato que, se o assunto A é relacionado ao assunto B, então, na busca, o utilizador deve ser informado sobre isso, independente do sistema.

Pelo exposto, evidencia-se que a semântica oferece subsídios para o desenvolvimento teórico e metodológico para a organização e representação da informação. Fez-se uma breve discussão sobre como as potencialidades intratextuais da linguagem são abordadas na recuperação da informação podem ser exploradas. A seguir, discorre-se acerca das relações semânticas entre os conceitos.

3.3 RELAÇÕES SEMÂNTICAS: CONCEITO E PROPRIEDADES

As relações semânticas são o elemento essencial na composição da estrutura de uma área do conhecimento, haja vista que aos objetos e fenômenos no mundo são atribuídos rótulos. Concomitantemente, observa-se o mundo de forma segmentada. Assim, Khoo e Na (2006) reiteram que

concepts are the building blocks of knowledge structures. We spend much of events, and processes so that the world has an understandable structure and predicatability. Our lives and work depend on the accuracy and richness of this knowledge structure and its web of relations. Relations are need for reasoning and inferencing (Khoo e Na, 2006, p.157).

Nesse excerto, convém destacar os aspectos referentes à estrutura, à previsibilidade, à precisão e à rede de relações, que implicam na inferência. Ainda, os autores expõem a importância de investigar relações semânticas e como elas são objetos de estudo nas áreas da Linguística, Psicologia, Filosofia, Computação e Ciência da Informação. Assim, Khoo e Nah (2006) apresentam a definição completa que abarca distintas abordagens.

Semantic relations are meaningful associations between two or more concepts, entities, or sets of entities. They can be viewed as directional links between the concepts/entities that participate in the relation. The concepts/entities are an integral part of the relation as a relation cannot exist by itself. Associations between concepts/entities can be categorized into different types, abstracted, conceptualized and distinguished from other associations, and can thus be assigned meaning. The meaning or type of an association can sometimes but not always be derived from the meanings of the concepts involved. Psychologists and philosophers have attempted to identify the main types of relations and their features. [...] Semantic relations can refer to relations between concepts in the mind (called conceptual relations), or relations between words (lexical relations) or text segments. [...] semantic relations can occur at higher levels of text-between phrases, clauses, sentences, and larger text segments, as well as between documents and sets of documents (Khoo e Nah, 2006, p. 2).

Assim, compreende-se por relação semântica, a relação entre conceitos, os quais não existem de forma independente uns dos outros, isoladamente

(Khoo & Na, 2006; Storey, 1993). Complementar a essa definição, Almeida, Teixeira, Coelho & Souza (2010) apresentam a definição de outros autores:

- Hjørland (2007) que distingue relações semânticas entre conceitos e relações lexicais entre palavras.
- Rector e Yunes (1975) que consideram as relações semânticas como relações associativas que ocorrem em um mesmo campo semântico, ou seja, dentro de um contexto onde os termos têm sentido e função.

Também, Green (2001, p.3 *apud* Khoo & Na, 2006) afirma que

relationship are involved as we combine simple entities to form more complex entities, as we compare entities, as we group entities, as one entity performs a process on another entity, and so forth. Indeed many things that we might initially regard as basic and elemental are revealed upon further examination to involve internal structure, or in other words, internal relationships.

Gonçalves e Souza (2013) apontam, na perspectiva da indexação relacional abordada por Farradane (1980), que o que identifica uma relação entre dois conceitos é a existência de algum significado entre eles na mente, convertido diretamente em uma notação de indexação. Assim, os conceitos podem se diferenciar caso a caso, mas a estrutura de relações é limitada e não varia. Farradane (1980, p. 268 *apud* Gonçalves e Souza, 2013) representa a dinâmica das relações ao representar as relações como um processo, classificadas pelas facetas espaço e tempo, assim como apresenta dois mecanismos principais para interconectar os conceitos: discriminação e associação, em que a última é formada na memória, com repetição da experiência. Se a repetição não ocorre quando é esperada, reconhece-se que sua ocorrência é ocasional, temporária, como, por exemplo, as relações de equivalência, pertença, associativa, causa e efeito.

Khoo e Na (2006), fundamentados em Chaffin e Heernann (1988) e Chaffin (1992), ao discutirem se as relações semânticas são conceitos, enfatizam que elas possuem as principais características dos conceitos, destacando quatro aspectos, a saber:

1. As relações podem ser analisadas em elementos ou características mais básicas;
2. Uma nova relação pode ser uma combinação de outras relações;
3. As relações têm estrutura graduada (algumas instâncias da relação, representada por pares de palavras, são mais típicas de uma relação particular do que outras);
4. As relações variam na facilidade com que podem ser expressas.

Dessa forma, as relações semânticas podem ser analisadas por várias perspectivas, seja textual ou lógica. Considerando a perspectiva da Semântica Lexical, baseada na concepção dos linguistas Murphy, Khoo e Nah (2006), que identificam as propriedades gerais das relações semânticas, apresentadas no quadro, a seguir:

QUADRO 17 - PROPRIEDADE GERAIS DAS RELAÇÕES SEMÂNTICAS

PROPRIEDADE DAS RELAÇÕES SEMÂNTICAS	
Produtividade	Novas relações podem ser criadas facilmente.
Binaridade	Algumas relações, por exemplo a antonímia, são binárias no sentido de que uma palavra pode ter apenas um verdadeiro antônimo, enquanto outras relações, por exemplo a sinonímia, podem relacionar um conjunto de palavras (isto é, uma palavra pode ter muitos sinônimos).
Variabilidade	As relações entre as palavras variam com o sentido da palavra usada e o contexto.
Prototípica e canônica	Alguns pares de palavras são melhores exemplos de uma relação do que outros, e alguns pares de palavras têm status especial como exemplos canônicos de uma relação (particularmente para antônimos).
Semi-semânticidade	Propriedades não-emocionais, como categoria gramatical, co-ocorrência no texto e semelhança na morfologia.
Incontabilidade	As relações semânticas são uma classe aberta e não podem todas ser listadas ou contadas.
Previsibilidade	As relações semânticas seguem certas regras e padrões gerais.
Universalidade	Os mesmos tipos de relações semânticas são usados em qualquer linguagem e os mesmos conceitos estão relacionados pelas mesmas relações semânticas em diferentes línguas.

Fonte: Khoo e Na (2006).

Complementar a essas propriedades, do ponto de vista da lógica, considerando o desenvolvimento automático ou manual de sistemas de organização da informação, as relações semânticas apresentam as seguintes propriedades lógicas: reflexividade, simetria, transitividade.

Mendonça (2015, p.156) esclarece:

1. Reflexividade: uma relação é reflexiva se as entidades relacionadas são idênticas e a relação é válida entre elas. Exemplo: *part_of* - cada objeto é parte de si mesmo ($x \text{ part_of } x$), por exemplo, o corpo de John é a própria parte de seu corpo (*John's body part_of John's body*).

2. Transitividade: uma relação é transitiva quando é possível inferir essa mesma relação em uma cadeia de entidades relacionadas. Exemplo: *part_of* - se temos três objetos x , y e z tal que $x \text{ part_of } y$ e $y \text{ part_of } z$ então x é também *part_of* z . Assim, por exemplo, se seu ventrículo direito é parte do seu coração e se seu coração é parte de seu corpo humano, então seu ventrículo direito é parte de seu corpo humano, pela propriedade transitiva.

3. Simetria: uma relação é simétrica quando ela é verdadeira entre as entidades relacionadas nos dois sentidos (direita-esquerda, esquerda-direita). Exemplo: *part_of* é uma relação assimétrica: se temos dois objetos x e y , onde $x \text{ part_of } y$ então y não pode ser parte de x simultaneamente (*not y part_of x*). Assim, por exemplo, o tórax de João é uma subdivisão do seu tronco, mas o tronco de João não é uma subdivisão do seu tórax.

Para fins de elucidação, segue a figura 25, adaptada de Stock (2010) e Storey (1993) sobre as propriedades lógicas das relações semânticas.

Quadro 18 - Propriedades lógicas

R	Reflexivity	$x \rho x$ “...is identical to...”	<ul style="list-style-type: none"> • Reflexivity: A relation R is reflexive if it can relate an entity to itself; $[x] \rightarrow (R) \rightarrow [x]$ is true for every x (e.g., the part-whole relation).
	Irreflexivity	$\neg(x \rho x)$ “...is the cause of...”	
S	Symmetry	$(x \rho y) \rightarrow (y \rho x)$ “...is equal to...”	<ul style="list-style-type: none"> • Symmetry: A relation R is symmetric if the two participants of the relation can occupy either slot; $[x] \rightarrow (R) \rightarrow [y]$ implies $[y] \rightarrow (R) \rightarrow [x]$ (e.g., synonymy).
	Asymmetry	$(x \rho y) \rightarrow \neg(y \rho x)$ “...is unhappily in love with...”	
T	Transitivity	$[(x \rho y) \ \& \ (y \rho z)] \rightarrow (x \rho z)$ “...is greater than...”	<ul style="list-style-type: none"> • Transitivity: A relation R is transitive if $[x] \rightarrow (R) \rightarrow [y]$ and $[y] \rightarrow (R) \rightarrow [z]$ implies $[x] \rightarrow (R) \rightarrow [z]$ (e.g., IS-A relation, and ancestor-descendent relation). • One-to-one relation: A relation R is one-to-one if, when one participant of the relation is known, the other participant is fixed; $[x] \rightarrow (R) \rightarrow [y]$ and $[z] \rightarrow (R) \rightarrow [y]$ implies $x = z$.
	Intransitivity	$[(x \rho y) \ \& \ (y \rho z)] \rightarrow \neg(x \rho z)$ “...is similar to...”	

Fonte: adaptada de Stock (2010) e Storey (1993).

Considerando que x , y e z são conceitos e ρ uma relação, a reflexividade em sistemas conceituais refere-se a como um conceito adere a si mesmo em uma relação. A simetria ocorre quando uma relação entre A e B também existe na direção oposta, entre B e A . Se existe uma relação entre dois conceitos A e B e, também, entre B e C e, novamente, entre A e C , chama-se de transitividade (Stock, 2010), conforme apresentam os quadros.

Essas propriedades têm implicações significativas na recuperação da informação. A reflexividade explica que uma relação R é uma relação reflexiva se a entidade se relaciona com ela própria, como, por exemplo, a relação parte-todo. A simetria expressa que a relação R é simétrica se dois participantes da relação (conceitos) podem ocupar qualquer *slot*, ou seja, quando a relação entre o conceito X e Y é a mesma entre Y e X , como, por exemplo, a relação de equivalência, sinônimo. Logo, a transitividade, importante aliada em abordagens probabilísticas, aponta que a relação R é transitiva se X está relacionado a Y e Y está relacionado a Z , de forma que X esteja relacionado a Z , como, por exemplo, nas relações hierárquicas, especificamente gênero/espécie ou hiperonímias.

Complementarmente, apresentam-se a seguir as propriedades lógicas no escopo das relações paradigmáticas.

Quadro 19 - Propriedades lógicas nas relações paradigmáticas

Reflexivity, symmetry, and transitivity of paradigmatic relations.			
	Reflexivity	Symmetry	Transitivity
Equivalence			
- Synonymy	R	S	T
- Gen-Identity	-R	-S	T
- Antonymy	-R	S	-T
Hierarchy			
- Hyponymy			
- simple Hyponymy	-R	-S	T
- Taxonomy	-R	-S	T
- Meronymy (unspecific)	-R	-S	?
- specific Meronymies	-R	-S	T
- Instance	R	-S	-T
Specific Relations			
- "see also"	-R	S	-T
- further Relations	Depending on the relation		

Pelo exposto, as relações semânticas e suas propriedades desenvolvem um papel fundamental na compreensão humana e de raciocínio, o que tem implicações diretas no desenvolvimento de sistemas de organização e de recuperação da informação. Para finalizar e ilustrar estas implicações, apresenta-se no quadro a qual discorre sobre situações reais, relações semânticas e a estrutura de um banco de dados (Storey, 1993).

Quadro 20 - Relações semânticas e situações reais

Real world situation	Semantic relationship	Database design implications
Classes of things	Class inclusion (<i>is-a</i> relationships, hierarchies, lattices)	Inheritance from supertype to subtype
Something along with its components, parts	Meronymic or part-whole relationships	Partial inheritance for some relationships; distinguished by: function, separable, homogeneous, contemporary
One thing included or surrounded by another	Spatial inclusion	Maximum cardinality of included entity type is 1
One thing possesses or owns another	<i>Possessor owns possesses/has</i> <i>Possession</i>	If possession definitional for possessor, then minimum cardinality of possessor is 1
One thing connected to another	Attachment	When thing attached is also a part, this is equivalent to aggregation or component/object relationship
A property or characteristic of something	Attribute	Both entity types and relationships can have attributes
Typical, opposite, and similar relationships	Case, antonym, and synonym relationships	Case relationships capture real-world situations; synonyms should be removed; antonym sets should be checked to determine if both are needed

Dessa maneira, evidencia-se como a variedade de relações semânticas impõe como objetos centrais nas tarefas cognitivas, para compreensão de discursos, inferências e resolução de problemas. Tendo em vista o conceito de relações semânticas e suas propriedades, evidencia-se também que há uma diversidade de tipos de relação cujo nome varia de acordo com a perspectiva, seja pela lógica (por exemplo, relação de inclusão, causa, possessão), seja pela semântica lexical (por exemplo, hiperonímia, meronímia), que coocorrem, conforme apresentado em alguns quadros e figuras.

No âmbito da indexação, Foskett (1996), acerca da função das relações semânticas, esclarece que a inclusão de uma rede de relacionamentos semânticos em um sistema de indexação permite melhorar a formulação de indexação ou pesquisa, sugerindo um termo substituto ou adicional que possa ser usado. Na fase de indexação, tenta-se prever formas em que os utilizadores desejarem encontrar um documento; no estágio de busca, tenta-se alcançar uma melhor combinação entre a formulação de uma consulta e os termos usados por indexadores ou autores.

Semantic relationships have therefore the improvement of recall as their primary function, but this may be at the expense of relevance. If we have to substitute another term for the one we first thought of, it may be an inferior substitute, though not necessarily; the term we first thought of may itself have been wide of the mark. It is often difficult to find exactly the right term to represent a concept, especially when it is remembered that the searcher is starting from a position of relative ignorance. If however we are certain of the right term, then even a synonym may be less acceptable; it may be argued that in the case of terms having connotative meanings there can be no exact synonyms. The merging of synonyms and other equivalence relationships, are both primarily means of ensuring that more documents are retrieved in response to a query, i.e. they are devices to improve recall. On the other hand, they may be used as devices improve relevance by enabling indexers or searchers to arrive at the optimum selection of terms to use in a given situation (Foskett, 1996, p.94).

3.4 TIPOLOGIA DAS RELAÇÕES SEMÂNTICAS

Antes de discorrer acerca dos tipos de relações semânticas, convém apresentar brevemente as normas ISO 25964 e a ANSI-NISO Z39:10-2005

utilizadas para na elaboração desta seção, pois, uma vez que os sistemas de organização da informação referentes a instrumentos que promovem a organização de estruturas semânticas de conhecimento e caracterizados pelos sistemas de classificação, listas de cabeçalhos de assunto, taxonomias, tesouros, ontologias, entre outros, exigem normas e metodologias adequadas de construção e atualização para finalidades e aplicações específicas.

Muitas vezes, a alta revocação, princípio abordado anteriormente, torna-se uma realidade quando se depara com um universo documental de informações acerca dos sistemas de organização do conhecimento, exemplificados pelos artigos científicos, livros, sites entre outras fontes que tratam sobre conceitos, estruturas, formatos e a contribuição que cada um deles traz na construção de outros sistemas para a recuperação em ambientes tradicionais e digitais (Bocato, 2011).

A relevância das normas decorre não apenas por causa da questão da interoperabilidade entre vocabulários controlados e da aplicação da análise facetada, mas também por causa da construção de vocabulários controlados multilíngues, com destaque para os tesouros.

A ISO 25964 - *Thesauri and interoperability with other vocabularies*⁴³ é composta por duas partes: *I - Thesauri for information retrieval* e *II Interoperability with other vocabularies* e seu objetivo específico no tocante ao tesouro é orientar a seleção dos termos utilizados na indexação e nas consultas. A primeira parte foi publicada em 2011, substitui a ISO 2788 e ISO 5964⁴⁴ e contempla aspectos voltados ao desenvolvimento de um tesouros monolíngue ou multilíngue, assim como, no âmbito da interoperabilidade das redes, inclui um modelo de dados e um esquema XML. É aplicável a recuperação da informação, considerando os tipos de recurso de informação,

⁴³ Disponível em: <<http://www.niso.org/schemas/iso25964/>>. Acesso em: 24 mai 2017.

⁴⁴ ISO 2788 - Guidelines for the establishment and development of monolingual thesauri - publicada em 1974 e atualizada em 1986, foi o primeiro padrão internacional para desenvolvimento de tesouros. Em 1985, foi complementada pela ISO 5964 - Guidelines for the establishment and development of multilingual thesauri. As normas podem ser obtidas em: <https://www.iso.org/obp/ui/>

independente dos meios utilizados (texto, som, imagem, multimídia), incluindo bases de conhecimento, portais, museus, bases de dados bibliográficos.

Publicada em 2013, a segunda parte acolhe outros aspectos, cujo objetivo principal é aprimorar a qualidade da recuperação da informação dos recursos em rede considerando distintos vocabulários. Descreve, compara e contrasta características específicas de determinado vocabulário que são requeridas e relevantes para a interoperabilidade. Dessa maneira, ela oferece recomendações para o estabelecimento, manutenção de mapeamento entre vários tesouros, considerado os conceitos e suas várias formas de uso.

Complementarmente, a ISO Z39.19 (2010) é padrão que apresenta diretrizes e convenções para a exibição de conteúdo, construção, testes, manutenção e um gerenciamento de vocabulários controlados e monolíngues. Essa norma centra-se em vocabulários controlados que são utilizados para a representação de conteúdos e objetos em sistemas de organização do conhecimento, incluindo listas, anéis de sinônimo, taxonomias e tesouros. Essa norma deve ser considerada como um conjunto de recomendações baseadas em técnicas e procedimentos preferidos. Processos opcionais são, no entanto, por vezes, descritos, por exemplo, para a apresentação de termos num vocabulário controlado. O objetivo principal do controle de vocabulário é conseguir consistência na descrição de objetos de conteúdo e facilitar a recuperação. O controle do vocabulário é realizado por três métodos principais: definir o escopo, ou significado, dos termos, utilizando a relação de equivalência para ligar termos sinônimos e quase sinônimos, distinguindo homógrafos. Neste sentido, adaptada de Bocatto (2011) apresenta-se no quadro a seguir uma simples comparação entre elas com algumas ocorrências.

Quadro 21 - Ocorrências significativas nas normas ISO 25964 e ANSI/NISO Z39.19-2005

OCORRÊNCIAS	ISO 25964	ANSI/NISO Z39:19-2005
Título	Information and documentation – Thesauri and interoperability with other vocabularies	Guidelines for the construction, format, and management of monolingual controlled vocabularies
Estrutura	Dois partes Parte 1 - Thesauri for information retrieval Parte 2 - Interoperability with other	Uma parte - trata da aplicação da padronização, controle de vocabulário monolíngue, tipos dos termos, composição,

	vocabularies	interoperabilidade, manutenção, teste e avaliação.
Vocabulários controlados	Sistemas de classificação, sistemas de classificação de negócios para gerenciamento de registros ⁴⁵ , listas de cabeçalhos de assunto, taxonomias, mapas conceituais, redes semânticas, terminologia, lista de autoridades, anéis de sinônimos.	Uma lista de termos que foram enumerados explicitamente. Esta lista é controlada e está disponível a partir de autoridade de registro de vocabulário controlado. Todos os termos em um vocabulário controlado devem ter uma definição inequívoca, não redundante. Aplica-se a tesouros e taxonomias, ontologias e redes semânticas.
Interoperabilidade	Capacidade de dois ou mais sistemas ou componentes para trocar informações e usar a informação que foi trocada ⁴⁶ .	A capacidade de dois ou mais sistemas ou componentes para trocar informações e usar a informação trocada sem esforço especial por parte de qualquer sistema. ⁴⁷
Relações semânticas	Relações paradigmáticas e relações sintagmáticas.	Equivalência, hierárquica e Associativa.

Fonte: ISO 25964; ANSI/NISO Z39.19-2005.

Destaca-se como pontos convergentes a interoperabilidade entre os vocabulários controlados, a aplicação da análise facetada na construção dos tesouros, os vocabulários controlados multilíngues e a definição de protocolos e formatos necessário para o intercâmbio de dados (Boccatto, 2011). Outra observação importante é a denominação acerca das relações semânticas no próprio âmbito da representação da informação: relações paradigmáticas e sintagmáticas, assim como, de forma mais precisa, as relações de equivalência, hierárquicas e associativas.

Portanto, baseando-nos nas normas e em outras pesquisas entende-se que, assim como o conceito de semântica assume diversas perspectivas, o mesmo ocorre com o conceito e tipologia das relações semânticas, comumente divididas em paradigmáticas ou sintagmáticas (Storey, 1993; Stock, 2010; Peter, 2009). A norma ISO 25964 aponta que

⁴⁵ Business classification schemes for records management: conjunto de termos organizados hierarquicamente, a partir da superordenação e subordinação existente entre si, para tratar, classificar e controlar informações geradas por uma corporação para sua recuperação, manutenção, eliminação de registros, entre outras informações corporativas (Bocatto, 2011, p. 178)

⁴⁶ Ability of two or more systems or components to exchange information and to use the information that has been exchanged NOTE Vocabularies can support interoperability by including relations to other vocabularies, by presenting data in standard formats and by using systems that support common computer protocols (ISO 25964, p.6).

⁴⁷ Interoperability: the ability of two or more systems or components to exchange information and use the exchanged information without special effort on the part of either system.

- Os relacionamentos que existem apenas porque os conceitos ocorrem juntos no contexto de um documento particular são chamados de relações sintagmáticas. Quando dois ou mais termos de indexação são atribuídos ao mesmo documento, este ato efetivamente registra uma relação sintagmática entre eles. Em um tesouro, no entanto, a inclusão de relações sintagmáticas não é recomendada.
- Relacionamentos que são válidos em quase todos os contextos, especialmente quando são inerentes aos conceitos que os termos representam, são conhecidos como relações paradigmáticas. É útil mostrar relacionamentos entre os conceitos do dicionário de sinônimos, uma vez que irão frequentemente orientar os utilizadores para conceitos estreitamente relacionados para os termos que eles pensaram primeiro

As referidas relações são esclarecidas na figura abaixo.

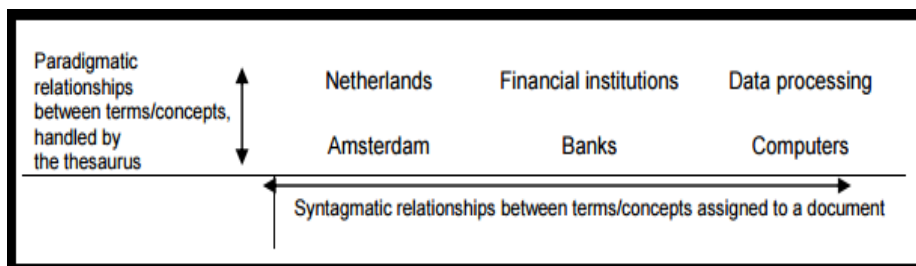


Figura 24 - Diferença entre relação sintagmática e paradigmática
 Fonte: ISO 25964:2011

As relações semânticas são diversas e sua denominação varia de acordo com a perspectiva ou objetivos pretendidos, entretanto não se apresentam divergência. Para elucidação, apresenta-se a figura 25 composta por diferentes nomenclaturas, de acordo com teóricos. Na figura, fica evidente que a proposta de Storey (1993) são denominadas na perspectiva da lógica matemática, enquanto as denominações usadas por Stock (2010) se aproximam das mais usuais em CI, as quais serão discutidas em 3.4.1, 3.4.2 e 3.4.3.

Além disso, explicar-se-á brevemente acerca das relações de inclusão, posse, anexo (*attachment*), atribuição, antônimos, sinônimos e causa.

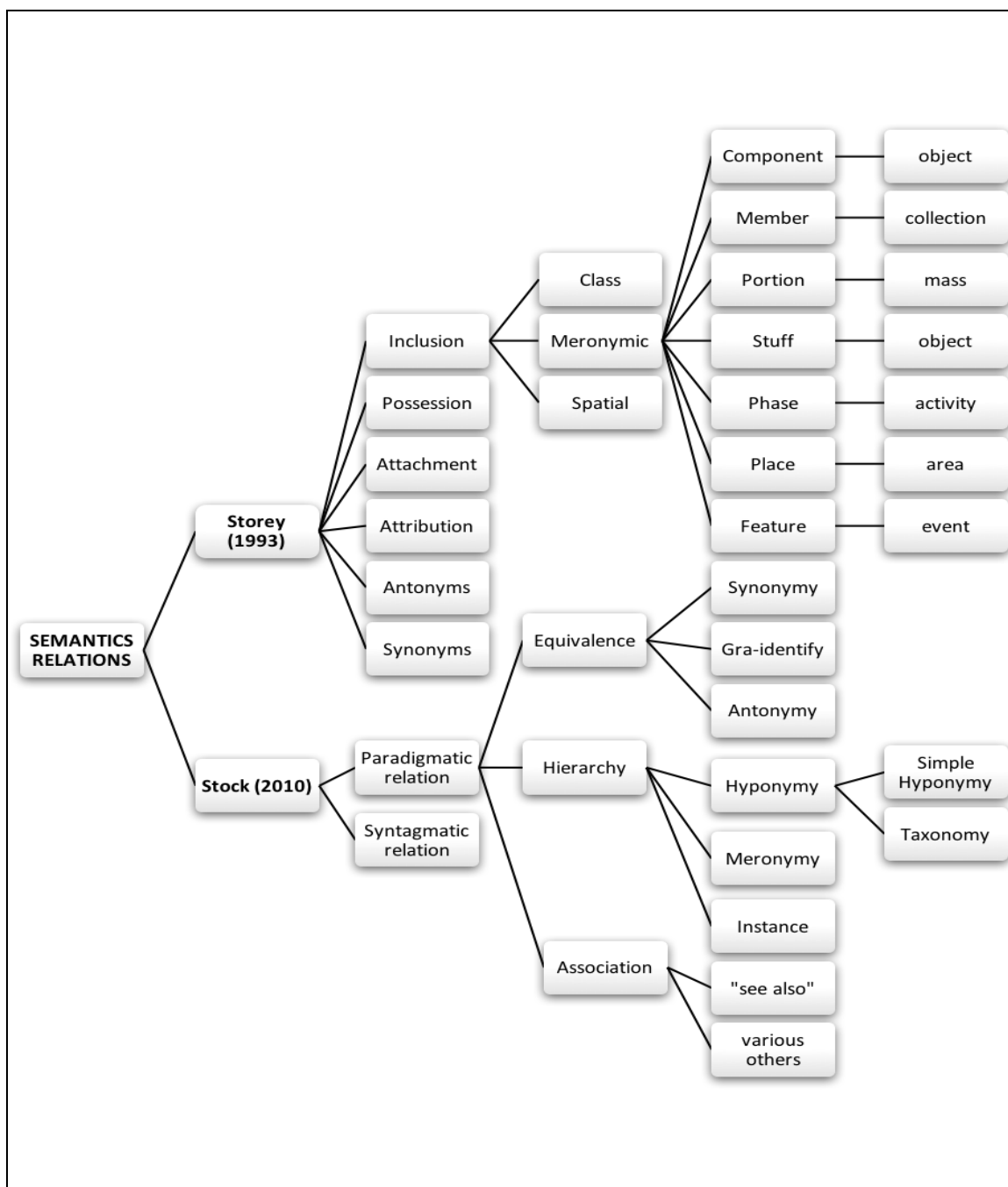


Figura 25 - Relações semânticas e distintas denominações

Fonte: Storey (1993) e Stock (2010)

- **Inclusão:** A relação de inclusão descreve situações em que um tipo de entidade compreende ou contém outros tipos de entidade. Ela compreende a inclusão de classes, meronímia e espacial. A primeira é uma hierarquia, um relacionamento de subtipo/supertipo, como, por exemplo: um carro é um veículo; rosa são flores. Nesse caso, a generalização hierarquizada ocorre quando um tipo de entidade é a união de subtipos não sobrepostos. As relações meronímicas ocorrem entre

algo e suas partes e as relações espaciais descrevem situações em que um objeto é cercado por outro, mas não é parte do que o rodeia, por exemplo: um cliente está em um território.

- **Posse:** trata-se de uma relação de propriedade. Dois casos podem ser apresentados:

Caso 1 – Uma pessoa possui licença de motorista.

Caso 2 – O motorista possui carteira de motorista.

No primeiro caso, a posse de uma licença de motorista é opcional para a pessoa. No segundo caso, o relacionamento é, em princípio, obrigatório, uma vez que um motorista deve ter uma carteira de motorista. Em ambos os casos a licença de motorista identifica exclusivamente o possuidor, entretanto, no primeiro caso a relação é opcional.

- **Anexo:** nesta relação, um tipo de entidade está conectada ou vinculada a outro. Por exemplo: um modem está anexado a um computador.
- **Atribuição:** trata-se da relação entre um objeto e seu atributo como, por exemplo: um carro tem cor. O atributo pode estar associado a um tipo de entidade ou a um relacionamento.
- **Antônimos:** podem ocorrer em atributos, entidades ou relacionamentos. Por exemplo, o atributo "status de trabalho" pode ter apenas os valores "tempo integral" ou "tempo parcial". Dois tipos de entidades serão mutuamente exclusivos se forem antônimos.
- **Sinônimos:** assim como os antônimos, sinônimos (ou quase sinônimos) podem ocorrer entre atributos, entidades ou relacionamentos. Por exemplo, os atributos "trabalho" e "função" do empregado podem ser sinônimos. Os tipos de entidade empregado e trabalhador são sinônimos.
- **Causa:** Este tipo de relacionamento semântico descreve coisas que são tipicamente verdadeiras no mundo real, por exemplo, coisas que os agentes usam ou atividades que realizam. Podem envolver instrumentos, agentes, objetos, ações.

3.4.1 RELAÇÕES DE EQUIVALÊNCIA

Quando o mesmo conceito pode ser expresso por dois ou mais termos, um deles é selecionado como o termo preferido. A relação entre termos preferidos e não preferidos é uma relação de equivalência na qual cada termo é considerado como referindo-se ao mesmo conceito. O termo preferido em efeito substitui outros termos que expressam conceitos equivalentes ou quase equivalentes (ISO 25964-2). Um bom sistema de organização da informação, sobretudo vocabulário controlado e ontologias, deve incluir termos que representam diferentes formas de fala e várias línguas, quando necessário (Harpring, 2016).

Neste mesmo seguimento, Simões (2008) esclarece que este tipo de relação caracteriza-se por estabelecer ligações entre descritores e não descritores, entre um termo preferido e um preterido. Assim, destaca que, na linguagem natural, um e outro representam o mesmo conceito, sendo equivalentes. Portanto, um determinado conceito, independente das formas que pode assumir na linguagem corrente, será representado em um sistema de organização da informação por um termo. Neste sentido, para além dos tesouros, a representação de um dado conceito sempre sob a mesma forma, origina uma consistência que, por um lado, assegura a coerência e a uniformidade na representação e posterior recuperação de recurso informacional de uma determinada área (Simões, 2008, p. 114).

No tocante às funções, estas relações controlam a sinonímia. A ISO Z39 19-2005 (R2010) aponta que a relação de equivalência deve cobrir cinco tipos básicos, a saber: sinônimos, variação lexical, quase sinônimos, *generic posting*⁴⁸ e *cross-reference*⁴⁹. Entretanto, neste texto, dedicar-se-á apenas aos sinônimos, variação lexical e quase sinônimos.

⁴⁸ *Generic posting* é uma técnica na qual o nome de classe e os nomes de seus membros são tratados como equivalentes, com o nome de classe mais amplo funcionando como o termo preferido. Por exemplo: Ceras X Ceras de plantas; Móvel X Cadeiras, Mesas, Camas. Esta técnica deve ser limitada às áreas periféricas de um campo de assunto ou usada quando o número de descritores de uma classe não garante que ela seja dividida em subclasses (ISO Z39 19-2005).

⁴⁹ *Cross-reference*: pode ser feita a partir de um termo composto para seus componentes nos casos em que a pós-coordenação dos componentes do termo individual é usada para

3.4.1.1 SINÔNIMOS

Sinônimos são termos cujos significados são considerados o mesmo ou quase o mesmo em uma grande variedade de contextos. Sinônimos verdadeiros são raros em linguagem natural. Embora os termos sejam intercambiáveis em muitas circunstâncias, a utilização pode variar como um resultado de fatores (ANSI/NISO Z39.19-2005).

Em conformidade, Simões (2008) também expõe que há duas categorias para sinônimos: linguísticos e documentais. O primeiro refere-se aos termos que tem uma tradução direta para um descritor, expressando exatamente a mesma noção, ou seja, são dois ou mais termos que expressam o mesmo conceito e que são traduzidos por um mesmo descritor. Já os sinônimos documentais referem-se a termos que têm significado diferente na língua corrente, mas por questão de redução do vocabulário são considerados sinônimos.

Assim, os exemplos relacionados no quadro são baseados nas normas ISO 25964 e apenas ilustram as diversas classes de sinônimos que podem ser encontradas na prática.

Quadro 22 - Classes de sinônimos na norma ISO 25964

Variedade de sinônimos	Exemplos
Sinônimos de origem linguística diferentes	Gatos/felinos Sódico/sódio Suor/transpiração
Sinônimos de nomes populares <i>versus</i> científicos	Aspirina/ ácido acetilsalicílico Gaivotas/ larídeos Sal/ cloreto de sódio
Sinônimos com nomes genéricos <i>versus</i> comerciais	Fotocópia/Xerox Geladeira/refrigerador
Nomes de variante para conceitos emergentes	Laptop/notebook
Termos substituídos ou em desuso	Poliomielite/paralisia infantil Países em desenvolvimento/ países subdesenvolvidos
Jargões ou Gírias	Coroa/pessoa idosa

Fonte: adaptado de Boccato (2011).

3.4.1.2 VARIAÇÃO LEXICAL

representar o conceito desejado. Deve ter cuidado para garantir que seja claro para indexadores e pesquisadores que vários termos devem ser usados. Por exemplo: Mineração de carvão X carvão e mineração; Filmes ferromagnéticos X materiais ferromagnéticos e filmes.

Segundo a ANSI/NISO Z39.19-2005 (2010), as variantes lexicais diferem dos sinônimos, na medida em que sinônimos são termos diferentes para o mesmo conceito, enquanto variantes lexicais são diferentes formas de palavras para a mesma expressão. Assim, as variantes lexicais podem resultar de diferenças ortográficas, variações gramaticais e abreviaturas, assim como termos em ordem invertida e natural, no plural e no singular.

Quadro 23 - Tipos e exemplos de variantes lexicais

TIPOS DE VARIANTES LEXICAIS	EXEMPLOS
Ordem direta versus inversa, variantes ortográficas, plurais irregulares	<ul style="list-style-type: none"> • Romênia/Roumania/Rumania • Ground water/ Ground-water/graoundwater • Peditrics/paediatrics • Rato/ Ratos • Water color/ watercolour/ watercolor/ aquarela/ aguarela/ color, water • Bordado (particípio passado de bordar) /Bordar/Bordado (Substantivo) • Laminação/Laminado • Matemática/matemático
Variantes de nome completo e abreviatura	<ul style="list-style-type: none"> • International Federation for Documentation/ FID • Polyvinyl chloride/ PVC

Fonte: ANSI/NISO Z39.19-2005 (2010).

Também, conforme consta no quadro, envolve a mesma forma gráfica (bordado) como particípio passado ou substantivo, respectivamente, técnica e processo. Similarmente, ocorre com descritores que ocorrem no particípio passado ou em forma adjetiva (Harpring, 2016).

3.4.1.3 QUASE SINÔNIMOS

Quase sinônimos ou sinônimos próximos são termos cujos significados são geralmente considerados diferentes, mas que são tratados como equivalentes para os propósitos de um vocabulário controlado. À medida em que são tratados como termos quase-sinônimos, dependem em grande medida do domínio coberto pelo vocabulário controlado e seu tamanho (ANSI/NISO Z39.19-2005). Podem incluir os antônimos, entretanto, Simões (2008, p.113) adverte que estes termos nem sempre deverão ser inseridos nessa tipologia, pois tal situação apenas fará sentido quando se encontrarem integrados como,

por exemplo, em um tesouro, nos quais representem conceitos periféricos. Em outros casos, deverão ser considerados descritores e, por isso, passíveis de se estabelecerem em relações associativas.

3.4.2 RELAÇÕES HIERÁRQUICAS

As relações hierárquicas referem-se a relacionamentos mais genéricos e mais específicos entre os registros lógicos, nos quais cada registro representa um conceito como, por exemplo, os termos genealógicos: filhos, irmãos, pai, avós. Nesses casos, existem vários tipos de relacionamentos: todo/parte, gênero/espécie e de exemplo (Harpring, 2016).

A Norma Z39.19-2005 considera que o relacionamento hierárquico é baseado em graus ou níveis de superordenação e subordinação, quando o termo superordenado representa uma classe ou um todo, e os termos subordinados se referem a seus membros ou partes. Por conseguinte, assim como as relações de equivalência, elas se caracterizam pela assimetria, na medida em que um descritor de um nível superior corresponderá a um nível inferior e vice-versa, assim como a hierarquia entre elas se estabelece por meio de relações binárias e reversíveis (Simões, 2008). Observa-se que os tipos de relações hierárquicas dividem-se em genéricas, todo/parte ou instância.

A relação genérica é a mais comum, aplicável a um grande número de assuntos e identifica a posição que os membros de uma classe ou categoria têm dentro dela. A relação parte/todo é caracterizada por identificar situações ou contextos nos quais o nome de uma parte ou espécie se encontra implícito no nome do todo ou gênero, ou seja, quando um descritor de nível inferior é uma parte particular do descritor de nível superior. E a relação de instância ou de exemplo ocorre mais frequentemente em vocabulários nos quais nomes próprios são organizados por categorias gerais de coisas ou eventos, quando, por exemplo, nomes próprios são organizados sob as categorias gerais (Simões, 2008; Harpring, 2016).

Apesar do estabelecimento das relações aparentemente não ser tão complexo como as relações associativas, Simões (2008), fundamentada em Foskett (1996), reitera a importância em determinar e obedecer critérios de ordenação coerentes e uniformes para escolha dos descritores e cita o modelo que assenta em oito níveis, cujo critério de aplicação depende da área na qual se inscreve os descritores: ordem cronológica, ordem de evolução, ordem de complexidade crescente, ordem de dimensão crescente e decrescente, ordenação espacial, ordem de preferência, ordem de tradição, ordem alfabética.

No tocante às relações poli-hierárquicas, estas se referem a quando os conceitos pertencem a mais de um contexto genérico, a mais de uma categoria, no qual cada registro existe somente uma vez no vocabulário, mas pode ser ligado a múltiplos pais e pode aparecer em múltiplas perspectivas hierárquicas. Os critérios para o seu estabelecimento devem ser sempre explícitos (Harpring, 2016; Z39.19-2005).

Ademais, conforme afirma Simões (2008), como acontece com as relações de equivalência, estas relações caracterizam-se pela assimetria, na medida em que um descritor de nível superior corresponderá a um de nível inferior e vice-versa. No tocante ao estabelecimento destas relações, ele contribui de forma indireta para o estabelecimento das demais relações, na medida que a sistematização dos termos em categorias conceituais pressupõe a divisão do tema geral em termos específicos e a respectiva análise (Simões, 2008, p. 119).

3.4.3 RELAÇÕES ASSOCIATIVAS

Conforme conceitua a ISO 25964-1 (2011):

The associative relationship covers associations between pairs of concepts that are not related hierarchically, but are semantically or conceptually associated to such an extent that the link between them needs to be made explicit in the thesaurus, on the grounds that it may suggest additional or alternative terms for use in indexing or retrieval.

E Hjørland (2007):

Associative relation: A relation which is defined psychologically: that (some) people associate concepts (A is mentally associated with B by somebody). Often are associative relations just unspecified relations. In thesauri are antonyms, for example, usually not specified but may be listed, along with terms representing other kinds of relations, under "associative relations" (Hjørland, 2007, p.30).

O primeiro conceito ressalta o vínculo entre os conceitos, para além da relação e hierárquica e o segundo complementa que a relação associativa definida na mente do sujeito algumas vezes não é especificada.

Pinho (2009) destaca que as linguagens associativas representam inovações no nível léxico como na sua organização, pretendendo que o utilizador tenha facilidade no uso do vocabulário. Considera que o princípio norteador de uma linguagem associativa é a estruturação horizontal e são nas relações associativas que se revela com mais materialidade a inclusão, entretanto, toda inclusão é complexa. Apoiado em Garcia Guitiérrez (1998, p. 2), Pinho (2009, p.106) enfatiza que

Para se formar uma Epistemografia⁵⁰ é necessário utilizar elementos da Semântica estrutural, da Linguística Documental e da Análise do Discurso operando nos seguintes níveis:

1. Na definição do aparato terminológico e conceitual, juntamente com modificações e os enquadramentos necessários.
2. Na transformação das relações hierárquicas e associativas convencionais em macroestruturas resultantes de relações binárias lógico-semânticas,
3. Na maneira de equiparar a estrutura de tesouros com outras organizações categoriais, resultando em um quadro macrocategorial,
4. Na construção de ferramentas necessárias junto ao sistema de descrição e a metodologia,
5. Na resolução de problema da redução dos elementos do vocabulário à determinada categoria léxica,
6. Na criação de macrorreglas e regras que afetem o vocabulário, a estrutura, o uso e as relações entre níveis.

⁵⁰ “[...] uma epistemografia supõe a representação formalizada, em uma base de conhecimento com fins documentários (de recuperação da informação), de um constructo cognitivo”.

Consoante ao que foi exposto, Harpring (2016) esclarece que os relacionamentos associativos existem entre registros que são conceitualmente parecidos, mas nos quais a relação não é de equivalência nem hierárquica. O tipo mais básico de relação associativa é simplesmente o "relacionado a". Em alguns vocabulários, podem ser designados tipos mais específicos de relações associativas.

Um exemplo de relação associativa que inicialmente pode causar confusão são os termos diabetes gestacional e gravidez de diabéticas. Segundo as definições apresentadas no DeCS (2017), a primeira trata-se de

Diabetes mellitus induzida por gravidez, porém resolvida no final da mesma. Não há mulheres previamente diagnosticadas como gravidez em diabéticas. A diabetes gestacional geralmente se desenvolve no final da gravidez quando os hormônios antagonistas da insulina produzindo seus níveis máximos levando à resistência à insulina, intolerância à glucose e hiperglicemia (DeCS, 2017).

Enquanto o segundo termo refere-se à

Gravidez em diabéticas: estado de gravidez em mulheres com diabetes mellitus. Isto não inclui nem o diabetes sintomático nem a intolerância à glucose induzida pela gravidez (diabetes gestacional), mas que é dissipada no final da gestação (DeCS, 2017).

Concernente a esse exemplo, Harpring (2016) esclarece que deverá haver relacionamentos associativos entre termos destinados a serem usados como conceitos separados, mas que podem ser confundidos pelos utilizadores. Ou seja, quando há necessidade de explicar a diferença entre conceitos, convém estabelecer a relação associativa.

As relações associativas podem ser estabelecidas entre registros da mesma hierarquia ou diferentes hierarquias. Pode haver relacionamentos entre irmãos sobrepostos ou outros termos, em que os significados são semelhantes e os termos são ocasionalmente (mas não geralmente) usados como sinônimos (Harpring, 2012).

Adicionalmente, Simões (2008, p. 121) fundamentada na ISO (2788:1986) reitera que as relações associativas caracterizam-se por não integrarem um mesmo conjunto de equivalências, não constituindo parte da mesma cadeia hierárquica, mas por estarem mentalmente associadas, caracterizando-se nesta perspectiva por uma analogia semântica. Todos os termos que constitui um sistema são passíveis de associação mental. Quando existe uma relação associativa entre um termo com um outro termo de grau superior hierárquico, esta é extensiva aos termos que de si dependem. Na prática, não é necessário estabelecer ligações entre todos os termos afins. Isso ocorrerá quando se observar uma sobreposição de sentido.

Termos que são mutuamente exclusivos não requerem relacionamentos associativos, particularmente quando eles não podem ser confundidos um com o outro, independentemente de compartilharem o mesmo pai ou não. Por exemplo: ligar a coluna balaústre e colunas espirais, pois não há motivo pelo qual um utilizador os confunda (Harpring, 2016).

As relações aplicam-se a dois tipos de termos, segundo Simões (2008, p.122):

- **Termos pertencentes à mesma categoria:** termos que têm o mesmo termo genérico e o significado sobrepõe-se (Exemplo: barco e navio). Neste caso, a aplicação de uma relação associativa concorre para que utilizadores e indexadores se inteirem da existência de um segundo termo, que pode ser utilizado para descritor.
- **Termos que pertencem a categorias diferentes:** são os termos que, pertencendo a tipos conceituais diferentes, estão semanticamente implicados com outros.

Além disso, os antônimos podem ser tratados como relacionamentos associativos, pois um vocabulário pode requerer um número substancial de relacionamentos associativos adicionais muito específicos. Esses tipos podem variar de um vocabulário para outro, dependendo da natureza dos termos e de como eles são destinados para o uso na recuperação. A exemplo disso são os relacionamentos entre termos genéricos que diferenciam-se de relacionamento

entre pessoas, que podem incluir relacionamentos familiares e profissionais (Harpring, 2016).

Um ponto relevante das relações associativas é que elas fornecem pistas sobre assuntos afins por meio dos termos relacionados, o que permite ao utilizador uma navegação horizontal, uma pesquisa intra ou intercategoria. Estes relacionamentos são complementares aos hierárquicos, uma vez que o utilizador, ao entrar numa categoria hierarquizada, pode realizar pesquisas genéricas e específicas. Também evitam a poli-hierarquia, na medida que pode substituir a dependência de um termo específico sob dois ou mais termos genéricos, por um termo relacionado (Simões, 2008, p.123).

O estabelecimento das relações associativas é complexo, uma vez que, em uma área de conhecimento, todos os termos possuem um tipo de relação, mas é necessário verificar e estabelecer de forma consistente a relação associativa, formulando regras e critérios para o seu estabelecimento, bem como verificar sua pertinência, pois deve-se evitar ligações desnecessárias entre conceitos relacionados.

Estas relações são sempre recíprocas e assimétricas (Simões, 2008; Harpring, 2016) e o seu estabelecimento também depende dos objetivos do tesouro, ontologia, do conhecimento de quem está construindo o sistema, das especificidades da área. A subjetividade, conseqüentemente, incide na fragilidade da sua consistência, principalmente entre as linguagens de indexação.

Simões (2008) elaborou um quadro comparativo, fundamentando-se em Van Slype, Lancaster, Chaumier, Foskett, Aitchson, Gilchrist e na norma 2788:1896. A autora observou que na realidade não há grande diversidade ou pluralidade de relações associativas. Suas conclusões apontam que

- Regra geral, observa-se uma similaridade quanto ao número e a nomenclatura dos critérios a ter em conta no estabelecimento de relações associativas e as seguintes variáveis: objeto, propriedade e causalidade.

- O mesmo se verifica em relação aos antônimos, cuja presença é observada em todos os textos.

Neste sentido, com base no quadro elaborado por Simões (2008), elaborou-se outro quadro incluindo as normas Z39.19-2005 e ISO 25964.

Quadro 24 - Relações associativas

Van Slype	Lancaster	Chaumier	Aitchson e Gilchrist	Foskett	ISO 2788:1986	Z39.19-2005	ISO 25964
<ul style="list-style-type: none"> • Relação de causalidade • Coisa e instrumentos • Relação espaço-tempo • Concomitância • Materiais constitutivos • Similaridade • Antônimos • Coisa/propriedade • Ação/processo, disciplina • Localização 	<ul style="list-style-type: none"> • Coisa/aplicação • Causa/efeito • Coisa/propriedade • Matéria/produto • Atividades complementares • Antônimos • Atividade/propriedade de • Atividade/agente dessa atividade • Atividade/produto dessa atividade • Coisas (indissociáveis) / partes 	<ul style="list-style-type: none"> • Objeto/propriedade • Objeto/objeto • Objeto/procedimento • Objeto/aplicação • Propriedade/processo 	<ul style="list-style-type: none"> • Disciplinas/objetos estudados • Operações-processo/agentes/instrumentos • Ocupações/pessoas • Ação/objeto • Partes de um todo com outro todo • Objeto/propriedade • Objetos/origens • Causa/efeito • Coisas-ações/anti-agentes • Matérias primas/produtos • Ações/ações associadas a essas ações • Antônimos 	<ul style="list-style-type: none"> • Relação de coordenação • Relação genética • Concorrente • Causa-efeito • Instrumentos • Materiais • Similaridade 	<ul style="list-style-type: none"> • Disciplinas/objetos estudados • Processo/operação/agente/instrumento • Conceitos relacionados com a sua origem • Causalidade • Coisa e seu antídoto • Ação/produto/resultado • Ação/sujeito passivo • Conceito/propriedades • Conceito e a sua unidade de medida • Locuções sintagmáticas e os nomes das subcategorias 	<ul style="list-style-type: none"> • Processo Agente • Processo/ Contra-agente • Ação/propriedade • Ação/produto • Ação/Alvo • Causa/Efeito • Conceito ou objeto/Propriedade • Conceito ou objeto/Origem • Conceito ou objeto/Unidades ou mecanismos de medida • Matéria-prima/Produto • Disciplina ou campo/Objeto ou praticante • Antônimos 	<ul style="list-style-type: none"> • disciplina ou campo de estudo e os objetos ou fenômenos estudados • operação ou processo e seu agente ou instrumento. • ação e o produto da ação • ação e seu destinatário • Objetos ou materiais e suas propriedades definidoras. • Artefato e suas partes, se eles não se qualificam para o relacionamento hierárquico inteiro • Conceitos ligados pela dependência causal • objeto ou processo e seu agente • conceito e sua unidade de medida. • Termo composto e o substantivo que é seu foco, se os dois não tiverem um verdadeiro relacionamento hierárquico. • organismo ou substância produzida ou derivada de outra.

Fonte: Fonte: Adaptado de Simões (2008), Z39.19-2005 e ISO 25964

Para finalizar esta abordagem acerca das relações semânticas, em decorrência do diálogo entre as áreas Organização da Informação e a Terminologia, contexto já abordado na seção 1. 3 deste texto, apresenta-se algumas considerações acerca das relações conceituais. A Terminologia considera que um termo se relaciona com todos os demais termos que formam parte do mesmo campo especializado, com os que constituem um campo conceitual ou campo nocional. A exemplo disso, o campo conceitual de química é o conjunto de conceitos que constituem essa disciplina, um conjunto de conceitos estruturado em subconjuntos mais restringidos que, por sua vez, também estão estruturados em outros subconjuntos mais específicos e assim sucessivamente. Conseqüentemente, o conjunto geral de conceitos de química está formado por noções referidas aos aparatos utilizados em laboratório, as substâncias com que se trabalha, os processos, operações e reações que se produzem, os componentes das substâncias, entre outros. Neste sentido, os distintos conceitos que formam parte de um mesmo campo nocional estão emparelhados entre si por dois grandes tipos de relações: as relações lógicas e as relações ontológicas (Cabré, 1993).

As relações lógicas se estabelecem entre os conceitos e se baseiam na semelhança que estes mantêm entre si, por possuir uma ou mais características comuns: essa semelhança se manifesta tanto quando se compara distintos conceitos como quando se combinam entre si. A comparação entre os conceitos leva a dois possíveis tipos de inter-relacionamentos (Cabré, 1993):

- O primeiro tipo de relação ocorre quando um conceito é mais geral que o outro, uma vez que o primeiro (genérico) cede todas suas características ao segundo (específico). O conceito específico possui traços cedidos pelo genérico, além de outros que o particularizam. Neste caso fala-se em hiponímia lógica.
- O segundo tipo de relação ocorre quando os conceitos específicos possuem as mesmas noções de conceitos de um genérico, mas

mantendo suas características particulares que os diferem um do outro. Neste caso, refere-se à coordenação lógica entre os termos dos conceitos específicos.

Cabré (1993) reitera que a combinação entre ambos os tipos de noções conduz ao conceito de campo semântico. Assim um campo hierárquico de conceitos constitui uma série vertical de noções que normalmente mantem uma relação de inclusão entre si. Em cada estado da série se encontra uma série horizontal de noções do mesmo nível, que mantem entre si relações de oposição. Como por exemplo: Mamífero – réptil, anfíbio/ Carnívoro – roedor. Também as relações hierárquicas mais importantes são as de inclusão e compreendem relações genéricas e partitivas.

Nas relações gênero-espécie os conceitos mais específicos formam parte da extensão do conceito genérico que os inclui. A relação de inclusão que apresentam os conceitos entre si se reflete na descrição de cada conceito através da explicitação de suas características. Um termo específico contém todas as características de seu genérico e alguma que é particular. Assim, nem todos os conceitos são sempre o resultado de simples combinação de características, mas em muitos casos são o produto completo da combinação de conceitos simples. Assim, deste ponto de vista sua descrição, os conceitos completos estão formados pela combinação de conjuntos estruturados de características, cada um com os quais descreve um conceito simples.

A combinação de conceitos simples integrando conceitos complexos que não apresentam entre si relações nem de hiponímia, nem de combinação lógica pode conduzir a três tipos de casos, todos baseados na transferência de significado (Cabré, 1993):

1. Da combinação de dois conceitos resulta um terceiro conceito, específico de uma das noções de partida, por exemplo: máquina – bater – máquina de bater. Neste caso fala-se em determinação de uma noção por outra.

2. Da combinação dos conceitos resulta um terceiro conceito, específico dos conceitos de partida, por exemplo: Calça pantalon. Neste caso, diz-se conjunção de noções.

3. Da combinação dos conceitos resulta um terceiro conceito caracterizado somente por traços diferenciadores das noções de origem e, portanto, não específico de nenhum dos conceitos de partida, por exemplo: indo – europeu – indoeuropeu. Neste caso fala-se em disjunção de noções (Cabré, 1993).

Quanto às relações ontológicas, na perspectiva da Terminologia, a ontologia se ocupa das relações chamadas ontológicas, que não se baseiam tanto na lógica, na semelhança entre os conceitos, mas na proximidade situacional dos elementos na realidade. Assim se distinguem dois tipos de relações ontológicas baseadas no contato dos conceitos (Cabré, 1993, p. 206):

A) As relações de coordenação (parte-todo), que descrevem todos os tipos de relações: as que estabelecem entre o todo e suas partes e as que se dão entre as distintas partes de um mesmo todo. Estas relações se baseiam na contiguidade dos objetos no espaço e, portanto, são relações que têm lugar simultaneamente.

B) As relações de encadeamento que se baseiam na sucessão dos objetos no tempo (causa e efeito) e são, portanto, sequenciais.

Por fim, acerca do estabelecimento das relações semânticas, ressalta a importância em começar pelas relações hierárquicas, uma vez que a sistematização dos termos em categorias conceituais pressupõe a divisão do tema geral em termos específicos e na respectiva análise. É neste procedimento analítico que, naturalmente, observa-se a sinonímia, a polissemia e, sobretudo, identificam-se não apenas os níveis de hierarquia entre os termos que constituirão as relações verticais, mas, também, em simultâneo, se poderão registrar os termos passíveis de constituir as relações horizontais – relações associativas, na medida em que se detecta o parentesco de homogeneidade e simetria semântica entre estes termos e outros pertencentes a categorias diversas (Simões, 2008, p. 119).

3.5 RELAÇÕES SEMÂNTICAS NO DOMÍNIO DA SAÚDE

Para além das relações hierárquicas, a seguir, apenas a título de exemplificação, apresenta-se o quadro composto por relações semânticas associativas acompanhadas dos substantivos compostos.

Quadro 25 - Relações semânticas em saúde

Relacionamentos semânticos	Substantivos compostos
Wrong parse Subtype	exhibit asthma, ten drugs, measure headache headaches migraine, fungus candida, hbv carrier, giant cell, mexico city, t1 tumour, ht1 receptor
Activity/physical process Produces (on a genetic level)	bile delivery, virus reproduction, bile drainage, headache activity, bowel function, tb transmission migraine relief, headache resolution headache induction, headache onset papilloma growth, headache transformation, disease development, tissue reinforcement polyomavirus genome, actin mrna, cmv dna, protein gene
Cause	asthma hospitalizations, aids death, automobile accident heat shock, university fatigue, food infection virus, diarrhoea virus, influenza infection
Characteristic Physical property Defect Physical Make up	receptor hypersensitivity, cell immunity, drug toxicity, gene polymorphism, drug susceptibility blood pressure, artery diameter, water solubility hormone deficiency, csf fistulas, gene mutation blood plasma, bile vomit
Person afflicted Demographic attributes	aids patient, bmt children, headache group, polio survivors childhood migraine, infant colic, women migraineur
Person/center who treats Research on Attribute of clinical study Procedure	headache specialist, headache center, diseases physicians, asthma nurse, children hospital asthma researchers, headache study, language research headache parameter, attack study, headache interview, biology analyses, biology laboratory, influenza epidemiology tumor marker, genotype diagnosis, blood culture, brain biopsy, tissue pathology
Frequency/time of Time of	headache interval, attack frequency, football season, headache phase, influenza season morning headache, hour headache, weekend migraine
Measure of Standard	relief rate, asthma mortality, asthma morbidity, cell population, hospital survival headache criteria, society standard

Instrument	aciclovir therapy, chloroquine treatment, laser irradiation, aerosol treatment, vaccine antigen, biopsy needle, medicine ginseng, heroin use, internet use, drug utilization
Object Misuse Subject	bowel transplantation, kidney transplant, drug delivery, drug abuse, acetaminophen overdose, ergotamine abuser, headache presentation, glucose metabolism, heat transfer
Purpose Topic Modal Material Bind Activator Activator Inhibitor	headache drugs, hiv medications, voice therapy, influenza treatment, polio vaccine time visualization, headache questionnaire, tobacco history, vaccination registries, health education, pharmacy database brain artery, tract calculi, liver cell, hospital beds, emergency surgery, trauma method, formaldehyde vapor, aloe gel, gelatin powder, latex glove, receptor ligand, carbohydrate ligand Activator, acetylcholine receptor, pain signals, headache trigger, headache precipitant, adrenoreceptor blockers, influenza prevention
Defect in localtion	lung abscess, artery aneurysm, brain disorder

Fonte: Adaptado de Rosario e Hearst (2001)

O quadro adaptado de Rosario e Hearst (2001) foi elaborado durante o desenvolvimento de uma pesquisa sobre sobre relações semânticas na área da saúde com substantivos compostos e no âmbito do Processamento de Linguagem Natural, Linguística Computacional. Também, utilizaram classificadores automáticos, com apoiados na perspectiva das redes neurais.

CAPÍTULO 4 - PERCURSO INVESTIGATIVO

A convergência entre a base tecnológica, a dinâmica da indústria e o crescimento da internet, advogada por Castells (1999), assim como as inovações produzidas pela digitalização da informação e pela conectividade, alteraram a natureza das relações entre os indivíduos nas organizações, modificando o modo como todos os indivíduos e profissionais da informação organizam o conhecimento (Lima, Souza & Dias, 2013). Nessa conjuntura, a folksonomia surgiu como um novo processo de representação do conhecimento na *web*, no qual observamos uma comunicação assimétrica existente entre a extensa variedade de perfis de utilizadores. Estes atribuem e compartilham conceitos e significados a partir de uma ação individual, a etiquetagem.

As investigações acerca das folksonomias devem considerar três elementos-chave: “as redes sociais, a linguagem e a colaboração” (Assis Pinto, 2011). Neste sentido, uma vez que privilegiam-se estudar as relações semânticas implícitas nas folksonomias, esta investigação possui caráter exploratório e é de natureza qualitativa.

Selltiz (1967, cit. por Campos, 2009) destaca que os estudos exploratórios podem ter outras funções que não a formulação de um problema para a investigação mais exata ou para a criação de hipóteses. Estudos desta natureza objetivam aumentar o conhecimento do pesquisador acerca do fenômeno que deseja investigar e qualificar estudos futuros. Para Gil (1999), pesquisas exploratórias são desenvolvidas com o objetivo de proporcionar uma visão geral, de tipo aproximativo, acerca de determinado fato, e que são realizadas quando o tema escolhido é pouco explorado e torna-se difícil formular hipóteses precisas e operacionalizáveis.

Nessa mesma perspectiva, Sampiere, Collado e Lúcio (2013) afirmam que estes estudos são realizados quando a revisão de literatura revelou que existem apenas orientações não pesquisadas e ideias vagamente relacionadas

com o problema de estudo, ou sobre temas e áreas a partir de novas perspectivas. Refere-se às pesquisas que pretendem analisar fenômenos desconhecidos ou novos, e servem para nos tornar familiarizados com fenômenos relativamente desconhecidos, obter informação sobre a possibilidade de realizar uma pesquisa mais completa, relacionada com um contexto particular, pesquisar novos problemas, identificar conceitos ou variáveis promissoras.

Para além, a abordagem da pesquisa é qualitativa, uma vez que utiliza uma “multiplicidade de métodos que abordam a problemática de forma naturalista e interpretativa” (Coutinho, 2011). Essa abordagem é utilizada para descobrir e aprimorar questões de investigação, no qual se examina o mundo social e, nesse processo, desenvolve-se uma teoria consistente para subsidiar a observação daquilo que acontece. As hipóteses não são testadas, mas construídas durante o processo e vão sendo aprimoradas conforme os dados são coletados, ou são resultados do estudo (Sampieri, Collado & Lúcio, 2013).

Pelo exposto, apresenta-se a seguir, as abordagens metodológicas, os instrumentos e procedimentos executados e uma breve descrição das etapas.

4.1 ABORDAGENS METODOLÓGICAS

Em decorrência do contexto complexo e multifacetado desta investigação, a metodologia será desenvolvida em etapas lógicas e sucessivas. Para tanto, utilizaremos as seguintes abordagens metodológicas.

4.1.1 ANÁLISE DE DOMÍNIO

Investigar fenômenos em uma área específica do conhecimento (domínio), comparar esquemas e modelos de representação. Por isso, optou-se pela Análise de Domínio, que é a delimitação e compreensão do conjunto de informações de um dado contexto, por meio do entendimento de padrões de comunicação e da relevância. Integra perspectivas individualistas e o contexto social das comunidades, buscando enfatizar aspectos práticos e semânticos, valorizando o papel da sociedade em diferentes grupos de pessoas e contextos. O entendimento de domínios do conhecimento, por meio do

pensamento, discurso e documentos, pode refletir o compartilhamento de conceitos entre os diferentes membros de uma comunidade.

Hjørland e Albrechtsen (1995) propõem a análise do domínio como um paradigma que visa melhorar a maneira de entender as informações:

The domain-analytic paradigm in information science (IS) states that the best way to understand information in IS is to study the knowledge-domains as thought or discourse communities, which are parts of society's division of labor. Knowledge organization, structure, cooperation patterns, language and communication forms, information systems, and relevance criteria are reflections of the objects of the work of these communities and of their role in society (Hjørland & Albrechtsen, 1995, p.400).

Assim, esta metodologia permite a construção de esquemas de representação do conhecimento a partir da investigação de determinadas características de domínios específicos do conhecimento. Lida com a contextualização dos conceitos (tags, termos e palavras-chave) de uma área de domínio, no qual o documento não deve ser interpretado como uma fonte isolada de conhecimento, mas como parte de uma área de conhecimento, uma contribuição a ela. Procura contextualizar a representação mais amplamente, considerando não só a terminologia, mas também o uso que é feito da informação produzida, ou seja, sua aplicação para a elaboração de serviços e produtos (Novellino, 1996).

A organização da informação e do conhecimento, a estrutura, cooperação, os padrões, a linguagem e formas de comunicação e critérios de relevância são reflexos dos objetos de trabalho de determinada comunidade e seu papel na sociedade. Neste sentido, o conhecimento, as necessidades de informação e a relevância são critérios que devem ser vistos nesta perspectiva. O paradigma de análise de domínio tem a concepção da Ciência da Informação como uma das ciências sociais, assim como uma abordagem funcionalista, uma vez que tenta entender as funções implícitas e explícitas da informação e comunicação e para rastrear os mecanismos subjacente à informação. Também compreende uma abordagem filosófica-realista, que tenta encontrar a base para a Ciência da Informação em fatores que são externos a

subjetividade individual do utilizador, em oposição aos paradigmas comportamentais e cognitivos (Hjørland, 2002).

De acordo com Hjørland (2002) um domínio pode ser conhecido por meio de:

1. Produção de guias de literatura.
2. Produção de classificações especiais.
3. Pesquisa na indexação e recuperação de especialidades.
4. Estudos empíricos de utilizadores.
5. Estudos bibliométricos.
6. Estudos históricos.
7. Estudos de documentos e gênero.
8. Estudos epistemológicos e críticos.
9. Estudos terminológicos, estudos discursivos.
10. Estudos em estruturas e instituições em comunicação científica.
11. Análise de domínio na cognição profissional e inteligência artificial.

4.1.2 ANÁLISE CONCEITUAL

A análise conceitual compreende a revisão dos atributos essenciais e da definição dos conceitos, sendo um processo que incorpora um exame completo dos elementos básicos que compõem um pensamento, ideia ou noção (Fernandes, Nóbrega, Garcia, Macêdo-Costa, 2011).

Geralmente utilizada para desenvolver linguagens de indexação/sistemas de organização do conhecimento, compreende um processo intelectual que, associado a outras técnicas, envolve a identificação, determinação, estruturação e validação dos conceitos e relações semânticas apropriadas para a representação do conhecimento (Campos, 2009). Trata-se da aplicação da Teoria do Conceito (Dahlberg, 1978), no qual o conceito é formado por três elementos: o referente (aquilo que se pretende conceituar), as características (soma dos enunciados verdadeiros sobre o referente) e a forma verbal – formando o triângulo do conceito. Assim, o conceito passa de um artefato mental, abstrato, para instrumentar a pesquisa e a manipulação para a

construção de sistemas de organização do conhecimento. Esta teoria fornece os seguintes elementos: a formação dos conceitos; a compreensão dos conceitos como sínteses de características; a origem das características dos conceitos; as relações conceituais; e a construção de sistemas de conceitos de acordo com critérios objetivos. Portanto, a análise conceitual apoia esta investigação, uma vez que, para compreender e explicitar os relacionamentos semânticos presentes nas folksonomias seja necessário perceber a relação existente entre as tags (conceito), considerando que um conceito isolado nada é.

Os estudos que versam sobre sistemas de organização do conhecimento, no âmbito do tratamento da informação, exigem uma delimitação temática. Sendo assim, a escolha temática “Diabetes” justifica-se pela necessidade de ampliar as investigações; pela não existência de um vocabulário controlado (ou mesmo de um dicionário terminológico); por abranger doenças de longa duração, multifatoriais, algumas com necessidade de acompanhamento constante; por estar associada à qualidade de vida; por ser considerado um problema de saúde pública; e ter, conseqüentemente, significativo impacto socioeconômico em nível mundial.

Apesar desta investigação apoiar-se na teoria do conceito advogada por Dalbergh, há outras estratégias metodológicas para análise conceitual, que apenas a título de ilustração encontram-se resumidas na tabela abaixo.

Quadro 26 - Tipos de análise conceitual

ANÁLISE CONCEITUAL	DESCRIÇÃO
<p>Análise conceitual de Walker e Avant</p>	<p>Esta análise conceitual compreende oito etapas:</p> <ol style="list-style-type: none"> 1. Seleção do conceito. 2. Determinação dos objetivos da análise conceitual. 3. Identificação dos possíveis usos do conceito. 4. Determinação dos atributos críticos ou essenciais. 5. Construção de um caso modelo⁵¹. 6. Desenvolvimento de outros casos. 7. Identificação de antecedentes e conseqüências do conceito. 8. Definição de referências empíricas para os atributos essenciais.
<p>Análise conceitual de Rodgers ou análise</p>	<p>Considera-se os conceitos dinâmicos, constitui um modo indutivo e descritivo de investigar o consenso relacionado</p>

⁵¹ Trata-se da elaboração de um exemplo baseado na vida real, do uso do conceito, que inclua atributos essenciais.

<p>evolucionista</p>	<p>ao conceito, o conhecimento histórico, assim como um instrumento para verificar área de concordância e discordância do uso do conceito entre as diversas disciplinas. O conceito adquire significado por meio do seu uso e aplicação a cada situação no tempo, influenciado por fatores internos e externos que ampliam, refinam e caracterizam adequadamente esse conceito.</p> <p>Etapas:</p> <ul style="list-style-type: none"> • Seleção do conceito de interesse. • Seleção do cenário apropriado para a coleta de dados e documentos a serem analisados. • Identificação dos atributos. • Análise do ciclo evolutivo do conceito. • Caso-modelo.
<p>Análise conceitual de Swartz-Barcott e Kim ou Modelo híbrido</p>	<p>Interliga análise teórica e observação empírica, sendo composto de três fases: fase teórica (seleção do conceito, busca e revisão de literatura, definição conceitual e operacional do conceito – instrumento de medida), fase de trabalho de campo (integração da análise teoria {as observações empíricas referentes ao fenômeno no contexto em que se manifesta), que envolve o estabelecimento do cenário do estudo, a seleção de casos, a coleta e análise de dados. Por fim, a fase analítica, que compreende o cotejamento entre os dados resultantes da fase teórica inicial e as observações empíricas, sendo o produto final a apresentação da definição do conceito, bem como a identificação de lacunas conceituais que direcionam o desenvolvimento de outras pesquisas. Esta análise não inclui a apresentação de caso que expresse um modelo do conceito.</p>

Fonte: adaptada de Fernandes, Nóbrega, Garcia, Macêdo-Costa (2011)

4.2 INSTRUMENTOS E PROCEDIMENTOS

Os instrumentos que foram utilizados na investigação associam-se aos seguintes procedimentos: coleta das indexações livres e tags, compreensão e definição de alguns conceitos (análise conceitual, análise de domínio), conseqüentemente consulta a literatura especializada, contato com o especialista e edição de ontologia.

Primeiramente, para composição do *corpus* e coleta das *tags*, foi explorado o CiteULike A escolha justifica-se por ter sido significativamente utilizado em investigações correlatas e ser considerado de fácil manipulação. Este sistema baseado em folksonomia foi criado em 2004, é um serviço gratuito para armazenar, organizar e compartilhar os trabalhos acadêmicos. A seguir, encontra-se a figura 26, que representa o resultado de busca com a

palavra-chave Diabetes e em destaque vermelho são as indexações livres e tags atribuídas pelos utilizadores.

✓ **HLA, infecções e inflamação nos estágios iniciais da aterosclerose em crianças com diabetes tipo 1.**
Acta diabetologica , vol. 55, n ° 1. (24 de janeiro de 2018), pp. 41-47
 por [Michal Odehmaršky](#), [Erkki Pesonen](#), [Timo Sorsa](#), [Åke Lernmark](#), [Pirkko J. Pussinen](#), [Petru Liuba](#)
 postado para [adolescentes](#) [aterosclerose cardiometabólica](#) [crianças](#) [cardiovasculares](#) [diabetes](#) [hla](#) [infecção imune](#) [type1](#) by [Rwebb1981](#) on 2018-10-01 07:49:32

■ Cópia ■ Abstrata ■

✓ **O papel e a função do HDL em pacientes com diabetes mellitus e o risco cardiovascular relacionado.**
Lipídios em saúde e doença , vol. 16, No. 1. (30 de outubro de 2017)
 por [Mark Fomak](#), [Anna Gluba-Bezák](#), [Aleksandra Cielkowska-Rysz](#), [Jacek Rysz](#)
 postou a [aterosclerose cardiometabólica](#) [cardiovascular](#) [diabetes](#) [HDL](#) [lipoproteína](#) [lipoproteínas](#) [type1](#) [type2](#) for [Rwebb1981](#) on 2018/10/01 07:48:00

■ Cópia ■ Abstrata ■

✓ **Karma Metabólico - O Legado Atherogenic do Diabetes: A Palestra do Prêmio Edwin Bierman 2017.**
Diabetes , vol. 67, n ° 5. (maio de 2018), pp. 785-790
 por [Mark Emmanuel E. Cooper](#), [Assam El-Osta](#), [Terri Jean J. Allen](#), [Anna Margareta Dorothea MD Watson](#), [Merlin Christopher C. Thomas](#), [Karin Agnes Maria AM. Jandeleit-Dahm](#)
 postado [aterosclerose cardiometabólica](#) [cardiovascular](#) [colesterol](#) [diabetes](#) [lipídios](#) [lipídios](#) [lipoproteínas](#) [lipoproteínas](#) [metabolismo](#) [revisão](#) [type1](#) [type2](#) by [Rwebb1981](#) em 2018-10-01 07:09:10

■ Cópia ■ Abstrata ■

Figura 26 - CiteULike - coleta das tags

Para auxiliar na compreensão dos conceitos e suas características, prioritariamente, utilizou-se como referência o vocabulário estruturado e trilingue DeCS (Descritores em Ciências da Saúde), criado pela Bireme (Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde), desenvolvido a partir do MeSH (*Medical Subject Headings da U. S. National Librar of Medicine (NLM)*). Secundariamente, quando surgiram dúvidas, foram realizadas pesquisas na PubMed. A seguir, encontra-se o resultado de pesquisa sobre o conceito obesidade no DECS.

1 / 8 **DeCS**

Descritor *Inglês*: **Obesity**

Descritor *Espanhol*: **Obesidad**

Descritor *Português*: **Obesidade**

Sinônimos *Português*: Tratamento da Obesidade

Categoria: [C18.654.726.500](#)
[C23.888.144.699.500](#)
[E01.370.600.115.100.160.120.699.500](#)
[G07.100.100.160.120.699.500](#)
[SP6.016.047](#)

Definição *Português*: Estado no qual o **PESO CORPORAL** está grosseiramente acima do peso aceitável ou ideal, geralmente devido a acúmulo excessivo de **GORDURAS** no corpo. Os padrões podem variar com a idade, sexo, fatores genéticos ou culturais. Em relação ao **ÍNDICE DE MASSA CORPORAL**, um IMC maior que 30,0 kg/m² é considerado obeso e um IMC acima de 40,0 kg/m² é considerado morbidamente obeso (**OBESIDADE MÓRBIDA**).

Nota de Indexação *Português*: [camundongos hiperglicêmicos](#) obesos se indexam com [CAMUNDONGOS OBESOS](#); não acrescenta **OBESIDADE** e nem **HIPERGLICEMIA** a menos que se discuta uma [doença](#) em particular; /terap: veja também [CONTROLE DA OBESIDADE](#)

Figura 27 - Busca por conceitos no DeCS

Por fim, para edição da ontologia, optou-se por utilizar o Protegé, desenvolvido pela Universidade de Standford e trata-se de uma plataforma livre, *open-source*, que proporciona a criação e edição de ontologias e bases de conhecimento, de forma rápida e intuitiva.

A seguir, encontram-se descritas as etapas executadas nesta investigação.

4.3 DESCRIÇÃO DAS ETAPAS

Para alcançar o objetivo geral proposto, foram realizadas sete etapas:

Etapa 1 - Coleta das indexações livres e tags: A coleta foi manual, sendo considerada também um elemento e atividade relevante no processo de assimilação dos conceitos e compreensão do domínio. Nesta coleta foram excluídos os resultados que apresentavam apenas uma tag atribuída ao recurso informacional. Também, foram eliminadas as tags referentes ao gênero textual, preposições, nome do utilizador, pronomes de tratamento, uma vez que não são pertinentes a esta investigação

Etapa 2 – Elaboração da taxonomia e identificação das tags mais expressivas: Esta etapa consistiu em compor uma lista de candidatos a descritores do domínio extraídas do CiteULike. Primeiramente, convêm fazer os seguintes esclarecimentos:

- As tags foram traduzidas para o português, com auxílio de dicionários e consulta ao DeCS.
- São considerados candidatos a descritores, e apenas candidatos, uma vez que para confeccionar e estabelecer como descritor é necessário a consulta a vários especialistas do domínio para validação e maior rigorosidade no controle de sinônimos, significado. Neste sentido, convergente a proposta da tese, em termos práticos, o especialista foi consultado para validar apenas os descritores mais expressivos e que foram utilizados para analisar as relações semânticas e compor as afirmações. O especialista consultado tem licenciatura em Ciências da

Nutrição, pós-graduação em Gestão de Unidades de Saúde e mestrado, a concluir, em Alimentação Coletiva.

- Apesar da pesquisa não se configurar como uma pesquisa quantitativa, para efeitos práticos, foram consideradas as tags mais expressivas àquelas com maior ocorrência no corpus.

Etapa 3 – Descrição e análise dos conceitos, características e atributos: esta etapa consistiu na consulta aos DeCS, dicionários e literatura especializada, considerando as definições das tags mais expressivas.

Etapa 4 – Estabelecimento das categorias: para o estabelecimento das categorias, apoiou-se na literatura e as manifestações foram com as tags, sem critérios quantitativos.

Etapa 5 – Sistematização das relações semânticas: esta etapa consiste em analisar as relações semânticas entre as tags mais expressivas.

Etapa 6 – Consulta ao especialista: consiste em apresentar ao especialista um protocolo composto pelo conceito, categoria e afirmações⁵². No apêndice 2 consta parte do protocolo apresentado ao profissional e, no apêndice 1 consta um exemplo do diálogo, o retorno do especialista referente ao protocolo, com realce nas observações e dúvidas.

Etapa 7 – Estruturação da ontologia: apenas para verificar um contributo, como uma mais valia da tese, com os conceitos e relações semânticas extraídos da folksonomia, editou-se preliminarmente uma estrutura de ontologia sobre diabetes. Ressalta-se que os modelos de construção de ontologias indicam outras etapas como reutilização de outras ontologias, elaboração de restrições, axiomas e avaliação. Entretanto, para cumprir um dos objetivos da tese, fez-se uma adaptação, de modo que o enfoque foi dado a conceitualização e relações entre os conceitos. Conforme reitera Mendonça (2015, p. 1997) fundamentado em (Pereira e Soares, 2008), a fase da conceitualização é uma das mais importantes para o sucesso de uma

⁵² O protocolo é apresentado no capítulo 5 como um contributo da investigação, assim como nos apêndices.

ontologia, pois é nela que se realizam atividades como a identificação, análise e negociação dos conceitos de um determinado domínio, em um determinado contexto de aplicação, e que serão incluídos como classes na ontologia (Pereira & Soares, 2008). Além disso, é nessa fase que ocorre a organização e estruturação do conhecimento por meio de artefatos de representação, que irão se transformar em relações, propriedades e restrições de uma ontologia. O resultado final dessa etapa é um modelo conceitual do domínio que englobe conceitos, suas relações e suas propriedades.

Pelo exposto, apresentou-se detalhadamente as etapas empreendidas para a execução da tese. Certamente, podem apresentar algumas limitações, mas foram etapas e critérios de escolhas considerados adequados diante das complexidades do domínio e da proposta da investigação.

CAPÍTULO 5 - O DOMÍNIO DA DIABETES: CARACTERIZAÇÃO DO SEU LÉXICO A PARTIR DE FOLKSONOMIA

O resultado apresentado contempla todas as ações empreendidas e as atividades cognitivas necessárias para compreensão do domínio da diabetes com o intuito de sustentar as análises e atingir o objetivo proposto: explicitar as relações semânticas implícitas nas folksonomias e como contribuem para a estruturação de uma ontologia.

Notadamente, seguem-se apontamentos e especificidades sobre o domínio, espaço no qual apresentam-se a lista de tags extraídas (189 candidatos a descritores), as áreas limítrofes ao domínio e uma breve contextualização com o intuito de expor a complexidade do assunto. Seguidamente, são apresentados os qualificadores, a exemplificação de categorias, as relações semânticas, com as afirmações e uma demonstração acerca dos contributos para a ontologia.

5.1 APONTAMENTOS E ESPECIFICIDADES SOBRE DIABETES.

Após realizar buscas nos CiteULike, foram selecionadas as indexações livres que possuíam mais de uma tag, conforme exemplificado no quadro.

Quadro 27 - Exemplificação das indexações livres

	Resultado da busca	Indexações livres	Exemplos das indexações
CiteULike	800 artigos	500	a) hbgm-trial hbgm_dmt2 b) diabetes diet food gastroparesis nutrition review type1

Fonte: autoria própria.

Sob a perspectiva da teoria do conceito, abordada por Dahlberg (1978), enfatizam-se os seguintes aspectos: as características dos conceitos são obtidas por meio dos enunciados, que têm elementos que são suas características e um conjunto de características determina um conceito, o qual é uma unidade de conhecimento.

Os conceitos e as características acerca do tema são essenciais para compreensão desse universo de conhecimento para fins de representação. Nesse sentido, Diabetes mellitus⁵³ consiste em um grupo de doenças caracterizado por altas concentrações de glicose no sangue resultantes de defeitos na secreção de insulina, ação da insulina ou ambos. Também estão presentes anormalidades no metabolismo de carboidratos, proteínas e gorduras. Os diabéticos têm organismos que não produzem ou não respondem à insulina, um hormônio produzido pelas células beta do pâncreas que é necessário para uso ou armazenamento de combustíveis corporais. Sem insulina eficiente, ocorre hiperglicemia, a qual pode levar a complicações do diabetes mellitus. Algumas complicações agudas referem-se a hipoglicemia, cetoacidose diabética e as complicações a longo prazo incluem doenças macrovasculares, microvasculares e a neuropatia (Mahan & Escott-Stump, 2005).

A diabetes inclui 4 classes clínicas (Medina, 1997):

1. Diabetes tipo 1: resulta da destruição das células pancreáticas produtoras de insulina (células beta), o que implica deficiência absoluta de insulina. É uma doença autoimune.
2. Diabetes tipo 2: resulta de uma anomalia progressiva da secreção de insulina associada a um estado de resistência à insulina. É uma doença relacionada com os estilos de vida moderna, constituídos por ingestão exagerada de calorias e vida sedentária (a maior parte destes doentes estão com excesso de peso)
3. Outros tipos específicos devido a outras causas, como: defeitos genéticos na função da célula beta, anomalias genéticas na ação da insulina, doenças do pâncreas exócrino e induzida por fármacos ou agentes químicos.
4. Diabetes mellitus gestacional: é uma forma de diabetes diagnosticada durante a gravidez.

⁵³ Grifos nosso

As palavras sublinhadas objetivam indicar uma terminologia do domínio, seguida de suas características. Nesse sentido, segue a lista de tags extraídas. Algumas estão acompanhadas por suas definições, pois foram utilizadas para análise de categorias, estabelecimento de relações semânticas e exemplificação.

Acerca das definições, convém ressaltar que, para o desenvolvimento de esquemas de representação da informação como tesouros e ontologia, elas auxiliam na compatibilização semântica, pois descrevem o conteúdo semântico de um termo. A descrição possibilita que agentes inteligentes possam entender o significado de um termo e estabelecer inferências sobre esses significados, pois a definição é composta de características de conceitos que são também conceitos que se relacionam formando o entendimento semântico dos termos em questão (Campos, 2012).

Quadro 28 - Lista de tags

Tags	Definições
A	
Acidente Vascular Cerebral (AVC)	Grupo de afecções caracterizadas por perda súbita, não convulsiva, da função neurológica, devido a isquemia encefálica ou hemorragias intracranianas. O acidente cerebral vascular é classificado pelo tipo de necrose de tecido, como localização anatômica, vasculatura envolvida, etiologia, idade dos indivíduos afetados e natureza hemorrágica versus não hemorrágica (DeCS, 2017).
Glicogênese	
Ácidos Graxos Não Esterificados (NEFA)	
Açúcar	Dissacarídeo não redutor composto por glucose e frutose ligados por intermédio dos seus carbonos anoméricos. É obtido comercialmente da cana de açúcar, beterraba e outras plantas (DeCS, 2017). Sacarose, ou em sentido mais amplo, glúcido ou glícido (Manuila & Almeida, 2004).
Albuminúria	
Alcoolismo	Doença crônica, primária, com fatores genéticos, psicossociais e ambientais influenciando seu desenvolvimento e manifestações. A doença é geralmente progressiva e fatal. É caracterizada pela falta de controle sobre a bebida, preocupação com a droga álcool, uso de álcool apesar das consequências adversas, e distorções no pensamento, negação notável. Cada um destes sintomas pode ser contínuo ou periódico (DeCS, 2017). Conjunto de perturbações associadas à ingestão repetida, habitualmente durante anos, de

	quantidade excessivas de álcool: perturbações mentais, nervosas, lesões orgânicas (Manuila & Almeida, 2004).
Amido	
Anemia	
Apolipoproteína	
Aterogênese	
Atrofia	
Anticorpos	
Apneia obstrutiva do sono	
Auto-gerenciamento	
Atividade física	Para exercício, atividade física geralmente regular e feita com a intenção de melhorar ou manter aptidão física ou a saúde. É diferente de esforço físico que é voltado principalmente para as respostas fisiológicas e metabólicas ao uso da energia (DeCS, 2017).
B	
Bariátrica	
Bem-estar	
Betabloqueadores	
Bomba de insulina	Referente a bomba: Qualquer mecanismo para uso da energia metabólica a fim de realizar o transporte ativo de uma substância (Stedman & Araújo, 1996). É um aparelho com comando eletrônico, pequeno, com peso de aproximadamente de 100g, que envia microdoses de insulina continuamente durante as 24 horas do dia de forma precisa e pré-programada para manter o controle glicêmico entre as refeições e ao longo da noite. ⁵⁴
Bypass	
C	
Calcificação	
Caloria	
Câncer	
Câncer de Mama	
Cardiomiopatia	
Célula Alfa	
Célula Beta	Células das Ilhotas de Langerhans localizadas no pâncreas, responsáveis por produzirem um hormônio designado insulina.
Cetoacidose diabética (DKA)	Sinônimo de acidocetose, refere-se a acidose observada essencialmente na diabetes e por vezes durante o jejum prolongado. Deve-se à acumulação no organismo de corpos cetônicos, produtos resultantes da desintegração dos lipídios utilizados em quantidade excessiva para

⁵⁴ Conceito disponível em: <http://www.medtronicdiabetes.com.br/acerca-da-diabetes/o-que-e-uma-bomba-de-insulina.html> Acesso: 04 de novembro de 2017

	<p>as necessidades energéticas, em substituição dos glicídios cujo metabolismo está perturbado (Manuila & Almeida, 2004).</p> <p>Em diabetes melito, a instalação de cetoacidose indica uma carência absoluta ou relativa de insulina frente aos hormônios cetogênicos, como a epinefrina, o glucagon, os esteroides e a somatotropina (hormônio de crescimento). Tal situação encontra-se essencialmente no coma diabético, onde aparece primeiro cetonúria sem acidose, depois uma baixa do pH (Rey, 2006).</p>
Cisteína	
Citocinas	
Colesterol	<p>- Substância orgânica presente na maioria dos tecidos e humores do organismo humano. O colesterol entra na constituição das lipoproteínas do sangue sob a forma livre ou de estéres (Manuila & Almeida, 2004).</p> <p>- Principal esteroide de todos os animais superiores, distribuído nos tecidos do corpo, especialmente no cérebro e na medula espinhal, e nas gorduras e óleos animais (DeCS, 2017)</p>
Colesterol HDL	Lipoproteína de alta densidade que controla a depuração do colesterol nas paredes vasculares. A sua concentração diminui em numerosas circunstâncias que têm em comum a frequente associação com uma doença ateromatosa: envelhecimento, tabagismo, diabetes (Manuila & Almeida, 2004).
Colesterol LDL	Lipoproteína de baixa densidade, rica em colesterol (Manuila & Almeida, 2004).
Comorbidade	Presença de doenças coexistentes ou adicionais com relação ao diagnóstico inicial ou com relação à doença índice que é o objetivo do estudo. A comorbidade pode afetar o desempenho de indivíduos afetados e até mesmo a sua sobrevivência. Pode ser usado como um indicador prognóstico para a duração da hospitalização, fatores de custo e de melhoria ou sobrevivência (DeCS, 2017).
Complicações macrovasculares	
Complicações microvasculares	
Controle glicêmico	
Cefaleia	
Coagulação	
D	
Densidade Mineral Óssea (BMD)	
Depressão	
Derrame	Ver AVC
Diabetes	<p>Nome dado a diversas doenças caracterizada pela emissão de urina anormalmente abundante e acompanhada de uma sensação de sede intensa. Empregado isoladamente, o termo diabetes significa geralmente diabetes açucarada, diabetes mellitus (Manuila & Almeida, 2004).</p> <p>Grupo de doenças metabólicas de diferentes etiologias e gravidade variável, tendo como</p>

	característica comum a eliminação de volume anormalmente elevado de urina (Rey, 2006)
Desordem alimentar	
Diabetes gestacional	Diabetes mellitus induzida por gravidez, porém resolvida no final da mesma. Geralmente, se desenvolve no final da gravidez quando os hormônios antagonistas da insulina atingem seus níveis máximos levando à resistência à insulina, intolerância à glucose e hiperglicemia (DeCS, 2017).
Diabetes juvenil	Ver Diabetes mellitus
Diabetes autoimune latente em adultos	Diabetes autoimune em adultos com falência progressiva lenta das células beta pancreáticas e a presença de autoanticorpos circulantes dirigidos contra antígenos das células das ilhotas pancreáticas (DeCS, 2017).
Diabetes mellitus	Grupo de transtornos heterogêneos caracterizados por hiperglicemia e intolerância à glucose (DeCS, 2017). Sinônimo de diabetes açúcarada, trata-se da infecção crônica caracterizada por poliúria com polidipsia e glicosúria persistente e hiperglicemia. Deve-se a uma insuficiência de produção de insulina pelo pâncreas. Na ausência de um tratamento regular, pode complicar-se de acidose grave com coma e de diversas lesões degenerativas graves (acidentes vasculares, cerebrais ou cardíacos, retinopatia, lesões renais, perturbações nervosas) A diabetes juvenil constitui uma forma particularmente grave e obriga a tratamento de insulina (Manuila & Almeida, 2004).
Diabetes parentais	
Diabetes Mellitus tipo 1	Subtipo de diabetes mellitus caracterizada por deficiência de insulina. Manifesta-se por um início repentino de intensa hiperglicemia, progressão rápida a cetoacidose diabética e morte, a menos que tratada com insulina. A doença pode ocorrer em qualquer idade, mas é mais comum durante a infância ou adolescência (DeCS, 2017).
Diabetes tipo 1 adulto atípico ou com tendência à cetose	
Diabetes tipo 2	Subclasse de diabetes mellitus que não é responsiva ou dependente de insulina. Inicialmente, caracteriza-se por resistência à insulina e hiperinsulinemia e finalmente, por intolerância à glucose, hiperglicemia e obviamente, diabetes. A diabetes mellitus tipo 2 não é mais considerado uma doença encontrada exclusivamente em adultos. Os pacientes, raramente desenvolvem cetose, porém com frequência exibem obesidade (DeCS, 2017).
Dieta	Método regular de ingestão de comida e bebida adotado por uma pessoa ou animal (DeCS, 2017). Alimento e bebida em geral. Um método prescrito de comer e beber, no qual a quantidade e o tipo de alimento, assim como o

	número de refeições que devem ser feitas, são regulados para objetivos terapêuticos (Stedman & Araújo, 1996). Regime alimentar especial que inclui, exclui ou limita certos alimentos ou bebidas (Manuila & Almeida, 2004).
Dieta lowcarb	
Dieta lowgi	
Disfunção erétil	
Disfunção sexual	
Doença arterial coronariana	
Doença autoimune	
Doença cardiometabólicas	
Doença cardiovascular	
Doença crônica	
Diagnóstico	
E	
Educação diabética	
Endocrinologia	
Enxaqueca	
Espessura íntima-média carotídea (IMT)	
Estatinas	
Estilo de vida	Ver qualidade de vida
Estradiol	
Estreptozotocina (STZ)	
Estress	
Estudo Coorte	
Euglicemia	
Exenatida	
Exercício	Ver atividade física
F	
Fator de risco	Um aspecto do comportamento pessoal ou do estilo de vida, da exposição ao meio ambiente, ou uma característica própria ou herdada do indivíduo que se sabe, tendo por base a evidência epidemiológica, estarem associados com condições importantes de se prevenir para proteger a saúde. O fator de risco representa uma probabilidade maior de ser atingido por determinada afecção ou dano. Fala-se de alto risco quando a intensidade de um fator ou soma de vários fatores elevam consideravelmente o risco, como o de pessoas que fumam muito e têm também hipertensão ou colesterolemia alta, formando um grupo de alto risco para infarto do miocárdio. A pobreza reúne por si só um grande número de fatores de risco, principalmente para doenças transmissíveis e morbimortalidade infantil (Rey, 2006).
Fator de transformação do crescimento beta (TGFB)	
Fatores de risco cardiometabólicos (CMRF)	

Folatos	
Frutose	Monossacarídeo encontrado em frutas doces e mel, solúvel em água, álcool ou éter. É empregado como conservante e em infusão intravenosa na alimentação parental (DeCS, 2017). Açúcar simples (não hidrolizável), hexose abundante nos frutos, nos quais coexiste frequentemente com a glicose (Manuila & Almeida, 2004).
Fibras	
G	
Gastroparesia	Retardo crônico do esvaziamento gástrico. A gastroparesia pode ser causada por disfunção motora ou paralisia dos músculos do estômago ou pode estar associada com outras doenças sistêmicas como diabetes mellitus (DeCS, 2017). Motilidade gástrica prejudicada, ela resulta em contrações retardadas ou irregulares do estômago, levando a vários sintomas gastrointestinais (Mahan & Escott-Stump, 2005).
Glicação	
Glicemia	Concentração de glicose no sangue (Manuila & Almeida, 2004). Glicose no sangue (DeCS, 2017).
Glicómetro	
Glicose	Açúcar simples não (não hidrolizável) que é a hexose (Manuila & Almeida, 2004). Fonte primária de energia dos seres vivos. É encontrado em frutas e outras partes das plantas em seu estado livre. É utilizada terapêuticamente na reposição de líquidos e nutrientes (DeCS, 2017)
Glicose em jejum (IFG)	
Globulina de ligação hormonal sexual (SHBG)	
Glucagon com péptido-1	
H	
HbA1c	Medida que indica o controle da glicemia nos últimos meses. É uma das análises mais importantes para avaliar o controle da diabetes. O termo HbA1c, ou simplesmente A1c, significa hemoglobina glicada (Dicionário Diabetes)
Hidratos de carbono	
Hiperglicemia	- Elevação da quantidade de glicose no sangue (Manuila & Almeida, 2004). - Nível anormalmente alto de glicemia (DeCS)
Hipertensão	Pressão arterial sistêmica persistentemente alta. Com base em várias medições, a hipertensão é atualmente definida como sendo a pressão sistólica repetidamente maior que 140mm Hg ou a pressão diastólica de 90 mm Hg ou superior (DeCS, 2017). Elevação persistente da pressão sanguínea no sistema circulatório (Rey, 2006).
Hipoglicemia	- Diminuição da quantidade de glicose no sangue (Manuila & Almeida, 2004). - Síndrome de nível anormalmente baixo de

	glicemia (DeCS, 2017)
Histona deacetilase	
Homocisteína	
Hormônio	
Hemodiálise	
I	
Índice de massa corporal (IMC/BMI)	Indicador da densidade do corpo que é determinado pela relação entre peso corporal e estatura (DeCS, 2017). Relação quantitativa entre peso e altura (Rey, 2006).
latrogenia	
Inalador de insulina	
Inflamação	
Insuficiência renal	
Insulina	- Hormônio pancreático de 51 aminoácidos que desempenha um papel fundamental no metabolismo da glucose, suprimindo diretamente a produção endógena de glucose e indiretamente de glucagon e a lipólise. A insulina é usada para controlar o Diabetes Mellitus tipo 1 (DeCS, 2017). - Hormônio peptídico, secretado pelas células beta das ilhotas de Langerhans, que promove a utilização da glicose, a síntese protéica e a formação e o armazenamento de lipídeos neutros, obtida de diversos animais e disponível em uma série de preparações (Stedman & Araújo, 1996). Hormona do pâncreas, segregada pelas células beta. Favorece a utilização do açúcar pelos tecidos e reduz a glicemia. O seu antagonista é o glucagon (Manuila & Almeida, 2004).
Insulina basal	
Insulina glargina	
Interleucina 6	
Insulinoressistência	
Inibidor da bomba de prótons (PPIS)	
Índice glicêmico (GI)	Sistema numérico para medir a taxa de glicemia por um determinado alimento quando comparado à glucose, que é igual a 100. Os alimentos com alto índice glicêmico geram oscilações maior es dos níveis de açúcar no sangue (DeCS, 2017).
J	
Jejum	Abster-se de todo alimento (DeCS, 2017).
L	
Lactose	
Leptina	
Lipase de lipoproteína	
Lipídios	Termo genérico para gordura e lipídeos, constituintes do protoplasma, solúveis em álcool e éter, e são insolúveis em água. Compreendem as gorduras, óleos graxos, óleos essenciais,

	<p>ceras, fosfolípídeos, glicolípídeos, sulfolípídeos, aminolípídeos, cromolípídeos e ácidos graxos (DeCS, 2017).</p> <p>Qualquer substância que hidrolisada, fornece ácidos carboxílicos com grande cadeia de carbonos (Rey, 2006).</p>
Lipídios hepáticos	
Lipoproteínas	
M	
Macronutrientes	
Magnésio	
Metabólica	
Metformina	<p>Hipoglicêmico biguanida utilizada no tratamento da diabetes mellitus não dependente de insulina que não responde às modificações na dieta. Metformina melhora no controle glicêmico aumentando a sensibilidade à insulina e diminuindo a absorção intestinal de glicose (DeCS, 2017).</p>
Múltiplas injeções diárias (MDI)	
Metabolismo	<ul style="list-style-type: none"> - Reações químicas em organismos vivos por meio das quais a energia é fornecida para os processos vitais e atividades e por que novos materiais são assimilados (DeCS, 2017). - A soma das alterações químicas que ocorrem no tecido, consistindo em anabolismo, as reações que convertem pequenas moléculas em grandes, e catabolismo as reações que convertem grandes moléculas em pequenas (Stedman & Araújo, 1996). - Conjunto das transformações químicas e físico químicas que ocorrerem em todos os tecidos do organismo, sofridas pelos constituintes da matéria viva (Manuila & Almeida, 2004).
Monitoramento contínuo em tempo real da glicose (RTCGM)	
Monitoramento contínuo da glicose (CGM)	
MicroRNAs circulantes (MIRS ou MIRNA)	
N	
Nefropatia	<p>Qualquer afecção dos rins (Manuila & Almeida, 2004).</p> <p>Processos patológicos do rim ou de componentes de seus tecidos (DeCS, 2017)</p>
Neuropatia	<ul style="list-style-type: none"> - Afecção do sistema nervoso, central ou periférico (Manuila & Almeida, 2004). - Termo clássico que designa qualquer distúrbio que acomete qualquer segmento do sistema nervoso (Stedman & Araújo, 1996). - No DeCS (2017) apresenta-se Neuropatia diabética: transtornos periféricos autônomos do nervo craniano associados com diabetes mellitus. Estas afecções normalmente resultam de lesão microvascular diabética envolvendo pequenos vasos sanguíneos que nutrem os nervos.

Nicotiamida	
Nutrição	<ul style="list-style-type: none"> - Conjunto de processos de assimilação e desassimilação dos alimentos no organismo, através dos quais são asseguradas a manutenção em vida, das funções fisiológicas e a energia de que necessita (Manuila & Almeida, 2004). - Uma função de vegetais e animais vivos, consistindo na ingestão e no metabolismo de material alimentar por meio do qual o tecido é formado, liberando energia (Stedman & Araújo, 1996). - No DeCS (2017) o conceito mais próximo associa-se a Fenômenos fisiológicos da Nutrição: Processo e propriedades de organismos vivos pelos quais eles assimilam e equilibram o uso de materiais nutritivos para energia, produção de calor, ou material construtor para o crescimento, manutenção ou reparo de tecidos, e as propriedades nutritivas dos alimentos.
O	
Obesidade	<ul style="list-style-type: none"> - Acumulação excessiva, mais ou menos generalizada, de tecido adiposo, que provoca um aumento de peso superior a 25% do peso estimado normal. Pode ser exógena, devida a superalimentação ou endógena, devida a perturbações metabólicas ou endócrinas (Manuila & Almeida, 2004). - Estado no qual o peso corporal está grosseiramente acima do peso aceitável ou ideal, geralmente devido a acúmulo excessivo de gorduras no corpo (DeCS, 2017). - Aumento anormal da gordura nos tecidos conjuntivos subcutâneos (Stedman & Araújo, 1996).
Oxidação	
Óxido nítrico	
Ômega 3	
P	
Pâncreas	<p>Órgão nodular no abdome que abriga uma mistura de glândulas endócrinas e glândulas exócrinas. A pequena porção endócrina consiste das ilhotas de Langerhans que secretam vários hormônios na corrente sanguínea. A grande porção exócrina é uma glândula acinas composta que secreta várias enzimas digestivas no sistema de ductos que desemboca no Duodeno (DeCS, 2017).</p> <p>O pâncreas é a glândula mista que por suas ilhotas de Langerhans, é de secreção interna (produção de insulina) e, pelo resto, uma glândula exócrina tubuloacinososa, semelhante a parótida, que elabora suco pancreático (Rey, 2006).</p>
Pé diabético	Problemas comuns no pé de pessoas com diabetes mellitus, causados por qualquer uma das combinações dos fatores, como neuropatias diabéticas, doenças vasculares periféricas e

	infecção. Com a perda da sensação e circulação deficitária, as lesões e infecções, com frequência, levam a sérias úlceras do pé, gangrena e amputação (DeCS, 2017).
Peso	
Pioglitazona	
Polineuropatia	
Poluentes orgânicos persistentes (POPs)	
Pressão sanguínea	Pressão do sangue nas artérias e de outros vasos sanguíneos (DeCS, 2017). Pressão exercida pelo sangue sobre as paredes das artérias, que depende da força de contração do ventrículo esquerdo, da resistência exercida pela rede arterial e capilar, da elasticidade das artérias, do volume e da viscosidade do sangue (Rey, 2006).
Pró-renina	
Proteína	<ul style="list-style-type: none"> - Nome genérico dos compostos orgânicos azotados, animais e vegetais, cuja molécula é composta por numerosos aminoácidos e que desempenham uma função essencial em toda matéria viva (Manuila & Almeida, 2004). - Macromolécula que consiste em longas sequencias de aminoácidos (Stedman & Araújo, 1996). - Polipeptídeos lineares sintetizados nos ribossomos e posteriormente podem ser modificados, entrecruzados, clivados ou agrupados em proteínas complexas com várias subunidades. A sequência específica de aminoácidos determina a forma que tomará o polipeptídeo, durante o dobramento de proteína e a função da proteína (DeCS, 2017).
Proteína C reativa	
Proteína de ligação	
Prevenção	Usado com doenças para aumento da resistência humana ou animal contra doença, ara controlo de agentes transmissores, para prevenção e controle de danos ambientais ou de fatores sociais que conduzam à doença. Inclui medidas preventivas em casos individuais (DeCS, 2017).
Q	
Questionário de frequência alimentar (FFQ)	
Qualidade de vida	Conceito genérico que reflete preocupação com a modificação e o aprimoramento dos componentes da vida, como por exemplo ambiente físico, político, moral e social, bem como saúde e doença (DeCs, 2017). Do ponto de vista subjetivo, é o grau em que as pessoas se sentem competentes para atuar física, emocional e socialmente, julgando suas vidas dignas de serem vividas. De um ponto de vista objetivo, compreende um conjunto das características genéticas, dos hábitos e costumes adquiridos e dos tipos de comportamento do indivíduo, que lhe asseguram alto grau de saúde física e mental, com mínimo

	de risco presentes ou futuros, proporcionando bem-estar, auto-satisfação e excelente relacionamento familiar e social (Rey, 2006).
R	
Recomendação nutricional	
Renina	
Resistência a insulina	Redução da efetividade da insulina em diminuir a glicemia: há a necessidade do uso de 200 unidades ou mais de insulina por dia para impedir a hiperglicemia ou a cetose (DeCS, 2017).
Retinopatia	Doença da retina resultante de uma complicação de diabetes mellitus. Caracteriza-se por complicações microvasculares e progressivas, como aneurisma, edema inter-retiniano e neovascularização patológica intraocular (DeCS, 2017). Toda e qualquer afecção da retina, seja qual for sua causa. Constitui causa importante de cegueira, sendo particularmente grave no diabetes melito insulino-dependente (Rey, 2006).
Risco	A probabilidade de que um evento ocorrerá. Ele abrange uma variedade de medidas e probabilidade de um resultado geralmente desfavorável (MeSH). Medida de dano potencial ou prejuízo econômico expressa em termos de probabilidade estatística de ocorrência e de intensidade ou grandeza das consequências previsíveis. Relação existente entre ameaça de um evento adverso ou acidente determinado e o grau de invulnerabilidade do sistema receptor a seus efeitos (DeCS, 2017).
S	
Sacarose	
Sal	
Sedentarismo	
Síndrome metabólica	
Sobrepeso	
Sódio	
Stress	
Sulfoniluréia	
Supressores de sinalização de citocinas	
T	
Tabagismo	Uso de tabaco com o detrimento da saúde de uma pessoa ou de sua função social. Está incluída a dependência de tabaco (DeCS, 2017) Dependência do tabaco e intoxicação pelos seus componentes, entre eles a nicotina (Manuila & Almeida, 2004).
Tasglutide	
Terapia de reposição de testosterona	
Testosterona	
Tiazida	

	<p>Teste para determinar a capacidade de um indivíduo em manter a homeostase da glicemia. Inclui a medida dos níveis de glicemia em jejum e em intervalos pré-estabelecidos antes e após a ingestão de glucose ou de uma infusão intravenosa (DeCS, 2017).</p> <p>Sinônimo de hiperglicemia provocada, trata-se da prova que consiste em sobrecarregar o organismo com glucose, acompanhando em seguida as variações consecutivas da glicemia (Manuila & Almeida, 2004).</p>
Tolerância a glicose	
Transportadores de glicose	
Triglicérides	
Troglitozona	
U	
V	
Vitaminas	<p>Nome genérico das substâncias indispensáveis ao crescimento e ao funcionamento dos órgãos, fornecidas, na sua maioria, em pequena quantidade pela alimentação, que o organismo não é capaz de sintetizar e cuja carência provoca perturbações características. As suas atividades são diversas. Certas vitaminas participam na elaboração de hormonais e enzimas, quer favorecendo a sua produção, quer entrando diretamente na sua composição química. As vitaminas de origem alimentar classificam-se como: vitaminas lipossolúveis (vitamina A, vitamina D, vitamina E e vitamina K) e vitaminas hidrossolúveis (vitaminas do grupo B e vitamina C) (Manuila & Almeida, 2004).</p>
Vitamina B12	
Vitamina D	<p>Vitamina que inclui tanto os colecalciferóis como os ergocalciferóis e que têm em comum o efeito de impedir ou curar o raquitismo em animais. Também pode ser analisada como um hormônio, já que sua formação ocorre na pele por ação dos raios ultravioleta sobre os precursores 7 de hidrocolesterol e ergosterol e atua sobre os receptores de vitamina D para regular o cálcio, contrapondo-se ao hormônio paratireóideo (DeCS, 2017).</p> <p>A vitamina D protege contra o raquitismo e exerce uma ação reguladora do metabolismo do fósforo e do cálcio, favorecendo a sua absorção intestinal e impedindo a eliminação excessiva pela urina (Manuila & Almeida, 2004).</p>
Vitamina B6	
Z	
Zinco	

Fonte: autoria própria.

Sem a pretensão de questionar ou avaliar as definições apresentadas, mas, considerando o papel da definição como imprescindível para auxiliar na compreensão do domínio e no desenvolvimento de instrumentos consistentes

de representação, é relevante, apenas a título de ilustração, expor algumas questões de ordem prática que ocorreram nesta atividade de consultar as definições, as quais são exemplificadas abaixo:

Sobre hipertensão, tag expressiva no corpus, as definições encontradas nos dicionários são:

- Pressão arterial sistêmica persistentemente alta. Com base em várias medições, a hipertensão é atualmente definida como sendo a pressão sistólica repetidamente maior que 140mm Hg ou a pressão diastólica de 90 mm Hg ou superior (DeCS, 2017).
- Elevação persistente da pressão sanguínea no sistema circulatório (Rey, 2006)

O mesmo ocorre com a definição de obesidade:

- Acumulação excessiva, mais ou menos generalizada, de tecido adiposo, que provoca um aumento de peso superior a 25% do peso estimado normal. Pode ser exógena, devido a superalimentação ou endógena, devido a perturbações metabólicas ou endócrinas (Manuila & Almeida, 2004).
- Estado no qual o peso corporal está grosseiramente acima do peso aceitável ou ideal, geralmente devido a acúmulo excessivo de gorduras no corpo (DeCS, 2017).
- Aumento anormal da gordura nos tecidos conjuntivos subcutâneos (Stedman & Araújo, 1996).

Nesse sentido, por exemplo, como determinar qual categoria: doença, sintoma, medida (para hipertensão) uma vez que a noção não está tão evidente, mencionada ou precisa nas definições? Também por não indicarem o gênero próximo e a diferença específica do conceito, seus componentes ou etapas. É nesta perspectiva que Campos (2010) esclarece que, nos dicionários, a definição é realizada de forma isolada, sem qualquer relação com outras palavras, e que, em ontologias, a unidade de informação é o conceito, cuja sequência de conceitos representa um tipo de hierarquia e essa hierarquia é partilhada por grupos de conceitos. A consistência dessa hierarquia depende da definição explícita das propriedades (atributos e características) que descrevem os conceitos, fornecendo um conceito do outro.

Ainda, com base nas tags, foi possível expor algumas áreas limítrofes acerca do tema diabetes, que consiste em áreas do conhecimento que compartilham conceitos e possuem uma relação estreita, conforme consta na figura 28.

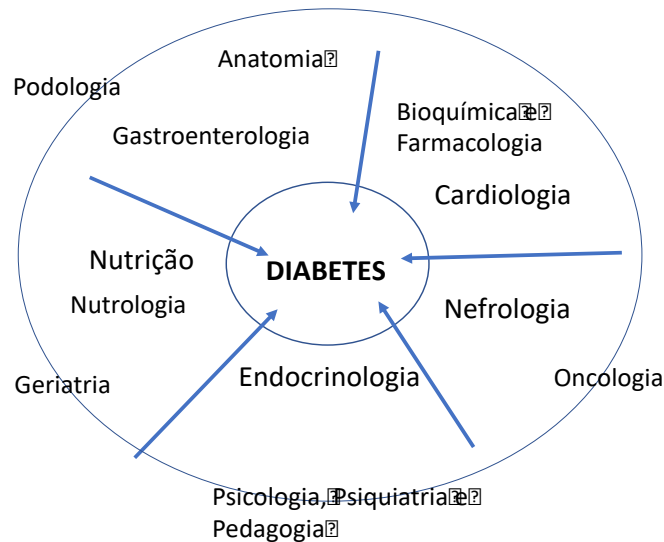


Figura 28 - Áreas limítrofes
Fonte: autoria própria.

Com base na figura 28, elucida-se a ocorrência dos conceitos por áreas do conhecimento:

- Patologia: pâncreas, fígado, rim.
- Bioquímica: metabolismo, glicogênese, secreção de insulina, transportador de glicose (GLUT2), stress oxidativo, insulina.
- Genética: microRNAs.
- Farmacologia: metformina, sulfoniluréia, insulina.
- Nutrição: dieta, metabolismo, obesidade, nutrientes, proteína.
- Endocrinologia: obesidade, câncer, diabetes mellitus, pé diabético, insulina.
- Epidemiologia: mortalidade, risco, fator de risco, prevenção.
- Nefrologia: insuficiência renal, nefrologias.
- Psicologia, Psiquiatria: depressão, obesidade, atividade física, alcoolismo.
- Oncologia: câncer.

- Cardiologia: doenças cardiovasculares, hipertensão.
- Neurologia: acidente vascular cerebral, neuropatias.

No tocante aos aspectos da representação da informação, com base na lista de tags, evidenciam-se como característica a **comorbidade e fator de risco**. O primeiro, com base nas menções a outras doenças, a exemplo: doença cardiovascular, hipertensão, acidente vascular cerebral, além de outras tags, como depressão, câncer e alcoolismo. A comorbidade ocorre quando uma entidade clínica distinta adicional ocorre ou vem a ocorrer durante a evolução de um paciente, e ela pode ser patogênica quando duas ou mais doenças estão etiologicamente relacionadas (Marques, Nardi, Figueira, Mendlowicz, Andrade, Camisao, & Versiani, 1994)

Adicionalmente, observou-se, para além da ocorrência da tag diabetes juntamente a outras doenças, a tag fator de risco. Rey (2006) esclarece que fator de risco consiste em um aspecto do comportamento pessoal ou do estilo de vida, da exposição ao meio ambiente, ou em uma característica própria ou herdada do indivíduo que se sabe, tendo por base a evidência epidemiológica, estarem associados com condições importantes de se prevenir para proteger a saúde. O fator de risco representa uma probabilidade maior de ser atingido por determinada afecção ou dano. Fala-se de alto risco quando a intensidade de um fator ou soma de vários fatores elevam consideravelmente o risco, como o de pessoas que fumam muito e têm também hipertensão, formando um grupo de alto risco para infarto do miocárdio.

Conforme Vaz, Pombo, Fantinato & Pecky (2007), em termos epistemológicos, fator de risco não é uma causa nem suficiente, nem necessária de uma doença; ele designa, sim, o que amplia as chances de adoecer. Todavia, no item 5.4, retoma-se a discussão acerca da tag risco e fator de risco.

5.2 QUALIFICADORES DO DECS E AS TAGS

Compreende-se que os qualificadores são termos que se agregam aos descritores de modo a definir aspectos particulares destes, os quais são utilizados de acordo com princípios já estabelecidos. Ou seja, são utilizados em associação a um descritor principal de modo a especificar ou delimitar um determinado assunto, sendo relevante para reconhecer e definir diferentes aspectos, conceitos e pontos de vista. Neste contexto, eliminação das ambiguidades causadas pela homonímia e polissemia advindas da linguagem natural são um importante recurso na representação e recuperação da informação (Boccatto, 2009).

Posto isso, abaixo demonstram-se os qualificadores indicados pelo DeCS para o descritor diabetes mellitus, os quais foram utilizados como referência para verificar sua incidência no CiteULike.

Quadro 29 - Qualificadores no DeCS e as tags

QUALIFICADORES DECS	TAGS
Cirurgia	
Congênito	
Diagnóstico	Presente
Economia	Presente
Enfermagem	
Epidemiologia	Presente
Etnologia	
Genética	Presente
Imunologia	
Líquido cefalorraquidiano	
Microbiologia	
Parasitologia	
Prevenção e controle	Presente
Radioterapia	
Sangue	Presente
Tratamento farmacológico	
Veterinária	
Classificação	Presente
Diagnóstico por imagem	
Dietoterapia	
Embriologia	
Enzimologia	
Etiologia	
Fisiopatologia	
História	Presente
Induzido quimicamente	
Metabolismo	Presente
Mortalidade	Presente
Patologia	
Psicologia	
Reabilitação	

Terapia
Urina
Virologia

Fonte: autoria própria.

Entre os 33 qualificadores, 10 são mencionados, sendo que o mais expressivo foi Epidemiologia (aproximadamente 19). No DeCS (2017), a definição indica que este conceito é usado com doenças humanas e animais para sua distribuição, fatores causais e características em população definidas. Inclui incidência, frequência, prevalência, surtos endêmicos e epidêmicos e levantamentos ou estimativas de morbidade em áreas geográficas e populações específicas. Usados também com descritores geográficos para a localização de aspectos epidemiológicos de uma doença.

Assim, pode-se destacar que o assunto diabetes, no sistema baseado em folksonomia utilizado, se impõe-se com ênfase em uma abordagem epidemiológica. Adicionalmente, destaca-se a significativa menção às tags risco e fator de risco. Entretanto, acerca da tag epidemiologia e aos conceitos a que está associada, discutir-se-á nos próximos itens.

Ademais, o DeCS apresenta uma estrutura hierárquica para os qualificadores. Tal estrutura indica ao indexador utilizar o termo mais específico para cobrir o conceito expressado em determinado recurso informacional. Por exemplo, o qualificador “epidemiologia” está subordinado ao qualificador “estatística & dados numéricos”.

Após o exposto, percebe-se que verificar estes qualificadores, auxilia algumas formas pelas quais o domínio pode ser compreendido. No âmbito de bibliotecas especializadas, Simões (2008) argumenta que pode não haver necessidade de recorrer a qualificadores se os termos homógrafos corresponderem a áreas diversificadas e apenas uma delas for coberta pelos fundos próprios. Contudo, compreender as imposições da interdisciplinaridade e reconhecer os meios informáticos e telemáticos disponibilizados de uma forma sempre crescente e, num raio alargadíssimo, o acesso a bases de dados bibliográficas de todas as áreas do saber, pelo que, tal como em uma biblioteca

geral, a qualificação poderá impor-se para facilitar a pesquisa e garantir a pertinência da informação.

5.3 CATEGORIZAÇÃO A PARTIR DAS TAGS E DO CONTEXTO

A categorização refere-se a analisar e a identificar recortes conceituais, os quais permitirão determinar a identidade dos conceitos que compõem este domínio.

Esta aplicação ocorre por meio do conceito, unidade de conhecimento, do qual se estabelece a relação do termo com o objeto/referente, fornecendo bases seguras para estruturar as relações lógicas. Tomado o conhecimento do método analítico-sintético proposto por Ranganathan incorpora-se o princípio de categorização como uma solução para a organização dos conceitos num sistema de conceitos, não importando a finalidade de aplicação e propõe princípios para o estabelecimento de relacionamentos associativos, conforme apontam Campos e Gomes (2006).

A autoras, apoiadas em Dahlberg (1978, p.34), enfatizam a importância da categoria na estruturação do conceito: as categorias não apenas estruturam todos os nossos elementos de conhecimento e unidades de conhecimento, mas também fornecem, ao mesmo tempo, por este meio, o esqueleto, os ossos e tendões para estruturar todo o nosso conhecimento. E, com o uso consciencioso, o corpo do nosso conhecimento pode se manter unido, pode se mover, pode ser manter flexível e pode crescer organicamente.

Entretanto, corrobora-se com Campos e Campos (2014), que Ranganathan não deixa claro qual a fundamentação que adota para a definição de cada categoria e a falta de explicitação mais precisa, aliadas a uma redação que dá margem, por vezes, a entendimentos ambíguos, têm levado a dificuldades na compreensão e na aplicação de sua teoria, dificuldade reconhecida por ele.

É neste sentido, considerando também a complexidade do domínio, é que se apresentam as questões que surgiram na elaboração de estruturas conceituais, nomeadamente no estabelecimento das categorias e

relacionamentos semânticos, evidenciando o contexto e o compromisso ontológico.

Quadro 30 - Categorias e manifestação

Categoria	Subcategoria	Manifestação
Doença	Doença endócrina	<i>Diabetes mellitus</i> <i>Diabetes mellitus tipo 1</i> <i>Diabetes mellitus tipo 2</i> <i>Obesidade</i> <i>Cetoacidose diabética</i> <i>Resistência a insulina</i>
	Doença macrovascular	<i>Pé diabético</i> <i>Acidente Vascular Cerebral</i>
	Doenças microvasculares	<i>Retinopatia</i> <i>Nefropatia</i> <i>Neuropatia</i>
Mecanismo comportamental		<i>Tabagismo</i> <i>Alcoolismo</i> <i>Depressão</i> <i>Sedentarismo</i>
Sintoma		<i>Hiperglicemia,</i> <i>Hipoglicemia</i> <i>Poliúria</i> <i>Sede</i>
Instrumento		<i>Bomba de insulina</i> <i>Inalador de insulina</i> <i>Glicosímetro</i> <i>Glucômetro</i> <i>Célula beta</i>
Substância química		<i>Insulina</i> <i>Glucagon</i> <i>Glicose</i> <i>Colesterol</i> <i>Vitamina D</i> <i>Carboidrato</i> <i>Sacarose</i>
Medicamento		<i>Sulfonilureias</i> <i>Metformina</i>
Processo		<i>Nutrição</i> <i>Metabolismo</i> <i>Resistência à insulina</i> <i>Secreção de insulina</i>
Análise		<i>HbA1c</i> <i>Tolerância à glicose</i> <i>Glicemia</i> <i>Índice glicêmico</i>
Tratamento	Tratamento médico	<i>Insulinoterapia</i> <i>Infusora Subcutânea Contínua com Insulina (CSII)</i> <i>Múltiplas injeções diárias (MDI)</i> <i>Controle glicêmico</i>
	Tratamento nutricional	<i>Dieta</i> <i>Atividade física</i>

Área do conhecimento		Endocrinologia Epidemiologia Nutrição
----------------------	--	---

Fonte: autoria própria.

As categorias foram desenvolvidas com base no conjunto de tags identificadas. É importante ressaltar que não tem como propósito abarcar de forma exaustiva todas as tags, mas apenas as mais expressivas. Ademais, é reconhecido-se que as categorias podem e devem receber mais desdobramentos em subcategorias. A seguir, faz-se uma explanação com o intuito de expor os critérios para os agrupamentos.

- **Doença endócrina:** categoria e conceito utilizado pelo DeCS (2017) e ICD 10. É pertinente e preciso para agrupar os tipos de diabetes e afecções associadas.

- **Doença macrovascular e doença microvascular:** a exemplo da ocorrência da tag doença vascular, além da utilização destes conceitos em textos que tratam especificamente do assunto, o desenvolvimento de complicações como macroangiopatias, cardiopatia isquêmica, doença vascular periférica e acidente vascular cerebral, assim como microangiopatias, como retinopatia diabética, nefropatia diabética e neuropatia sensitiva distal. Neste sentido, considerou-se pertinente a adoção das duas categorias, afim de evidenciar a relação da diabetes com outras doenças (Mahan & Escott-Stump, 2005; Scheffel, Bortolanza, Weber, Costa, Canani, Santos & Tschiedel, 2003).

- **Mecanismo comportamental:** considera-se como um conjunto de elementos, ações, transtornos de ordem comportamental, psíquica e mental. Alguns descritores podem estar disponíveis na literatura como transtorno e doença, como depressão, alcoolismo.

- **Sintoma:** considera o termo sintoma como alguma alteração da percepção normal que uma pessoa tem de seu próprio corpo, metabolismo, sensações, podendo consistir ou não em um início de doença (Simões, 2016). Trata-se, etimologicamente, de um fenômeno que coincide, índice do processo patológico subjacente do qual é efeito (conexão causal), apontando para uma alteração mórbida sem traduzi-la. Também, o termo costuma ser empregado em circunstâncias menos definidas, como traço característico que

isoladamente ou associado a outros configura um dado padrão de manifestação, fenômeno, sem a possibilidade imediata de confirmação independente ou de identificação do processo mórbido subjacente (Banzato, 2000). Acerca dos sintomas de diabetes mellitus, Mahan & Scott-Stump (2005) descrevem como sintomas hiperglicemia, sede excessiva, micção frequente, perda de peso significativa. Considera-se que, para este domínio, a categoria sintoma seja pertinente para agrupamento e estabelecimento de relações semânticas.

- **Análise:** esta categoria abarca um conjunto de técnicas, exames e índices aplicados para fins de monitoramento da doença e diagnósticos.

- **Instrumento:** considera-se que esta categoria poderá abarcar os objetos, aparelhos, artefatos, meio ou coisa que auxilia, executa, realiza processos, técnicas e procedimentos.

- **Substância química:** considera-se que esta categoria deverá abarcar substâncias e compostos químicos, semelhante à categoria apresentada no DeCS (2017).

- **Medicamento:** esta categoria compartilha a mesma definição exposta no Decreto-Lei n.º 20/2013, de 14 de fevereiro, referente à sétima alteração ao Decreto-Lei n.º 176/2006, de 30 de agosto, que estabelece o regime jurídico dos medicamentos de uso humano, transpondo a Diretiva n.º 2010/84/UE do Parlamento Europeu e do Conselho, de 15 de dezembro de 2010⁵⁵, no qual considera medicamento toda a substância ou associação de substâncias apresentada como possuindo propriedades curativas ou preventivas de doenças em seres humanos ou dos seus sintomas ou que possa ser utilizada ou administrada no ser humano para fins de estabelecer um diagnóstico médico ou, exercendo uma ação farmacológica, imunológica ou metabólica, restaurar, corrigir ou modificar funções fisiológicas.

- **Processo:** esta categoria deverá abarcar conceitos relacionados a métodos, sistemas, modo de fazer algo, assim como um conjunto de manipulações para obter um resultado. Ela inclui os processos biológicos, fisiológicos, moleculares, celulares, entre outros associados a doença diabetes. Por exemplo, os processos fisiológicos necessários para o funcionamento dos

⁵⁵ Disponível em: http://www.infarmed.pt/documents/15786/1068535/035-G1_DL_20_2013_1ALT.pdf acessado em 8 de setembro de 2017.

tecidos integrados vivos, órgãos e organismos, como a digestão, metabolismo, nutrição, envelhecimento. Ou, ainda processos moleculares como glicogênese.

- **Teste laboratorial:** esta categoria deverá abarcar um conjunto de testes e exames clínicos.

- **Tratamento médico e Tratamento nutricional:** esta categoria refere-se ao modo de proceder, tratar, cuidar e alimentar. Na área da Saúde, tratamento refere-se a um conjunto de meios de qualquer tipo, sejam higiênicos, farmacológicos, cirúrgicos ou físicos, com o objetivo de obter a cura ou o alívio de enfermidades ou sintomas (Comel, 2013). Dependendo do contexto, pode optar-se para as categorias Tratamento farmacológico, Tratamento de emergência, como é exposto no DeCS (2017), especificamente na categoria Técnicas e Equipamentos analíticos, diagnósticos e terapêuticos. Ou, ainda optar pelo termo Terapêutica. Nesta tese, apoiando-se em Mahan & Escott-Stump (2005), optou-se pelas categorias Tratamento Médico e Tratamento Nutricional.

- **Áreas do conhecimento:** esta categoria contempla ramos de conhecimento associados ao assunto diabetes.

As categorias são dependentes do contexto e auxiliarão no estabelecimento e identificação dos relacionamentos interconceituais.

5.4 RELAÇÕES SEMÂNTICAS NO DOMÍNIO DA DIABETES

As relações hierárquicas, conforme apresentado ao longo deste texto, caracterizam-se por apontar alguma subordinação entre os conceitos, enquanto as relações associativas necessitam de maior fundamentação conceitual e decorrem da experiência, do empirismo, do campo conceitual e são dependentes do contexto, o qual apoia os elos e as associações interconceituais (Dodebei, 2002; Mostafa & Cruz, 2012). As relações associativas não estão sempre explícitas, quantificadas e detalhadas para as distintas e diversas áreas do conhecimento.

Lancaster (1986, p.46) aponta que não é possível estabelecer regras precisas o bastante para relacionamentos associativos do mesmo modo que se

pode fazer para as relações hierárquicas. O melhor que se pode dizer é que dois termos quaisquer cujos significados são relacionados, mas que aparecem em diferentes hierarquias, são possíveis candidatos a uma relação associativa.

No DeCS (2017), os termos relacionados indicados para o descritor diabetes mellitus são: Diabetes insípido; Dieta para diabéticos; Gastroparesia; Intolerância à glucose, Produtos Finais de Glicosilação Avançada; Estado Pré-Diabético; Escleredema de adulto. Nesse sentido, considerando como um exercício que auxilia a compreensão do domínio em tela, descrevem-se os termos relacionados, acompanhados pela definição e afirmações lógicas, as quais são relevantes para compreensão e detecção da relação associativa.

Quadro 31 - Relações associativa no DeCS para o termo diabetes

TR – DECS	DEFINIÇÃO; CONTEXTO; RELAÇÃO
Diabetes Insípido	<p>Definição: Doença caracterizada por micção frequente, excreção de grandes quantidades de urina diluída e sede excessiva. A etiologia do diabetes insípidus compreende deficiência do hormônio antidiurético como HAD (Hormônio antidiurético) ou VASOPRESSINA secretado pela NEURO-HIPÓFISE, resposta do rim ao HAD e regulação hipotalâmica da sede prejudicadas. Também é classificada como Doenças Uretais.</p> <p>Contexto: Diabetes mellitus e insipidus são duas patologias diferentes com a única semelhança que é o próprio diabetes, isto é, poliúria estabelecida (Figueiredo & Rabelo, 2009).</p> <p>Sugestão para relação: Oposição.</p>
Dieta para diabéticos	<p>Definição: Dieta prescrita para pacientes que limita a quantidade de comida com alto índice glicêmico. É classificado como processo ou técnica.</p> <p>Contexto: ingestão de dieta adequada é praticamente duas vezes mais efetiva que o tratamento farmacológico no controle do DM. (Costa, Balga, Alfenas, & Cotta, 2011). Recomenda-se aos pacientes diabéticos tipo 2 atenção quanto ao controle de peso, considerando que o tratamento dietético dirigido à redução do peso corporal.</p> <p>Sugestão para relação: Ação.</p>

Gastroparesia	<p>Definição: Retardo crônico do esvaziamento gástrico. Pode ser causada por disfunção motora ou paralisia dos músculos do estômago ou pode estar associada com outras doenças sistêmicas como Diabetes Mellitus (DeCS, 2017).</p> <p>Contexto: Diabetes causa Gastroparesia. A diabetes pode, por vezes, provocar danos no nervo vago devido aos altos níveis de açúcar no sangue durante um prolongado período de tempo. Um nervo transporta informações de e para o cérebro (acerca do organismo). O cérebro não comunica devidamente com o estômago quando o nervo vago está danificado, o que implica que os alimentos permaneçam no estômago por tempo demasiado antes de passarem ao intestino delgado.</p> <p>Sugestão para relação: Causa/efeito.</p>
Intolerância à glucose	<p>Definição: Estado patológico no qual o nível de glicemia é menor que aproximadamente 140mg/100ml de plasma em jejum, e acima de 200 mg/100ml aos 30, 60 e 90 minutos durante um teste de tolerância a glucose. Esta afecção é encontrada frequentemente em diabetes mellitus, mas também ocorre com outras doenças e desnutrição</p> <p>Contexto: De acordo com um relatório da Organização Mundial de Saúde, "IGT denota um estado de maior risco de progressão para o diabetes, embora também foi observado que muitos retornam ao normal. A diferença entre intolerância à glicose e diabetes em toda a regra é que os níveis de intolerância à glicose de açúcar no sangue são altos, mas não o suficiente para fazer um diabético.</p> <p>Sugestão para relação: Risco.</p>
Produtos Finais de Glicosilação Avançada	<p>Definição: Produtos derivados da reação não enzimática da glucose e proteínas in vivo que exibem uma pigmentação amarelo-amarronzada e uma habilidade em participar das ligações cruzadas entre proteínas. Estas substâncias estão envolvidas em processos biológicos relacionados ao turnover de proteínas e acredita-se que seu acúmulo excessivo contribui para as complicações crônicas da Diabetes mellitus.</p> <p>Contexto: Entre os mecanismos pelos quais os Produtos Finais de Glicosilação Avançada (AGEs) podem contribuir para o desenvolvimento e a progressão das complicações vasculares do diabetes, a interação desses compostos com receptores presentes na superfície de diversos tipos celulares merece papel de destaque.</p> <p>Os AGEs são considerados importantes mediadores patogênicos das complicações do diabetes, convencionalmente classificadas em micro ou macroangiopatias.</p> <p>Sugestão para Relação: Risco para doenças micro e macrovasculares em diabéticos.</p>

Estado Pré-Diabético	<p>Definição: Período antes do desenvolvimento do diabetes sintomático. Por exemplo, pode-se observar alguns fatores de risco em indivíduos que desenvolverão resistência à insulina como no diabete mellitus tipo 2</p>
	<p>Contexto: Pré-diabetes é risco potencial de desenvolvimento da diabetes</p>
	<p>Sugestão para relação: Risco para Diabetes</p>
Escleredema do adulto	<p>Definição: Endurecimento difuso, simétrico, semelhante à madeira, não depressível da pele de etiologia desconhecida, tipicamente começando na face, cabeça ou pescoço e espelhando-se progressivamente para comprometer os ombros, braços, tórax e às vezes locais extracutâneos, usualmente precedido por qualquer um dentre vários processos infecciosos, especialmente uma infecção estafilocócica. Ele ocorre em associação com diabetes mellitus na maioria dos casos, predominantemente em mulheres, em geral desaparece espontaneamente dentre de 6 de meses a 2 anos.</p>
	<p>Contexto: Pode estar associada com infecções bacterianas ou virais, alterações hematológicas e diabetes melito (DM). De etiologia desconhecida, pode estar associada com infecções bacterianas ou virais, alterações hematológicas e diabetes melito (DM). Em metade dos casos publicados de EB, há associação com DM. As características mais comuns nesses pacientes são: adultos do sexo masculino, DM de longa duração, mal controle metabólico e presença de complicações específicas do DM.</p>
	<p>Sugestão para relação: Diabetes é risco para Escleredema.</p>

Fonte: autoria própria.

Relativamente ao quadro apresentado acima, o aspecto que deve ser destacado refere-se à importância do contexto para analisar e interpretar os conceitos. Considera-se que o critério de escolha e determinação das relações semânticas deve se apoiar-se no contexto de uso das palavras. Certamente, as sugestões para relacionamentos expostas aqui podem variar, mas representam mais um exercício de interpretação. Isto posto, segue abaixo um quadro com exemplos de relações semânticas sobre diabetes.

Quadro 32 - Exemplos de relações semânticas sobre diabetes

RELAÇÕES SEMÂNTICAS	MANIFESTAÇÃO
Termos hierárquicos	Diabetes / Tipo 1 e tipo 2 Dieta/ Dieta lowcarb Hormônios/ Insulina e Guclagon
Termos equivalentes	Diabetes juvenil/Diabetes Mellitus Diabetes/Diabete Quilocaloria/Caloria Tensão arterial/Pressão do sangue Sacarose/Açúcar Atividade física/Exercício Resistência a insulina/Insulinoressistência Acidente Vascular Cerebral/Derrame Bomba de Insulina/ Bomba Infusora de Insulina Pré-diabetes/Intolerância á glicose
Ação/resultado da ação	Dieta/redução do peso Dieta/Controle glicêmico
Agente (ou instrumento) / processo (ou operação)	Pâncreas/produção de insulina Célula Beta/Secreção de insulina Insulina/Glicogênese Célula Alfa/ Produção do glucagon Sulfonilureia/ Aumenta a produção de insulina Amiláse/Digestão
Causa/Efeito	Resistência à insulina/hiperglicemia Insulina/Hipoglicemia Guclagon/Hiperglicemia Hiperglicemia/Cetoasidose diabética Retinopatia diabética/Cegueira Hipoglicemia ou hiperglicemia/Coma
Risco	Obesidade/Diabetes Hiperglicemia/Pé diabético Pré-diabetes/Diabetes Diabetes gestacional/ Diabetes Diabetes/Hipertensão
Processo/etapas do processo	Glicogênese/síntese, catalisação, hidrólise. Diálise/Difusão, ultrafiltração, convecção
Objeto/Mecanismo de medida	Proteína/Quilocaloria Hidratos de carbono/Quilocaloria Glicemia/Miligramas por decilitro (mg/dl.) Peso/Índice de Massa Corporal (IMC)
Matéria-prima/produto	Fruta/Frutose Leite/Lactose Biguanidina/Metfomina Fibras/Cereais
Oposição	Sedentarismo/Atividade física Guclagon/Insulina Hiperglicemia/Hipoglicemia

Fonte: Autoria própria

Conforme já abordado, é evidente que a tag Diabetes ocorre com tags referentes a outras afecções, demonstrando uma interrelação justificada pela característica de comorbidade das doenças crônicas. Em alguns casos, as

doenças são um risco ou fator de risco para desenvolver a diabetes, como nas manifestações apresentadas no quadro acima.

Adicionalmente, as indexações livres (as tags atribuídas no CiteULike) indicam essa relação, conforme os exemplos abaixo:

Quadro 33 - Tag diabetes, risco e fator de risco

Diabetes, Risco e Fator de risco
cancer diabetes risk type1 type2
cardiometabolic cardiovascular diabetes qol risk type1 type2
assessment cardiometabolic cardiovascular diabetes risk type1
cardiovascular diabetes mortality risk

Fonte: Autoria própria

Nesse sentido, considerando a tag risco/fator de risco como uma das mais expressivas no corpus e presentes nos contextos, questiona-se: qual a possibilidade de inserir Risco como uma relação estruturante no domínio da diabetes, considerando o “caráter probabilístico que instrui o raciocínio causal em saúde”? (Ayres, 2011).

O conceito de risco é abordado em distintas áreas do conhecimento com as suas devidas perspectivas. No âmbito das Ciências Sociais tem sido abordado sob a perspectiva de como o indivíduo percebe as situações de risco, seja como cidadão, trabalhador, sem ignorar fatores subjetivos (éticos, morais, culturais) que direcionam suas opções. Em Ciência Econômicas segue a perspectiva de transformar as incertezas em probabilidades, quantificando os riscos para avaliar custos e possíveis perdas, enquanto a Engenharia analisa o impacto da introdução de modernas tecnologias na sociedade, seja por meio de um método quantitativo, como medições ambientais, relação custo-benefício, seja por meio da discussão do gerenciamento do risco (Guilam, 1996). Estes exemplos realçam que risco e fator de risco incorporam aspectos quantitativos e não-quantitativos, atingindo todas as dimensões da vida.

O termo risco começou a aparecer no Código Sanitário em 1961, empregado equivocadamente como risco de saúde e tem sido usado com significado distintos: como perigo ou ameaça (por exemplo, ao efeito danoso dos agrotóxicos, raticidas e inseticidas) e como causa de risco (Costa, 1999).

Apesar de difícil precisão acerca de risco e fatores de risco, consolidou-se o uso da palavra fator de risco. Neste ínterim, Rey (2006) define fator de risco como um aspecto do comportamento pessoal ou do estilo de vida, da exposição ao meio ambiente, ou uma característica própria ou herdada do indivíduo que se sabe, tendo por base a evidência epidemiológica, estarem associados com condições importantes de se prevenir para proteger a saúde. O fator de risco representa uma probabilidade maior de ser atingido por determinada afecção ou dano. Fala-se de alto risco quando a intensidade de um fator ou soma de vários fatores elevam consideravelmente o risco, como o de pessoas que fumam muito e têm também hipertensão ou colesterolemia alta, formando um grupo de alto risco para infarto do miocárdio. A pobreza reúne por si só um grande número de fatores de risco, principalmente para doenças transmissíveis e morbimortalidade infantil.

Essa definição converge com o que foi observado nas indexações livres: a ocorrência da tag diabetes com as tags risco, mortalidade, prevenção, epidemiologia, revelando que o assunto em tela, no CiteULike tem sido abordado maioritariamente no escopo da Epidemiologia.

A Epidemiologia é entendida como o estudo da distribuição e dos determinantes das doenças ou condições relacionadas à saúde e às populações especificadas, assim como a aplicação desses estudos para controlar problemas de saúde (Lima-Costa & Barreto, 2003).

Almeida-Filho e Coutinho (2007) expõem que a Epidemiologia, designada inicialmente como um "método", pode ser compreendida como uma ciência que se baseia num tipo peculiar de aplicação setorial da lógica formal. Trata-se de um modo de raciocinar caracterizado por uma ontologia particular – baseada nas categorias de conjunto e de elementos – e por uma lógica operativa padronizada – que obedece a regras específicas de validação e de inferência – o determinismo causalista.

Complementarmente, considerando seus momentos históricos⁵⁶, a

⁵⁶ Rodrigues Silva (1990) citado por Guilam (1996) sugere 3 momentos histórico na evolução da Epidemiologia: epidemiologia da constituição pestilencial e dos miasmas, relacionadas às

epidemiologia define risco como a probabilidade de um membro de uma determinada população desenvolver uma dada doença em um período de tempo. É nesta perspectiva que o Almeida-Filho (1989) propõe um modelo de determinação causal das doenças e que, a partir da utilização do conceito de risco, a Epidemiologia abarcaria as doenças crônicas, como hipertensão, câncer, entre outras, uma vez que era restrita ao âmbito das doenças infecciosas agudas, segundo Guilam (1996).

A autora ainda ressalta que, na medida que a Epidemiologia incorporou a noção de risco e seus frutos, estimulou a discussão sobre prevenção de doenças através do controle de fatores inerentes ao indivíduo ou a grupos de indivíduos e que a utilização deste conceito não enfatiza mais a causa, mas a associação de determinados fatores com as patologias.

Complementarmente,

o estudo das doenças não-transmissíveis trouxe um conjunto mais amplo de “causas” à “multicausalidade”, ao ponto de se acrescer mais um nível de intervenção, o “primordial”, voltado ao controle de “modelos” ou “padrões de vida” ditados pela condição social, econômica ou cultural das populações. O “estilo de vida” do sujeito passou a ser o alvo de prevenção e um “estilo” admite vários fatores passíveis de correção (Beaglehole, Bonita & Kjellstroem, 1993). Como os epidemiologistas foram capazes de separar estes aspectos como “condições determinantes” da causação de outros tipos como “fatores causais específicos”, configurou-se um quadro de “causa e contexto”. Com isso, ganhou relevância geral o uso do termo “fatores de risco” como forma de configurar o contexto, uma noção que já estava em uso na epidemiologia, especialmente para se estudar fatores que condicionam a ocorrência e evolução de doenças crônicas (Minayo & Miranda, 2002, p.89)

Os conceitos de Risco e as práticas que lhe correspondem no campo da Saúde podem ser reunidos em três grupos (Almeida-Filho & Coutinho, 2007):

teorias pré-pasteurianas; Epidemiologia dos modos de transmissão, cujo início coincide com o da era bacteriológica, constituindo o modelo explicativo dominante desde o final do séc. passado, explorando fortemente o comportamento das doenças infecto-contagiosas; e por fim, Epidemiologia dos fatores de risco, como um modelo hegemônico nos nossos dias, que se propõe a dar conta das doenças crônico-degenerativas.

- **Risco como indicador de causalidade** (ou resíduo da probabilidade): Trata-se de reconhecer e reafirmar sua base indutiva. Esse conceito particular subsidia modelos de prevenção de doenças ou eventos mórbidos.
- **Risco como perigo estruturado**: Tal conceito subsidia largamente modelos de intervenção nos campos da Saúde Ambiental e Ocupacional.
- **Risco como emergência**: Trata-se, nesse caso, de explicitar a base filosófica da contingência, articulada como processos de emergência em modelos de complexidade, subsidiando os modelos de Promoção da Saúde e Vigilância em Saúde.

Certamente, o conceito de risco e fator de risco implica questões complexas e paradoxais como, por exemplo, o uso equivocado do conceito pela mídia, as generalizações dos estados de quase-doença, o conceito sustenta-se como aplicação da lógica indutiva e da lógica dedutiva, respectivamente, a problemas particulares e problemas gerais de determinação de fenômenos de saúde-doença (Vaz, Pombo, Fantinato & Peclly, 2007).

Pertinente a Organização da Informação, especificamente sobre as teorias de classificação, a noção de risco e fator de risco, em certa medida, retoma fundamentos das categorias causalidade e contingência de Aristóteles (Almeida-Filho & Coutinho, 2007).

Isto posto, orientando-se pelas tags mais expressivas, pela ocorrência entre os conceitos de risco e fator de risco, pelos contextos de uso, nesta tese, considera-se Risco como uma possibilidade de relação associativa, que implica raciocínio e probabilidade.

Após o exposto, apresenta-se a síntese das relações semânticas. O quadro é composto pelo conceito, categoria, relacionamento, afirmação e propriedades, que são os verbos que indicam o relacionamento.

Quadro 34 - Síntese das relações semânticas sobre diabetes

Conceitos	Categorias	Relacionamentos	Afirmações	Propriedades
Acidente vascular cerebral	Doença macrovascular	Risco	Diabetes é um risco para Acidente Vascular Cerebral .	Possibilita

Bomba de insulina	Instrumento	Ação	Bomba de insulina mantém o controle glicêmico .	Mantém
Célula beta	Instrumento	Processo	Célula beta secreta insulina .	Secreta
Colesterol	Substância química	Risco	O colesterol em excesso é risco para diabetes . O colesterol em excesso pode causar diabetes .	Possibilita
Controle glicêmico	Tratamento médico	Processo	O controle glicêmico controla diabetes mellitus .	Controla
Dieta	Tratamento nutricional	Ação	Dieta regula a glicemia Dieta previne doenças micro e macrovasculares	Regula Previne
Doença cardiovascular	Doença	Risco	Diabetes é risco para doença cardiovascular	Possibilita
Glicemia	Medida	Ação	Glicemia controla a diabetes	Controla
Hemoglobina glicada	Técnicas	Ação	Hemoglobina glicada avalia o controle glicêmico	Avalia
Hiperglicemia	Sintoma	Causa/Efeito	Hiperglicemia causa Cetoasidose diabética Hiperglicemia é sintoma de diabetes	Causa
Hipertensão	Doença macrovascular	Risco	Diabetes é um risco para hipertensão	Possibilita
Hipoglicemia	Sintoma	Causa/Efeito	Hipoglicemia é um sintoma de diabetes	
Infusora subcutânea contínua com insulina	Tratamento	Processo	Infusora subcutânea contínua com insulina controla a glicemia	Controla
Índice glicêmico	Medida	Ação	Índice glicêmico mede a glicose	Mede
Insulina	Substância química – hormônio	Processo	Insulina controla glicemia	Controla
Nefropatia	Doença macrovascular	Risco	Diabetes é um risco para nefropatia	Possibilita
Neuropatia	Doença	Risco	Diabetes é um risco para neuropatia	Possibilita
Obesidade	Doença nutricional e endócrina	Risco	Obesidade é um risco para diabetes	Possibilita
Pé diabético	Doença macrovascular	Causa/efeito	Diabetes causa pé diabético	Causa
Resistência á insulina	Processo e doença	Causa/efeito	Resistência a insulina causa diabetes	Causa
Retinopatia	Doença Microvascular	Risco	Diabetes é risco para retinopatia	Possibilita

Tipo 1	Doença nutricional endócrina	Hierárquica	Diabetes tipo 1 é um tipo de Diabetes Mellitus	É
Tipo 2	Doença nutricional endócrina	Hierárquica	Diabetes tipo 2 é um tipo de Diabetes Mellitus	É
Vitamina D	Substância química	Ação	Vitaminas previne diabetes, doenças micro e macrovasculares.	Previne
Epidemiologia	Área do conhecimento			Estuda Investiga
Nutrição	Área do conhecimento			Estuda Investiga

Fonte: autoria própria.

No quadro exposto, verifica-se a predominância das as relações associativas, verifica-se a predominância da relação Risco (8), Ação (8), Processo (4) e Causa/efeito (4) e Área do conhecimento (2). Quanto ao inventário de relações semânticas, Green (2001) argumenta que ele inclui tanto um conjunto fechado de relacionamentos (incluindo principalmente relações hierárquicas e de equivalência) e um conjunto aberto de relacionamentos. Toda vez que um novo verbo é inventado, por exemplo, surge o potencial para a introdução de uma nova relação conceitual.

Relativamente à estruturação da ontologia, a conceitualização, ou seja, o processo de criação de um modelo conceitual, no qual o modelador faz abstrações para representar parte da realidade que lhe interessa, privilegia características mais importantes para os objetivos que se buscam e, neste sentido, estas características podem variar de um grupo para o outro (Teixeira, 2009). Entende-se por modelo conceitual os processos intelectuais e metodológicos necessários para a definição de um sistema de conceitos e relações para a formação de um sistema terminológico (Gonçalves & Souza, 2013).

Neste sentido, retoma-se as perspectivas de Gruber (1993) que argumenta que, quando o conhecimento de um domínio é representado por um formalismo declarativo, o conjunto de objetos e relações entre eles, passíveis de descrição, são espelhadas no vocabulário representacional com o qual o sistema baseado em conhecimento representa o conhecimento. Pode-se descrever a ontologia de um programa ao definir um conjunto de termos

representacionais. As definições associam nomes de entidades do universo do discurso (classes, relações, funções ou outros objetos) com textos legíveis para pessoas, os quais descrevem os nomes que se deseja representar e axiomas formais que restringem a interpretação e a formação desses nomes. Isto posto, apresenta-se na figura 29 edição e a formalização preliminar da ontologia no software Protégé. As categorias, apresentadas no quadro 31, foram definidas como classes e subclasses para estabelecer a relação hierárquica entre os conceitos, conforme consta na figura 30.

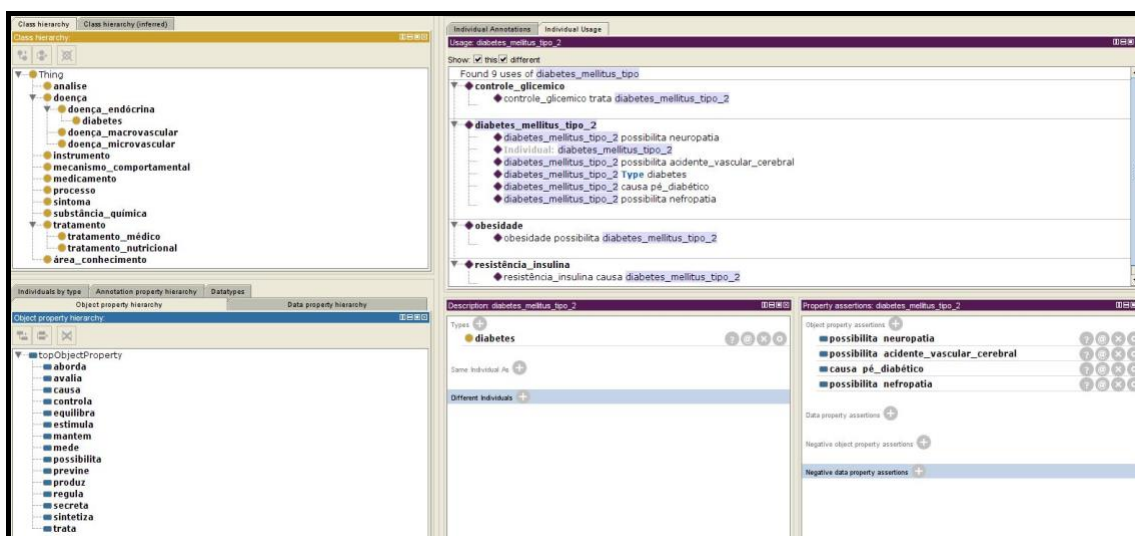


Figura 29 - Edição preliminar da ontologia
Fonte: Autoria própria

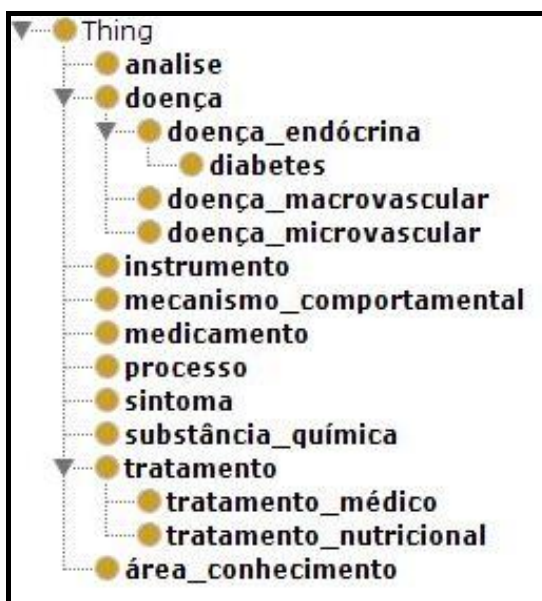


Figura 30 - Ontologia: Classes e subclasses
Fonte: Autoria própria

As manifestações exemplificadas nos quadros 31 e 33 foram definidas como indivíduos, conforme consta na figura 31. Já as relações entre diferentes indivíduos, com exceção da relação hierárquica, estão definidas como *object properties*, conforme a figura 32.



Figura 31 - Ontologia: indivíduos
Fonte: Autoria própria.



Figura 32 - Ontologia: propriedades
Fonte: autoria própria

Por fim, apresenta-se a figura 33, que expõe como as diferentes relações, triplos semânticos, foram definidas. Por exemplo: sujeito (obesidade), predicado (possibilita), objeto (diabetes_mellitus_tipo_2).



Figura 33 - Triplos semânticos

Fonte: Autoria própria.

Certamente, reconhece-se que, no tocante a ontologia, o que foi apresentado trata-se da conceitualização, ou seja, uma composição de conceitos e relações fundamentais para a estruturação.

CONSIDERAÇÕES FINAIS

A folksonomia existe em decorrência de uma nova conjuntura social e tecnológica referente ao armazenamento, tratamento, disseminação e acesso a informação. Ela integra distintos objetos digitais, promove o cruzamento de informações, a partir de um processo pessoal de etiquetagem, atendendo a interesses e perspectivas individuais, mas que promove uma abordagem coletiva, colaborativa e dinâmica para a organização e recuperação da informação.

Desenvolver um estudo que articule os contributos da folksonomia, da ontologia e a explicitação de relações semânticas no âmbito da Saúde representa um desafio que se converte em motivação. Primeiramente, porque explorar as folksonomias e expor seus possíveis contributos implica em lidar com suas fragilidades e limitações, frente aos estabelecidos e tradicionais sistemas de organização da informação e suas práticas. Posteriormente, apesar de haver sistemas de organização da informação consolidados na área da saúde, compreender o domínio e sua terminologia corresponde a uma atividade complexa.

As motivações para investigar as relações semânticas assentam-se também na necessidade de explicitar a natureza da relação entre os conceitos do domínio da diabetes, considerando a relação semântica um dos elementos comuns à folksonomia e à ontologia. Neste sentido, enfatizaram-se os aspectos de compreensão e explicitação de relacionamento e não ao desenvolvimento de uma ontologia, deixando sua continuidade para trabalhos futuros.

Um aspecto relevante é que a literatura e os sistemas de organização expõem e indicam a relação associativa (termo relacionado) entre os conceitos, mas poucos explicitam e exemplificam a relação que, por vezes, também está restrita a um universo de conhecimento. Complementarmente, ainda não há bases teóricas suficientes para relações associativas (Cintra, 2002)

De forma similar, o mesmo ocorre no estabelecimento de categorias, em que há pouca fundamentação para defini-las, o que leva a um entendimento ambíguo e reforça a importância de se amparar no contexto de uso e no compromisso ontológico.

Nesse sentido, perceber **em que medida as relações semânticas implícitas nas folksonomias permitem estruturar uma ontologia sobre diabetes, faz depreender que**

- Embora a folksonomia não seja uma **única e abrangente** fonte para compilação exaustiva da variedade de conceitos que integrem um sistema de organização da informação em ambientes tradicionais e especializados, fornece subsídios para estruturação e configuração de uma ontologia (hipótese 1), adequada ao seu público. Nesta investigação, por exemplo, com base nos resultados, foi possível supor o desenvolvimento de uma ontologia cujo objetivo seja apoiar o paciente.
- Em convergência a isso, há um núcleo essencial de relações semânticas implícitas nas folksonomias que são basilares para estruturar ontologias (hipótese 2).
- O estabelecimento da relação semântica, se dá também a partir de informação contextual. Dias e Naves (2007) ao tratarem sobre a extração de conceitos, enfatizam que conceito, assunto e contexto são interdependentes, e que a informação contextual pode incluir informação sobre o ambiente de recuperação no qual o documento existe. Isso ficou evidente na interpretação, explicitação e estabelecimento das relações semânticas associativas.

Concernente aos objetivos propostos, é correto afirmar que foram realizados satisfatoriamente. Referente à sistematização de um núcleo de relações semânticas na perspectiva da CI, isso foi abordado com maior profundidade e de forma bastante consistente no capítulo 3, com amparo das

normas específicas para construção e manutenção de sistemas de organização da informação, assim como recorreu-se à Linguística, à Semântica, à Lógica e à Terminologia, revelando uma diversidade nas denominações acerca da tipologia das relações semânticas (relação de antonímia, relação de oposição ou relações paradigmáticas, relação de inclusão, relação hierárquica), porém adequadas à perspectiva teórico-metodológica de cada área do conhecimento.

Relativo aos aspectos da representação da informação sobre diabetes, a lista de tags apresentada demonstra seus conceitos e terminologia. Pelo esboço das áreas limítrofes, pelo exercício de categorização, verificação das tags mais expressivas e reconhecimento dos qualificadores, foi possível perceber os aspectos mais significativos da representação da informação sobre diabetes em contexto digital, a partir do sistema baseado em folksonomia.

Quanto à identificação dos padrões de relações semânticas sobre diabetes, ela fundamentou-se nas tags mais expressivas no corpus. Tanto a categorização quanto as relações são dependentes do contexto, não são estáticas, mas, sim, dinâmicas. A exemplo disso, a sugestão de adotar risco como uma relação entre os conceitos assenta-se na ocorrência da tag diabetes com outras doenças. Neste caso, consideram-se outras alternativas como criar a categoria Complicação e uma explicitação desse relacionamento como, por exemplo, acidente vascular cerebral pode ser uma complicação de diabetes. Entretanto, sugeriu-se risco como relação devido a sua ocorrência expressiva e pela perspectiva de probabilidade de um evento ocorrer, distinguindo-se da relação de causa/efeito, que são mais evidentes. Assim, é pertinente, no escopo da OI, discutir o conceito de risco como uma relação estruturante no domínio da Saúde, revisitando os aspectos da causalidade, da lógica e o pensamento de Aristóteles.

Ademais, há poucas abordagens acerca da explicitação das relações semânticas associativas. Neste sentido, o conteúdo apresentado nesta tese referente a explicitação dessas relações contribui também para esclarecer e demonstrar os processos cognitivos, os elementos e o raciocínio empírico empreendido no estabelecimento das relações associativas, pois estes

aspectos têm implicações determinantes no desenvolvimento, estruturação e na qualidade de sistemas de organização da informação, assim como nas demais abordagens práticas que envolvam o processamento de linguagem natural.

Referente aos contributos para estruturação de uma ontologia a partir de um núcleo de relações semânticas extraído da folksonomia evidencia-se a importância do processo cognitivo e da percepção humana para representações artificiais da realidade. Certamente, uma ontologia não se faz apenas com conceitos, relações e definições, porém são os elementos fundamentais para o recorte de um universo de conhecimento, de forma a garantir o compromisso ontológico.

Concernente ao percurso metodológico, considera-se que as abordagens foram adequadas ao desenvolvimento da pesquisa, na medida em que viabilizou delimitação e compreensão de um determinado contexto, o qual integrou o compartilhamento de conceitos, seguidos da identificação de padrões e atos de comunicação. Também, atendeu aos propósitos de uma investigação de caráter empírico.

Acerca das limitações da pesquisa, reconhece-se que seria conveniente realizar alguns procedimentos de forma automática, como a coleta das tags e sua quantificação, pois auxiliariam na visualização e na análise dos dados. Todavia, considera-se que este procedimento manual foi oportuno para conhecer o domínio e se familiarizar com os conceitos e relações.

As principais dificuldades encontradas para o desenvolvimento desta investigação referem-se à compreensão das especificidades terminológicas do domínio, nomeadamente, no estabelecimento de categorias, de relacionamentos, na busca por definições claras e precisas que fundamentassem as escolhas, mesmo consultando materiais de referência. Tal dificuldade foi reduzida com a consulta ao especialista.

Após o exposto, seguem algumas sugestões para possíveis desdobramentos e trabalhos futuros:

- Desenvolver e aplicar uma ontologia de apoio aos diabéticos.
- Discutir a harmonização entre os conceitos, definições e categorias nos sistemas de organização da informação estabelecidos, como DeCS.
- Analisar os verbos adequados para indicar os relacionamentos.
- Reconhecer métricas (quantificar atributos, propriedade ou características para uma entidade) para ontologias no domínio da diabetes, com vistas a padronização.
- Correlacionar a organização da informação e comportamento informacional e literacia dos diabéticos.
- Ampliar o domínio para doenças crônicas não-transmissíveis.

Por fim, as contribuições mais significativas desta investigação referem-se aos procedimentos metodológicos para analisar as relações semânticas entre os conceitos do domínio da diabetes, assim como a explicitação e exemplificação dos relacionamentos. Espera-se que os dados levantados, nomeadamente as indexações livres, possam ser reutilizadas para continuidade e andamento a outras investigações.

REFERÊNCIAS

- Aitchison, J., & Clarke, S. D. (2004). The thesaurus: a historical viewpoint, with a look to the future. *Cataloging & classification quarterly*, 37(3-4), 5-21.
- Aitchison, J., & Gilchrist, A. (1987). *Thesaurus construction: a practical manual*.
- Aitchison, J., Gilchrist, A., & Bawden, D. (2000). *Thesaurus construction and use: a practical manual*. Psychology Press.
- Al-Khalifa, H. S., & Davis, H. C. (2007, September). Towards better understanding of folksonomic patterns. In *Proceedings of the eighteenth conference on Hypertext and hypermedia* (pp. 163-166). ACM
- Almeida Filho, N. D. (1989). *Epidemiologia sem números*. Rio de Janeiro: Campus.
- Almeida-Filho, N., & Coutinho, D. (2007). Causalidade, contingência, complexidade: o futuro do conceito de risco. *Physis-Revista de Saúde Coletiva*, 17(1).
- Almeida, M. B., & Bax, M. P. (2003). Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. *Ciência da Informação, Brasília*, 32(3), 7-20.
- Almeida, M. B., Teixeira, L. M., Coelho, K. C., & Souza, R. R. (2010). Relações semânticas em ontologias: estudo de caso do Blood Project. *Liinc em Revista*, 6(2), 384-410.
- Aranalde, M. M. (2009). Reflexões sobre os sistemas categoriais de Aristóteles, Kant e Ranganathan. *Ciência da Informação*, 38(1).
- Araújo, C. A. Á. (2009). Fundamentos teóricos da classificação. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis*, (22), 117-140.
- Assis Pinto, J. (2011). *Indicadores de qualidade da informação em sistemas baseados em Folksonomia: uma abordagem semiótica*. Dissertação, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais. Brasil.
- Ayres, J. R. D. C. M. (2011). Desenvolvimento histórico-epistemológico da Epidemiologia e do conceito de risco. *Cadernos de Saúde Pública*, 27(7), 1301-1311.
- Bamidis, P. D., Kaldoudi, E., & Pattichis, C. (2009). From taxonomies to folksonomies: a roadmap from formal to informal modeling of medical concepts and objects. In *Information Technology and Applications in Biomedicine, 2009. ITAB 2009. 9th International Conference on* (pp. 1-4). IEEE.

Barbisan, L. B. (2013). Semântica Argumentativa. *Semântica, semânticas: uma introdução*, 1, 19-30.

Barcellos Almeida, G. M. (2003). O percurso da Terminologia: de atividade prática à Consolidação de uma disciplina autônoma. *Tradterm*, 9, 211-222

Barcellos Almeida, G. M. (2009). A Teoria Comunicativa da Terminologia e a sua prática. *ALFA: revista de Linguística*, 50(2).

Barros, L. A. (2006). Aspectos epistemológicos e perspectivas científicas da terminologia. *Ciência e cultura*, 58(2), 22-26

Basso, C. A. M., Ferreira, J. M., & da Silva, S. R. P. (2009). Uma Proposta para a Evolução de Ontologias a partir de Personomias em Sistemas Baseados em Tagging. In XXIX Congresso da Sociedade Brasileira de Computação/VII Encontro Nacional de Inteligência Artificial.

Bean, A., & Green, R. (2001). Relationships in the Organization of Knowledge (Vol. 2). Springer Science & Business Media.

Beghtol, C. (1986). Semantic validity: concepts of warrant in bibliographic classification systems. *Library resources & technical services*, 30(2), 109-125

Biblioteca Virtual em Saúde. (2016). DECS – Descritores em Ciências da Saúde. Disponível em: <http://decs2014.bvsalud.org/P/decsweb2014.htm>

Binzabiah, R., & Wade, S. (2012, June). Proposed method to build an ontology based on Folksonomy. In *Information Society (i-Society), 2012 International Conference on* (pp. 441-446). IEEE.

Biolchini, J. C. D. A. (2001). Semântica e cognição em bases de conhecimento: do vocabulário controlado à ontologia. *Datagramazero*, 2 (3).

Blair, D. C. (2003). Information retrieval and the philosophy of language. *Annual review of information science and technology*, 37(1), 3-50.

Bocato, V. R. C. (2009). A linguagem documentária vista pelo conteúdo, forma e uso na perspectiva de catalogadores e usuários. *A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias*. São Paulo: Cultura Acadêmica, 119-135.

Bonnevie, E. (2001). Dretske's semantic information theory and meta-theories in library and information science. *Journal of documentation*, 57(4), 519-534.

Brandt, M. B. (2009). *Etiquetagem e Folksonomia: uma análise sob a óptica dos processos de organização e recuperação da informação na Web*. (Dissertação de Mestrado), Universidade de Brasília, Brasília, DF.

Brandt, M. B., & Medeiros, M. B. B. (2010). Folksonomia: esquema de representação do conhecimento? *TransInformação*, 22 (2), p. 111-121.

Brooks, T. A. (1998, October). The semantic distance model of relevance assessment. In *PROCEEDINGS OF THE ANNUAL MEETING-AMERICAN SOCIETY FOR INFORMATION SCIENCE* (Vol. 35, pp. 33-44). LEARNED INFORMATION (EUROPE) LTD

Broughton, V. (2004). Essential classification.

Broughton, V. (2006). Essential thesaurus construction. Neal-Schuman.

Broughton, V. (2006). The need for a faceted classification as the basis of all methods of information retrieval", *Aslib Proceedings*, Vol. 58 Iss 1/2 pp. 49 – 72

Broughton, V. (2007). Meccano, molecules, and the organization of knowledge: the continuing contribution of SR Ranganathan. *presentation, International Society for Knowledge Organization UK chapter, London, 2.*

Buchanan, B. (1979). Theory of library classification.

Budd, J. M. (2011). Meaning, truth, and information: prolegomena to a theory. *Journal of Documentation*, 67(1), 56-74.

Cabré Castellví, M. T. (2003). Theories of terminology: Their description, prescription and explanation. *Terminology*, 9(2), 163-199.

Cabré, M. T. (1993). *La terminología: teoría, metodología, aplicaciones*. Editorial Empúries.

Café, L. M. A., & Medeiros, M. B. B. (2011). Organização do Conhecimento: teorias semânticas como base para estudo e representação de conceitos.

Café, L., & Brascher, M. (2008). Organização da informação ou organização do conhecimento. *Encontro Nacional de Pesquisa em Ciência da Informação*, 9.

Campos, J. A. G., & Souza, R. R. (2009). Análise conceitual sobre as relações semânticas em Ciência da Informação: contribuições para o desenvolvimento de ontologias. *Perspectivas em Ciência da Informação*, 14(3), 243-243.

Campos, J. A. G., & Souza, R. R. (2009). Análise conceitual sobre as relações semânticas em Ciência da Informação: contribuições para o desenvolvimento de ontologias. *Perspectivas em Ciência da Informação*, 14(3), 243-243.

Campos, M. L. A. (2001). *Linguagem documentária: teorias que fundamentam sua elaboração*. Eduff.

Campos, M. L. A., & Gomes, H. E. (2006). Princípios de organização e representação do conhecimento na construção de hiperdocumentos. *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia*, 1(1).

Campos, M. L. A., & Gomes, H. E. (2010). Taxonomia e classificação: o princípio de categorização. *DataGramZero-Revista de Ciência da Informação*, 5, 9.

Cançado, M. (2013). Semântica Lexical: uma entrevista com Márcia Cançado.

Carlan, E., & Medeiros, M. B. B. (2011). Sistemas de Organização do Conhecimento na visão da Ciência da Informação. *Revista Ibero-Americana de Ciência da Informação*, 4(2).

Castells, M (1999). *A Era da Informação: economia, sociedade e cultura*, vol. 3, São Paulo: Paz e terra, 439p.

Castells, M. (2003). *A galáxia Internet: reflexões sobre a Internet, negócios e sociedade*. Rio de Janeiro: Jorge Zahar. 27

Catarino, M. E. Baptista, A. A. (2007). Folksonomia: um novo conceito para a organização dos recursos digitais na Web. *DataGramZero*, 8(3).

Chaffin, R. (1992). The concept of a semantic relation. *Frames, Fields and Contrasts*, 253-288.

Chaffin, R., & Herrmann, D. J. (1984). The similarity and diversity of semantic relations. *Memory & Cognition*, 12(2), 134-141

Chaffin, R., & Herrmann, D. J. (1988). Effects of relation similarity on part-whole decisions. *The Journal of general psychology*, 115(2), 131-139.

Cintra, A. M. et al. (2002). *Para entender as linguagens documentárias*. 2 ed. São Paulo: Polis.

Collier, N., Doan, S., Kawazoe, A., Goodwin, R. M., Conway, M., Tateno, Y., ... & Shigematsu, M. (2008). BioCaster: detecting public health rumors with a Web-based text mining system. *Bioinformatics*, 24(24), 2940-2941.

Costa Medeiros, W. M., de Oliveira, L. A. H. G., & de Sousa, L. G. Uso de Ontologias para Acesso a Informações de Saúde Armazenadas em Bases de Dados Heterogêneas.

Costa Medeiros, W. M., de Oliveira, L. A. H. G., & de Sousa, L. G. Uso de Ontologias para Acesso a Informações de Saúde Armazenadas em Bases de Dados Heterogêneas.

Costa, J. D. A, Balga, R. S.M., Alfenas, R. D. C.G., & Cotta, R.M.M. (2011). Promoção da saúde e diabetes: discutindo a adesão e a motivação de indivíduos diabéticos participantes de programas de saúde. *Ciência & Saúde Coletiva*, 16(3).

Coutinho, C.P. (2011). *Metodologia de Investigação em Ciências Sociais e Humanas: Teoria e Prática*. Coimbra: Almedina.

Coyle, D., & Doherty, G. (2008, September). Towards ontologies for technology in mental health interventions. In *Ontologies in Interactive Systems, 2008. ONTORACT'08. First International Workshop on* (pp. 18-26). IEEE

Currás, E. (1995). *Tesauros: linguagem terminológica*. Brasília: IBICT.

Dahlberg, I. (1976). Teoria da classificação, ontem e hoje. In *Conferência Brasileira de Classificação Bibliográfica*.

Dahlberg, I. (1978). Fundamentos teórico-conceituais da classificação. *Revista de Biblioteconomia de Brasília*, 6(1), 9-21

Dahlberg, I. (1978). Fundamentos teórico-conceituais da classificação. *Revista de Biblioteconomia de Brasília*, 6(1), 9-21.

Dahlberg, I. (1978). Teoria do conceito. *Ciência da informação*, 7(2).

Dahlberg, I. (2006). *Knowledge organization*. Consultado em: junho 01, 2015, Retrieved from:

http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/knowledge_organization_Dahlberg.htm

Dahlberg, I. Teoria do conceito. (1978). *Ciência da Informação*, Rio de Janeiro, v. 7, n. 2, p. 101- 107.

Dahlberg, I. Teoria do conceito. (1978). *Ciência da Informação*, Rio de Janeiro, v. 7, n. 2, p. 101- 107.

Dias, C. A. (2000). Terminologia: conceitos e aplicações. *Ciência da informação*, 29(1), 90-92.

Dixon, B. E., & McGowan, J. J. (2010). Enhancing a taxonomy for health information technology: an exploratory study of user input towards folksonomy.

Dodebei, V. L. D. (2002). Tesauro: linguagem de representação da memória documentária. Niterói; Rio de Janeiro: Intertexto. *Interciência*, 120.

Doerr, M. (2006). Semantic problems of thesaurus mapping. *Journal of Digital information*, 1(8)

Duarte, E. A. (2010). Classificação facetada: um olhar sobre a construção de estruturas semânticas Faceted classification: a look at the construction of semantic structures. *Revista Digital de Biblioteconomia e Ciência da Informação*, 7(2), 46-58.

Evangelista, I. V., Simões, M. D. G. M., & Guimarães, J. A. C. (2016). A exaustividade e a especificidade como valores éticos no processo de indexação: uma análise baseada na literatura disponibilizada em Portugal. *Páginas a&b*, 58-75.

Evens, M. W. (2009). *Relational models of the lexicon: Representing knowledge in semantic networks*. Cambridge University Press

Farradane, J. Relational Indexing: part I. *Journal of information science*, n.1, p. 267-276, 1980b

Figueiredo, D. M., & Rabelo, F. L. A. (2009). Diabetes insipidus: principais aspectos e análise comparativa com diabetes mellitus. *Semina: Ciências Biológicas e da Saúde*, 30(2), 155-162.

Floridi, L. (2010). *Information: A very short introduction*. Oxford: Oxford University Press, 2010, 130 p.

Foskett, A. C., & de Lemos, A. A. B. (1973). *A abordagem temática da informação*.

Frohmann, B. P. (1983). An investigation of the semantic bases of some theoretical principles of classification proposed by Austin and the CRG. *Cataloging & classification quarterly*, 4(1), 11-27

Frost, J. H., & Massagli, M. P. (2008). Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another's data. *Journal of medical Internet research*, 10(3).

Gil Urdiacaín, B. (1998). Orígenes y evolución de los tesauros en España. *Revista general de información y documentación*, 8(1), 63.

Gil, A.C. (1999). *Métodos e técnicas de pesquisa social*. São Paulo: Atlas.

Golder, S. A., & Huberman, B. A. (2006). Usage patterns of collaborative tagging systems. *Journal of information science*, 32(2), 198-208

Gomes, H. E. (1990). *Manual de elaboração de tesauros monolíngues*. Programa Nacional de Bibliotecas das Instituições de Ensino Superior.

Gomes, H. E., Motta, d. F. D., & Campos, M. L. D. A. (2006). Revisitando Ranganathan: a classificação na rede. *Rio de Janeiro*.

Gonçalves, J., & Souza, R. R. (2013). Relações e conceitos em ontologias: teorias de Farradane e Dahlberg. *Seminário de pesquisa em ontologia no Brasil. Universidade Federal Fluminense. Departamento de Ciência da Informação. Niterói*.

Gonzalez, M., & Lima, V. L. (2003). Recuperação de informação e processamento da linguagem natural. In *XXIII Congresso da Sociedade Brasileira de Computação* (Vol. 3, pp. 347-395).

Guedes, R. D. M., Moura, M. A., & Dias, E. J. W. (2011). Indexação social e pensamento dialógico: reflexões teóricas. *Informação & Informação*, 16(3), 40-59.

Guilam, M. C. R. (1996). O conceito de risco: sua utilização pela epidemiologia, engenharia e ciências sociais. *Rio de Janeiro, RJ: Dissertação de Mestrado-ENSP/FIOCRUZ*.

Hammwöhner, R., & Kuhlen, R. (1994). Semantic control of open hypertext systems by typed objects. *Journal of Information Science*, 20(3), 175-184.

Hedlund, T., Pirkola, A., & Järvelin, K. (2001). Aspects of Swedish morphology and semantics from the perspective of mono-and cross-language information retrieval. *Information Processing & Management*, 37(1), 147-161

Hjørland, B. (1998). Information retrieval, text composition, and semantics. *Knowledge organization*

Hjørland, B. (2007). Semantics and knowledge organization. *Annual review of information science and technology*, 41(1), 367-405.

Hjørland, B., & Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the Association for Information Science and Technology*, 46(6), 400-425.

Ilari, R., & Geraldi, J. W. (1985). *Semântica*. Ática.

Khoo, C., & Myaeng, S. H. (2002). Identifying semantic relations in text for information retrieval and information extraction. In *The Semantics of Relationships* (pp. 161-180). Springer Netherlands

Khoo, Christopher SG; NA, Jin-Cheon. (2006). *Semantic relations in information science*.

Khoo, Christopher SG; NA, Jin-Cheon. (2006). *Semantic relations in information science*.

Kristine Ådland, M., & Lykke, M. (2012). Chapter 5 Social Tagging in Support of Cancer Patients' Information Interaction. In *Social information research* (pp. 101-128). Emerald Group Publishing Limited.

Lancaster, F. (2004). Indexação e resumo: teoria e prática. Ver, ampl. e atual. Tradução de Antonio Agenor Briquet de Lemos

Langridge, D. W. (1977). Classificação: Abordagem para estudantes de biblioteconomia. Tradução de Roseli P. Fernandez. *Interciência Rio de Janeiro. RJ*, 126pp

Lee, K., Kim, H., Shin, H., & Kim, H. J. (2009, May). Folksoviz: A semantic relation-based folksonomy visualization using the wikipedia corpus. In *Software*

Engineering, Artificial Intelligences, Networking and Parallel/Distributed Computing, 2009. SNPD'09. 10th ACIS International Conference on (pp. 24-29). IEEE.

Lévy, P. (2009). *Cibercultura*. São Paulo: Editora 34.

Lima-Costa, M. F., & Barreto, S. M. (2003). Tipos de estudos epidemiológicos: conceitos básicos e aplicações na área do envelhecimento. *Epidemiologia e serviços de saúde*, 12(4), 189-201.

Lima, G. Â. B. (2004). O modelo simplificado para análise facetada de Spiteri a partir de Ranganathan e do Classification Research Group (CRG). *Información, cultura y sociedad*, (11), 57-72.

Lima, I.F., Souza, R.R. & Dias, G.A. (2013). O uso de análise de conteúdo na construção de instrumento para avaliação de bibliotecas digitais. In Aquino M. A. (org), *Experiências metodológicas em Ciência da Informação*. (139-148). João Pessoa: UFPB.

Lima, J.L.O. & Alvares, L. (2012). Organização e representação da informação e do conhecimento. In L. Alvares (org), *A organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações* (21-47). São Paulo: B4.

Limpens, F., Gandon, F., & Buffa, M. (2009). Linking folksonomies and ontologies for supporting knowledge sharing: a state of the art.

Linehan, C., Doughty, M., Lawson, S., Kirman, B., Olivier, P., & Moynihan, P. (2010, April). Tagliatelle: social tagging to encourage healthier eating. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems* (pp. 3331-3336). ACM.

López Rodríguez, C. I., Sánchez, M. T., & Faber Benítez, P. (2006). Gestión terminológica basada en el conocimiento y generación de recursos de información sobre el cáncer: el proyecto Oncoterm. *RevistaeSalud. com*, 2(8).

Maldonado-Martínez, Á. (2010). Los tesauros como herramienta fundamental en la descripción y recuperación de documentos electrónicos

Maldonado-Martínez, Á. (2010). Los tesauros como herramienta fundamental en la descripción y recuperación de documentos electrónicos

Mantelo Cecilio, H. P., Oliveira de Arruda, G., Ferraz Teston, E., Lima Santos, A., & Silva Marcon, S. (2015). Comportamentos e comorbidades associados às complicações microvasculares do diabetes. *Acta Paulista de Enfermagem*, 28(2).

Marlow, C. et. al. *Position paper, tagging, taxonomy, flickr, article, toread*. 2006. Disponível em: <<http://www.danah.org/papers/WWW2006.pdf>>. Acesso em: 3 jul. 2012.

Marques, C., Nardi, A. E., Figueira, I., Mendlowicz, M., Andrade, Y., CAMISAO, C., & Versiani, M. (1994). Comorbidade: conceito e implicacoes na pesquisa clinica em psiquiatria. *J Bras Psiquiatr*, 117-121.

Marteleteo, R. M. (2005) A metodologia de análise de redes sociais (ARS). In: Valentim, Marta Lígia Pomim (Org.). *Métodos qualitativos de pesquisa em Ciência da Informação*. São Paulo: Polis.

Medeiros, M. B. B., & Café, L. M. A. (2013). Organização da informação ou organização do conhecimento?.

Melo Fernandes, M. D. G., Lima da Nóbrega, M. M., Ribeiro Garcia, T., & de Freitas Macêdo-Costa, K. N. (2011). Análise conceitual: considerações metodológicas. *Revista Brasileira de Enfermagem*, 64(6).

Mendes, P. R., dos Reis, R. M. & Maculan, B.C.M.S (2015). Tesouros no acesso à informação: uma retrospectão Thesaurus in information access: a retrospection. *Revista ACB: Biblioteconomia em Santa Catarina*, 20(1), 49-66.

Mendonca, F. M., Cardoso, A. M. P., & Drumond, E. (2009). *Ontologia de aplicação no domínio de mortalidade: uma ferramenta de apoio para o preenchimento da declaração de óbitos* (Doctoral dissertation, Dissertação (Mestrado). Belo Horizonte: Pontifícia Universidade Católica de Minas Gerais (PUC-MG), Programa de Pós-graduação em Informática).

Mendonça, R., Rosa, A. F., Oliveira, J. L., & Teixeira, A. (2013, June). Ontology-based Health Information search: Application to the neurological disease domain. In *Information Systems and Technologies (CISTI), 2013 8th Iberian Conference on* (pp. 1-6). IEEE

Mika, P. (2007). Social Networks and the Semantic Web. *Semantic Web And Beyond Computing for Human Experience*, vol. 5

Minayo, M. C. D. S., & Miranda, A. C. D. (2002). *Saúde e ambiente sustentável: estreitando nós*. Editora Fiocruz.

Moreira, M. P., & Moura, M. A. (2006). Construindo tesouros a partir de tesouros existentes: a experiência do TCI–Tesouro em Ciência da Informação. *DataGramaZero*, 7(4).

Moreira, V., & Pimenta, H. (2012). *Gramática de português: 3. o ciclo do ensino básico e ensino secundário*. Porto editora.

Mostafa, S. P., & NOVA CRUZ, D. V. D. (2012). David Hume e as relações associativas: da teoria ao tesouro.

Motta, D. F. D. (1987). Método relacional como nova abordagem para a construção de tesouros.

Moura, M. A. (2009). Folksonomias, redes sociais e a formação para o tagging literacy: desafios para a organização da informação em ambientes colaborativos virtuais. *Informação & Informação*, 14 (1), p.25-45.

Moura, M. A. (2009). Folksonomias, redes sociais e a formação para o tagging literacy: desafios para a organização da informação em ambientes colaborativos virtuais. *Informação & Informação*, 14 (1), p.25-45.

Nascimento, G. F. C. L. (2009). Folksonomia como estratégia de indexação dos bibliotecários no Del. icio. us. *Informação & Sociedade: Estudos*, 19(1).

National Information Standards Organization (US). (2005). *Guidelines for the construction, format, and management of monolingual controlled vocabularies*. NISO Press

National Library of Medicine. (2016). Fact Sheet Medical Subject Headings (MeSH). Disponível em: <https://www.nlm.nih.gov/pubs/factsheets/mesh.html>

Naumis Peña, C. (2007). Los tesauros documentales y su aplicación en la información impresa, digital y multimedia. *Buenos Aires: Alfagrama*.

NAVES, M. M. L., & Dias, E. W. (2007). *Análise de assunto: teoria e prática* (Vol. 3). Thesaurus Editora.

Novellino, M.S.F. (1996). Instrumentos e metodologias de representação da informação. *Informação & Informação*, 1(2). Retrieved from <http://www.uel.br/revistas/uel/index.php/informacao/article/viewFile/1603/1358>

O'Reilly, T. (2012). What is Web 2.0. 2005. See <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>.

Peters, I. (2009). *Folksonomies: indexing and retrieval in Web 2.0* (Vol. 1). Walter de Gruyter.

Peters, I. (2009). *Folksonomies. Indexing and retrieval in Web 2.0*. Walter de Gruyter.

Peters, I., & Weller, K. (2008). Paradigmatic and syntagmatic relations in knowledge organization systems. *Information Wissenschaft und Praxis*, 59(2), 100-107

Pinto, V. B., & Soares, M. E. (2010). *Informação para a área de saúde: prontuário do paciente, ontologia de imagem, terminologia, legislação e gerenciamento eletrônico de documentos*. UFC

Pinto, V. B., Campos, H. D. H., Ferreira, J. L. O., & Freitas, E. M. D. (2013). Ontologia de imagens do domínio da nefrologia a partir dos atributos visuais e verbais dos laudos e prontuários de pacientes.

Pombo, O. (1998). Da classificação dos seres à classificação dos saberes. *Revista da Biblioteca Nacional de Lisboa*, (2), 19

Quintarelli, E. (2005). *Folksonomies: power to the people*. Milano: incontro ISKO Italia UniMIB

Rafferty, P., & Hilderley, R. (2007, July). Flickr and democratic indexing: dialogic approaches to indexing. In D. Bawden (Ed.), *Aslib Proceedings* (Vol. 59, No. 4/5, pp. 397-410). Emerald Group Publishing Limited. Peters, I. (2009). *Folksonomies: Indexing and retrieval in Web 2.0* (Vol. 1). Walter de Gruyter.

Ranganathan, S. R. (1967). Prolegomena to library classification. *The Five Laws of Library Science*

Rector, M., & Yunes, E. L. M. (1980). *Manual de semântica* (Vol. 7). Ao Livro Técnico.

Rezende (2004). Fundamentos da terminologia médica. *Linguagem Médica*, 3a. ed., Goiânia, AB Editora e Distribuidora de Livros Ltda. Disponível em: <http://www.jmrezende.com.br/terminologia.htm>

Ribeiro, C.J.S.(2012). Entendimento de requisitos de sistema com abordagem orientada ao domínio. *DataGramaZero*, 13(2). Retrieved from http://www.dgz.org.br/abr12/Art_01.htm

Rosario, B., & Hearst, M. (2001, June). Classifying the semantic relations in noun compounds via a domain-specific lexical hierarchy. In *Proceedings of the 2001 Conference on Empirical Methods in Natural Language Processing (EMNLP-01)* (pp. 82-90).

Sampieri, R. H., Collado, C. F., & Lúcio, P. B. (2013). *Metodologia de Pesquisa* (5th ed.). Porto Alegre: Penso.

Santiago, M. S., & Krieger, M. G. (2009). Terminologia a serviço da informação: rede de palavras-chave para artigos de divulgação científica da Medicina. *Calidoscópio*, 7(3), 237-242

Schiessl, M., & Shintaku, M. (2012). Sistemas de organização do conhecimento. In L. Alvares (org), *A organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações* (49-118). São Paulo: B4.

Schreiner, H. B. (1976). Considerações históricas acerca do valor das classificações bibliográficas. In: *Conferência brasileira de classificação bibliográfica* (Vol. 1, pp. 190-207)

Sharif, A. (2009, August). Combining ontology and folksonomy: an integrated approach to knowledge representation. In *IFLA 2009 Satellite Meeting in Florence, Italy*.

Sharif, A. (2009, August). Combining ontology and folksonomy: an integrated approach to knowledge representation. In *IFLA 2009 Satellite Meeting in Florence, Italy*.

Silva, A. R., & Lima, G. A. B. O. (2011). A teoria clássica de categorização e os princípios. In *xii enancib: políticas de informação para a sociedade*

Silva, D. L. D., Souza, R. R., & Ramos, J. A. D. A. (2010). *Representação da informação e do conhecimento em sistemas de informações: contribuições de campos interdisciplinares*.

Silva, F. S. D. (2012). Uma Ontologia para as Doenças Tropicais Negligenciáveis-NTDO.

Silva, R. E. D., Santos, P. L. V. A. D., & Ferneda, E. (2013). Modelos de recuperação de informação e web semântica: a questão da relevância. *Informação & Informação*, 27-44

Silva, R. E., da Costa, P. L. V. A., & Ferneda, E. (2013). Modelos de recuperação de informação e web semântica: a questão da relevância. *Informação & Informação*, 18(3), 27-44.

Silva, s. F., Castellví, m. T. C., & Aymerich, j. F. La variación denominativa en terminología desde una perspectiva cognitivo-discursiva

Simões, M. D. G. (2007). *Relaciones conceptuales en los tesauros: normativa internacional y aplicación práctica* (Master's thesis)

Simões, M. D. G., & Ferreira, C. (2013). Resumo e abstract: evolução morfológica e semântica. In *Biblos: Revista da Faculdade de Letras da Universidade de Coimbra* (Vol. 11). Imprensa da Universidade de Coimbra.

Simões, M. G. (2008). Da abstracção à complexidade formal: Relações conceptuais num Tesouro. *Coimbra: Almedina*.

Simões, M. G. (2011). *Classificações bibliográficas: percurso de uma teoria*. Almedina.

Siqueira, J. C. (2012). A noção de folksonomia: uma abordagem terminológica. *Tradterm*, 20, 129-145

Song, M., & Galardi, P. (2001). Semantic Relationships between Highly Cited Articles and Citing Articles in Information Retrieval. In *Proceedings of the ASIST Annual Meeting* (Vol. 38, pp. 171-81)

Souza, R. R. (2006). Sistemas de recuperação de informações e mecanismos de busca na web: panorama atual e tendências. *Perspectivas em ciência da informação*, 11(2), 161-173.

- Souza, R. R., & Alvarenga, L. (2004). A Web Semântica e suas contribuições para a ciência da informação. *Ciência da Informação, Brasília*, 33(1), 132-141.
- Sowa, J. F. (1983). Conceptual structures: information processing in mind and machine.
- Specia, L. ; Motta, E. (2007). Integrating Folksonomies with the Semantic Web. *The Semantic Web: Research and Applications (4519)*, pp. 624-639. Retrieved 1, June, 2009, from www.eswc2007.org/pdf/eswc07-specia.pdf
- Stock, W. G. (2010). Concepts and semantic relations in information science. *Journal of the Association for Information Science and Technology*, 61(10), 1951-1969.
- Stokolova, N. A. (1977). Elements of a semantic theory of information retrieval—I. The concepts of relevance and information language. *Information Processing & Management*, 13(4), 227-234.
- Storey, V. C. (1993). Understanding semantic relationships. *The VLDB Journal—The International Journal on Very Large Data Bases*, 2(4), 455-488.
- Storey, V. C. (1993). Understanding semantic relationships. *The VLDB Journal—The International Journal on Very Large Data Bases*, 2(4), 455-488
- Sun, J. T., Chen, Z., Zeng, H. J., Lu, Y. C., Shi, C. Y., & Ma, W. Y. (2004, November). Supervised latent semantic indexing for document categorization. In *Data Mining, 2004. ICDM'04. Fourth IEEE International Conference on* (pp. 535-538). IEEE
- Svenonius, E. (2000). *The intellectual foundation of information organization*. MIT press.
- Tjondronegoro, D., & Spink, A. (2008). Web search engine multimedia functionality. *Information Processing & Management*, 44(1), 340-357.
- Trant, J. (2009). Studying social tagging and folksonomy: A review and framework. *Journal of Digital Information*, 10(1).
- Vander Wal, T. (2007). Folksonomy. *online posting, Feb, 7*.
- Vander Wal, T. (2008). Getting to know collective and collaborative. PersonallInfoCloud.com
- Vaz, P., Pombo, M., Fantinato, M., & Pecky, G. (2007). O fator de risco na mídia. *Interface-Comunicação, Saúde, Educação*, 11(21).
- Vickery, B. C. (1997). Ontologies. *Journal of information science*, 23(4), 277-286.

Wang, S., Wang, W., Zhuang, Y., & Fei, X. (2015). An ontology evolution method based on folksonomy. *Journal of applied research and technology*, 13(2), 177-187

World Health Organization. (2016). International Classification of Diseases (ICD). Disponible em: <http://www.who.int/classifications/icd/en/>

Yedid, N. (2013). Introducción a las folksonomías: definición, características y diferencias con los modelos tradicionales de indización. *Información, cultura y sociedad*, (29), 13-26.

Zhang, L., Wu, X., & Yu, Y. (2006). Emergent semantics from folksonomies: a quantitative study. In *Journal on Data Semantics VI* (pp. 168-186). Springer Berlin Heidelberg

APÊNDICES

APÊNDICE 1 - INDEXAÇÕES LIVRES PARA O TERMO DIABETES

Indexações livres	
1.	2010 apo_b_100 diabetes dm2 metformin
2.	4 diabetes in mirs
3.	5 diabetes in mirs
4.	ada cardiovascular complications diabetes review type1
5.	apoptosis betacell diabetes review type1
6.	being_overweight bmi encdr_journal_club obesity t1dm youth
7.	cardiovascular dcct diabetes edic epidemiology ischaemic type1
8.	cardiovascular diabetes macrovascular microvascular nephropathy neuropathy retinopathy review type1 type2
9.	cardiovascular diabetes mortality type1
10.	costs diabetes health_care_costs
11.	desarrollo diabetes
12.	diabetes diet guidelines
13.	diabetes dpp menopause
14.	diabetes hbgm
15.	dm factorsrisc
16.	gwas ppis readlater
17.	<u>muscle</u>
18.	2 children diabetes in type
19.	2 diabetes exercise loss nutrition protein type weight
20.	2 child diabetes genes mellitus mody type youth
21.	2 d diabetes type vitamin
22.	2 diabetes onset young
23.	2 diabetes pediatric self-management type
24.	2000 dm1 edic nejm
25.	2002 cardiovascular_risk_factors epidemiolog diabetes_care dm2 metformin troglitazone
26.	2002 coagulation_factors diabetes_journal metformin
27.	2003 apoptosis beta_cell diabetes dm2
28.	2003 diabetes_care dm2 exenatide metformin sulfonylureas
29.	2004 care diabetes dm2 exenatide sulfonylureas
30.	2004 diabetes dm2 microalbuminuria nejm trial
31.	2010 cardiovascular_events diabetes diabetes_care hypoglycemia
32.	2011 care diabetes dm2 elderly_mass_loss metformin
33.	2011 beta_cell diabetes_care dm2 insulin_intensive remission
34.	2012 ada clinical diabetes guidelines practice
35.	2012 csii diabetes_care dm2 remission
36.	2012 diabetes diabetes_care diet food mnt
37.	2012 diabetes diabetic_medicine perioperative
38.	2012 diabetes editorial glp1
39.	2013 diabetes_care dm2 exenatide glp1 taspoglutide
40.	2014 ada diabetes mellitus

41. 3 diabetes in mirs
42. 5- articulo diabetes
43. absenteeism diabetes
44. activation diabetes foxp3
45. actividad atp canales diabetes metabolico
46. actividad diabetes hipertension oxido proteinas
47. activity behavior diabetes epidemiology interaction mortality physical sedentary
48. activo diabetes
49. ada cardiovascular complications diabetes review type1
50. ada diabetes diet food guideline nutrition recommendations type1 type2
51. ada diabetes diet food nutrition review type1 type2
52. adherence diabetes insulin
53. adherencia diabetes
54. adolescents age diabetes glycation type1
55. adolescents cohortstudy diabetes longitudinal pilot search type1 type2
56. adolescents csii diabetes hypoglycaemia insulin pump safety type1
57. adolescents csii diabetes insulin mdi pump retinopathy type1
58. adolescents diabetes glycaemiccontrol qol type1
59. adolescents diabetes support type1
60. adults antibodies autoimmune diabetes gad high lada latent of titer titre
61. adults calcification coronaryarterydisease diabetes hba1c
62. adults diabetes hypoglycaemia physicalactivity type1
63. age diabetes diagnosis hba1c predictors
64. age carotid diabetes glycation intimamediathickness ldl oxidation type1
65. aii dn normoalbuminuria
66. alcohol chronological_change cohort diabetes lifestyle
67. algorithm diabetes therapeutic type2
68. altitude diabetes
69. alzheimers diabetes
70. and diabetes dx particle size
71. and diabetes obesity
72. and diabetes oxidative stress
73. and diabetes stress
74. and diabetes stress
75. androgen dn gender human
76. anesthesia dm gsk rat
77. antigen_specific_t_cell diabetes
78. apo-a1 diabetes hdl lipid lipids lipoproteins type1
79. apolipoprotein diabetes insulin lipid lipidomics longitudinal
80. apoptosis autoimmune cause diabetes review type1
81. apoptosis betacell diabetes mouse type1
82. apoptosis betacell diabetes review type1
83. apoptosis betacell diabetes review type1 type2
84. apoptosis diabetes dna-repair metabolism mtor review warburg
85. aromatase diabetes estradiol men obesity testosterone

86. arterialstiffness bloodpressure cardiometabolic cardiovascular diabetes hypertension nephropathy retinopathy type1
87. arterialstiffness diabetes endothelial type1
88. arteriogenesis diabetes hyperglycemia pad rabbit
89. ashkenazi autosomal diabetes factor hepatocyte hnf4 nuclear
90. ashkenazi autosomal diabetes factor hepatocyte hnf4 nuclear
91. aspirin diabetes prevention
92. assessment cardiometabolic cardiovascular diabetes risk type1
93. assessment cardiometabolic cardiovascular diabetes risk type1 type2 nephropathy
94. atherosclerosis cardiovascular diabetes hypoglycaemia severe type1 vascular
95. atherosclerosis cardiometabolic cardiovascular cholesterol diabetes hdl lipid lipids lipoproteins type1
96. atlas epidemiology risk
97. atrophy complications diabetes neuropathy type1 type2
98. atypical diabetes flatbush ketosis prone
99. autoantibodies diabetes risk stratification type1 epidemiology
100. autoimmune diabetes genetics ptpn22 type1
101. autoimmune beta cell diabetes insulin latent reserve
102. autoimmune controls diabetes latent tcf7l2 type2
103. autonomic diabetes hypoglycaemia review type1 type2
104. autonomic cardiovascular csii diabetes insulin pump type1
105. autonomic_dysfunction autonomic_ns can diabetes review_article
106. autonomic_dysfunction dan diabetes neuropathy review_article
107. autonomic_dysfunction diabetes hyperglycemia neuropathy sudomotor_function
108. autonomic_ns diabetes
109. autosomal diabetes hnf4a norwegian
110. b12 b6 diabetes folate homocysteine nutrition t2dm
111. background cytokine diabetes immune
112. bariatrics diabetes
113. barriers diabetes management
114. barriers diabetes management
115. barriers diabetes management
116. barriers diabetes medications
117. bcells-mouse db-db-mouse mafa tesis
118. behavior diabetes
119. betablocker diabetes thiazide
120. betacell cause diabetes environment ilet review rotavirus type1
121. betacell diabetes immune review type1
122. betacell cohortstudy diabetes dipp enterovirus environment finland type1 virus
123. betacell diabetes review type1
124. betaell diabetes pathogenesis review type1 type2
125. biology diabetes
126. biomarkers diabetes mrna retinopathy type1 type2
127. biomarkers_db bld-monocytes cd36 hu_pbmc_isol insulin-resistance mf-diabetes pbmc
128. birth cause cesareansection children diabetes genetics polymorphism type1

129. birth_weight diabetes genetics intrauterine nephropathy obesity polygenic_risk_score
130. blood cardiovascular disease nephropathy glucose hba1c heart
131. blood diabetes fat low sugar vegan
132. blood cholesterol diabetes fructose lipid
133. blood crp cytokine diabetes grant human il6 inflammation obesity serum study thesis tnfa
134. bloodpressure cardiometabolic cardiovascular diabetes epideomiology lipid lipids lipoproteins risk type1
135. bloodpressure cardiovascular diabetes hypertension type1
136. bloodpressure csii diabetes insulin pump type1
137. bloodpressure diabetes hypertension kidney nephropathy type1 type2
138. bloodpressure cardiovascular diabetes diet finndiane food hypertension mortality nutrition salt sodium type1
139. bloodpressure csii diabetes type1
140. bm cmrf cost-effectiveness diabetes jmcp_manuscript
141. bm cmrf cost-effectiveness diabetes jmcp_manuscript
142. bm cmrf diabetes jmcp_manuscript
143. bmd diabetes human men testosterone
144. bmi chronological_change diabetes lifestyle obesity weight
145. bone cohortstudy diabetes fracture type1
146. bone density diabetes mineral review type1
147. brca diabetes metabolic_syndrome_nephropathy
148. breaking breaks epideomiology diabetes metabolic risk sitting time
149. breast_cancer diabetes
150. brown-fat
151. bypass diabetes surgery
152. caffeine coffee diabetes glucose insulin
153. can dan diabetes neuropathy polyneuropathy review_article
154. cancer diabetes
155. cancer diabetes risk type1 type2
156. cancer nephropathy risk diabetes
157. cancer diabetes diabetes_cancer
158. cancer diabetes review
159. carbohydrates diabetes diet glycaemiccontrol macronutrient nutrition type1 type2
160. cardiac cardiomyopathy cardiovascular complications conventionalinsulintherapy dcct diabetes edic glycaemiccontrol intensiveinsulintherapy type1
161. cardiac diabetes metabolism
162. cardiometabolic cardiovascular cohortstudy diabetes risk type1 type2
163. cardiometabolic cardiovascular diabetes gol rick type1 type2
164. cardiometabolic cardiovascular diabetes hipertenobesity
165. cardiometabolic cardiovascular diabetes review risk type1
166. cardiovascular cvd diabetes human men replacement testosterone
167. cardiovascular dcct diabetes edic epidemiology ischaemic type1
168. cardiovascular diabetes metabolomics phenotype type1 vascular
169. cardiovascular diabetes mortality risk
170. cardiovascular cohortstudy complications diabetes mortality pittsburgh type1

171. cardiovascular cohortstudy complications diabetes pittsburgh type1
172. cardiovascular csii diabetes insulin mdi mortality observational pump type1
173. cardiovascular diabetes estrogens hrt
174. cardiovascular diabetes metabolic-syndrome risk
175. cardiovascular_disease diabetes
176. cardiovascular_disease diabetes
177. cardiovascular_disease diabetes nihr_app_refs
178. care diabetes medical standards
179. casecontrolstudy cause children codliveroil diabetes environment type1
180. cause children codliveroil diabetes environment typ1
181. cause diabetes enterovirus environment immune review type1 virus
182. cause diabetes enterovirus review type1 virus
183. cause diabetes environment sunlight type1 vitamind
184. ccm chronic-care-model diabetes
185. cefalea-migrana diabetes
186. cgm diabetes dmt1 hbgm
187. cgm hypoglycemia
188. chapter dm insulin-resistance pathophysiology review treatment
189. chapter dm insulin-resistance pathophysiology review treatment
190. <u>chetoacidosi</u>
191. child diabetes incidence mellitus us youth
192. children csii denmark insulin pump
193. children diabetes epidemiology kuwait type1
194. children diabetes pollutants type1
195. cholesterol diabetes ezetimibe simvastatin
196. cholesterol crp diabetes dialysis hemodialysis_patients ldl
197. cholesterol diabetes hdl-cholesterol
198. cholesterol diabetes hdl-cholesterol
199. chronic_physical_conditions complex_interventions diabetes education fidelity implementation self-management
200. chronicdisease diabetes
201. chronological_change diabetes intervention lifestyle
202. cis-regulatory_element diabetes eqtl gwas human islets
203. clinical-trials diabetes pro-renin
204. <u>closed-loop</u>
205. closed-loop diabetes insulin pregnancy type1
206. cognitive diabetes impairment ulceration
207. cohort_study diabetes obesity vitamin_d
208. community diabetes screening
209. comorbidity diabetes medications
210. complications csii diabetes insulin kroc methodology pump retinoplathy type1
211. complications dcct diabetes epidemiology macrovascular microvascular type1
212. complications diabetes glycaemiccontrol review type1
213. complications diabetes retinopathy review type1

214. complications csii diabetes insulin kroc microvascular nephropathy pump retinopathy type 1
215. complications diabetes epidemiology global retinopathy risk type1 type2
216. complications diabetes erectiledysfunction neuropathy sexualdysfunction type1 type2
217. complications diabetes kidney nephropathy review type1 type2
218. complications diabetes retinopathy review type1
219. compund_heterozygote diabetes exome
220. computer diabetes gluconet
221. connfunding diabetes stroke
222. consecuencias diabetes
223. continuous glucose monitoring real-time rt-cgm type-2-diabetes
224. control_theory diabetes systems_biology
225. coordination diabetes
226. coronaryarterydisease diabetes nefa type1
227. cost diabetes
228. cost diabetes improvement
229. cost-effectiveness diabetes zt
230. c-peptide diabetes insulin
231. crossover csii diabetes insulin lispro pump rct type1
232. crossover diabetes insulin lispro nph type1
233. crp diabetes human inflammation men testosterone
234. csii diabetes insulin pump type1
235. csii csii-insulin diabetes insulin technology
236. csii diabetes insulin mdi pump qol type1
237. csii diabetes insulin pump socialmedia type1
238. csii diabetes insulin pump type1
239. csii diabetes insulin pump type1 weight
240. csii-exercise csii-hypoglycaemia
241. cv diabetes mortality register uk
242. cvd diabetes hospital_admissions hse_intro
243. cvd diabetes obesity risk sitting
244. cybersecurity diabetes e-health security
245. d diabetes vitamin
246. dairy diabetes
247. dairy diabetes health
248. dbdb dm epigenetics h3k9 h3k9me h3k9me1 h3k9me2 h3k9me3 histone hmt il il6 inflammation mcp1 metabolic-memory mir125 mirna mouse paperkmup papersuv suv39h1type2 vsmc
249. dc diabetes master
250. dc diabetes regulatory_t_cells vitamind
251. dcct diabetes gol type1
252. dcct diabetes intensiveinsulintherapy rct type1 weight
253. dcell-apoptosis calcium-channel calcium-flux calcium-singaling diabetes dm-treatment review-modulation-ionic-channels verapamil
254. dc-peptide dm dn

255. definition diabetes
256. depression diabetes
257. diabetes diabetesuk diet food guideline nutrition recommendations review type1 type2
258. diabetes diet exercise hypoglycaemia nocturna nutrition physicalactivity rct type1
259. diabetes diet food gastroparesis nutrition review type1
260. diabetes diet lowcarb rct type2
261. diabetes diet lowgi metaanalysis nutrition type1
262. diabetes economic me
263. diabetes elderly hypoglycaemia risk type1
264. diabetes epidemiology type1
265. diabetes f21 overweight-obesity risk factors
266. diabetes gender prevelence type1
267. diabetes genetics iatrogenic medication
268. diabetes genetics lifestyle metformin pharmacogenomics randomized_controlled_trial response
269. diabetes glargine insulin phase3 rct type1
270. diabetes glucosevariability glycaemiccontrol hba1c review type1
271. diabetes glut4
272. diabetes gut-microbiota obesity
273. diabetes hsqc
274. diabetes human insulin men replacement testosterone trt
275. diabetes ifg igt ngt progression t2dm
276. diabetes inflammation review type1
277. diabetes inpatient management
278. diabetes insulin
279. diabetes juglans-regia
280. diabetes lipids lipoproteins type1
281. diabetes metabolomics type1
282. diabetes methods statistics study-design
283. diabetes mice
284. diabetes microna nephopathy tgfb type1 diabetes
285. diabetes nihr_app_refs
286. diabetes pollution
287. diabetes rct type2 vitamind
288. diabetes review subclasses type1
289. diabetes roux-em-y-gastric-bypass t2dm
290. diabetes statins
291. diabetes taskshiftdiabetes
292. diabetes vitamin_d
293. diabetes diabetes_and_anemia hba1c
294. diabetes diabetes_genes tulip
295. diabetes diabetes_genes tulip
296. diabetes diabetes_prevention metabolic_syndrome prevention type_2_diabetes
297. diabetes diabetes-classification

298. diabetes diagnosis
299. diabetes diagnosis gwa insulin-secretion molecular risk
300. diabetes diet epidemiology ffq nutrition type1
301. diabetes diet epidemiology nutrition prevalence triggers type1
302. diabetes diet fibre food glycaemiccontrol lipid lipids metabolism nutrition type1 type2
303. diabetes diet glycaemiccontrol nutrition
304. diabetes diet loss weight
305. diabetes disease_management_program
306. diabetes disordered-eating
307. diabetes disorders eating
308. diabetes dpp human metformin sex steroid testosterone women
309. diabetes economic me
310. diabetes economic me
311. diabetes economic me
312. diabetes economic me
313. diabetes education spanish
314. diabetes egg nutrition risk t2dm
315. diabetes empowerment
316. diabetes endophenotype
317. diabetes epidemiology lifeexpectancy qol scotland type1
318. diabetes epidemiology neuropathy type1 type2 uk
319. diabetes epidemiology triggers type1
320. diabetes epigenetics
321. diabetes epilepsy glibenclamide katp kir62 physiology tolbutamide
322. diabetes estreptozotocina model
323. diabetes ethnicity
324. diabetes exercise
325. diabetes exome
326. diabetes exome
327. diabetes family support
328. diabetes family_conflict parenting
329. diabetes fruits vegetables
330. diabetes gene_therapy tagged_by_labonte
331. diabetes gene-expression insulin-resistance lipid-metabolism muscle obesity
332. diabetes genetics hla ptpn22 teddy type1
333. diabetes genetics insulingene risk type1
334. diabetes gestational guideline pregnancy
335. diabetes gi
336. diabetes glargine me
337. diabetes glargine me
338. diabetes glargine me
339. diabetes glargine me
340. diabetes glucosa proteinas
341. diabetes glucosa proteinas vascular

342. diabetes glucosa vascular
343. diabetes glucosa vascular
344. diabetes glucose
345. diabetes glucose
346. diabetes glucose leptin
347. diabetes glucose low
348. diabetes gpc
349. diabetes groupvisits healthyheart
350. diabetes groupvisits healthyheart
351. diabetes groupvisits healthyheart
352. diabetes gwas
353. diabetes gwas
354. diabetes hdaci review
355. diabetes healthyheart
356. diabetes healthyheart
357. diabetes hepatic lipids lipoproteinlipase lipoproteins triglycerides type1
358. diabetes hipertension
359. diabetes hipertension proteinas
360. diabetes hipertension vascular
361. diabetes history review type1 type2
362. diabetes hnf4 mapping
363. diabetes hnf4a mir
364. diabetes ikb ikk nfkb type1diabetes
365. diabetes improvement rct
366. diabetes in mirs
367. diabetes in_vitro pdf vitamin_d
368. diabetes incidence
369. diabetes inflammation wbc
370. diabetes informatics management
371. diabetes inhaler
372. diabetes insomnia insulin sleep
373. diabetes insomnia sleep_maintenance
374. diabetes insuficiencia_renal
375. diabetes insulin
376. diabetes insulin ir resistance
377. diabetes insulin neonatal
378. diabetes insulin prediabetes proinsulin
379. diabetes insulin review type1
380. diabetes insulin review type1
381. diabetes insulina
382. diabetes insulina stz
383. diabetes intervention self-management web-based
384. diabetes intervention web-based
385. diabetes intrauterine
386. diabetes intrauterine

387. diabetes italy prevalence type2
388. diabetes knhnes prevalence stat
389. diabetes lipid lipids lipoprotein lipoproteins review type1 type2
390. diabetes lipid management
391. diabetes lipids lipoproteins type1
392. diabetes low-carb
393. diabetes magnesium
394. diabetes magnesium
395. diabetes management rpm software
396. diabetes maternal meta_analysis offspring
397. diabetes maternal offspring parental paternal
398. diabetes me
399. diabetes me
400. diabetes me
401. diabetes me pediatrics type_1_diabetes
402. diabetes me pediatrics type_1_diabetes
403. diabetes me pediatrics type_1_diabetes
404. diabetes men metabolic shbg testosterone
405. diabetes menarche
406. diabetes metabolic syndrome hipertension
407. diabetes metabolico nitrico oxido
408. diabetes metabolism mouse
409. diabetes metabolites type1
410. diabetes metabolomics type1
411. diabetes metagenomics microbiome obesity
412. diabetes methods sample-size statistics
413. diabetes methods statistics study-design
414. diabetes methods statistics study-design
415. diabetes mh_chang_pub
416. diabetes mhealth
417. diabetes micronutrients
418. diabetes mindfulness
419. diabetes mitochondria
420. diabetes model
421. diabetes mri stroke
422. diabetes national register
423. diabetes naturalhistory pathogenesis review type1
424. diabetes navajo
425. diabetes negative vitamin_d
426. diabetes neighborhood ses social-environment spatial
427. diabetes neurod2
428. diabetes neurons
429. diabetes neurons
430. diabetes neuropathy review type2
431. diabetes nhanes_iii vitamin_d

432. diabetes nicotinamide
433. diabetes hipertension obesity
434. diabetes obesity-child
435. diabetes omega-3
436. diabetes parenting
437. diabetes parenting
438. diabetes paternal prenatal risk_factor
439. diabetes pdac
440. diabetes peds searhh study
441. diabetes periodontal
442. diabetes persistant_organic_pollutants pops
443. diabetes power statistics tcf712 type2
444. diabetes pregnancy presentation
445. diabetes prevention
446. diabetes prevention
447. diabetes prevention
448. diabetes prevention review type1
449. diabetes proteinas
450. diabetes proteinas
451. diabetes qol
452. diabetes qol scale type1
453. diabetes reimbursement self_management
454. diabetes review streetlighteffect type1
455. diabetes review type1 type2
456. diabetes richtlijnen
457. diabetes rpm
458. diabetes screening
459. diabetes self_management
460. diabetes self_mgmt
461. diabetes senegal senegaldiabetes
462. diabetes senegal senegaldiabetes
463. diabetes social_netowrk
464. diabetes standard
465. diabetes statins
466. diabetes stroke
467. diabetes t1d
468. diabetes technologies
469. diabetes telomeres
470. diabetes typ2
471. diabetes type_2_diabetes hipertension
472. diabetes vascular
473. diabetes vascular
474. diabetes vascular
475. diabetes vascular
476. diabetes ventriculo

477. diabetes vitamin_d
478. diabetes vitamin_d vitamin_d_deficiency
479. diabetes vpt
480. diabtes diabetes_clinic hipertension glycemia
481. diabtes vitamins
482. dka yorkshire
483. dm genetics
484. dm image mouse ros stz
485. dm socs
486. <u>dm_aging</u>
487. <u>dmt1</u>
488. dmt1dcct dmt1-edic
489. dmt1-epidemiology transition-child-adult
490. dn prediction
491. dyslipidemia fatty-acids human human-pancreas palmitate pancreatic-islets transcript
492. endocrinology generalinternalmedicine
493. exercise insulin pdf
494. export-3-1-16-b insulina-nasal
495. genetics human
496. <u>gestational_diabetes</u>
497. <u>glenn</u>
498. glucose metformin production type-2-diabetes
499. hamoa_ir
500. hbgm-trial hbgm_dmt2
501. hbgm-trial hbgm_dmt2
502. hepatic-insulin-resistance insulin-resistance insulin-secretion model rat
503. in_vivo metabolic_syndrome nitric_oxide obesity uv
504. <u>liver_mrs</u>
505. meta_g metabolomics
506. muscle pgc1
507. <u>ncv</u>
508. nhv2009 ovriga
509. <u>osa</u>

APÊNDICE 2 - PROTOCOLO PARA VALIDAÇÃO PRELIMINAR DOS CONCEITOS

A tese intitulada **Análise das relações semânticas implícitas em folksonomias: contributos a uma ontologia sobre Diabetes** é desenvolvida no âmbito do programa doutoral **Informação e Comunicação em Plataformas Digitais**, (Faculdade de Letras da Universidade do Porto e Universidade de Aveiro) e tem como objetivo central perceber como as relações semânticas (relações significativas entre dois conceitos ou mais) implícitas nas folksonomias (tags atribuídas a recursos informacionais de forma aberta e compartilhada em ambientes digitais) podem ser incorporadas na estruturação de uma ontologia (conjunto de conceitos e relacionamentos formalizados) sobre diabetes. Para sua execução, é necessário um conjunto de descritores categorizados e relacionados sobre o assunto. Entretanto, mesmo fundamentados na literatura específica, considera-se relevante o diálogo e a consulta aos especialistas na área da saúde. Neste sentido, contamos com a vossa colaboração para revisão dos conceitos, relações e afirmações, apresentados conforme o exemplo a seguir.

Conceitos	Categorias	Afirmações	Observação
Nefropatia	Doença macrovascular	Diabetes é fator de risco para nefropatia	
Obesidade	Doença nutricional e endócrina	Obesidade é fator de risco para diabetes	
Diabete tipo 2	Doença nutricional endócrina	Diabetes tipo 2 é um tipo de Diabetes Mellitus	
Insulina	Substância química - hormônio	Insulina controla glicemia	
Pâncreas	Anatomia	Pâncreas produz insulina	

Grata pela colaboração.

Jacqueline A. Souza

APÊNDICE 3 - PROTOCOLO PARCIAL – RETORNO DO ESPECIALISTA

Conceitos	Categorias	Afirmações lógicas
Célula beta	Casuística	Célula beta secreta insulina Célula beta secreta glucagon
Dieta	Procedimento preventivo	Dieta mantém o controle glicêmico ⇔ Dieta controla o índice glicêmico Dieta previne as complicações ⇔ Dieta regula a glicemia Dieta trata a diabetes mellitus -> não trata previne as complicações, diabetes são doenças crônicas.
Doença cardiovascular	Complicações clínicas	Diabetes é fator de risco para doença cardiovascular
Epidemiologia	Área do conhecimento	Diabetes é uma doença epidemiológica? Não querias dizer a prevalência da doença é uma epidemia, tem a ver com os dados, a sua prevalência?
Glicemia ⇔ índice glicêmico	Medida	Glicemia é o nível de glicose no sangue
Índice glicêmico	Medida	Nível de glicose no sangue
Hemoglobina glicada	Técnicas de diagnóstico e/ou monitorização	Hemoglobina é técnica de diagnostica das diabetes
Hiperglicemia	Sintoma	Hiperglicemia causa Cetoacidose diabética Hiperglicemia é sintoma de diabetes
Hipertensão	Complicação Clínica	Diabetes é um fator de risco para hipertensão? Não há causalidade entre diabetes e hipertensão Hipertensos são propensos a ter diabetes? Possuem fatores de risco em comum? Hipertensão aumenta o risco de diabetes? - Diabetes é fator de risco para hipertensão (essa é a frase)
Obesidade	Doença nutricional e endócrina	Obesidade é fator de risco para diabetes tipo 2 Obesidade causa resistência a insulina Obesidade eleva o colesterol e a pressão sanguínea
Pé diabético	Complicação clínica	Doença microvascular e/ou macro vascular Diabetes Fator de risco para o pé diabético Pé diabético causa amputação
Resistência a insulina	Processo e doença	Resistência a insulina causa diabetes Resistência a insulina provoca hipertensão
Retinopatia	Complicação clínica	Diabetes é risco para retinopatia Doença Microvascular Retinopatia causa cegueira
Tipo 1	Doença nutricional? – Somente - endócrina	Diabetes tipo 1 é um tipo de Diabetes Mellitus Caracteriza-se por deficiência da secreção da insulina.
Tipo 2	Doença comportamental (relacionada com estilo de vida e comportamento alimentar ao longo da vida) endócrino	Diabetes tipo 2 é um tipo de Diabetes Mellitus Diabetes tipo 2 caracteriza-se por resistência à insulina