



UNIVERSIDADE FEDERAL DA BAHIA – UFBA
LABORATÓRIO NACIONAL DE COMPUTAÇÃO CIENTÍFICA – LNCC/MCT
UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA – UEFS
UNIVERSIDADE DO ESTADO DA BAHIA – UNEB
INSTITUTO FEDERAL DE EDUCAÇÃO CIÊNCIA E TECNOLOGIA DA BAHIA – IFBA
FIEB/SENAI/CIMATEC
FACULDADE DE EDUCAÇÃO – FACED – UFBA – Sede
INSTITUTO DE HUMANIDADES, ARTES E CIÊNCIAS – IHAC – UFBA – Co-Promotor.

DOUTORADO MULTI-INSTITUCIONAL E MULTIDISCIPLINAR EM DIFUSÃO DO CONHECIMENTO

HILDENISE FERREIRA NOVO

**Análise conceitual e cognitiva: Modac - um modelo dinâmico para
auxiliar a construção de Sistemas de Organização do
Conhecimento (SOC)**

Salvador

2014

**UNIVERSIDADE FEDERAL DA BAHIA – UFBA
LABORATÓRIO NACIONAL DE COMPUTAÇÃO CIENTÍFICA – LNCC/MCT
UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA – UEFS
UNIVERSIDADE DO ESTADO DA BAHIA – UNEB
INSTITUTO FEDERAL DE EDUCAÇÃO CIÊNCIA E TECNOLOGIA DA BAHIA – IFBA
FIEB/SENAI/CIMATEC
FACULDADE DE EDUCAÇÃO – FAGED – UFBA – Sede
INSTITUTO DE HUMANIDADES, ARTES E CIÊNCIAS – IHAC – UFBA – Co-Promotor**

**DOUTORADO MULTI-INSTITUCIONAL E MULTIDISCIPLINAR EM DIFUSÃO
DO CONHECIMENTO**

Hildenise Ferreira Novo

**Análise conceitual e cognitiva: Modac - um modelo dinâmico para
auxiliar a construção de Sistemas de Organização do
Conhecimento (SOC)**

Tese apresentada ao Curso de Doutorado Multi-institucional e Multidisciplinar em Difusão do Conhecimento, da Universidade Federal da Bahia, como requisito parcial para a obtenção do título de doutor em Difusão do Conhecimento. **Área de Concentração: Modelagem da Geração e Difusão do Conhecimento. Linha 01 – Construção do Conhecimento: Cognição, Linguagens e Informação.**

Orientadores: Prof. Dr. José Garcia Vivas Miranda
Profa. Dra. Teresinha Fróes Burnham

Salvador

2014

UNIVERSIDADE FEDERAL DA BAHIA – UFBA
LABORATÓRIO NACIONAL DE COMPUTAÇÃO CIENTÍFICA – LNCC/MCT
UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA – UEFS
UNIVERSIDADE DO ESTADO DA BAHIA – UNEB
INSTITUTO FEDERAL DE EDUCAÇÃO CIÊNCIA E TECNOLOGIA DA BAHIA – IFBA
FIEB/SENAI/CIMATEC
FACULDADE DE EDUCAÇÃO – FACED – UFBA – Sede
INSTITUTO DE HUMANIDADES, ARTES E CIÊNCIAS – IHAC – UFBA – Co-Promotor
DOUTORADO MULTI-INSTITUCIONAL E MULTIDISCIPLINAR EM DIFUSÃO DO
CONHECIMENTO

Hildenise Ferreira Novo

**Análise conceitual e cognitiva: Modac - um modelo dinâmico para auxiliar a construção
de Sistemas de Organização do Conhecimento (SOC)**

Tese apresentada como requisito parcial para obtenção do título de doutor em Difusão do Conhecimento, na Universidade Federal da Bahia, à seguinte banca examinadora:

Prof. Dr. José Garcia Vivas Miranda (Orientador) _____

Doutor em Ciências Ambientais - Professor do Instituto de Física da UFBA

Profa. Dra. Teresinha Fróes Burnham (Co-orientadora) _____

Doutora em Filosofia - Professora da FACED - UFBA

Henriette Ferreira Gomes (examinadora) _____

Doutora em Educação – UFBA - Professora do ICI/UFBA

Hernane Borges de Barros Pereira (examinador) _____

Doutor em Engenharia Multimídia - Professor da UNEB - Departamento de Educação e SENAI CIMATEC.

Maria Luiza de Almeida Campos (examinadora) _____

Doutora em Ciência da Informação – Professora do Departamento de Ciência da Informação e do Programa de Ciência da Informação da UFF

Silvia Maria Gomes Caldeira (examinadora) _____

Doutora em Ciência Cognitiva e Linguagem - Professora UFBA

Salvador, 16 maio 2014.

Para

Hildete Ferreira Novo, minha mãe;

Hilton Pereira Novo, meu pai;

Eduardo Novo Terra e Mariana Novo Souza, meus filhos;

Théo Pereira Novo Terra e Bernardo Pereira Novo Terra, meus netos, por tudo...

AGRADECIMENTOS

Este trabalho é fruto de uma caminhada profissional e de estudo que embora tenha tido momentos de solidão eu não realizei sozinha, mas sim com a contribuição de muitos, portanto não posso deixar de agradecer:

A Deus pela força infinita;

Aos meus pais Hilton e Hildete pelo apoio incondicional às minhas conquistas na trajetória de vida;

Aos meus filhos Eduardo e Mariana por entender a minha dedicação à profissão e aos estudos;

Aos meus netos Theo e Bernardo que mesmo diante da minha ausência me presenteiam com um sorriso largo e inocente as poucas vezes que chego em “casa”;

À minha nora Kenia, que guarda e cuida dos nossos tesouros;

À minha Irma Hilanilza, nossa querida Hila, por cuidar de todos na minha ausência;

À Marcela Novo, minha sobrinha, uma inspiração para caminhar na pesquisa e a Alysson e Bruno, meus sobrinhos, por serem “irmãos dos meus filhos”;

Aos meus amigos e amigas em especial: Maria Eduarda Serpa por ter disponibilizado sua biblioteca pessoal para os meus estudos e por ter me escutado muitas vezes; a Marise Sanches pelo companheirismo e dedicação ao CAOS; a Leila Bárbara Menezes que mesmo residindo tão distante de Salvador foi presença constante, me incentivando com palavras que me deram força para caminhar; A Ivana Lins, que elevava minha autoestima e a Francisco Pedroza que me apresentou o DMMDC.

Aos colegas professores do ICI/UFBA pelo esforço conjunto que viabilizou o meu afastamento para à conclusão do doutorado;

Às alunas e alunos do ICI/UFBA, pela oportunidade de aprender continuamente;

Aos professores do DMMDC, pela partilha, meu reconhecimento;

A Hélio José de Carvalho, secretário do DMMDC, pela presteza;

Aos colegas do Doutorado em Difusão do Conhecimento pelos momentos inesquecíveis...;

Aos colegas do CAOS, pelo entendimento da proposta de aprendizado colaborativo;

Aos membros da banca, pela contribuição nesta tese desde a qualificação;

À Maria Luiza de Almeida Campos por ter compartilhado a “régua e o compasso” para os estudos em Organização do Conhecimento;

E principalmente aos meus orientadores: José Garcia Vivas Miranda, pelos ensinamentos, companheirismo, paciência, dedicação e bom humor e a Teresinha Fróes Burnham, exemplo de determinação e coragem, que além de nos conduzir no universo da AnCo, nos faz lembrar a todo o momento que toda construção é colaborativa.

Mama Palavra

[...]

Aquela que não funcionar
Palavra-falha
Aquela que não se juntar
Vira palavra-tralha
Tralha

Quando tudo fala igual
Palavra-palha
Pra tudo que é marginal
Palavra que batalha
Palavra que batalha

Aquela que não funcionar
Palavra-falha
Aquela que não se juntar
Vira palavra-tralha
Tralha

Arnaldo Antunes

Resumo

A organização e representação do conhecimento pressupõe o trabalho com modelização de domínios de conhecimento, na intenção de construção de bases de relacionamentos entre categorias de termos que facilitam sua compreensão e permitem compartilhamento conceitual e semântico, com objetivo de difundir conhecimento. Na construção da base de modelização de domínios de conhecimento aliam-se metodologias e ferramentas. Este trabalho de tese no Doutorado Multi-Institucional e Multidisciplinar em Difusão do Conhecimento DMMDC, apresenta uma proposta de Modelo Dinâmico de Análise Conceitual (Modac) para organização do conhecimento que poderá contribuir na construção de Sistemas de Organização do Conhecimento (SOC). O modelo está baseado em análise cognitiva (AnCo) e análise do discurso representativo do domínio meio ambiente, a partir das subáreas e especialidades das Geociências, domínios inter e multidisciplinares. Além disso, indica princípios teóricos, epistemológicos e metodológicos de autores seminais da Ciência da Informação (CI), Filosofia, Linguística e Ciência da Computação (CC), assim como modelos de construção de ontologias desenvolvidos por autores (as) contemporâneos (as). A tese caracteriza-se como pesquisa qualitativa e aplicada. O estudo apontou caminhos, para a segmentação de textos em redes semânticas, de dissertações e teses com objetivo de identificar macroproposições semânticas e tipos de conceitos que podem declarar assuntos atinentes a domínios de conhecimento. O modelo proposto está amparado por Teorias como as do Conceito (TC), desenvolvida por I. Dahlberg e Teoria da Classificação Facetada (TCF) de S. R. Ranganathan e sustentado na conjunção de modelos triádicos de construção de ontologias e análise conceitual. O Modac, objetiva, auxiliar a construção de taxonomias e ontologias de domínios.

Palavras chave: Análise Cognitiva. Análise Conceitual. Análise do Discurso Representativo. Modelo Dinâmico de Análise Conceitual. Organização do Conhecimento. Representação do Conhecimento.

Abstract

The organization and representation of knowledge working with modeling of knowledge domains in order to construct bases of relationships between categories of terms. This is done to facilitate their understanding and enable conceptual and semantic sharing to spread knowledge. In the construction of a base modeling of knowledge domains are combined both methodologies and tools. In this dissertation, a proposal of a Dynamic Model of Conceptual Analysis for organization of knowledge that can contribute to build Knowledge Organization Systems (KOS) is presented. The model is based on both cognitive analysis and speech representative analysis of the environmental domain from subfields and specialties of Geosciences, inter and multidisciplinary areas. Furthermore, this work indicates theoretical, epistemological and methodological principles of pioneer authors of Information Science, Philosophy, Linguistics and Computer Science, as well as models of ontology construction developed by contemporary authors. This dissertation is characterized as qualitative and applied research. This study shows pathways for segmentation of texts in semantic networks from dissertations and theses in order to identify semantic macropropositions and types of concepts that can declare subjects relating to knowledge domains. The proposed model is supported by theories such as Concept Theory (CT), developed by I. Dahlberg and Faceted Classification Theory (FCT) of SR Ranganathan. In addition, it is also supported by conjunction of triadic models of ontology construction and conceptual analysis. The main objective of the proposal model is to support the construction of taxonomies and ontologies of domains.

Keywords: Cognitive Analysis. Conceptual Analysis. Dynamic Model of Conceptual Analysis. Knowledge Organization. Knowledge Representation. Speech representative Analysis.

LISTA DE ILUSTRAÇÕES

LISTA DE ILUSTRAÇÕES

FIGURAS

Figura 1 – Exemplo da categorização conceitual – Conceito: água fluvial	37
Figura 2 - Triângulo Conceitual de Dahlberg	46
Figura 3 Modelo para construção dos conceitos	47
Figura 4 - Ciência geral das coisas	53
Figura 5 – Diagrama conceitual do campo ampliado das Ciências Cognitivas	60
Figura 6 - Fluxo da construção das redes/sub-redes de termos	87
Figura 7 – Recurso de pesquisa – Acervo: produção científica	88
Figura 8 - Texto transformado em TXT das teses/dissertações	89
Figura 9 – Conjunto de programas para redes semânticas	91
Figura 10 – Filtragem dos termos e tratamento do corpus no Excel	92
Figura 11– Trabalho de marcação dos termos por cores no laboratório de dados	94
Figura 12 – Visualização e tratamento da rede/sub-rede no Gephi.....	96
Figura 13 – Visão geral da sub-rede – termo orla	98
Figura 14 – Grau de nós da sub-rede – termo orla	99
Figura 15 – Rede de palavras do texto/discurso	99
Figura 16 – Sub-rede do termo orla	106
Figura 17 – Exemplificação do modelo triádico de relações categoriais Fundamentais.....	108
Figura 18 – Tríade de análise terminológica.....	109
Figura 19 – Análise conceitual do termo vegetação no domínio da Geografia	111
Figura 20 – Análise conceitual do termo vegetação no domínio da Geoquímica	112
Figura 21 - Diagrama do fluxo de análise no Modac	119
Figura 22 - Caminhos da segmentação dos conceitos da rede conceitual	127
Figura 23- Sub-redes de palavras do termo clima – domínios: Geografia e Geoquímica	130
Figura 24 - Esquema das relações conceituais e da assimilação conceitual	137

Figura 25 – Sub-rede do termo clima com categorias formais.....	140
Figura 26 – Sub-rede do termo clima e suas relações	141
Figura 27 – Sub-redes do termo bactéria	145

QUADROS

Quadro 1 – Cânones para o trabalho no Plano das Ideias	43
Quadro2 - Exemplificação dos elementos fundamentais na construção de ontologias....	74
Quadro 3 – Definição formal dos conceitos da TAT	115
Quadro 4 - Apresentação dos requisitos do modelo (processos)	121
Quadro 5 - Relações entre entidades	125
Quadro 6 - Termo origem e Categoria formal e fundamental no domínio da Geografia	131
Quadro 7 - Termos relacionados e categorias no domínio da Geografia	132
Quadro 8 - Termos relacionados ao termo origem e Categorias no domínio da Geoquímica	134
Quadro 9 - Processamento da linguagem natural – quadro resumo	136
Quadro 10 – Identificação dos conceitos úmido e clima em relação ao domínio Meio Ambiente	143

LISTA DE ABREVIATURAS E SIGLAS

AC – Análise Conceitual

AnCo – Análise Cognitiva

CAOS - Conhecimento: Análise Cognitiva, Ontologia e Socialização

CC - Ciência da Computação

CI – Ciência da Informação

CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico

CVJ – Crença Verdadeira Justificada

DC - Difusão do Conhecimento

DMMDC - Doutorado Multi-institucional e Multidisciplinar em Difusão do Conhecimento

DSI – Disseminação Seletiva da Informação

IES – Instituição de Ensino Superior

IFE – Instituição Federal de Ensino

IGEO – Instituto de Geociências

ISKO - International Society for Knowledge Organization

MOBI – Modelo de Ontologia Baseado em Instâncias

Modac – Modelo dinâmico de análise conceitual

OC – Organização do Conhecimento

OI – Organização da Informação

OWL – Web Ontology Language

PDF - Portable Document Format

PMEST – Personalidade, Matéria, Energia, Espaço e Tempo

RC - Representação do Conhecimento

RI – Representação da Informação

RSD - Rede Semântica do Domínio

SOC - Sistemas de Organização e Representação do Conhecimento

SRI - Sistemas de Recuperação de Informação

TAT – Tríade de Análise Terminológica

TC – Teoria do Conceito

TCF – Teoria da Classificação Facetada

TIC - Tecnologias de Informação e Comunicação

UFBA – Universidade Federal da Bahia

SUMÁRIO

1 INTRODUÇÃO	11
2 ORGANIZAÇÃO DO CONHECIMENTO: PRINCÍPIOS TEÓRICOS E METODOLÓGICOS PARA REPRESENTAÇÃO DE CONHECIMENTO	23
2.1 A TEORIA DA CLASSIFICAÇÃO FACETADA: PRINCÍPIOS PARA O TRABALHO COM CATEGORIAS CONCEITUAIS.....	40
2.2 A TEORIA DO CONCEITO: UMA PROPOSTA DE FUNDAMENTAÇÃO TEÓRICA E METODOLÓGICA.....	45
2.3 ANÁLISE DO DISCURSO REPRESENTATIVO: UM PROCESSO PARA ORGANIZAÇÃO DE DOMÍNIOS.....	50
2.4 ANÁLISE COGNITIVA: UMA METODOLOGIA PARA ANÁLISE CONCEITUAL .	55
2.5 REDES SEMÂNTICAS/CONCEITUAIS: O PONTO DE PARTIDA PARA ANÁLISE DO CORPUS SELECIONADO.....	65
3 TAXONOMIAS E ONTOLOGIAS DE DOMÍNIOS: PARA QUE FORMALIZAR CONCEITOS?	72
4 CAMINHOS DA PESQUISA: DA TEORIA AO MODAC	80
4.1 MATERIAL E MÉTODO: O DOMÍNIO A SER REPRESENTADO	82
4.2 CORPUS SELECIONADO	84
4.2.1 Corpus para análise: a construção das redes	85
4.2.2 Aplicação do modelo de raciocínio em redes semânticas	100
4.2.3 O conceito como objeto de representação	101
4.3 PRINCÍPIOS PARA ANÁLISE CONCEITUAL: UM MODELO BASEADO EM ANÁLISE COGNITIVA.....	103
4.4 O MODELO DINÂMICO DE ANÁLISE CONCEITUAL (MODAC) – BASE PARA SOC	104
4.4.1 Análise de termo em rede semântica no domínio da Geografia física	105
4.4.2 Análise de termos em redes semânticas nos domínios da Geografia física e Geoquímica	111
4.4.3 Proposta de modelo dinâmico de análise para organização do conhecimento	115

5 APRESENTAÇÃO E APLICAÇÃO DO MODELO DINÂMICO DE ANÁLISE CONCEITUAL - Modac	116
5.1 REDES DE CONCEITOS – DADOS NÃO ESTRUTURADOS.....	124
5.2 GERAÇÃO DE CONCEITOS A PARTIR DA AMOSTRA DA REDE – DADOS ESTRUTURADOS.....	138
5.3 RESULTADOS E DISCUSSÃO SOBRE O MODELO DINÂMICO DE ANÁLISE CONCEITUAL – MODAC.....	149
6 CONSIDERAÇÕES FINAIS.....	155
REFERÊNCIAS	161
APENDICE A – REFERÊNCIAS DOS TRABALHOS QUE COMPÕEM O CORPUS (CP) - SELECIONADOS PARA ANÁLISE NA QUALIFICAÇÃO DO DOUTORADO.....	171
APÊNDICE B – FONTES PARA DEFINIÇÃO DOS TERMOS	173

1 INTRODUÇÃO

A linguagem fala. Isso significa primeiro e antes de mais nada: a linguagem fala. A linguagem? Não o homem? O que a frase nos acena agora não será ainda mais provocador? Queremos negar o homem enquanto ser dotado de fala? De modo algum. Não negaremos isso como também não negaremos a possibilidade de subordinar os fenômenos da linguagem à rubrica ‘expressão.’ Perguntaremos então: em que medida fala o homem? Perguntamos: o que é falar? (HEIDEGGER, 2011, p.15).

Este estudo e pesquisa, que proporcionou esta tese de doutorado teve sempre a intenção de entender a “fala” humana, no sentido verdadeiro do verbo “falar” e que se apresenta em uma linguagem possível e imanente ao sujeito que percebe e se expressa em bases de dados e redes que disponibilizam acesso à informação.

Para tanto, escolhemos pensar a possibilidade de uma análise de “linguagem possível” através da organização e representação do conhecimento que se dá por conceitos. Sabedores de que lidar com a linguagem é lidar também com o pensamento e com abstrações propomos um modelo dinâmico de análise conceitual, que passamos a chamar de Modac.

Na concepção do modelo encontra-se a proposta de trabalhar com variações de tipos de conceitos construídos por um *continuum* imanente de um todo possível, que pode mudar constantemente a depender da argumentação textual e da expressão de um enunciado. Entretanto, não pretendemos discutir sobre uma proposta nova de linguagem documentária nem tampouco, uma filosofia da linguagem, apenas um modo de analisar conceitos para possibilitar organização de domínios de conhecimento que lidam com aspectos de representações, de tal forma que possam atender as demandas de informação dos usuários.

O domínio da Organização do conhecimento (OC), objeto de estudo desta tese de doutorado há muito vem sendo contemplada em pesquisas, principalmente nas áreas da Ciência da informação (CI), Ciência da Computação (CC) e no campo da Difusão do Conhecimento (DC)¹. Citamos essas duas áreas e campo para delimitar em que “lugar” esta pesquisa tem sua principal sustentação. A CI, a CC e a DC, por utilizarem práticas

¹ No âmbito desta pesquisa: 1 Área de conhecimento: como construção através de diálogos entre subáreas compreendida em certos limites de estudo como apresentado na tabela de área do conhecimento do CNPq. <http://www.cnpq.br/documents/10157/186158/TabeladeAreasdoConhecimento.pdf>

interdisciplinares e colaborativas em suas pesquisas forneceram bases teóricas e metodológicas que facilitaram a construção deste trabalho.

A CI enquanto ciência social aplicada tem por objetivo estudos no âmbito da gênese e natureza da informação como apresentado por Le Coadic (2004), assim como processos de construção de artefatos para comunicação e uso da informação. Para tanto necessita do apoio de Tecnologias de Informação e Comunicação (TIC), principalmente para as questões relacionadas ao armazenamento, acesso e difusão da informação, em variados Sistemas de Recuperação de Informação (SRI), que atendem às necessidades informacionais de usuários diversificados.

Convém esclarecer, que às questões relacionadas ao uso da informação são objetos de pesquisa na CI e estão imbricados em estudos de gestão da informação e do conhecimento e ainda aprendizagem informacional. Os estudos com aprendizagem informacional têm como um de seus objetivos a inclusão do cidadão na sociedade da informação e envolvem questões culturais, comportamentais e de competências infocomunicacionais.

Não intencionamos neste trabalho estudos aprofundados sobre tais questões. A nossa preocupação é entender como o sujeito processa dinamicamente e cognitivamente a informação e a utiliza, para que possamos elaborar mecanismos de análise conceitual que viabilizem um modelo igualmente dinâmico para auxiliar a construção de Sistemas de Organização e Representação do Conhecimento (SOC).

Os SOC abrangem esquemas de Organização e Representação do Conhecimento (OC e RC) como: classificações, tesouros, taxonomias e ontologias. Sistemas de organização de natureza classificatória e temática padronizam terminologias e dentre suas principais funções, encontra-se o estabelecimento de relacionamentos semânticos entre conceitos. A padronização terminológica facilita o trabalho de indexadores de informação e conseqüentemente viabilizam melhores “respostas” às buscas dos usuários em bases de dados referenciais e às que disponibilizam textos completos.

A padronização terminológica deve atender ao consenso do que o sujeito compreende como informação, ou seja, o significado que essa informação tem para o sujeito cognoscente. Sujeito esse que ressignifica a informação atribuindo-lhe sentido para o seu uso.

A ação de ressignificação se concretiza no movimento da informação nos espaços formais e informais compreendidos pelo conhecimento registrado e o conhecimento adquirido ao longo da vida do sujeito que percebe, e assim, transforma, altera ou modifica o conhecimento.

Cabe aos profissionais que trabalham com sistemas de organização e representação da informação e do conhecimento entender a importância da ação de ressignificação, e através de processos técnicos concernentes ao trabalho de representação, aliados a tecnologia mediar possibilidades de informação. Para tanto necessitamos de sistemas que viabilizem tais processos e procedimentos e somente um trabalho interdisciplinar poderá produzi-los, como os desenvolvidos na CI, CC e DC.

A CC enquanto área vem se ocupando nas últimas décadas, de estudos de formalização de representações e operações matemáticas para ambientes digitais, com metodologias e técnicas para softwares e modelagem de dados aplicáveis a qualquer área de conhecimento, principalmente às relacionadas na atualidade às questões de redes de computadores, Internet e Web. No caso da organização de informação e conhecimento, destacamos os estudos de ontologias que vem sendo desenvolvidos pela CC.

O campo da DC, nos últimos anos vem se expandindo, principalmente porque profissionais que trabalham produzindo conhecimento, organizando e disseminando informação entenderam que o conhecimento construído na academia, resultado de pesquisas desenvolvidas colaborativamente através de práticas inter, multi ou transdisciplinares precisam ser difundidas a todo o cidadão que tem direito ao acesso a informação.

Convém esclarecer que os conceitos de disseminação e difusão são entendidos muitas vezes como sinônimos, pois ambos objetivam a democratização da informação. Quando o conceito de disseminação se amplia com a Disseminação Seletiva da Informação (DSI) entende-se que o serviço se torna especializado, aliam-se ao serviço, práticas de estudo de usuários, personalizando a informação para um determinado público. Preferimos falar de difusão da informação e do conhecimento, em uma perspectiva contemporânea de levar a informação a todos de forma dirigida ou não, seguindo os caminhos e evolução dos meios de comunicação.

Nesta ótica, a difusão precisa de estudos voltados para linguagens que facilitem a busca de informação, seja natural, controlada ou livre. Linguagens que atendam as expectativas não só dos pesquisadores dos domínios de conhecimento, mas, sobretudo, aos que anseiam selecionar, encontrar e obter informações atinentes às suas necessidades nas bases de dados. Sendo assim, os estudos de modelagem de conhecimento e Análise Cognitiva (AnCo) são fundamentais para o campo da DC. Trabalho como o de Rios; Fróes Burnham (2012) esclarece sobre os processos da difusão do conhecimento.

No propósito de encontrar um caminho para a difusão, alguns domínios de conhecimento como: os da linguística e análise cognitiva e práticas como análise do discurso não poderiam ficar a margem nesta tese, pois colaboram com lastros que garantem a base epistemológica deste trabalho. Estudos nos domínios citados nos forneceram suporte para apontarmos teorias de fundamentação para o modelo dinâmico, que apresentamos como possibilidade de análise conceitual.

A organização do conhecimento pressupõe o trabalho com modelização de domínios de conhecimento na intenção de construção de bases de relacionamentos entre categorias de termos e designação de conceitos que facilitem sua compreensão e permitam compartilhamento conceitual.

Na construção do Modac, o estudo e entendimento de metodologias, teorias e ferramentas utilizadas em modelização foram fundamentais. Este modelo está baseado em análise conceitual e na organização dos conceitos advindos de domínios inter e multidisciplinares e pretende auxiliar na construção de taxonomias e ontologias de domínios de conhecimento.

A organização do conhecimento se pauta ainda, no trabalho com categorias, por entender que desde a época aristotélica o conhecimento é representado por uma classificação de algum modo. Afinal, classificação é um processo mental, que ordena as coisas a partir da percepção do indivíduo dando a elas sentidos. As categorias são responsáveis pela ordenação conceitual e consequentemente delimitará o uso do conceito a partir dos moldes estabelecidos nas comunidades discursivas.

Categorizar não significa engessar ou limitar a OC. As categorias objetivam a interpretação particular de um determinado domínio, uma interpretação conceitual parcial, ou seja, do que significa o conceito em tempo e espaço, ou ainda no tempo atual da pesquisa, e consequentemente para o domínio de conhecimento.

Através do modelo de análise para construção de taxonomias e ontologias, do trabalho com relações categoriais e das relações associativas é possível verificar a manifestação dos conceitos que traduzem as pesquisas desenvolvidas nos domínios de conhecimento. Acreditamos que a construção de taxonomias e ontologias de domínios depende de teoria de fundamentação e de um modelo dinâmico de representação conceitual, pois ambas trabalham a todo o momento com a complexidade inerente a tarefa de definição, categorização e organização de conceitos que emergem nos domínios.

Taxonomias são estruturas classificatórias que vão além das utilizadas em esquemas de classificação, como: os filosóficos, bibliográficos e sistemáticos, pois existe uma preocupação esquemática na ordenação taxonômica em que a estrutura hierárquica dos termos corresponda aos fenômenos por eles representados, ou seja, de acordo com pontos de vista, sejam: científicos, tecnológicos, filosóficos, artísticos, históricos e culturais.

As classificações filosóficas serviram aos esquemas de classificações bibliográficos e sistemáticos, dado que sua elaboração tem como base a filosofia do conhecimento. Criada por filósofos, epistemólogos, enciclopedistas, às classificações objetivam hierarquizar o conhecimento de forma geral. Trabalhos na área da CI e Biblioteconomia, como o de Barbosa (1969) esclarecem sobre os objetivos das classificações.

Os esquemas de classificações bibliográficos foram desenvolvidos para atender a organização das coleções de bibliotecas, reunindo os livros por assuntos, com o objetivo de dar um lugar a esses livros na coleção, mesmo que de forma relativa. Uma classificação sempre será relativa, pois é dependente de um contexto, se adéqua à abrangência dos assuntos tratados em determinada coleção, atendendo às necessidades de uma determinada comunidade que a utiliza. As classificações ordenam objetos por assuntos, semelhanças e contrastes ou diferenças e requererem sensibilidade e entendimento sobre a entidade a ser classificada e categorizada.

As categorias não devem ser reduzidas às classificações, pois a categorização acrescenta e agrupa informação. Como no caso das taxonomias, que embora sejam listas estruturadas de categorias de conceitos, para o caso da organização do conhecimento, podem ser definidas como estruturas classificatórias não estáticas que mapeiam o conhecimento e agregam dados.

As taxonomias representam através de conceitos o mapeamento de um dado domínio de conhecimento, onde termos são organizados a partir das relações hierárquicas e partitivas de acordo com as necessidades de ambientes específicos sendo consideradas como “espinha dorsal” de ontologias.

Ontologias de domínio, enquanto conceito utilizado atualmente nas ciências que se encarregam de organização de conhecimento, como a Ciência da Informação e da Computação, principalmente, surgem da necessidade de representações formais de conceitos, com finalidade que vai além da mera definição dos conceitos, mas principalmente, com a intenção de revelar acontecimentos de um dado contexto. Têm as ontologias o objetivo do compartilhamento do conhecimento e compatibilização semântica.

Existem na literatura vários registros de tipos de ontologias ou de reuso de ontologias, assim como metodologias para construção e estudos relacionados. Ontologias de domínios têm por objetivo apresentar além de um vocabulário controlado de conceitos, seus relacionamentos sobre determinadas atividades ou ainda acordo ontológico. No decorrer dos últimos anos surgiram nas universidades brasileiras grupos de estudos voltados para a construção de taxonomias e ontologias que buscam soluções para OC. Dissertações e teses foram defendidas neste âmbito nos domínios abarcados pela CC, CI e DC.

Embora existam pesquisas com diferentes enfoques no domínio da OC que envolvem construção de taxonomias e ontologias de domínios, principalmente na área da CI, observamos a partir de leituras realizadas nos estudos desenvolvidos na última década por mestrandos e doutorandos e autores contemporâneos à importância de estudos dessa natureza e ainda a seguinte lacuna: ainda persiste a carência de estudos e aplicações e modelos de construção de SOC, para taxonomias e ontologias, que sustentem a organização de relações conceituais dinâmicas que ocorrem nos domínios de conhecimento principalmente quanto aos

domínios inter e multidisciplinares. Citamos trabalhos que evidenciam de uma forma ou de outra a nossa afirmação: (CAMPOS, 2004, CAMPOS et. al., 2009, FRANCELIN; KOBASHI, 2011, GOMES, 2009, MOREIRA; LARA, 2011, SOUZA JUNIOR; CAFÉ, 2012, VITAL; CAFÉ, 2011).

Em trabalho recente desenvolvido por Jorge (2012) no DMMDC o autor propõe um Modelo de Ontologia Baseado em Instâncias (MOBI) cujo objetivo principal é a estrutura “*botton-up*”, ou seja, parte da especialidade do termo para a sua generalização. Esse modelo é importante porque dá a possibilidade de entender um conceito que pode ter várias definições ou classificações. A partir deste entendimento um conceito pode ser reutilizado por diferentes domínios.

Embora o modelo de Jorge (2012) tenha sua base na CC existe a mesma preocupação que o modelo apresentado nesta tese a de entender que mesmo quando generalizamos um conceito, quando abstraímos sobre o objeto representado, a nossa intenção é utiliza-lo enquanto especificidade.

Os modelos para construção de SOC devem ser capazes de possibilitar ao sistema de recuperação de informação e de organização do conhecimento, como preconizado no The Léon Manifesto (2007), a liberdade para os usuários optarem em mudar de uma perspectiva para outra, e de perceberem a multidimensionalidade do pensamento complexo.

Assim pensamos: quais os caminhos teóricos e epistemológicos e metodológicos a serem percorridos para esta construção? A partir deste questionamento apontamos a questão norteadora desta pesquisa:

- um modelo dinâmico de análise conceitual baseado na Teoria do Conceito (TC) e Teoria da Classificação Facetada (TCF) e em Análise Cognitiva (AnCo) contribuirá para minimizar a lacuna de modelos para construção de SOC para domínios inter e multidisciplinares? Assim apresentamos algumas proposições e pressupostos a serem investigados:

Questões complementares de pesquisa:

- Como representar conceitos concebidos nos domínios inter e multidisciplinares?

- Como construir um modelo dinâmico de análise conceitual para domínios de conhecimento inter e multidisciplinares?
- As TC e TCF e AnCo se adequam a modelos dinâmicos de análise conceitual para domínios inter e multidisciplinares?
- Qual o tipo de análise deverá ser empregada para a construção de modelos dinâmicos de análise conceitual?

Pressupostos:

- A construção de um modelo de análise conceitual em um domínio de conhecimento inter ou multidisciplinar pressupõe o entendimento das diversidades discursivas de cada comunidade de discurso.
- Duas ou mais comunidades discursivas em um mesmo domínio de conhecimento inter ou multidisciplinares possuem níveis de compatibilização semântica diferenciados, a partir do enfoque referencial, teórico metodológico aplicado em seus estudos.

Foram traçados objetivos, para nortear a metodologia empregada na tese:

Objetivo geral: Construir um modelo dinâmico de análise conceitual para auxiliar na construção de SOC, especificamente, taxonomias e ontologias de domínios para ambientes inter ou multidisciplinares.

Objetivos específicos:

- Apresentar teorias e base epistemológica para fundamentação e construção de modelos dinâmicos para Sistemas de Organização do Conhecimento (taxonomias e ontologias de domínios);
- Apresentar relações conceituais com base nos argumentos textuais de autores;
- Comprovar que a análise conceitual é uma metodologia de interpretação e análise cognitiva.
- Apresentar um Modelo dinâmico de análise conceitual.

A justificativa para o desenvolvimento deste trabalho se deve ao tema central da pesquisa: proposta de análise conceitual baseada em análise cognitiva, princípios teóricos para construção de modelos para Sistemas de Organização do Conhecimento (SOC), e pretende contribuir com pesquisas em desenvolvimento no âmbito da Ciência da Informação, Ciência da Computação e Difusão do Conhecimento e continuar o trabalho iniciado no ano de 2005, que culminou em dissertação de mestrado defendida no domínio da representação, em taxonomia de domínios interdisciplinares.

No término do mestrado em Ciência da Informação, no ano de 2007, quando desenvolvemos um estudo de princípios classificatórios para taxonomias, percebemos que os estudos em ontologias de domínios avançavam devido a constante aceleração das tecnologias digitais, e que pesquisas desta natureza precisavam caminhar continuamente. Afinal, é fundamental garantir uma representação e organização do conhecimento que atenda às expectativas de busca do usuário. Isso só é possível quando se criam ferramentas que possibilitam a sua classificação, análise conceitual e de conteúdo e controle terminológico.

Além disso, ao longo dos últimos 10 (dez) anos os estudos voltados para a temática da organização e representação do conhecimento têm se expandindo, mas estudiosos desses domínios ainda vêm procurando estabelecer bases teóricas, abordagens epistemológicas e metodológicas, adequadas ao desenvolvimento de SOC.

Sendo a OC um domínio inter e multidisciplinar com temas variados de pesquisas relacionados aos SOC, principalmente, os que dizem respeito às questões semânticas e tecnológicas voltadas à recuperação de informações dispostas nas redes de comunicação, fica evidente a necessidade de estudos cooperativos e colaborativos entre áreas e domínios como a CI, CC e DC. Em pesquisa recente sobre tendências da pesquisa em OC, Gomes (2009) comprova esta evidência.

Construir ferramentas que possam auxiliar na indexação e recuperação de itens documentais, que possam contribuir para a difusão e resgate da memória, que divulguem acervos e coleções, que atendam aos anseios dos pesquisadores/usuários depende de métodos e modelos de construção de SOC bem como teorias que sustentem essa construção, mas que acima de tudo mantenham sua função prática. Assim acreditamos nessa escolha e a justificamos.

Para alcançarmos o objetivo desta pesquisa nos debruçamos sobre estudos de natureza semelhante como: dissertações, teses, periódicos, material bibliográfico e documental. Após leitura preliminar desses materiais e com vista à identificação de bases teóricas, e por já termos a experiência anterior do estudo desenvolvido no mestrado em CI optamos por uma teoria que nos pareceu fundamental para construção do modelo proposto: a Teoria do Conceito desenvolvida por Ingtraut Dahlberg na década de 70 do século XX e utilizada por

autores contemporâneos para trabalhos com OC e RC. Sem, contudo, excluir a Teoria da Classificação Facetada, de Shiyali Ramamrita Ranganathan, (1967), que nos permitiu um pensar dinâmico sobre assuntos dos documentos e a estruturação das categorias fundamentais.

A abordagem empírica para aplicação do modelo de análise que propomos para organização do conhecimento de domínios deu-se nas áreas das Ciências Exatas e da Terra, subáreas da Geografia, Geologia e na especialidade da Geoquímica. Elegemos domínios que se articulam interdisciplinarmente e multidisciplinarmente nessas áreas para atender às pesquisas no âmbito das questões ambientais, com foco no discurso das comunidades acadêmico científicas.

As áreas, subáreas e especialidades aqui determinadas seguem as especificações da tabela de áreas do conhecimento determinada pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), após estudos desenvolvidos por equipe interdisciplinar de especialistas, incluindo pesquisadores da Ciência da Informação.

O estudo desenvolvido nesta tese, seguindo a linha proposta teve como base metodológica a Análise do Discurso, a partir das concepções de Michel Foucault para processos de representação. Como também, a Análise Cognitiva (AnCo), campo em construção, objeto de estudos do grupo de pesquisa: Conhecimento: Análise Cognitiva, Ontologia e Socialização CAOS e do Doutorado Multi-institucional e Multidisciplinar em Difusão do Conhecimento – DMMDC. Trabalho de Lage; Fróes Burnham; Michinel (2012) esclarece sobre o campo da AnCo.

Modelos desenvolvidos para organização do conhecimento de domínios inter e multidisciplinares têm por objetivo atender questões de representação do mundo da experiência, onde são organizados os conceitos advindos de domínios de conhecimento que são expostos a condições que os tornam possíveis.

Estas condições podem ser entendidas a partir da análise das relações entre os conceitos, esses conceitos muitas vezes originam-se em outros domínios de conhecimento que se articulam com o domínio a ser representado por uma ontologia, por exemplo, por questões de trabalhos inter ou multidisciplinares. Nasce assim uma classe especial de representação,

sem limite de ordenação ou duração, mas das consequências possíveis desses relacionamentos.

O ponto de partida para aplicação do modelo dinâmico de análise conceitual foi as redes conceituais, organizadas a partir das redes semânticas. Em seguida construímos uma estrutura terminológica com relações associativas entre os conceitos baseada em taxonomias e ontologias de domínios, tal estrutura demonstrou a eficácia do modelo, sendo assim caracterizamos esta pesquisa como qualitativa e aplicada.

Em estudo realizado sobre análise de redes semânticas, baseada em títulos de periódicos científicos, a relevância no uso dessas redes para processos cognitivos foi comprovada por Fadigas et. al (2009) como fundamento teórico. Acreditam os autores que as redes podem ser usadas para representar conhecimento. Seguindo a representação estabelecida na rede, a utilizamos como um construto que possibilitou a análise dos argumentos textuais através das redes de conceitos. A partir dessas redes construímos o Modac e o apresentamos como nossa proposta de tese.

O trabalho de pesquisa que realizamos e que resultou nesta tese de doutorado foi organizado em capítulos apresentados no sumário. O conteúdo abordado em cada capítulo ora descrevemos de forma resumida:

No capítulo 1 (um), Introdução - apresentamos uma visão geral da pesquisa bem como: definição do problema; questões de pesquisa; pressupostos; objetivos: geral e específicos, justificativa e relevância da pesquisa.

No Capítulo 2 (dois), o referencial teórico da tese. Discorremos sobre Organização do Conhecimento: princípios teóricos e metodológicos para Representação do Conhecimento. Apresentamos a Teoria da Classificação Facetada que nos permitiu trabalhar a categorização que possibilitou o entendimento das relações entre entidades. Apresentamos também: a Teoria do Conceito; Análise do Discurso Representativo, Análise Cognitiva e redes semânticas objetivando uma proposta metodológica de raciocínio para a análise conceitual.

No Capítulo 3 (três), abordamos a temática das SOC: Taxonomias e Ontologias de domínios: para que formalizar conceitos?

No capítulo 4 (quatro) a proposta metodológica: os caminhos da pesquisa: da base teórica ao Modac; material e método: o domínio representado; o corpus selecionado; corpus para análise: a construção das redes; aplicação do modelo de raciocínio em redes semânticas; o conceito como objeto de representação; princípios para análise conceitual: um modelo baseado em análise cognitiva; o modelo dinâmico de representação – base para SOC; Análise de termos em redes semânticas e a proposta de modelo dinâmico de análise para organização do conhecimento.

No capítulo 5 (cinco) apresentamos o Modac e a sua aplicação, partimos dos dados não estruturados, as linguagens naturais a partir das redes conceituais. Em seguida demonstramos a aplicação do modelo e ainda os resultados e discussão.

No capítulo 6 (seis) as considerações finais da tese.

Apresentamos após o capítulo conclusivo, as referências das obras que serviram de base para o estudo e pesquisa e as que possibilitaram as definições dos termos analisados, (apêndice A e B). Como também as que foram citadas e utilizadas ao longo do trabalho de tese.

Por fim chegamos a um resultado que apontou um modelo cognitivo e dinâmico de análise de domínios para auxiliar a construção de taxonomias e ontologias de domínios e esperamos contribuir para pesquisas e trabalhos voltados para Sistemas de Organização e Representação do Conhecimento (SOC).

2 ORGANIZAÇÃO DO CONHECIMENTO: PRINCÍPIOS TEÓRICOS E METODOLÓGICOS PARA REPRESENTAÇÃO DE CONHECIMENTO

Talvez agora tenhamos condições de compreender exatamente que realização a fala representa para nós, de que maneira prolonga e de que maneira transforma a relação muda com o outro. Num certo sentido, a fala do outro não atravessa nosso silêncio, nada pode nos dar além de seus gestos: a dificuldade é a mesma de compreender como podem palavras dispostas em preposições nos significar outra coisa que não nosso próprio pensamento [...] (MERLEAU-PONTY, 2002, p.173).

Pensamento é um termo de difícil compreensão por ter uma extensão abrangente no que diz respeito à atividade cerebral humana, cujo limite não se pode alcançar. Ou pelo menos, não podemos observá-lo diretamente, a não ser pela análise do comportamento ou manifestação do pensamento através das palavras.

Esta manifestação difere em grau de complexidade e organização, pois se estendem ao conhecimento humano e às atividades mentais do indivíduo que percebe o mundo fenomênico. Desde o cientista que realiza pesquisas até a criança nos seus primeiros contatos com esse mundo, o que há de comum em todos nós indivíduos é que estamos empenhados em pensar e conhecer, seja qual for o grau de profundidade ou de organização mental. O importante é o ato de pensar que consideramos para esta pesquisa como uma atividade mental em constante movimento.

A natureza do pensamento pertence ao domínio da Filosofia e não é a nossa intenção discorrer sobre tais aspectos. Procuramos nos ater a ação do pensamento que nos guiará ao nosso objetivo o da **percepção conceitual**.

Para compreender tal percepção é preciso entender também o motivo pelo qual, o indivíduo pensa sobre o objeto/conceito, nesse caso, acreditamos que seja a **necessidade de informação** que inicia e determina a direção do pensamento.

E por fim precisamos nos ater aos **processos do pensamento** como: identificação, seleção, eliminação e organização que variam do pensamento associativo ao pensamento dirigido aos propósitos de busca para obtenção de itens documentais que levam ao conhecimento. Precisamos conhecer os usuários que utilizam as bases de dados e principalmente os conteúdos que eles querem encontrar e como selecionam os documentos a partir dos conceitos que são indexados como descritores de textos.

Sabemos que o conhecimento precisa ser organizado para ser difundido, para tanto é preciso colocar o indivíduo à frente de todo processo, para por fim promover o encontro do indivíduo com o item de materialização do conhecimento, o documento. Esse encontro é possível graças aos instrumentos de busca como: buscadores e base de dados, que se configuram como canais de comunicação. Mas para que esses canais atendam aos seus propósitos é preciso organizar o conhecimento que está contido nos itens documentais e disponíveis nos repositórios de informações.

Além disso, pelo que expomos neste início de capítulo, é preciso declarar que no processo de Organização do Conhecimento (OC) existe um trabalho de análise do “pensamento”, assim, acreditamos que como não há conhecimento sem conceituação, o trabalho com OC prescinde um trabalho de análise cognitiva e conceitual.

A OC se configura como domínio de conhecimento nas áreas da Ciência da Informação (CI) e Ciência da Computação (CC) assim como para os campos de estudos Representação do Conhecimento (RC) e Difusão de Conhecimento (DC). Organizar conhecimento pressupõe um trabalho teórico e metodológico para organização de conceitos advindos desses domínios que culminará em uma representação de domínios de conhecimento.

Sabemos que conhecimento pressupõe um sujeito que conhece intuitivamente e a ideia conhecida do objeto, provocando com isso uma individualidade advinda da vontade e da necessidade. Esta individualidade por vezes empobrece a ação da difusão. É preciso socializar o conhecimento, torná-lo representação ou manifestação nítida dos objetos, onde esses objetos tenham um sentido evidente para o indivíduo. Para Foucault (1999, p.72) “[...] não há conhecimento verdadeiro senão pela intuição, isto é, por um ato singular da inteligência pura e atenta, e pela dedução que liga entre si as evidências.”

A definição de conhecimento vem sendo discutida ao longo dos séculos pela filosofia. Platão, na obra Teeteto apresenta um diálogo sobre a natureza do conhecimento, um confronto entre realidade e relativismo, ou seja, defende que o conhecimento que o sujeito tem sobre um objeto (representação mental desse objeto) deva ser justificada. Para o filósofo, o

conhecimento seria a opinião verdadeira acompanhada de razão. Platão defendia a Teoria da Crença Verdadeira Justificada (CVJ).

Em artigo sobre o discurso filosófico de Platão, Santos (2001) nos apresenta uma leitura do pensamento Platônico acerca de conhecimento. Destacamos aqui uma breve síntese do pensamento desse Filósofo sobre a temática do conhecimento verdadeiro apresentado no texto da autora:

Toda especulação desenvolvida a partir do *Fedro* e, de modo especial, no *Sofista*, sobre as condições do verdadeiro conhecimento, articula-se à rigorosa investigação dos procedimentos metodológicos, simultaneamente à proposição de uma ‘metafísica’ da linguagem que possibilite a definição adequada das coisas. [...]
 Todo conhecimento é conhecimento de algo (Platão, *Rep.*, 476d-477a), tem por objeto um ser que ultrapassa a dimensão do sujeito cognoscente e que funciona como valor-base de indiscutível objetividade (SANTOS, 2001, p. 274-297).

Na contemporaneidade Gettier, (1963), como crítico da Teoria CJV apresenta outra concepção com relação a conhecimento e verdade. Para Gettier, o conhecimento vai além da concepção tripartida (crença + verdade + justificação), esta seria quadripartida. Ou seja, existe para ele, além desses três elementos uma condição extra, ou condições necessárias suplementares para validar o conhecimento.

Tanto a concepção de Platão como a de Gettier foram analisadas e debatidas por estudiosos como: pensadores, epistemólogos e filósofos. A concepção de verdade, não é nosso objeto de estudo na tese, nem tampouco é nossa intenção aprofundarmos sobre as questões filosóficas sobre o conhecimento. Preferimos dizer, que o trabalho com a verdade na representação conceitual diz respeito à questão do consenso em determinado domínio de conhecimento, sobre como definir um determinado objeto, formando um tipo de generalização que será útil à representação desse objeto, ou como dito por Habermas (1976) em uma situação ideal de discurso.

Preferimos dizer que existe uma possibilidade de organização e representação do conhecimento de um domínio, quando adequamos um conjunto de condições e formulamos um modelo possível de análise conceitual que viabilizará a representação. Um conjunto dinâmico com condições flexíveis, capazes de satisfazer a exigência de cada domínio de conhecimento.

A representação do conhecimento pretende: a apreciação por parte do indivíduo sobre o movimento conceitual, ou da ação que o conceito provoca em um domínio quando se associa ou se relaciona a outro conceito. O sujeito pode passar a conhecer o domínio de conhecimento quando convive com os objetos do espaço representado, e optar pela sua adoção enquanto conceito, ou não. Nas considerações de Maturana:

O conhecimento é uma apreciação de um observador sobre a conduta do outro, que pode ser ele mesmo. No momento em que se vê isto desta forma, por um lado, descobre-se que o conhecimento é sempre adquirido na convivência. Descobre-se que se aprende a ser de uma ou outra maneira na convivência com outros seres humanos. Por outro lado, descobre-se que o conhecimento tem a ver com as ações. Tem a ver com ações consideradas adequadas em um domínio particular. Desse modo, se quero que alguém adquira conhecimento em um certo espaço, o que tenho que fazer é convidá-lo a viver em um certo domínio; se ele aceita, terá esse conhecimento. (MATURANA, 2001, p. 123).

Comparamos a convivência do indivíduo no espaço representacional como a de uma criança em um espaço novo como a escola por exemplo. Por vezes lhe falta compreensão daquilo que seus sentidos percebem, com a maturação do aprendizado e a convivência no espaço, essa criança vai se tornando capaz de perceber e interpretar o que recebe nesse espaço e passa a compreendê-lo de tal modo a formar e organizar conceitos. Através da convivência passamos a selecionar o que é significativo em um determinado espaço, conceituando e conhecendo.

O conhecimento resulta, assim, de uma ecologia regida pela interação social e os instrumentos de registro, acesso e processamento das informações que representam uma cadeia formada pela inter-relação de acontecimentos antecessores, enfim é um ato humano que se apoia nos recursos tecnológicos de extensão da memória. (GOMES, 2008, p.2).

Mas o Conhecimento para ser socializado prescinde de formalizações lógicas fornecendo aos fatos e ideias, aspectos semânticos que os aproximam da realidade em que vão ser representados, permitindo mecanismos de inferência por parte dos usuários que deduzindo ou ressignificando informações as tornam úteis aos seus propósitos. Ressaltamos aqui a importância da semântica na composição e especificações de modelos de organização e representação, como no caso de ontologias de domínios. Essa temática será abordada em seção própria (2.5) como também em capítulo específico (3), de forma a elucidar as relações entre tais modelos de representação.

O termo representação já foi discutido em inúmeros trabalhos e nas mais variadas áreas de conhecimento. Não é objetivo deste trabalho uma abordagem aprofundada do tema, apenas apresentar alguns pontos de vista. A citação de Kobashi resume o caminho trilhado na tese, principalmente em relação ao objeto, sua percepção e apreensão.

Na tradição filosófica a palavra ‘representação’ refere-se aos diversos modos de **apreensão de um objeto**. Assim no sentido aristotélico a representação é assimilada à fantasia intelectual e ou sensível, enquanto para os estóicos ela está relacionada à impressão direta ou indireta. Para Descartes, ela é imaginação, para Spinoza, apreensão sensível, enquanto em Kant ela é **apreensão intuitiva ou conceptual**. No sentido da percepção, ela está presente em Leibniz, sendo a representação o mesmo que ideia, em Locke e Hume. (KOBASHI, 1996, p.11, grifo nosso).

É incontestável que para a apreensão de um objeto é preciso dar-lhe sentido, ter o contato com o objeto ou uma situação em que o conceito se exprime. Para representar é preciso generalizar, diferenciar, abstrair, simbolizar e porque não dizer enunciar. Destacamos aqui o que julgamos mais importante para representarmos conceitos, mas evidentemente existem caminhos e linhas de pensamentos a seguir e não é tarefa fácil a representação de conhecimento.

As bases teóricas e epistemológicas de um trabalho dessa natureza há muito vem sendo objeto de pesquisas da Ciência da Informação. Os trabalhos de representação são pautados e entrelaçados com estudos da: Filosofia, Psicologia, Educação, Ciência da Computação, Linguística, dentre outras áreas. Ampliam assim possibilidades de aplicações, principalmente na atualidade, para atender as formas de representações da web, objetivando o acesso à informação.

Os estudos desenvolvidos no âmbito da representação, na Ciência da Informação, segundo Francelin; Kobashi, (2011) são apoiados em amplos referenciais teóricos. O que não é diferente para as demais áreas. Acompanhando esta ideia apresentamos neste texto teorias para fundamentação e bases epistemológicas que sustentam um modelo dinâmico de organização e representação conceitual, objeto deste estudo de tese em OC.

Para Dahlberg (2006) a OC é uma ciência que estrutura e sistematiza unidades de conhecimento, os conceitos, de acordo com características comuns. Para a ISKO (International Society for Knowledge Organization), uma nova disciplina com fundamentação científica e

teórica, um ramo da ciência que se propõe, em conjunto com cientistas, organizadores de conhecimento e terminólogos, a definição e sistematização de conceitos de todos os campos disciplinares, através de classificação e categorização. Nas considerações de Campos e Gomes:

Organização, no contexto da ciência da informação/documentação, implica no conceito de classificação e nenhuma outra área do conhecimento buscou desenvolvimento de bases teóricas da classificação como ali. A classificação está presente na organização automática dos menus/diretórios, que se caracterizam por classificação de assuntos; está presente na classificação automática dos recursos eletrônicos, mais especificamente na construção dos índices das ferramentas de busca, baseados na varredura de textos completos, utilizando técnicas desenvolvidas nos anos 50 e 60 do século passado; está presente nas ontologias, voltadas para a Inteligência Artificial. E na Rede Semântica, proposta pelo Consórcio WWW3. (CAMPOS; GOMES, 2003, p.151)

A falta de padronização na prática da organização do conhecimento favorece a uma compreensão equivocada dos conceitos advindos dos domínios, dificultando também uma base terminológica apropriada, ou ainda um vocabulário que facilite a recuperação de informações relevantes para usuários e que possa ser compartilhado entre domínios como é o caso das ontologias.

A importância de se organizar a produção do conhecimento de um dado domínio se explica pela medição de como tal produção influencia em pesquisas atuais e futuras, de como se percebe e se recupera tais pesquisas através dos conceitos emanados dos domínios e ainda de como são assimiladas essas novas experiências conceituais por parte dos usuários.

Pois se o mundo é pensado e entendido por representações conceituais, há de se trabalhar com sujeito e objeto (conceito) sem disjunção, observando atentamente as condições do pensamento ou suas abstrações e porque não dizer da intuição. Nessa perspectiva Schopenhauer teceu considerações sobre as representações abstratas:

“[...] a maior diferença a assinalar entre as nossas representações é a do estado intuitivo e do estado abstrato. As representações de ordem abstrata formam apenas uma única classe, a dos conceitos, apanágio exclusivo do homem neste mundo.” (SCHOPENHAUER, 2001, p.12).

Para alcançarmos o objetivo da organização do conhecimento dos domínios será necessário analisar os conceitos que surgem em seu interior a partir das pesquisas

desenvolvidas e do seu discurso, ou ainda, das ideias apresentadas nas produções dos pesquisadores desse universo, afinal para Ranganathan (1967, p.81) o conhecimento é definido como “[...] totalidade de ideias conservadas pela humanidade”, ou seja, “universo de ideias”. Os assuntos que surgem nos domínios de conhecimento são, portanto “conjuntos de ideias” que são organizados e sistematizados.

As ideias são expressas através da linguagem, pelas palavras, do dizer e não dizer sobre as coisas, que devem ser organizadas para serem difundidas. Para Maturana (2002, p.20) “[...] a linguagem está relacionada com coordenações de ação, mas não com qualquer coordenação de ação, apenas com coordenação de ações consensuais.” O trabalho de organização do conhecimento e da informação para alcançar a ordenação dessas ideias, lança mão de instrumentos como ontologias e taxonomias que se configuram essencialmente de estruturas conceituais e relações consensuais entre conceitos.

Os instrumentos citados, ontologias e taxonomias, surgem na Inteligência artificial visando à manipulação de dados de um domínio de conhecimento, principalmente pelas questões atuais de acesso automatizado à informação, e que depende de organização do conhecimento advindo das pesquisas desenvolvidas nos domínios de conhecimento.

A expressão organização da informação se aplica às bases referenciais, enquanto organização do conhecimento passa a incluir a possibilidade de utilização de mecanismos que manipulam textos integrais e multimídia, que são formas atuais de representação do conhecimento. E todos esses recursos não prescindem de uma organização para que possam ser melhor explorados e recuperados. (CAMPOS; GOMES, 2003, p. 2)

A OC está presente na representação das redes semânticas ou conceituais, quando evidencia as relações entre os termos e palavras, que ao serem analisados (as) e interpretados (as) enquanto conceitos conferem movimentos de geração de um universo de ideias conectadas. A representação expressa nas redes de conceitos de domínios de conhecimento pode facilitar a detecção da produção de conhecimento de um dado domínio.

Tanto a RI quanto a Organização da Informação (OI) tem em seus estudos e práticas a mesma intenção: difundir informação; enquanto na OC e RC, o propósito é a construção de mecanismos de estruturação de bases de dados conceituais. Objetivam a RI e OI, levar o conhecimento à pesquisadores e usuários, em uma proposta de propagação, onde se possa

atender a todos os indivíduos em seus anseios de informação, possibilitando a construção de um conhecimento que venha a colaborar com o desenvolvimento cidadão.

A espécie humana em sua essência há muito, vem insistentemente agindo sobre o meio em que habita na tentativa de transformá-lo, num contínuo processo dinâmico, desenvolvendo também formas de registrar seus inventos e sua produção. Os registros intelectuais, produzidos a princípio, cresciam de forma gradual, mas, na atualidade, com o desenvolvimento tecnológico, esta produção passou a ser publicada em meio digital e eletrônico quase que simultaneamente à sua construção intelectual.

A velocidade das publicações, principalmente nas redes de informação como a Internet, vem instigando os profissionais que trabalham organizando conhecimento a encontrar meios de acesso à informação de forma rápida e eficaz, principalmente no que concerne a conteúdo temático. A principal preocupação está em como disponibilizar tais registros de forma atinente (concernente) às necessidades informacionais do usuário.

A percepção é de que a necessidade de pesquisas voltadas para organização, representação e recuperação de informação foi potencializada na era digital e virtual, mas embora crescendo e tentando acompanhar a velocidade com que as informações são disponibilizadas, e mais ainda, a atender as demandas de pesquisa, ainda persiste a demanda por estudos que aperfeiçoem os instrumentos de OC e RC. O que é justificável, pois, o que interessa aos profissionais cujo trabalho é organizar conhecimento para representá-lo é “poupar o tempo do usuário” e que esse usuário tenha acesso ao conteúdo dos textos. Como consequência, o trabalho de OC facilita os profissionais que trabalham com sistemas de recuperação de informação.

É importante frisar que não basta apenas traduzir o conteúdo temático de um texto ou os assuntos nele tratados para uma linguagem artificial como, por exemplo, a dos vocabulários controlados. Não é só atender a certos princípios de recuperação, dando ao documento pontos de acesso por assunto, mas acima de tudo, atender à expectativa do autor intelectual do documento e do usuário que busca uma informação, facilitando a interpretação dos conteúdos dos documentos.

Analisar os assuntos ou temas transcritos nos documentos e traduzi-los através de conceitos, resultado da organização do conhecimento, não é uma tarefa simples, pois a base da descrição autoral ou intelectual se faz por palavras carregadas de aspectos sintáticos e semânticos próprios da linguagem natural. Além disso, é preciso entender o discurso impresso no texto, para que enfim se possa representar o conhecimento e alcançar o objetivo maior o da recuperação e acesso a informação.

A organização do conhecimento se fundamenta nos estudos dos níveis lógicos, epistemológicos, ontológicos e conceituais, uma síntese do que foi apresentado por Brachman (1979), passamos a elencar:

- no nível lógico o foco está na análise da ação do pensar;
- o nível epistemológico é o da estruturação conceitual;
- o ontológico com base nas formalizações conceituais restringe os números de possibilidades de interpretação do conceito, garantindo a representação do contexto e do domínio de conhecimento no qual o conceito se insere;
- no conceitual, parte-se da premissa de que “a priori” o conceito possui uma interpretação definida, independente do formalismo.

Os níveis apresentados por Brachman estão de uma forma ou de outra, contemplados na Teoria do Conceito que apresentamos neste capítulo na seção 2.2, principalmente no que concerne a estruturação do conhecimento de forma sistematizada a partir do contexto do domínio; apontamos como fundamentais os níveis apresentados para sistemas de organização e representação do conhecimento.

Atentamos também que para o trabalho com os níveis aqui apresentados insere-se a análise cognitiva em seu aspecto de processos cognitivos informacionais e comunicacionais, mas que também se aproxima da dimensão do aspecto de socialização do conhecimento, quando auxilia no processo de OC que busca difundir conhecimento. Buscamos uma análise que dê lugar ao sujeito que percebe e fala.

OC para o caso desta pesquisa deve ser considerada como um ato de categorizar conceitos como também analisar as relações conceituais de um domínio com base em

princípios epistemológicos, teóricos e metodológicos que facilitará a Representação do conhecimento.

Representação que envolve a organização intelectual e cognitiva, com base nos conceitos, “unidades de pensamento” que carregam características únicas e quando associadas a outros conceitos, em suas relações, se tornam entidades interdependentes para traduzir um determinado contexto ou domínio. Sobre as relações conceituais, Campos (2001a), aborda a importância das características do conceito e enfatiza:

Os conceitos se relacionam entre si por existirem características comuns entre eles. São elas que determinam o tipo de relação que se evidencia entre dois ou mais conceitos. As características do conceito são, então, fator primordial para o estabelecimento das relações entre conceitos e seu posicionamento no sistema de conceitos. Além disso, auxiliam na identificação do conceito, porque são estabelecidas a partir da seleção das propriedades relevantes de um dado objeto que no nível do conceito, é conhecido como referente do conceito. (CAMPOS, 2001a, p.121)

A representação conceitual está presente em uma rede de conceitos, que tece elementos heterogêneos, mas quando associados são inseparáveis para compor um universo fenomênico do uno e do todo em um múltiplo aparentemente desordenado, incerto.

E se o sistema de organização do conhecimento vem para colocar ordem neste tecer é preciso ter o cuidado de não eliminar os aspectos de possíveis relações integrativas e sistêmicas, pois segundo o pensamento de Ranganathan (1967) os conceitos são formados a partir das percepções puras e compostas e da associação delas. No estudo realizado sobre a elaboração de taxonomia, Novo (2007, p.33), após análise de sua obra relatei: “O conhecimento é um conjunto de ideias e os assuntos são ideias organizadas e sistematizadas em um universo de conhecimento definido por ele [Ranganathan] como um contínuo de diferentes domínios.” Segundo Campos,

[...] dessa forma, é a partir da formação dos conceitos que se vai produzir na mente do ser humano um quadro de identidade com o mundo que o cerca. Em um momento posterior à formação dos conceitos, isto é, a partir da existência de um padrão conceitual já estabelecido, pode ocorrer a assimilação de novas experiências o que leva ao que Ranganathan denomina de apercepção. O conjunto destas apercepções depositadas na memória se dá, então a partir dos conceitos já presentes na memória, com o acréscimo da

assimilação de perceptos recentemente recebidos e conceitos recentemente formados. (CAMPOS, 2001a, p.39)

Quando uma ciência está em desenvolvimento e um consenso entre especialistas sobre uma definição conceitual não foi totalmente estabelecida é preciso escolher entre possibilidades alternativas de tradução para os conceitos que serão representados, ou seja, defini-los com base no que o conceito representa para o domínio e em que contexto está sendo empregado.

O ponto de partida pode ser a semelhança entre termos, mas as características conceituais são fundamentais, pois algumas semelhanças podem ser superficiais, a ponto de prejudicarem as relações entre categorias de conceitos, base para organização do conhecimento. Essa organização necessita de um método que permita esclarecer propriedades essenciais e inter-relações fenomênicas que manifestam algumas propriedades essenciais dos conceitos e suas inter-relações ontológicas ou associativas que descreverão: objetos, classes, atributos e relacionamentos.

Tais relações ou inter-relações podem ser demonstradas através de representações conceituais ou representação do conhecimento, em um tecer que envolve a linguagem, o sujeito e o seu discurso. Tecer relações é dar aparência aos objetos recobertos de interpretações que podem ocultar ou revelar significados. O que está em questão é a apropriação de sentidos, que permitirá ao indivíduo reconhecer a materialidade do objeto conhecimento.

O objeto de conhecimento é objeto que muda que tem uma história inscrita na história da ciência da qual é objeto, na confrontação de suas teorias, nas práticas específicas que o caracterizam, assim como nas condições históricas que produziram essa história, essas confrontações, essa prática. (HENRY, 1992, p.16).

E se o objeto muda é porque não há nada estático no mundo fenomênico, nem no plano das ideias, tudo está em movimento, tudo é dinâmico assim como é dinâmica a representação da produção do conhecimento de um dado domínio. Não existe um todo unificado, mas um entrelaçar de conceitos que se deslocam de um domínio para outro em relações categoriais igualmente dinâmicas, o que é todo passa a ser parte e o que é parte passa a se integrar a um novo todo. Assim se caracteriza o movimento em uma rede conceitual,

onde um objeto (conceito) pode modificar o estado das coisas dando um novo sentido ao contexto, portanto é complexo o trabalho de OC.

[...] um tecido (complexus: o que é tecido junto) de constituintes heterogêneas inseparavelmente associadas: ela coloca o paradoxo do uno e do múltiplo. [...] o tecido de acontecimentos, ações, interações, retroações, determinações, acasos, que constituem nosso mundo fenomênico. [...] a complexidade se apresenta com os traços inquietantes do emaranhado, do inextricável, da desordem, da ambiguidade, da incerteza [...] (MORIN, 2007, p. 13).

É preciso ter atenção com a incerteza das palavras que possuem diferentes significados, a depender da análise semântica. O importante é identificar o conceito e o seu significado para o domínio, situar o sentido reconstituído ao contexto aplicado. Talvez não se chegue à fórmula de um modo tão preciso, pois nada é definitivo, o importante é lidar com as relações entre os conceitos e as suas superestruturas, ou seja, o arranjo que se dará através de categorias formais, a partir da posição do conceito de acordo com apropriação científica observada no domínio ou o que pode ser representado. É importante estabelecer relações de equivalências não comprometendo assim o significado conceitual, o que possibilitará, em uma ontologia, por exemplo, reuso em outras aplicações.

A tarefa de OC se associa à análise do discurso e tem por missão atender a difusão do conhecimento. Apoiando-se sobre o discurso do domínio, seja no tempo, que atende ao mundo científico, seja no espaço, ou o que responde aos anseios de uma comunidade discursiva. Sendo assim pressupõe uma análise cognitiva/interpretativa das informações que circulam nos domínios. No caminho da análise existirá sempre um olhar atento aos fenômenos e às coisas ou à manifestação de ambos, para que se possa alcançar o entendimento conceitual, traçando uma direção que transitará do mundo sensível ao mundo inteligível.

Do mundo sensível, podemos formar um conceito empírico por intermédio da intuição na medida em que nossa sensibilidade é afetada; já do mundo inteligível, temos apenas uma representação por meio do pensamento, pois ultrapassaríamos o limite estabelecido ao conhecimento humano se quiséssemos perscrutá-lo pela intuição, visto que toda intuição sempre se refere a fenômenos e nunca às coisas como são em si mesmas. (SOUZA, 2009, p.39).

Pode ocorrer que um conceito seja interpretado por um analista cognitivo de um modo e por outro analista de outro modo, pois sabemos o quanto subjetivo é o espaço da interpretação. Por isso precisamos fundamentar a análise do objeto a ser representado tornando-o inteligível. Isso é possível através da definição conceitual e observação empírica (evidência) do objeto, ou de como se aplica o uso do conceito pelo domínio que se quer organizar e representar. Mas isso não quer dizer afastamento do indivíduo, pelo contrário, propomos uma análise que aproxime significados das comunidades científicas, mas que sejam declarados de forma a alcançar o público em geral que tem direito a qualquer tipo de informação.

A análise cognitiva e conceitual, no caso desta pesquisa, pauta-se em elementos essenciais que caracterizam o raciocínio humano, portanto, nas relações conceituais e de significação, nas necessidades de informação, nas especificações explícitas dos conceitos e suas relações no âmbito de domínios de conhecimento. Só após o entendimento dessas relações, avança-se para o nível de representação ontológica com as devidas definições conceituais, ou o nível da significação.

O trabalho com OC e RC não é tarefa fácil, devido ao grau de subjetividade existente quando descrevemos conteúdo, porque tentamos traduzir conceitos nos termos de uma linguagem e o trabalho com signos linguísticos vai além de um simples esquema de categorias, pois a dimensão de uma superestrutura de organização temática ou macroestrutura compreende aspectos sintáticos, semânticos e pragmáticos. Portanto requer uma interpretação diferenciada, pois cada parte representada possui a informação do todo de suas partes.

Sobre a questão da interpretação dos signos, Santaella evidencia:

É só na relação com o interpretante que o signo completa sua ação como signo. É apenas nesse ponto que ele age efetivamente como signo. Entretanto quando o signo é interpretado, esse ato embute os outros dois aspectos do signo: o do seu fundamento e o da relação com o objeto. Quando interpretamos signos – aliás, algo que estamos fazendo continuamente, sem descanso -, nossas interpretações são intuitivas e não nos damos conta da complexidade das relações que estão implícitas nesse ato. Contrariamente a isso, ao analisarmos signos, temos de tornar essas relações explícitas. É por isso que a análise dos interpretantes deve estar alicerçada na leitura cuidadosa tanto dos aspectos envolvidos no fundamento do signo como nos aspectos envolvidos nas relações do signo com seu objeto. (Santaella, 2012, p. 37).

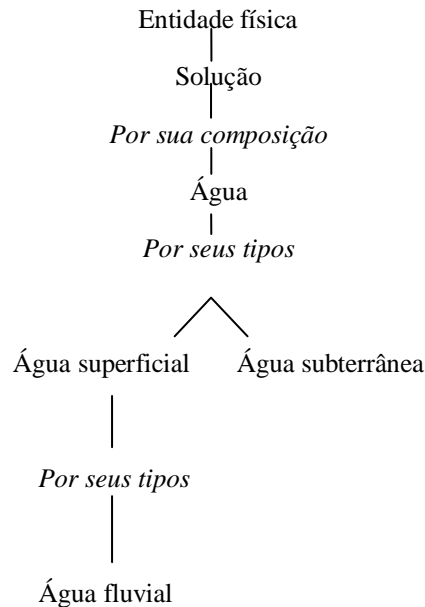
A questão do signo nesta pesquisa está relacionada ao objeto, uma tipologia de conceitos pode ser de objetos. Não vamos nos ater às questões da semiótica, mas somente os aspectos da explicitação dos conceitos e de suas relações com outros conceitos. Para tanto trabalhamos com categorização dos conceitos.

Quando categorizamos dividimos o universo de um domínio em entidades (objetos materiais) onde cada parte de uma entidade tem similaridades dentro de um contexto, estamos interpretando. Buscando exemplificar, vamos considerar o termo: **água fluvial**. Quando o definimos dizemos: que é a água proveniente de rio, no entanto a água de rio é um tipo de água superficial, pois ocorre na atmosfera acima da superfície topográfica, diferentemente da água subterrânea.

De qualquer forma tanto a água fluvial, quanto a água subterrânea são tipos de água, portanto uma solução que se configura por um sistema que pode envolver sólidos, líquidos ou gases. Os dois conceitos são entidades físicas, pois carregam propriedades e podem ser usados em algum tipo de atividade. Em uma taxonomia, de acordo com os postulados da Teoria da Classificação Faceta (TCF) ver seção 2.1, desenvolvida por Shiyali Ramamrita Ranganathan na década 1930, com suas séries de renques e cadeias² e a Teoria do Conceito (TC), seção 2.2, de Ingtraut Dahlberg, estaria assim hierarquizado:

² Cadeias são séries verticais de conceitos e renques são conceitos subordinados a um mesmo conceito de forma coordenada, numa série horizontal. CAMPOS, Maria Luiza de Almeida; GOMES, Hagar Espanha; MOTTA, Dilza Fonseca. Elaboração de tesauro documentário: tutorial. Disponível em: <http://www.conexaorio.com/bit/tesauro/index.htm>

Figura 1 – Exemplo da categorização conceitual – Conceito: água fluvial



Fonte: construção da autora

Em se tratando de um sistema de OC ou RC para domínios de meio ambiente como exemplo, a Geoquímica ambiental, esta categorização atende totalmente as questões de: **sintaxe** (do conceito em seu discurso); **semântica** (significado e interpretação do conceito no contexto) e **pragmática** (fatores contextuais de comunicação).

No entanto para atingir a demanda de informação de um usuário de uma comunidade ribeirinha que precisa saber das condições da água fluvial em sua região sobre os aspectos relacionados, por exemplo, a contaminação de um rio ou simplesmente sobre períodos de cheia e vazante, outros fatores deverão ser pensados quando na utilização de formas hierárquicas e associativas de organização e que revelarão conteúdos que atendam a demanda do sujeito desta comunidade em suas intenções de busca. Outras relações precisam ser explicitadas e relacionadas para atender a demanda de informação do usuário em geral.

Sendo assim não basta apenas categorizá-los como entidades, mas, sobretudo entender partes e todo que revelam sentidos para um determinado contexto de um domínio de conhecimento e suas práticas, e, sobretudo o que e a quem a mensagem de um discurso

precisa atingir, afinal, o objetivo do modelo que propomos para auxiliar aos SOC é não só disseminar seletivamente uma informação, mas, especialmente difundir conhecimento. Encontramos nas considerações de Chiavegatto sobre contexto e sentido uma tradução que se aproxima da nossa intenção:

Considerando que essas **entidades relacionais são signos**, podemos dizer que **um contexto é um constructo** (espaço mental) **‘grávido’ de sentidos**, que estão semioticamente disponíveis a partir das percepções tornadas experiências armazenadas nos domínios cognitivos. Mesmo as **entidades arquivadas na memória já são interpretantes dos objetos** aos quais foram relacionados. **Representados, abrem-se para novas relações e novas significações.** (CHIAVEGATTO, 2002, p. 128, grifo nosso).

O estudo das relações de ordem sintática, semântica e pragmáticas requer regras combinatórias de estrutura temática e de sintaxe, pois neste caso existem relações recíprocas que se dão entre signos, ou seja, signos que se inserem em sequência de outros, para Almeida (2011, p. 57) “as estruturas sintáticas determinam a forma e o sentido, são regras de interpretação que interessam à estrutura gramatical”.

Além disso, de acordo com as considerações apresentadas por Moreiro González (2011, p.27), no estudo semântico a atenção deve estar voltada para as questões da “semântica referencial (relação com a realidade extralinguística) e a semântica orientada ao conteúdo” que são imanentes ao sistema de signos, ligados ao conteúdo global do discurso e a temática abordada no texto. Neste ponto insere-se também a pragmática que está ligada a interpretação do discurso considerando a relação do signo com sua origem, além da função comunicativa.

É importante salientar que a importância da função comunicativa é prioridade nesta tese, pois acreditamos que os espaços de significação para os sujeitos merecem uma atenção prioritária nos estudos voltados para organização e representação do conhecimento, pois usualmente se observa a não valorização dos conhecimentos apropriados em espaços de construção ditos “periféricos”, ou seja, não construídos nos espaços acadêmicos.

Os modelos de representação construídos, poucos levam em conta contextos aparentemente distanciados destes espaços “periféricos”, como os de comunidades fora do contexto acadêmico, conseqüentemente excluem visões e práticas que poderiam contribuir

para evolução de pesquisas e desenvolvimento destas “fontes” que reconhecemos como geradoras de informação.

Diante do exposto pensamos em como lidar com uma estrutura de significação tão complexa como um sistema de ideias se não pelas TCF e TC e encontrar um aporte teórico para entender a organização conceitual de domínios inter e multidisciplinares, em uma perspectiva que chamamos de possível, dado ao alto nível de abstração que ocorre nas relações conceituais desses domínios.

Por vezes ordenamos categoricamente os conceitos em uma relação lógica de implicação. Em um outro olhar parecemos desordenar para possibilitar relações possíveis entre os conceitos estabelecendo pontos de ligação entre domínios que se comunicam. Segundo Morin, (2000, p. 114) “[...] A missão da ciência não é mais afastar a desordem de suas teorias, mas estudá-la. Não é mais abolir a ideia de organização, mas concebê-la e introduzi-la para englobar disciplinas parciais.”.

A OC e a RC que propomos tem o ponto de partida nos signos linguísticos, responsáveis pela representação das ideias e suas associações de significantes (conceitos), nas considerações de Aguiar,

A linguagem humana pode ser vista como um fenômeno em que signos linguísticos com significados próprios são organizados de forma a gerar uma estrutura com significado diferente da soma de cada unidade linguística. Isto significa que é possível analisar o fenômeno linguístico como um fenômeno complexo, a partir da emergência de uma propriedade global originada da compreensão de uma construção sintática (a sentença) proveniente da interação e organização de unidades linguísticas (as palavras). Da mesma forma que a compreensão de uma única sentença de um texto não é capaz de refletir toda a mensagem deste texto. (AGUIAR, 2009, p. 18)

Assim sendo, pretendemos, com as redes de conceitos geradas a partir das redes semânticas de textos completos, demonstrar a análise e compreensão da mensagem que os textos expressam sobre o domínio de conhecimento, refletindo as pesquisas desenvolvidas no domínio. A compreensão das mensagens dos textos, através das relações conceituais encontradas, nos possibilitou a construção do modelo de análise.

O processo de organização e de representação do conhecimento abrange e encerra elementos que se apresentam criativamente no espaço do pensamento, que une coisas

semelhantes e ao mesmo tempo tão diferentes, na perspectiva de todo e partição com possibilidades de construção e reconstrução de ideias, em um jogo de múltiplas possibilidades.

As possibilidades de uma modelização conceitual ou de representação de domínios inter e multidisciplinares são carregadas de subjetividades e abstrações próprias das condições ideacionais do sujeito, que percebe e se relaciona com o contexto dos domínios, contexto esse que se articula em múltiplas relações para atender a propósitos particulares como: soluções de problemas e comunicação de suas pesquisas, portanto tal modelização se caracteriza como tarefa árdua, no sentido da dificuldade de se encontrar uma estrutura conceitual ideal.

Os efeitos provocados pelo uso de determinados conceitos em um discurso provocará uma interpretação diferenciada a depender das conexões estabelecidas em uma rede conceitual, que nem sempre são previsíveis e dependerá da análise cognitiva do sujeito que as percebem ou de significantes e significados.

Quando enfim pretendemos uma “leitura” e interpretação de uma realidade que se traduzirá por conceitos e suas relações associativas que revelam os efeitos de uma linguagem e de sua sintaxe na lógica discursiva, através de uma organização do conhecimento, precisamos antes das questões práticas, de teorias, para que a nossa proposta se fundamente e conseqüentemente possa se estabelecer.

2.1 A TEORIA DA CLASSIFICAÇÃO FACETADA: PRINCÍPIOS PARA O TRABALHO COM CATEGORIAS CONCEITUAIS

Apresentamos nessa seção a Teoria da Classificação Facetada e justificamos a sua inclusão nesta tese, pelo viés do pensamento dinâmico que a Teoria favorece. O que já foi comprovado em estudos realizados na CI e principalmente o que realizamos na pesquisa de mestrado quando elaboramos uma taxonomia para domínios interdisciplinares e apresentamos princípios classificatórios com base na TCF.

Além disso, como evidenciamos na próxima seção na base dos estudos de Ingetraut Dahlberg para a Teoria do Conceito encontra-se a TCF, assim para o entendimento da TC uma breve explanação dessa Teoria é necessária.

A TCF foi desenvolvida por Shiyali Ramamrita Ranganathan e seus postulados encontram-se em quatro obras básicas: *Five Laws of Library Science*, 1931, reimpressa em 1963; na *Colon Classification*, 1933, 6.ª ed. 1963; *Philosophy of Book Classification*, 1951 e *Prolegomena to Library Classification*, 1967.

A Teoria da Classificação Facetada, embora tenha sido desenvolvida para organização de bibliotecas, propõe princípios para SOC e recuperação da informação, reconhecendo variedades de aspectos em um assunto específico, e a relação entre assuntos pode ser infinita em um conhecimento reconhecido como multidimensional.

Este reconhecimento é declarado na proposta de síntese da Classificação Facetada evidenciado pelo símbolo de dois pontos empregado para relacionar assuntos, e ainda, na possibilidade de análise por facetas dos assuntos do documento. Faceta aqui entendida como reunião de termos que se relacionam com um assunto geral. A TCF é dita como dinâmica por teóricos da classificação como Kumar que evidenciou:

A dynamic theory is one which is able to provide a sound methodology for the designing of a scheme for library classification. Such a theory enables one to organize new subjects and already-known subjects in their proper places in a scheme, without disturbing the helpful sequence. (KUMAR, 1981, p. 82)

A análise a partir das facetas é fundamental na modelagem conceitual quando facilita a seleção de conceitos relevantes ao domínio para, por fim, categorizá-los em uma sequência útil aos propósitos do domínio. Dentro de cada categoria ou faceta, os termos podem ser reagrupados de acordo com as características terminológicas divisionais, originando novas facetas, porém nessa nova faceta o termo não se sobrepõe a formação de conceitos compostos.

A TCF permite construir um modelo de natureza dedutiva, considerando-se em primeiro lugar o domínio e o contexto em que o termo esteja inserido, para em seguida trabalhar com os níveis conceituais em categorias, formando classes conceituais que permitirão a análise e a declaração conceitual, como evidenciados em taxonomias e ontologias, ver capítulo 3.

A Teoria de Ranganathan está fundamentada através de um método científico demarcado nos planos: da ideia, verbal e notacional. E ainda no desenvolvimento de cânones

de trabalho para os planos e nos princípios básicos normativos, chamados de leis: Lei de interpretação, Lei da imparcialidade, Lei de simetria, Lei da parcimônia, Lei de variação local e Lei de osmose. Não nos ateremos a descrevê-las, recomendamos o texto de Gomes; Motta; Campos (2006), que consta nas referências desta tese.

O plano das ideias da TCF é o lugar onde são encontrados os conceitos; o plano verbal onde se expressam verbalmente os conceitos e o plano notacional a fixação abstrata dos conceitos em sinais e letras, possibilitando uma sequência útil dos assuntos nas estantes de bibliotecas. Dahlberg descreveu esses planos como níveis e enfatizou sua importância:

Essa distinção em três níveis auxiliou consideravelmente a tornar mais claro o que pode ser considerado como objeto da ciência da classificação: é o conceito único e sua capacidade de combinação para representar o conhecimento que o homem tem do mundo que, desde Ranganathan, pode ser considerado como elemento característico dos sistemas de classificação. Isto pressupõe a disponibilidade das expressões da linguagem natural para sua descrição (plano verbal) e utiliza notações para sua representação em uma forma semiótica. (DAHLBERG, n.p., 1979)

Segundo Dahlberg, (1979) a distinção em três níveis seria a primeira grande contribuição de Ranganathan. Para o trabalho que realizamos nesta pesquisa encontramos no plano das ideias os princípios norteadores para recortar um domínio de conhecimento em conceitos. Nesse plano estabelece-se o princípio da análise do pensamento e análise conceitual, forma concreta que pode levar à informação.

Ranganathan desenvolveu cânones para o plano das ideias, apresentados por Gomes; Motta e Campos (2006), no quadro 1 (um) que segue.

Quadro 1 – Cânones para o trabalho no Plano das Ideias

Característica	Sucessão de características	Renque de classes	Cadeia de classes	Sequência de filiação
Diferenciação	Concomitância	Exaustividade	Extensão decrescente	Classes subordinadas
Relevância	Sucessão relevante	Exclusividade	Modulação	Classes coordenadas
Verificabilidade	Sucessão consistente	Sequência útil		
Permanência		Sequência consistente		

Fonte: Gomes; Motta; Campos (2006)

Segundo as Autoras: no conjunto de cânones das características objetiva-se o conteúdo do assunto; no conjunto de cânones de sucessão de características, encontra-se a orientação à sucessão de classes em um universo de assuntos; no conjunto de cânones dos renques e cadeias a formação e ordem dos elementos (conceitos) e no conjunto de cânones de sequência de filiação a construção de hierarquias consistentes.

Os Cânones são importantes na condução da organização de conceitos de domínios de conhecimento, e na abordagem analítica sintética que identifica assuntos. Ranganathan ainda apresentou estudos sobre a natureza dos domínios e se propôs a recortá-los através das categorias fundamentais: Personalidade, Matéria, Energia, Espaço, Tempo, (PMEST). Sobre a sua contribuição, Dahlberg teceu algumas considerações:

A segunda contribuição de Ranganathan à moderna teoria da classificação é a sua abordagem analítico sintética para a identificação dos assuntos. Isto implica em que a classificação de cada documento exige uma análise de seu título ou um enunciado descritivo de seu conteúdo nos termos dos conceitos que formam os componentes da ciência à qual esse documento pertence. Após a análise e a orientação desses elementos nas facetas (que são representativas dos tipos de conceitos em áreas especializadas do conhecimento), é possível sintetizá-los em expressões combinatórias que formam as classes, construídas analítico-sinteticamente, de um determinado tópico de documento. Tudo isso é realizado segundo as fórmulas de facetas mencionadas acima, e sua fórmula generalizada, a sequência PMEST (Personalidade, Matéria, Energia, Espaço e Tempo) isto serviu para: 1. a

representação dos assuntos, e 2. a ordenação dos conceitos de uma disciplina em classes formais, de acordo com as categorias existentes nessa disciplina. (DAHLBERG, n.p., 1979)

De forma resumida evidenciamos que a manifestação da categoria **Personalidade** se dá quando os conceitos não se enquadram por suas características fundamentais em outras categorias, Ranganathan, (1967) propôs para esta categoria o “método do resíduo”, “não é isso, não é isso”. Na categoria **Matéria** suas manifestações são de duas espécies: material e propriedade. Na categoria **Energia** a manifestação é de ação e processo. Na categoria **Espaço**, o entendimento do local de pertencimento de um dado objeto. Na categoria **Tempo**, o que entendemos por tempo.

A partir das categorias identificamos e categorizamos conceitos de domínios de conhecimento e em seu interior evidências dos cânones das características, sucessão de características, renques, cadeias e a sequência filial. No cânone de características ocorre o trabalho com os assuntos e a evidência dos atributos dos conceitos.

Nos renques ocorre a demonstração das classes, a partir da característica de divisão em série horizontal. Enquanto que nas cadeias a demonstração da sequência de classes em séries verticais de conceito, com características descendentes ou ascendentes.

Para modelar a formação dos conceitos nos domínios Ranganathan, (1967) previu trabalhos de:

- Dissecção: recortando um universo de entidades em partes de mesmo nível, podendo representar um universo de assuntos básicos ou de ideias isoladas. Quando enfileiradas essas partes formam um renque;
- Laminação: construção de faceta sobre faceta por superposição;
- Desnudação: diminuindo a extensão e aumentando a intensão de uma ideia isolada, podem-se formar cadeias;
- Reunião ou agregação livre: reúne duas ou mais ideias ou assuntos básicos ou compostos;

- Superposição: reúne duas ou mais ideias isoladas pertencentes ao mesmo universo de ideias.

A aplicação da TCF, principalmente com o foco no plano das ideias possibilita a construção de taxonomias e a organização e modelização de conceitos encontrados nos domínios de conhecimento. Conceitos reunidos através de um arranjo sistematizado permitem o entendimento de assuntos que se encontram no interior do texto.

Após o entendimento da Teoria da Classificação Facetada passamos a apresentar a Teoria do Conceito na seção que segue.

2.2 A TEORIA DO CONCEITO: UMA PROPOSTA DE FUNDAMENTAÇÃO TEÓRICA E METODOLÓGICA

Elegemos como teoria para fundamentação do modelo de análise proposto e do método de organização de estruturas conceituais, a Teoria do Conceito, desenvolvida por Ingtraut Dahlberg nos anos 70 do século passado, pioneira nos estudos de organização do conhecimento. Dahberg fundou com outros pesquisadores em: 1974 *The Journal International Classification*; 1977 *the Society for Classification*; 1989 *the International Society for Knowledge Organization*, e *Knowledge Organization*, 1993.

Os princípios de sua teoria desenvolvidos no campo da terminologia se aproximam aos da Teoria da Classificação Facetada (TCF) desenvolvida por Ranganathan (1931 (reimpressão 1963), 1933 (reimpressão 1963), 1951, 1967), e seus postulados sobre o universo do conhecimento e as categorias fundamentais por ele apresentadas conhecidas como PMEST. Segundo Dahlberg:

Uma categorização formal dos conceitos que tem importância na formação dos sistemas e na combinação dos mesmos conceitos pode ser a seguinte: Objetos — Fenômenos — Processo — Propriedades — Relações. Podem ser caracterizados da seguinte maneira:
 A Objetos. Ex.: plantas, produtos, papel, etc.
 B Fenômenos. Ex.: crescimento, chuva, tráfego, etc.
 C Processos. Ex.: imprimir, sintetizar, etc.
 D Propriedades. Ex.: cego/cegueira, suave/suavidade, etc.
 E Relações. Ex.: causalidade, necessidade.
 F Dimensão. Ex.: espaço, tempo, posição, etc.
 As classificações facetadas utilizam esta espécie de categorização dos conceitos. (DALBERG, 1978c, p. 105).

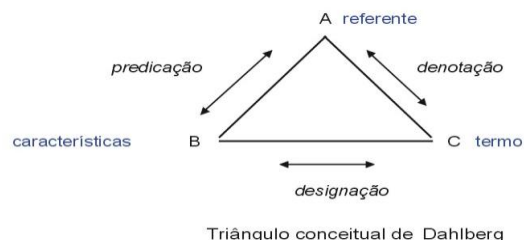
A proposta da Teoria é que os termos se definem uns em relação aos outros formando um sistema. Tanto a TCF quanto a Teoria da Terminologia, (Wüster, 1971, 1981) auxiliaram nos princípios para a construção das relações conceituais da TC, posicionando o conceito em um sistema de conceitos, onde:

O conceito não é mais apenas um elemento de significação do termo: o termo acaba sendo um elemento do próprio conceito, o ‘terminum’ que sintetiza o conceito como um todo e permite a comunicação, neste caso, verbal. (CAMPOS, 2001a, p.100)

A Teoria do Conceito possibilita um método segundo Campos, (2001a, p.100) para “[...] a fixação do conteúdo do conceito [...] Assim, tratar o termo como representante de um referente, com suas características, é dar a ele um tratamento terminológico.” Ou seja, para Dahlberg (1978a, p. 148) existe uma “equivalência entre o termo (definiendum) e as características necessárias de um referente, de um conceito (definiens), com o propósito de delimitar o uso do termo em um discurso” e assim, Dahlberg propõe que o conceito seja definido como “unidade de conhecimento”.

Esta unidade de conhecimento, o conceito, é formada segundo Dahlberg, por três elementos que o compõem: referente, características e forma verbal. Sua formação compreende três passos: referencial, predicacional e representacional. Quando selecionamos um item de referência atribuímos a ele predicados, designamos características e por fim denotamos formalmente o conceito. O modelo proposto por Dahlberg nos permite a construção e aplicação dos conceitos nos domínios, sua forma de análise está representada nas figuras 2 e 3:

Figura 2 - Triângulo Conceitual de Dahlberg

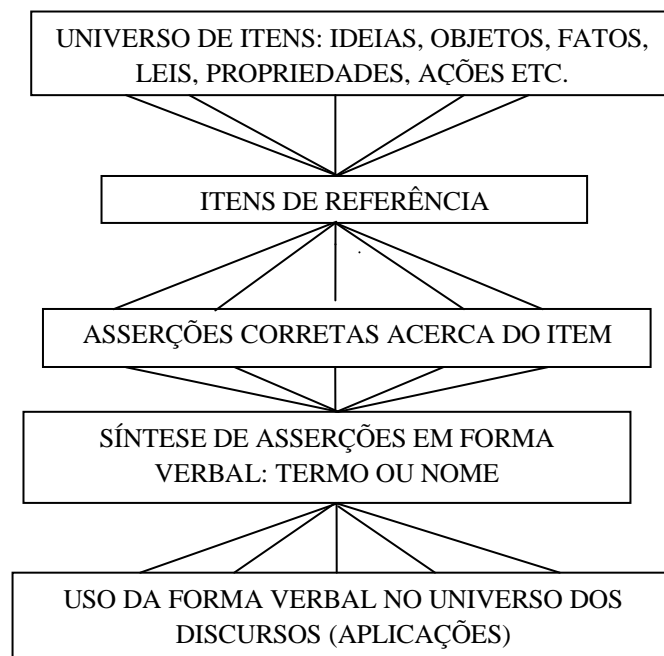


Fonte: Dahlberg (1978c)

O Objeto (referente) dependerá de um contexto, que determinará suas características e se traduzirá através de um nome (signo), não somente numa relação sintagmática entre

categorias que evidenciarão a demanda, função entre objetos e o mundo fenomênico, mas, sobretudo em uma relação associativa, ou o que se pensa cognitivamente. Nesta perspectiva, para Dahlberg (1978d) o conceito é “[...] a compilação de enunciados verdadeiros sobre determinado objeto, fixada por um símbolo linguístico.” Gonçalves e Souza (2008, p. 3) complementam, “[...] a tríade proposta por Dahlberg apresenta o conceito como a soma dos enunciados verdadeiros e essenciais sobre um referente, e o termo/léxico é a forma comunicável e representável do conceito.” Dahlberg apresenta ainda um modelo para construção de conceitos representado na Figura 3 a seguir.

Figura 3 Modelo para construção dos conceitos



Fonte: Dahlberg, 1978c

A partir do modelo de construção dos conceitos e do triângulo conceitual (Figuras 2 e 3), pode-se implementar categorias de conceitos e organizá-los em uma determinada área, com vista a elucidar a natureza do conceito e por fim dar forma as estruturas conceituais e suas relações, objeto das ontologias de domínios. As categorias possibilitam a sistematização do conhecimento, determinada pelo conceito a partir de suas definições que se traduzirão por uma forma verbal (termo ou nome).

As asserções corretas acerca do item de referência, ou conceito, de um determinado domínio dependem de suas propriedades, ou características conceituais (essenciais ou acidentais) ou ainda gerais e consecutivas. Uma ordenação sistemática (relacionamentos associativos) entre conceitos torna possível construir melhores definições conceituais.

Seguindo a síntese de atribuição das características que são dependentes das relações de espécies diferentes, as relações são classificadas como: quantitativas e qualitativas. A quantitativa compara conceitos do ponto de vista formal (identidade, inclusão, interseção e disjunção), enquanto a qualitativa evidencia os aspectos formais e materiais, objeto das relações formal/categorial. Tal formalização depende da espécie do item analisado dentro de uma categoria.

Além dessa comparação formal e material, novas possibilidades de relações são construídas, segundo Dahlberg, (1978c, p.18) tais como: “relação formal/categorial; relação material/paradigmática; relação sintagmática funcional”. Essas relações formam uma rede de relações ou ainda o sistema conceitual.

A primeira relação, **formal/categorial**, depende da espécie do item e suas combinações: qualidades, quantidades, processos, fenômenos, objetos materiais dentre outros. Reúnem conceitos de mesma natureza em uma categoria.

A segunda relação, **material/paradigmática**, organizam os conceitos no interior de uma categoria, são as relações hierárquicas, partitivas e de oposição. As hierárquicas são: relação de abstração (gênero-espécie) e lateral (renques de conceitos, ou horizontais). As partitivas são em relação ao todo e suas partes onde as partes podem possuir relações entre si. As de oposição são (contraditórias-oposição contrária, positiva-indiferente-negativa).

A terceira relação **funcional-sintagmática** é encontrada, entre conceitos de diferentes categorias, um processo pode ser suplementado por uma ação ocorrida em outra categoria. Para Dahlberg, (1978c, p.34) as categorias são estruturantes, fornece a união de todo o “corpo do conhecimento”, mas “pode se mover, pode se manter flexível e pode crescer organicamente.” Neste sentido a teoria do conceito, proposta por Dahlberg, vai ao encontro do modelo dinâmico de organização do conhecimento que apresentamos nesta tese.

O conceito é um complexo de natureza intencional relacionada ao seu significado, dependente de sua definição, e a sua extensão ou conjunto de elementos, de um determinado domínio onde o conceito se aplica. Existem dois tipos de definição conceitual: a nominal e ou intensional³. A intensão como termo define propriedades que compõem um conceito. De acordo com Dahlberg:

A primeira distinção que se costuma fazer é a separação entre definições nominais e definições reais. Mas existem outros tipos de definições. A *definição nominal* tem por fim a fixação do sentido de uma *palavra*, enquanto que a definição real procura delimitar a intensão^[3] de determinado conceito distinguindo-o de outros com idênticas características. A definição nominal relaciona-se com o conhecimento contido na linguagem. (DAHLBERG, 1978d, p.106)

O conhecimento contido na linguagem e impresso em um texto pode ter várias interpretações a depender dos domínios de conhecimento, sendo assim a preocupação primeira dessa tese diz respeito à tradução do conhecimento através dos conceitos emanados nos domínios, que dependerá do discurso das comunidades acadêmico/científicas que os utilizam em contextos inter e multidisciplinares, pois existem conceitos que transitam entre domínios sem pertencer a um único. Como afirmam Francelin e Kobashi:

A maneira como as coisas são vistas e representadas difere de cultura para cultura, como também de indivíduo para indivíduo. Assim, a construção do conceito depende basicamente de um referente, da emissão de juízos sobre o referente, de uma forma verbal (um termo ou um nome) e de uma maneira de usar essa forma verbal em um universo discursivo. (FRANCELIN; KOBASHI, 2011, p.210)

Dessa forma reforçamos que para organização do conhecimento de domínios inter e multidisciplinares que reúne conceitos de múltiplos espaços discursivos necessitamos de flexibilidade e dinamicidade para o trabalho de categorização de conceitos.

Para demonstrarmos um modelo dinâmico de representação e um método de raciocínio, igualmente dinâmico para categorização foi necessário traçar princípios de análise conceitual, assim elegemos a Teoria do conceito como base para categorizarmos os termos a serem trabalhados no universo discursivo do domínio de conhecimento. A Teoria do Conceito

³ Intensão: palavra parônima de intenção – (atividade, intensidade, força). Fonte: <http://www.dicio.com.br>

nos possibilitou maiores possibilidades de combinações nas categorias formais, facilitando o entendimento das relações conceituais.

Exemplo apresentado por Dahlberg (1978d, p.105): **Categoria A** Objetivos - ex.: plantas, produtos, papel, etc. **Categoria B** Fenômenos - ex.: crescimento, chuva, tráfego, etc. **Resultado da relação:** $A + B =$ crescimento das plantas.

A Teoria do Conceito em conjunto com princípios de análise cognitiva forneceu o alicerce para a construção do modelo dinâmico de análise. Acreditamos que o modelo possibilitará construir relações associativas mais consistentes para os domínios que se articulam dinamicamente. Ou ainda, a elaboração de modelos de representação em diferentes níveis que abarquem conhecimento abstrato advindo do sujeito que percebe e busca informação, mais coerentes com a sua realidade, é o que postulamos neste trabalho e defendemos.

A concepção em tríade proposta na TC coaduna com o modelo triádico de construção de ontologias, já apresentado em estudos desenvolvidos por pesquisadores contemporâneos da Ciência da Informação, como os de Sales, (2006, 2008, 2012). E ainda em trabalhos de modelagem conceitual e modelização de domínios de: Campos, (2001b, 2003, 2007); Campos; Campos, (2012); Gonçalves; Souza, (2008); Nonato, (2009) e Nonato; Lima, (2007). A teoria facilitou a apresentação do método de raciocínio empregado no modelo que apresentamos nesta tese.

A análise conceitual do universo discursivo, a partir do corpus selecionado como amostra para a implementação do modelo dinâmico de análise conceitual, nossa proposta de tese, está fundamentado em teorias, metodologias e práticas como: a análise do discurso e análise cognitiva as quais apresentamos no seguimento deste capítulo.

2.3 ANÁLISE DO DISCURSO REPRESENTATIVO: UM PROCESSO PARA ORGANIZAÇÃO DE DOMÍNIOS

O discurso impresso nas teses e dissertações é amparado por ideias já discutidas em outros escritos, por exemplo, com o aporte teórico de autores seminais da área em que a pesquisa tenha sido executada. Esse empréstimo discursivo natural nos escritos de domínios de conhecimento se caracteriza como um amparar no que já foi dito e cancelado nas áreas de

conhecimento e garantem ao autor o reforço ao seu pensamento e mais ainda, fornecem a garantia literária⁴ (o que está documentado, publicado) que valorizará a impressão de novos conceitos que por ventura lançará, pois apesar de “novos” se apoiam em pesquisas anteriormente desenvolvidas e experimentadas.

As práticas discursivas das comunidades acadêmicas e científicas transversam entre domínios multidisciplinares. No caso particular do discurso da área da Geografia, percebe-se nitidamente o uso de conceitos que transitam nas áreas das Ciências Humanas, Biológicas, Sociais Aplicadas, Geológicas dentre outras. O que nos demonstra a importância da categorização conceitual para melhor representar as relações associativas que ocorrem e se manifestam entre conceitos aparentemente deslocados de um e para um contexto de domínio de conhecimento.

Os discursos das teses e dissertações apresentadas nos programas de pesquisa vão da intenção ao jogo das representações metaenunciativas, um modelo de desdobramento do dizer ou de nomear respostas alcançadas nas pesquisas, para conseguir o objetivo do convencimento, tornando as palavras atravessadas pelo discurso de outrem, um jogo de palavras que se impõem ao produtor do discurso.

No processo de organização do conhecimento se faz necessário à interpretação explícita do comentário, da palavra que convém dizer, e que possivelmente possa se transformar em um conceito novo ou um novo conceito, uma análise discursiva e representativa do domínio.

São as causas do dizer sobre algo que servirão de base para a busca da informação, propriedades do referente (objeto), compatibilidade entre “as palavras e as coisas”. Um dizer refletido sobre o enunciado e facilmente interpretável quando no plano explícito do termo, para que de fato não só a disseminação seletiva da informação aconteça, mas principalmente a difusão do conhecimento alcance a todos, e que a linguagem científica, não se faça por uma “língua estrangeira”, deve-se na RI e OC atentar para a prática de uma mediação discursiva para enfim divulgar e disseminar pesquisas.

⁴ Garantia literária: princípio cunhado por Humel, E. W. em 1911. HULME, E.W. Principles of book classification: chapter III – on the definition of class headings, and the natural limit to the extension of book classification. *Library Association Record*, n.13, p. 444-449, 1911.

A divulgação científica (doravante D.C.) é classicamente considerada como uma atividade de disseminação, em direção ao exterior, de conhecimentos científicos já produzidos e em circulação no interior de uma comunidade mais restrita; essa disseminação é feita fora da instituição escolar-universitária e não visa à formação de especialistas, isto é, não tem objetivo de estender a comunidade de origem. As representações subentendidas nessa atividade – importante nos diversos meios da mídia, na medida que o exterior a ser alcançado é ‘a coletividade como um todo’ são as de uma função tida socialmente como necessária para o desenvolvimento das ciências. [...] através de uma disseminação desse saber no conjunto da coletividade. (AUTHIER-REVUZ, 1998, p. 107).

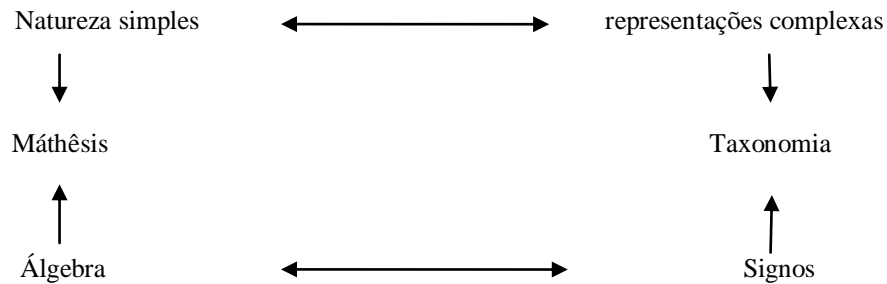
Para atender à coletividade, quando representamos para difundir ou disseminar, a mediação da informação deve ser reconhecida na OC. Quando os termos são organizados não se pode pensar apenas nas regras impostas para SOC, sobretudo, deve-se disponibilizar uma linguagem real para o usuário, que revele um enunciado esclarecedor, como pronunciado por Foucault, (1999, p. 55) “[...] em fazer tudo falar. Isto é, em fazer nascer, por sobre todas as marcas, o discurso segundo o comentário.”.

De fato a representação é dependente da OC, de uma análise conceitual criteriosa, designando as coisas através de descritores, ou seus elementos, decompondo-as até a origem terminológica, mas também, e, sobretudo, combinando os termos através de suas associações com outros termos. Segmentando textos, através de macro-proposições semânticas e levando em consideração os argumentos, para enfim se estruturar um vocabulário.

Em Foucault, (1999, p.89) “[...] a representação é sempre perpendicular a si mesma: é ao mesmo tempo indicação e aparecer; relação a um objeto e manifestação de si.” A representação é sempre dependente de outra a que estará ligada e dela mesma, através das ligações, como uma abstração da ideia primeira ou de onde o conceito foi originado. Para Varela; Thompson e Rosch (2001, p.42) “[...] atitude abstrata é o traje espacial, o acolchoamento feito de hábitos e pressuposições, a armadura com a qual uma pessoa habitualmente se distancia de sua experiência.”.

Em uma representação sobre a origem das coisas e a ciência geral das coisas, Foucault (1999, p. 100) enfatiza que existe uma álgebra nas relações, “um método para conferir signos às naturezas simples e para operar sobre esses signos” apresentamos o modelo por ele proposto na figura 4:

Figura 4 - Ciência geral das coisas



Fonte: Foucault, (1999. p. 100).

O modelo apresentado por Foucault explica as percepções humanas e o movimento do pensamento a partir do “cálculo das igualdades e a gênese das representações”, que muito embora configurem o saber da época clássica, ainda perdura quando se pensa na ordenação das coisas para a organização do conhecimento. Ou ainda na lógica aristotélica dos enunciados verbais, com suas possibilidades de verdade ou inverdade, e na ontologia, para que a interpretação ocorra nos sistemas de proposições contemporâneas como os das ontologias de domínios. Como afirmaria o autor sobre as questões de análise e representação, na concepção de Foucault (1999, p. 105) “[...] o limite do saber seria a transparência perfeita das representações nos signos que as ordenam”.

A figura 4 (quatro) aponta a ligação da natureza geral com uma ciência também geral, como se fosse capaz de explicar e ordenar tudo sem levar em conta o objeto (conceito) dentro de um contexto, o que discordamos. Quanto às representações complexas Foucault apresenta as taxonomias e os signos, neste aspecto o que apontamos é que uma taxonomia muito embora categorize os conceitos numa hierarquia, o que pode parecer rígido em demasia, é a base para construção das relações associativas entre os conceitos como os que ocorrem nas ontologias.

Para Foucault, (1999), a taxonomia é responsável pelo contínuum das coisas, o que se aproxima da Teoria da Classificação Facetada, desenvolvida por Ranganathan (1967) e tão

bem empregada em estudos no domínio de representação e organização do conhecimento na área da Ciência da Informação, para taxonomias e ontologias de domínios.

Não é por acaso que a base para a construção de ontologias é a taxonomia, permitindo a organização dos termos de forma hierárquica para somente após encontrar a sua origem através das relações lógicas (gênero e espécie) ou relações de identidade (parte-todo) percebidas imediatamente pelos sentidos, construir as relações ontológicas ou (associativas) entre os objetos (conceitos).

Estas relações entre os elementos significantes é que representarão as designações coletivas dos termos pelos domínios, que suprem as necessidades dos domínios e, por conseguinte, os limites do nosso conhecimento. São elas que provocam a síntese do diferente, a possibilidade de ligação entre as representações e o conhecimento possível. Assim o sistema clássico das taxonomias dá lugar a um espaço aberto ao termo, onde a representação conceitual estabelecida pelas relações entre os objetos, ou as coisas, por fim nomeia através do termo.

A análise discursiva representativa dos domínios de conhecimento que se entrelaçam, dependem da decomposição das ideias que esclarecem o saber, conhecendo e reconhecendo o interior e o exterior do texto através da análise, e recompondo pela síntese, constituindo e reconstituindo o “a priori” que torna possível a descrição das experiências constatadas empiricamente e descritas através de conteúdos discursivos das comunidades de práticas acadêmicas.

Assim se busca a formalização conceitual e um consenso ontológico nas práticas das ontologias de domínios, formalizando a linguagem para que se possa “controlar” uma linguagem natural; segundo Foucault, (1999 p. 414) “[...] interpretar e formalizar tornaram-se as duas grandes formas de análise de nossa época: na verdade, não conhecemos outras.”.

A prática da categorização conceitual é fundamental porque permite evidenciar a inter e multidisciplinaridade dos domínios. A OC em uma prática interdisciplinar, neste trabalho, abraça a Análise Cognitiva (AnCo) como aporte metodológico, e assim, enriquece o trabalho que há muito vem sendo feito nas áreas da CI e CC. Mas não se pode correr o risco de

categorizações rígidas e fechadas que não levem em conta, por exemplo, as questões de ordem cultural da linguagem.

2.4 ANÁLISE COGNITIVA: UMA METODOLOGIA PARA ANÁLISE CONCEITUAL

O aporte metodológico empregado em AnCo com base na perspectiva multirreferencial vem buscando nos processos de construção do conhecimento, múltiplas análises; uma delas está evidenciada nesta pesquisa, quando ao implementarmos uma prática de análise conceitual, utilizada nos processos de OC, assumimos a proposta de entender a multidimensionalidade do trabalho com redes conceituais e ainda quando propomos considerar os pontos de vistas da CC, CI, Filosofia e da Linguística.

O trabalho multirreferencial está explicitado no trabalho de Fagundes; Fróes Burnham (2001, p. 49) quando afirmam que “[...] A multirreferencialidade pode ser entendida como uma pluralidade de olhares dirigidos a uma realidade e uma pluralidade de linguagens para traduzir esta mesma realidade e os olhares dirigidos a ela.”.

Ou ainda podemos nos apropriar da concepção de Pereira⁵ (2013, p. 33) quando declara que “[...] a análise multirreferencial é uma análise compreensiva, interpretativa, uma forma hermenêutica de análise, que se volta à apreensão mais global da complexidade, buscando a sua *compreensão* e não a sua explicação.” Para nossa pesquisa, a compreensão interpretativa do objeto e a apreensão da complexidade do conjunto das relações conceituais.

Os métodos utilizados em análise conceitual por profissionais como Bibliotecários, há um tempo vinham sendo feitos apenas através de práticas de bom senso, quando utilizavam uma leitura do documento para fins de indexação e retiravam palavras-chave, com vista à indexação do material analisado em base de dados referenciais e textuais. Essa prática, muitas vezes, não levava em conta o ponto de vista do leitor (usuário da informação), era efetuada apenas com base no atendimento às necessidades da instituição a qual o centro de documentação ou a biblioteca estava vinculada.

⁵ Os sobrenomes Pereira e Lage pertencem à mesma autora, no caso Ana Lúcia. Referenciamos os dois sobrenomes conforme apareceram nos trabalhos que utilizamos, seguindo os padrões estabelecidos em normalização segundo ABNT e AACR2.

Mesmo aparentemente atendendo ao contexto em que a biblioteca estivesse inserida, variáveis como as do discurso, linguagem e ideologia, frequentemente não eram analisadas. Conceitos eram dispostos em base de dados com pouco rigor, o que prejudicava a contribuição que a descrição temática e conceitual pode ter na revelação de sentido/conteúdo para se chegar a uma recuperação adequada de informação.

Acreditamos que necessitamos de uma AC em que o analista atente para o conteúdo de um texto que pode ter informações úteis não só para o produtor de conhecimento (autor), mas também indivíduos em geral. É preciso observar a distribuição das ocorrências conceituais encontradas em textos, procurando atender a demanda de informação do sujeito. Precisamos de uma linguagem de representação que seja um produto advindo de uma análise do discurso, do saber e da pesquisa, que melhor determine as funções e categorias dos conceitos.

Sendo assim, entendemos que necessitamos de estudos que vão além das questões estruturais propostas em linguagens automatizadas de recuperação de informações, sobretudo, um estudo que possa unir a estrutura da linguagem com estruturas do pensamento, revelando o sentido da união dessas estruturas que levará ao conhecimento. Isso requer um trabalho multirreferencial.

Estamos em busca de uma análise em dois níveis: a do texto e da análise conceitual. A primeira análise, a textual, o foco está no discurso do produtor do conhecimento (autor (a)) e a segunda concentra-se na identificação das macroproposições semânticas traduzidas por palavras/conceitos. Sendo assim o analista documentário se configura como um analista cognitivo/interprete cognitivo.

A análise cognitiva que aqui propomos está evidenciada quando identificamos a organização discursiva do produtor do texto, através da segmentação de textos integrais por uma rede de palavras, que interligadas são tradutoras de ideias do discurso. Convém esclarecer que os textos são produzidos em um universo de conhecimento, neste caso o domínio meio ambiente, apresentado no capítulo 4 (quatro) seção 4.

Em seguida escolhemos palavras da rede, selecionadas a partir de parâmetros quantitativos estimados pela frequência e pela proporção de coocorrência de pares das palavras no texto (chamado de incidência-fidelidade). O conceito de incidência-fidelidade foi

definido na pesquisa de mestrado de Teixeira (2007) e no trabalho de Teixeira et al. (2010) como um índice que está associado a essa frequência.

A partir das palavras selecionadas construímos uma sub-rede de palavras/termos que objetiva condensar o conteúdo expresso pelo autor no texto e demonstre os resultados e conclusões por ele obtidas, as propostas explícitas e por vezes implícitas, mas reveladas pelas palavras e suas relações com outras palavras a serem analisadas enquanto conceitos e que configurará como descritor do documento.

Os traços descritivos do texto/contexto revelados pelos conceitos poderão compor um SOC do domínio. Na etapa de construção do SOC está inserida mais uma etapa de análise cognitiva, quando da ordenação conceitual em estruturas lógicas e ontológicas. Esta etapa se constitui por combinações de dispositivos de reformulação e assimilação.

Na etapa de reformulação: em um grupo de conceitos, pode incorporar-se um conceito externo ao domínio, dando a essa etapa um resultado advindo da interpretação ou de construção do conceito novo, ou seja, precisamos declarar o seu significado.

Na etapa de transformação por assimilação: conceitos distintos unem-se através de um processo de síntese de argumentos, pois em uma rede ou sub-rede de palavras, um conceito está ligado a outro conceito de forma coordenada ou subordinada. Uma coerente interpretação da carga ideológica e semântica em relação aos conceitos pode nos levar a uma hipótese discursiva significativa.

As etapas descritas são efetuadas pelo interprete cognitivo, em um movimento que chamamos de construção de estruturas conceituais, por esse motivo entendemos que a AC é um processo que abarca análise cognitiva e esse interprete se configura como analista cognitivo. As etapas aqui descritas estão evidenciadas no capítulo 5 (cinco).

As bases para construção das estruturas conceituais estão ligadas ao pensamento e ao conhecimento do ser humano, portanto devem ser entendidas como funções comunicativas e constituídas da linguagem. Para Maturana:

[...] o resultado desta condição de constituição da linguagem é que nós, seres humanos, existimos como observadores na linguagem, e quaisquer distinções que fazamos são operações na linguagem, em conformidade com

circunstâncias que surgiram em nós na linguagem. (MATURANA, 2001, p. 131)

É necessário entender como mente e linguagem desempenham seus papéis no processo cognitivo responsável por organizar o sistema conceitual, que se reflete na linguagem científica e cotidiana da pesquisa e na percepção intuitiva e declarada da realidade, capaz de perceber similaridades e diferenças entre objetos/conceitos.

Além do trabalho de análise semântica, contido na AC, necessitamos analisar o conceito como usuários do domínio de conhecimento, pois uma das formas de construção do conceito se encontra na interpretação da experiência do sujeito, ou de como ele organiza significados sobre o objeto.

É fundamental que o analista cognitivo, no trabalho com AC, se desloque do “lugar” do analista para o “lugar” do usuário da informação, para compreender o papel desempenhado pelo conceito no cotidiano desse usuário. Segundo Varela (2007, p. 38) “[...] na transmissão da informação, não há somente um gerador, um transmissor; há também um receptor, uma consciência receptora, uma necessidade humana.” Portanto, entender a formação do conceito no domínio é entender a formação do indivíduo e a experiência vivida por ele, que organiza mentalmente e confere significado ao conceito.

O conceito é a forma objetiva de declarar o que se sabe, o que se faz e o que se deseja, do conhecimento, por isso ele o representa e precisa ser analisado enquanto representação e vontade do sujeito.

Analisar passa a ser acompanhar o processo, compreendê-lo, apreendê-lo mais globalmente através da familiarização, nele reconhecendo a relativamente irremediável opacidade que o caracteriza. Passa a ser produzir a explicitação, a elucidação (e não a explicação racional), sem procurar interromper o seu movimento, mas realizar esta produção ao mesmo tempo em que tal processo se renova, se recria, na dinâmica intersubjetiva da penetração na sua intimidade, na multiplicidade de significados, na possibilidade de negação de si mesmo, que caracteriza o sujeito das relações sociais. (FRÓES BURNHAM, 1998, p. 41).

Além da visão do sujeito que pesquisa, precisamos encontrar uma linguagem de interação, assim baseamos-nos na linguagem e na forma como a ciência se expressa por essa linguagem nos domínios de conhecimento, para entendermos como os conceitos se articulam

e se relacionam com outros conceitos, e para traduzir o discurso dos cientistas que estão impressos nos textos, resultados de suas investigações. Segundo, Maturana:

Nós vivemos na experiência, na praxis de viver de seres humanos no fluir de sermos sistemas vivos na linguagem, como algo que acontece em nós e a nós à medida que linguajamos. Este é o motivo pelo qual, ao explicarmos como cientistas nossa experiência como seres humanos, reformulando-a com seus elementos através da aplicação do critério de validação das explicações científicas, nos encontramos gerando a ciência como um domínio cognitivo que não nos coloca fora da experiência, e que nos mantém na linguagem. (MATURANA, 2001, p. 127).

No caso deste trabalho de pesquisa buscamos uma linguagem expressa em textos, com limites discursivos. Uma linguagem dependente do agrupamento das palavras e das relações existentes entre elas, capazes de assumir funções a depender do lugar que ocupa em relação às demais. Mas não podemos deslocar as palavras da sua situação de enunciação que objetiva o discurso do domínio de conhecimento, não podemos nos afastar da estrutura do pensamento, que se traduz por conceitos. Para tanto precisamos de uma análise conceitual, para tanto necessitamos de AnCo.

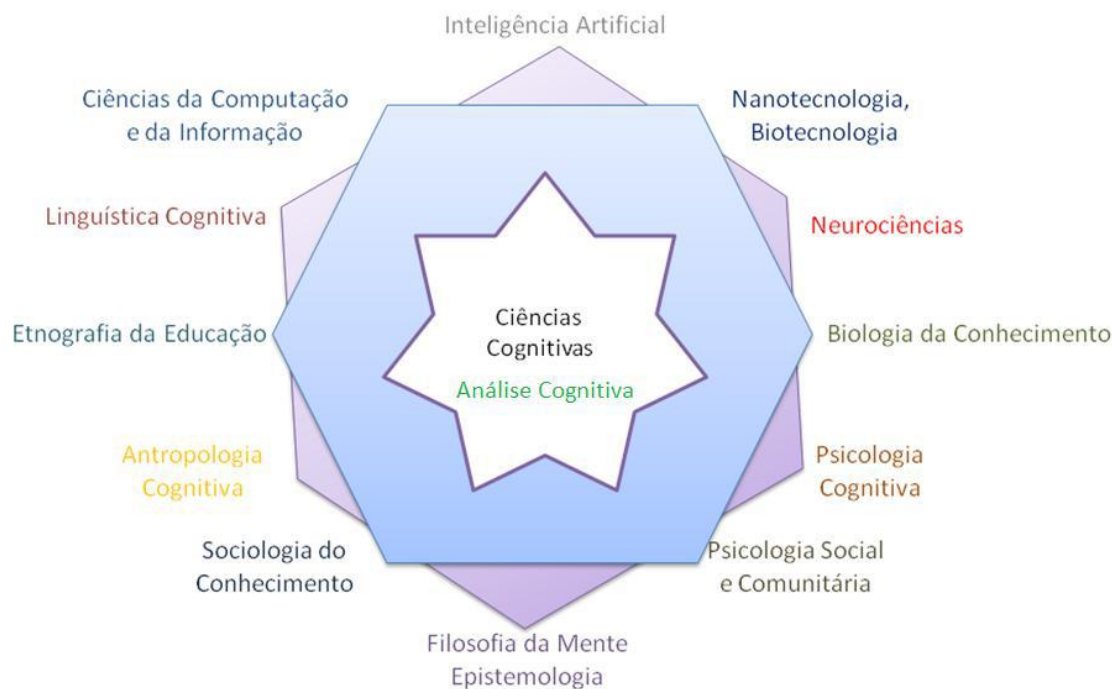
A Análise Cognitiva (AnCo) vem sendo objeto de pesquisa no grupo de pesquisa CAOS, vinculado ao DMMDC, e até o momento estudos revelaram que a AnCo, vem se configurando como um campo de conhecimento em construção que possibilita aporte epistemológico, metodológico e teórico para melhor compreender os estágios de apreensão do sujeito nos processos de aprendizagem formais e informais, presenciais ou virtuais, no contexto em que esteja inserido. Sobre as incursões realizadas no âmbito da AnCo em artigos científicos Fróes Burnham declara:

- 1) uma grande diversidade de áreas do conhecimento instituídas que a [análise cognitiva] empregam – Psicologia, Neurociência, Ciências da Computação e Engenharia, Antropologia, Saúde, Linguística, Artes (Música), Humanidades, Filosofia, Ciências Biológicas, Direito, Economia, entre outras;
- 2) uma ampla gama de focos de conteúdos – abrangendo estudos de gênero, personalidade, comportamento de líderes, testemunhas em tribunais, pacientes atendidos em unidades de saúde, tratamentos neurofarmacológicos, analgesia hipnótica, formação de professores, desempenho cognitivo de estudantes e de pessoas com necessidades especiais, contatos intergrupais, análise de tarefas, de informação e de requisitos ergonômicos, para citar apenas uns poucos exemplos;

3) uma extensa dispersão de significados atribuídos – construção de modelos analíticos, testagem de modelos teóricos no campo empírico, linguagens estruturadas para comunicação de processos de interoperabilidade, técnicas de organização de tarefas, integração de métodos e sistemas na engenharia de software, processos de raciocínio na resolução de problemas e de avaliação de desenvolvimento cognitivo e aprendizagem complexa, metodologias para acompanhar a construção de noções sociais. (FRÓES BURNHAM, 2011, p.58).

Para demonstrar a ampliação das diferentes abordagens epistemológicas da cognição e das Ciências Cognitivas, Lage; Fróes Burnham; Michinel, (2012) apresentaram um diagrama conceitual “[...] baseado no hexágono de Sloam (1978) e no mapa polar de Varela (1996)”, como se pode observar na figura que segue:

Figura 5 – Diagrama conceitual do campo ampliado das ciências cognitivas



Fonte: Lage; Fróes-Burnham; Michinel (2012, p.95)

Diante da afirmativa de Fróes Burnham quanto a diversidade de áreas do conhecimento que a AnCo mantém articulações e observando o hexágono conceitual, afirmamos que esta pesquisa tem seu lastro teórico na CI.

Convém esclarecer a partir deste ponto sobre o aspecto cognitivo, principalmente, se intencionamos considerar a forma como o indivíduo pensa e manipula a informação para apreender conhecimento, o que vem sendo estudado pela Ciência Cognitiva. Sendo assim apresentamos uma breve descrição sobre a área.

Os estudos em Ciência Cognitiva tiveram sua origem na década de 1950 do século XX, nos encontros realizados em Cambridge e Dartmouth com as pesquisas de Herbert Simon, Noam Chomsky, Marvin Minsky e John McCarthy. Para uma ciência cognitiva moderna, mais ampla e interdisciplinar, ampliaram-se estudos nas ciências sociais e biológicas, além dos estudos da ciência da computação. Segundo Varela; Thompson e Rosch:

[...] a hipótese cognitivista implica uma afirmação muito vigorosa a respeito das relações entre sintaxe e semântica. [...] não há nenhum lugar em que as manifestações do cognitivismo, sejam mais visíveis que na Inteligência Artificial (IA), que é a implementação literal da hipótese cognitivista. [...] cognição e intencionalidade (representação formam um par inseparável) [...] (VARELA; THOMPSON; ROSCH, 2001, p. 58-65).

Seguindo os passos do cognitivismo, as primeiras experimentações e regras do processamento simbólico de informações, estava amparado em modelos sequenciais de conexões que começavam em um estado inicial arbitrários, depois, o processo caminhou para redes dinâmicas conexionistas.

O aspecto do conexionismo é importante para a tarefa de modelagem conceitual, pois segundo Rozados, (2003) admite a conexão e o contexto, o que é fundamental na representação conceitual e simbólica da realidade de um domínio de conhecimento. Sendo assim é útil para o trabalho de análise das redes semânticas ou conceituais e nos processos de apreensão do significado do conceito.

Com a evolução dos sistemas computacionais, as operações numéricas nas redes dinâmicas conexionistas tomam lugar das operações simbólicas, para atender aos padrões complexos de unidades de uma rede, um conexionismo, um “paradigma subsimbólico” principalmente utilizados em redes neurais. Sobre princípios cognitivos desse paradigma, Varela; Thompson e Rosch (2001, p. 112), descreveram: “[...] a estratégia, como dissemos, é construir um sistema cognitivo sem começar com símbolos e regras, mas com componentes simples que se conectariam intensamente uns com os outros de maneira dinâmica.”.

O paradigma subsimbólico, seria para a representação como um aspecto de generalização, enquanto que no paradigma simbólico os conceitos de um sistema de representação são independentes, mas parte de uma representação conceitual pode reconstituir outra parte. O que é importante para perceber a similaridade entre conceitos. Tais paradigmas são úteis, principalmente quando usados em conjunto, de forma híbrida, em sistemas de processamento de sentenças que visam processamento de linguagem.

Assim nos perguntamos se a atribuição de significado ao conceito, no nível da representação se torna claro ao reconhecimento e ao aprendizado, no “paradigma subsimbólico”? O que buscamos responder é que a interpretação conceitual é complexa quanto ao seu significado simbólico, pois a denotação atribuída a um conceito em determinado domínio não é totalmente independente do conceito emergente. Talvez seja possível tratar questões relacionadas a categoriais conceituais, produzindo um modelo de raciocínio lógico do domínio, produzindo inferências baseadas em regras simbólicas, no entanto tal modelo é de um nível macro de representação ou meta-representação, ou seja, aproximado do seu significado.

A intuição central por detrás do cognitivismo é que a inteligência - incluindo a inteligência humana - assemelha-se à computação em suas características essenciais de que a cognição pode ser efetivamente definida como computações de representações simbólicas. (VARELA; THOMPSON; ROSCH 2001, p.56).

Se buscarmos, uma representação abrangente ou de nível inferior, podemos então utilizar apenas uma forma abrangente de análise, no entanto para estruturas que representam a linguagem humana, uma descrição em alto nível necessita de um nível simbólico e subsimbólico para traduzir conceitos e demonstrar representações sistemáticas.

A AnCo se caracteriza nesta pesquisa como uma metodologia de análise e avaliação do processo cognitivo que auxilia a construção do modelo dinâmico de análise conceitual. Tal modelo intenciona contribuir com processos de modelização de domínios de conhecimento e atender as respostas que os usuários precisam encontrar e obter.

É sabido que em um processo que envolve construção terminológica cujo resultado será tesouros; taxonomias ou ontologias, ferramentas fundamentais em sistemas de busca à

informação, deve-se levar em conta a experiência do usuário, bem como suas expectativas em encontrar a informação e o seu conhecimento sobre o contexto do domínio deve ser levado em conta.

O processo de construção da rede de conceitos que propomos basear-se-á, portanto em três passos:

- analisar conceitos de acordo com a necessidade de informação dos usuários dos domínios que utilizam sistemas de recuperação de informação;
- organizar e representar conceitos que revelem o domínio de conhecimento;
- apresentar relações entre conceitos, para que o usuário possa apreender novos conceitos.

Para realizar a análise conceitual, em cada passo do processo propomos que se deverá ter o cuidado de, a todo o momento, persistir nas perguntas:

- O que o usuário precisa encontrar sobre o domínio?
- O que o usuário precisa obter do domínio?
- O que o usuário pode conhecer sobre o domínio?

Neste passo da análise o modelizador de domínios, observador da realidade, mais do que em qualquer outro passo precisa compreender o domínio a ser representado pelos conceitos e fundamentalmente à comunidade a que a representação atenderá.

No entanto, se quero explicar o observador, tenho que colocá-lo nesses termos. Tenho que propor um mecanismo gerativo que, como resultado de seu funcionar, me dê à experiência do observar, e das coerências operacionais implícitas nesse mecanismo devo deduzir outro fenômeno, sem experienciá-lo. Se faço isso, tenho a explicação científica do observar. Interessante que isso possa ser feito. Vejam que, no momento em que assumo isso, certamente vai mudar minha filosofia, e ela terá que mudar porque não posso deixar de assumir o fato de que a universalidade da ciência não está em sua referência a um universo, mas está na configuração de uma comunidade humana que aceita esse critério explicativo. (MATURANA, 2001, p. 48).

É importante que o modelizador tenha como garantia um sistema de armazenamento das buscas dos usuários e dos documentos recuperados por esses usuários, para entender as

possíveis buscas futuras. Assim, o trabalho com organização dos termos poderá ser mais coerente com as necessidades dos usuários, além de possibilitar o encontro de novos itens documentais.

Deve-se também pensar como os conceitos emergentes dos domínios, poderão ser categorizados e organizados em um SOC para compor uma base de dados que atenderá às consultas dos usuários. Deseja-se um índice de precisão que atenda às necessidades desses usuários, que torne funcional a base de dados e a interface de pesquisa proposta por essa base. Assim precisamos também observar:

- O que o conceito expressa sobre o domínio de conhecimento?
- O que o conceito explica sobre o domínio?
- O que o conceito revela do texto?
- O que o texto propõe?

Entender o comportamento do usuário nas bases de busca, quando pretende encontrar, obter e conhecer um domínio é fundamental para organizar conceitos em SOC de domínios de conhecimento. Cabe ao modelizador de domínios atuar como analista cognitivo e mediador entendendo as reações dinâmicas do meio informacional. Como dito por Gomes, (2008, p.2) “[...] nas atividades de acesso ao conhecimento estabelecido, ao conhecimento registrado, os educadores, os espaços informacionais e seus agentes são os mediadores que transmitem as informações disponíveis, realizando assim práticas informacionais.”.

A AnCo será útil ainda para avaliar o processo de criação de ontologias de domínios, além de contribuir para o projeto de uma interface de busca para bases de dados referenciais. A metodologia permitirá trabalhar com o conhecimento que o usuário possui e que precisa ter para chegar à informação útil, a sua questão de busca, pois nem sempre interfaces de pesquisa fazem o elo entre o que o usuário sabe e o que ele precisa descobrir.

As descobertas mais interessantes e valiosas podem acontecer por acaso. Esperamos que o usuário não desista das buscas e acreditamos que conceitos organizados a partir da análise dos textos dos domínios de conhecimento possam contribuir para que usuários descubram o esperado e o inesperado.

A análise conceitual proposta nesta tese inicia-se com a análise do discurso representativo do domínio, amparada pela AnCo. O discurso analisado a princípio nas redes semânticas, através das palavras originadas das dissertações e teses selecionadas no domínio, para em seguida após a síntese dos termos selecionados, propor uma base de análise para a estrutura ontológica dos conceitos. Para Teixeira, (2007, p. 27) “O conceito é o objeto de estudo de investigação da memória semântica”.

Sendo assim o amparo que a AnCo nos dá diz respeito à forma como interpretamos os conceitos, analisando-os sob a ótica do pensamento e percepção do usuário, como o sujeito reconhece e compreende o conceito em bases referencias e textuais para solucionar problemas que afetam o cotidiano de suas práticas e conhecimento. O pensamento que aqui nos interessa é o que está em andamento no ato da busca, que varia de acordo com fatores pessoais e ambientais. Pensamento que se caracteriza por unidade conceitual, integração e continuidade.

Procuramos através de hipóteses, apresentar e representar conceitos que guiarão as informações e documentos, que auxiliarão o usuário, informação promotora do desenvolvimento de suas atividades. O analista cognitivo, portanto, para o caso desta pesquisa, é o interprete de termos advindos dos domínios e que facilitarão as pesquisas dos usuários que pretendem: procurar, identificar, selecionar e por fim obter informações através de ferramentas de busca cujas bases de organização do conhecimento são os conceitos.

Passamos então a entender as redes semânticas, o primeiro passo para análise dos movimentos conceituais do domínio de conhecimento.

2.5 REDES SEMÂNTICAS/CONCEITUAIS: O PONTO DE PARTIDA PARA ANÁLISE DO CORPUS SELECIONADO

As interações conceituais se configuram em redes quando nas relações internas provocadas pelo discurso impresso em domínios de conhecimento, pode-se observar nas interações conceituais uma representação desse domínio, é onde o conceito nasce, ou ainda se reconfigura, segundo Foucault, (1999, p.69) “[...] lugar de emergência dos conceitos [...] em um emaranhado de compatibilidade e incompatibilidades conceituais.” Configuram-se por si

só em redes de relações possíveis, pois trabalham com conceitos enredados e cuja representação não se apresenta de forma linear.

A representação da informação e do conhecimento tem nas redes semânticas uma ferramenta fundamental de notação gráfica de nodos interconectados, e de suporte a sistemas automatizados de recuperação de informação. O objetivo de tais redes é demonstrar a construção de ligações, dados dois nós e um rótulo de ligação, com a intenção de declarar o significado dos signos, sob a definição do termo semântica.

Nesta pesquisa de doutorado não nos atemos a discutir o conceito de rede semântica, apenas utilizamos o termo por estar cunhado na literatura. O que pode ser confirmado quando autores contemporâneos como Fadigas et. al. (2009); Marcondes (2011); Moreira Gonzáles (2011a; 2011b); Sowa (1987); Teixeira (2007); dentre outros, utilizam o termo em seus trabalhos.

O que nos interessa é o estudo do sentido da palavra e a relação que essa palavra tem com outra, que facilita o entendimento das relações entre conceitos. O principal objetivo de uma rede semântica a nosso ver está na contextualização da palavra, que tem um significado para um determinado domínio por isso, semântica.

A rede semântica se aproxima do nosso objetivo de tese, quando nos possibilita analisar tematicamente o discurso impresso no texto, ou o seu conteúdo. Conteúdo esse, considerado em um conjunto de ligações que se dão entre signos (palavras) e que possibilita uma organização em macroestrutura, a partir das estruturas do texto e que serão representadas por uma linguagem conceitual, controlada artificialmente. Sobre redes semânticas Moreira González declara:

Estudo do significado das relações que existem entre os signos e os objetos ou fenômenos que designam. Considera-se o signo em relação ao que significa. Pode-se chegar a um conceito mais amplo, que distinga entre uma semântica referencial (relação com a realidade extralinguística) e uma semântica orientada ao conteúdo, ambas entendidas como relações imanentes ao sistema, que se dão entre o conteúdo dos signos. (MOREIRA GONZÁLEZ, 2011b, p. 27)

Os signos são representações que dependem das estruturas mentais, sociais e culturais do sujeito, segundo Simões, (2002, p. 139) “[...] são sinais que podem ser interpretados por

evocarem algo que ali não está e que, por isso, representam. Nessa perspectiva, o signo faz-se notar como *representamen*". E esta representação é simbólica e se configura através da linguagem e a linguagem humana depende das palavras, que carregam inúmeros significados conceituais a depender do contexto que se expressam e de como esses conceitos são traçados por uma rede semântica.

A mais antiga das redes semânticas, tem origem como afirmou (Sowa, 1993) na árvore de Porfírio. Baseada numa estrutura hierárquica a partir das categorias formais de Aristóteles, objetivava implementar definições conceituais e padrões de relações para um esquema de classificação. Esta proposta utilizada entre as décadas de 1970 e início de 1980 serviu de base para a KL-One (Knowledge Language One), uma linguagem gráfica onde o conhecimento é representado por diagramas, as relações conceituais são demonstradas por subtipo e supertipo, tendo ao centro as questões de propriedade. Sobre a origem da rede, Freitas afirma:

Para além do desenvolvimento da teoria da suposição, baseados na árvore de Porfírio, os lógicos medievais, como Pedro Hispano, desenvolveram uma teoria ontológica formal. Podemos afirmar sem dúvida que aquilo a que chamamos hoje em dia redes semânticas têm a sua origem na árvore de Porfírio e são uma forma de representar o conhecimento. (FREITAS, 2013 p.6)

Em se tratando da representação gráfica das redes semânticas Moreiro González exemplifica:

Estas redes representam-se graficamente mediante nós orientados, com um nó inicial e um nó final. A relação existente entre os conceitos se define numa direção determinada. Assim podem-se visualizar os conceitos macroestruturais. Sua base é que o conhecimento representa-se por frases simples e estruturadas: - Sujeito - verbo - predicado; - Como uma associação: conceito - relação – conceito. (MOREIRO GONZÁLEZ, 2011a, p. 154)

Como o ponto fundamental desta tese, se encontra na organização e representação do conhecimento, preferimos trabalhar redes conceituais a partir das redes semânticas, por nos garantir, melhor entendimento de associações entre termos. Para Aguiar:

Em particular, a psicologia cognitiva sugere modelos que representem a arquitetura cognitiva a partir do conhecimento declarado por indivíduos. Uma dessas formas de representação computacional ou matemática da estrutura cognitiva dos indivíduos é conhecida como rede associativa ou rede semântica. (AGUIAR, 2009, p. 15)

As redes semânticas podem ser entendidas como redes conceituais, essas por vezes se caracterizam pelo uso de mapas conceituais, onde os conceitos são dispostos em nodos linkados por linhas etiquetadas e propositivas que representam as relações conceituais. Neste tipo de rede podem-se dividir os nós em subredes e ainda indicar links entre elas. Em sua organização a fundamentação está na hierarquia que vai do conceito mais abrangente até os mais específicos, para tanto são estabelecidas proposições de conceitos interligados por palavras que os relacionam. Sobre uso de mapa conceitual Lima evidencia:

O mapa conceitual tem sido utilizado em diversas áreas do conhecimento como técnica formal ou semi-formal de diagramação. Na área de educação, ciência política, linguística e filosofia da ciência, essa técnica tem sido usada para apresentar visualmente a estrutura do conhecimento e suas formas de argumentação. (LIMA, 2004, p.137).

O propósito da rede conceitual é apreender abordagens conceituais advindas de domínios de conhecimento, observando relações conceituais internas e associações com outros domínios e compreender enunciados contextualizados. Seria como um sistema de ideias, pois os conceitos emanam do plano das ideias, com capacidade de adaptações e articulações entre subsistemas.

Nas redes desta natureza o principal propósito é um refinamento técnico objetivando a web semântica, porém as questões ligadas à inteligência artificial, do ponto de vista de executar funções inteligentes ligadas aos significados das palavras, atribuindo sentidos aos conteúdos perceptíveis ao computador, são limitadas às operações lógicas para máquinas.

Para recursos de representação do conhecimento são necessários mecanismos mais avançados, que abarquem um tratamento terminológico condizente com a complexidade que é lidar com conceitos científicos ou com sistemas de organização e representação do conhecimento e ainda dependentes da cognição humana.

Mesmo nos dias atuais quando nos é dada a oportunidade de pesquisas na Internet, em buscadores como o Google, citando o mais utilizado, que usam robôs para cobrir a grande extensão de informações contidas na rede, em listas de palavras associadas a partir de URL (Uniform Resource Locator), bem como critérios de indexação de assuntos a partir da

frequência de uso de palavras chave utilizadas por usuários, percebemos que ainda perdemos muito tempo com textos disponibilizados que não atendem as nossas intenções de informação.

O Google vem tentando minimizar esse problema evoluindo para questões cognitivas, sobre os avanços do Google, Monteiro afirma:

O Google é um exemplo de múltiplas sintaxes, seja em seu buscador original, seja no avanço que faz no sentido de diversificar suas operações de busca no ciberespaço, indo em direção à ubiquidade (digitalizando e indexando conteúdo analógico e múltiplos setores), conceito necessário para construir uma verdadeira memória digital e, conseqüentemente, a busca perfeita, que consiste em conhecer o comportamento de busca e as preferências do usuário. (MONTEIRO, 2009, p. 85)

O Google vem buscando atender a busca no sentido da semântica, uma das propostas está no Knowledge Graph de 2012, diferente dos metadados advindos de ontologias e web semântica, pretende-se a aproximação com processos cognitivos de interpretações, aproximação de interpretantes e interpretes. A pesquisa é realizada através de um mapa de conhecimento que aproxima as informações relevantes relacionadas à questão primeira do usuário.⁶

Além disso, uma importante ferramenta de busca é o Google Acadêmico, sobre esta ferramenta e como especificado em páginas na web:

O Google Acadêmico classifica os resultados de pesquisa segundo a relevância. Como na pesquisa da web como o Google, as referências mais úteis são exibidas no início da página. A tecnologia de classificação do Google leva em conta o texto integral do artigo, o autor, a publicação em que o artigo saiu e a frequência com que foi citado em outras publicações acadêmicas. (http://infojardins.blogspot.com.br/2011_03_01_archive.html)

Entendemos os critérios de relevância da ferramenta Google Acadêmico como seletiva no que concernem os parâmetros de autoria, veículo de publicação e uso, das referências. Mas acreditamos que o parâmetro de citação é o que mais se aproxima da nossa intenção de pesquisa, pois quando citamos uma referência é porque essa reafirma o nosso discurso e o valida.

⁶ <http://www.google.com/insidesearch/features/search/knowledge.html>

Além disso, o uso da informação nas redes, principalmente no sentido da socialização da informação, a web semântica. A Web social provocou novos construtos para a Engenharia do conhecimento no sentido de modelos de representação. O que não quer dizer abandono das linguagens documentárias tradicionais, mas, sobretudo existe nos dias atuais a preocupação da rapidez tanto no processo de organizar o conhecimento quanto e principalmente na recuperação das informações.

Esta preocupação foi potencializada pela Web social, pois todos querem colaborar. Neste afã de colaboração surge a folksonomia, onde usuários colaboram com metadados, ou atribuição de tags. O conceito de folksonomia foi criado por Thomas Vander Wal, na primeira década do século XXI, em uma espécie de analogia, juntou as palavras folk (povo) e taxonomia (classificação). Uma linguagem comum (natural) realizada por pessoas comuns. Dziekaniak; Pacheco; Kern elucidam a lógica da folksonomia:

A lógica abordada no uso da folksonomia é a de que usuários que utilizam a mesma linguagem, por estarem acessando conteúdos de mesmo, interesse sejam conhecedores o suficiente dos conteúdos para elegerem termos representativos sobre o objeto digital (som, texto, imagem). Com isso, eles auxiliam no processo de classificação do conteúdo digital disponibilizado na Web, através do uso da linguagem natural, de maneira rápida e prática; criando uma espécie de linguagem de especialidade desenvolvida pelas comunidades. (DZIEKANIAK; PACHECO; KERN, 2011, p.206).

Atualmente devido à atuação mais efetiva dos usuários na rede quanto à busca da informação, a potencialização do uso da linguagem natural é evidente, sendo assim, profissionais das áreas da CI e CC tentam encontrar um caminho híbrido, entre a representação formal que parte de estruturação de conceitos como as utilizadas em taxonomias e ontologias e a linguagens utilizadas pelos usuários.

Sendo assim, sistemas de representação e organização do conhecimento que utilizam taxonomias e ontologias para estruturação conceitual, objetivando através de ordenações hierárquicas, partitivas e associativas entre conceitos de um determinado domínio explicitar significados a partir da interpretação dos contextos dos domínios de conhecimento, devem ter o cuidado de atender as comunidades acadêmicas e científicas, mas tendo em mente que indivíduos que por uma questão particular utilizam documentos advindos desses domínios, precisam entender os princípios conceituais.

Esperamos encontrar um caminho de análise conceitual que possa facilitar o trabalho de pesquisadores que procuram uma saída de organização e representação do conhecimento e que atenda a todos os usuários que buscam informação de domínios de conhecimento.

No próximo capítulo algumas considerações sobre taxonomias e ontologias são apresentadas, com intuito de esclarecer a contribuição que o modelo de análise proposto pode oferecer na construção de SOC.

3 TAXONOMIAS E ONTOLOGIAS DE DOMÍNIOS: PARA QUE FORMALIZAR CONCEITOS?

O que especifica uma ciência é ter um objeto. Pode-se sustentar que uma ciência é especificada por um objeto definido, pelo menos por um certo nível de operação, reproduzível, que chamamos de experiência. Mas devemos ser muito prudentes porque esse objeto muda, e singularmente, no curso da evolução de uma ciência. (LACAN, 1998, p.15).

As mudanças ocorridas nos estudos da CI e da CC nos últimos anos, nas questões relacionadas à OC e RC vêm acompanhando a evolução da ciência, principalmente no que concerne a experiência humana no universo da representação. O ser humano muda de acordo com a sua intenção de informação e também de acordo com os espaços aonde esta informação encontra-se armazenada, portanto tudo é mutável, no mundo nada é estático, assim a organização do conhecimento precisa ser dinâmica.

O domínio da OC, através de pesquisas no âmbito das ontologias de domínios, busca sistematizar e estruturar o conhecimento advindo das pesquisas científicas de domínios específicos, com o objetivo de difundi-las e torná-las públicas, em bases de dados ou na rede mundial de informação, possibilitando o alcance da informação a todos aqueles que tenham acesso às bibliotecas, centros de informação e principalmente à web.

Os primeiros estudos publicados sobre ontologias datam da década de 1990 do século XX e se expandiram até o século atual, como instrumento essencial na representação do conhecimento. Podemos citar autores seminais que nortearam as nossas intenções discursivas sobre a temática como: Almeida (2002, 2006); Almeida; Bax, (2003); Ding, (2002a, 2002b); Campos, (2004, 2005, 2010); Guarino, (1995,1997, 1998); Gruber, (1993, 1995); Smith, (2003); Vickery, (1997); dentre outros.

Resumidamente, podemos dizer que se classificam ontologias quanto ao conteúdo como: de domínios, tarefas, aplicações, genéricas e de representação. As de domínios descrevem o vocabulário de um domínio de conhecimento como, por exemplo: geologia ambiental; as de tarefas, conceituações sobre resoluções de problemas, usadas em qualquer domínio; as de aplicação descrevem conceitos de domínios e tarefas particulares; as genéricas descrevem conceitos gerais independentes de um domínio em particular. Através dos

formalismos de representação de conhecimento, as ontologias buscam compromissos ontológicos.

Quanto ao uso em variados domínios, Almeida; Bax (2003, p. 9) apontam: “[...] projetos de domínios como gestão do conhecimento, comércio eletrônico, processamento de linguagens naturais, recuperação da informação na *Web*, de cunho educacional, entre outros.”.

Ontologia foi definida por Gruber em 1993 como uma “especificação explícita de uma conceitualização”, que visa um consenso acerca do conhecimento de um domínio específico. Nesse tipo de ontologias de domínios, definições são associadas às entidades e aos discursos. Para Guarino, (1998), ontologia é um vocabulário intencional de uma realidade, carregado de pressupostos explicitados por uma lógica de primeira ordem, consensualmente aceito por uma comunidade. Tal vocabulário é composto por termos estruturados em classes, relações entre conceitos de um domínio; axiomas que restringem interpretação e uso dos termos e instâncias que representam objetos específicos.

Os aspectos ontológicos são representados em ontologias por um complexo conjunto de conceitos gerados por uma combinação de métodos: automáticos ou semiautomáticos e concentração de correlação sintática formal. Sob a preservação de uma estrutura semântica, esses métodos combinados garantem uma representação estável de domínios de conhecimento.

Mas na organização do conhecimento, no caso de construção de ontologias, cujo objetivo é revelar relacionamentos entre conceitos, para formalizar conceitos de domínios de conhecimento cujo processamento se dará através de “máquina”, esses conceitos precisam ser explicitados formalmente, como afirma Sales; Campos; Gomes.

[...] a necessidade de uma linguagem formal para o processamento da máquina faz com que as relações devam ser explicitadas, por exemplo, em forma de verbos (ex. affects), ou verbos proposicionais (ex. interacts_with). Portanto, além de identificar os elementos da relação, é necessário explicitar a própria relação. (SALES; CAMPOS, GOMES, 2008, p. 67)

Relações categoriais e formais são fundamentais na construção de ontologias como especificado por Sales e Campos, no mesmo trabalho com base em estudo de Sales, 2006 esclarecem:

As relações categoriais são aquelas que revelam duplas de categorias, por exemplo: coisa-processo, material-produto, etc.

As relações formais são aquelas que revelam o tipo de relação existente entre as duplas de categorias por exemplo: causado_ por, ocorre, etc. (SALES; CAMPOS, 2008, p. 68)

A linguagem formal para Web Ontology Language (OWL) foi definida como afirmado por Jorge (2012) pelo Word Wide Web Consortium (W3C), e pode incluir descrição de classes e relacionamento. A partir das afirmações das autoras Sales; Campos; Gomes e da leitura do trabalho de Jorge, exemplificamos no quadro 2 (dois) elementos fundamentais para a construção de ontologias baseados nos elementos apresentados por Breitman (2005):

Quadro2: Exemplificação dos elementos fundamentais na construção de ontologias

Classes de conceitos	Hierarquia	Relacionamento	Função	Instância	Axiomas
Inseto Barbeiro (mosquito)	Tipo de invertebrado Exoesqueleto quitinoso	É um vetor _de ...	Barbeiro é um tipo de _ mosquito exoesqueleto quitinoso propriedade _ dos invertebrados	Mosquito é um tipo _ de inseto	Barbeiro é um tipo de mosquito Todo inseto é um invertebrado que possui exoesqueleto quitinoso
Protozoário Tripanossomo Cruzii	Microorganismo do reino protista	Protozoário causa _ doença	Tripanossomo é um tipo de _ protozoário	Tripanossomo pode infectar _ insetos	Barbeiro (Mosquito) é um vetor do protozoário Tripanossoma
Doença de Chagas	Doença Infecciosa	Doença Infecciosa transmitida por _ inseto	Doença de Chagas é um tipo de _ Doença infecciosa transmitida por mosquito	Doença de Chagas é _ uma doença infecciosa	Doença de Chagas é transmitida pelo Barbeiro (mosquito) contaminado por Tripanossomo Cruzii

Fonte: Construção da autora

A partir do que foi apresentado do quadro 2 (dois), pode-se concluir:

- as classes de conceitos podem ser subdivididas em subclasses (subgrupos conceituais);
- a Hierarquia conceitual é fundamental para situar o pertencimento do conceito, ou como esse conceito está classificado no domínio de conhecimento;
- a classificação inicia-se na taxonomia;
- o relacionamento diz respeito à formalização entre categorias conceituais;
- a relação funcional revela a relação do conceito com o mundo fenomênico, ou se preferirmos com o contexto (especificando o conceito);
- a instância é a materialização da classe;
- os axiomas por fim determinam o que é verdadeiro sobre o conceito.

Podemos afirmar, a partir da exemplificação demonstrada no quadro 2 (dois), que ontologias são ferramentas de OC e seus estudos contribuem para encontrar caminhos teóricos e metodológicos de modelização de domínios, a partir dos elementos apresentados.

E ainda que seja difícil enumerar possibilidades de relações conceituais em domínios de conhecimento, devido à complexidade dessas relações, que muitas vezes fogem ao domínio que se pretende representar, pois um simples conceito pode se relacionar a um domínio discursivamente e não pertencer originariamente ou primitivamente a este domínio, como no caso do termo barbeiro que pode ter significados semânticos diferentes a depender do contexto que seja empregado.

Além disso, uma afirmação de Jorge (2012) chama a atenção, para a questão da unificação conceitual através de taxonomia.

A modelagem de um domínio para a construção de bases de conhecimento pautada em ontologias transcende a junção de dados e a dificuldade de lidar com os mesmos em estruturas indefinidas, demandando uma taxonomia unificada dos conceitos de uma área de conhecimento. (JORGE, 2012, p. 19).

Para a estruturação de uma ontologia é necessário compor sua espinha dorsal, o que se fará através da taxonomia, compondo uma ordenação hierárquica dos termos/conceitos que

representam um dado domínio de conhecimento. Essa ordenação só é possível quando se estabelecem princípios.

Os princípios para uma taxonomia são definidos com base em teorias e a que mais se adequa a condução do raciocínio lógico de interação dos conceitos e revelam as pesquisas dos domínios é a Teoria da Classificação Facetada (TCF), desenvolvida por Shialy Ramamrita Ranganathan na década de 1930 do século XX, apresentada no capítulo 2 (dois) desta tese.

Esta teoria é dita como dinâmica por possuir princípios norteadores para elaboração de esquemas flexíveis, fundamentado em um método científico demarcado no “plano das ideias”, proposto por Ranganathan na parte E do *Prolegomena to Library Classification* (1967), e que estabelece um esquema de classes, a saber: características; sucessão de características; renques de classes; cadeias de classes e sequência filial, fundamental para categorização de conceitos em uma taxonomia. Além das classes Ranganathan atribuiu regras, ou cânones que segundo Novo, (2007, p.47), “[...] conduzem o classificacionista na organização de conceitos em domínios de conhecimento”.

A teoria dinâmica proposta na TCF possibilita uma metodologia para a construção de uma taxonomia, baseada nas leis estabelecidas na teoria, nos cânones do plano das ideias e por fim nas categorias fundamentais de Personalidade, Matéria, Energia, Espaço e Tempo (PMEST), ou na ampliação dessas categorias a depender da necessidade de cada domínio, como proposto por Novo (2007), quando além das cinco categorias fundamentais, optamos por incluir na taxonomia do domínio da Geoquímica ambiental mais duas categorias.

A ampliação das categorias para a taxonomia foi necessária, para ordenar conceitos interdisciplinares, e melhor trabalhar suas facetas, assim criamos para o domínio as categorias: **Propriedade**, e não tratá-la como uma manifestação das categorias **Personalidade** ou **Energia** e a categoria **Áreas Periféricas** onde incluímos 2 (duas) facetas: **área formal e estrutural** e **áreas geológicas e cosmológicas**.

Os princípios para modelização de domínios como os de uma taxonomia dependem de um método de raciocínio (dedutivo/indutivo) para entendimento do objeto de representação (conceito), das relações entre os conceitos (categorial hierárquica e partitiva). A partir desta primeira fase de modelização estará composta a espinha dorsal de uma ontologia que se

concretizará nas relações associativas dos conceitos e, conseqüentemente, na sua formalização que mapeará um determinado domínio de conhecimento.

A formalização conceitual se tornou necessária no âmbito científico, objetivando evitar problemas relacionados à utilização da linguagem natural e conseqüentemente ambigüidades inerentes às palavras utilizadas nessa linguagem. Além disso, e principalmente para as questões ligadas ao compartilhamento semântico e “leitura” por máquinas, tornando a linguagem legível para computadores. Sobre linguagem natural, discorreremos algumas considerações no capítulo 5 (cinco).

Assim foram estabelecidos por autores seminais e contemporâneos, principalmente na área da Ciência da Computação, como Gruber (1993), Gómez-Pérez; Fernandez; Vicente, (1996) passos fundamentais para tal formalização como: organização das classes conceituais através de uma taxonomia; atributos de classes; instâncias para objetos de conceitos; atributos de instâncias; relações entre classes; constantes de inferência de conhecimento; termos; axiomas de interpretação dos conceitos; regras para atributos.

A partir dos passos para a organização da ontologia se pode por fim através de categorização definir as relações entre os conceitos e formalizar a representação conceitual de um dado domínio. A representação se dá por uma tríade (Categoria antecessora+tipo de relação+Categoria sucessora=relação conceitual), como apresentado nos trabalhos de Sales (2006, 2008), Sales; Sayão (2012) e o de Marcondes (2013, p.19), que afirma, “[...] relações são, portanto o elemento essencial do esquema de representação do conhecimento proposto. São expressas por três elementos: dois *relata* e um *tipo de relação*.”.

Exemplificando a representação por relações categoriais entre elementos de uma sentença linguística, para o domínio meio ambiente:

ex.: <Queimada/ causa incêndio/ na mata atlântica>. Queimada relata uma antiga prática utilizada na agricultura, portanto, um conceito advindo de um **processo**, manifestação da **categoria Energia (de acordo com a PMEST-TCF)**, causa incêndio é o **tipo de relação** (neste caso de **causa-efeito**, com termos (ação e ocorrência) na construção semântica da relação, **segundo a TC**) e por fim Mata Atlântica é um tipo de bioma (conjunto de diferentes

ecossistemas), portanto um conceito advindo de uma **faceta** da Biosfera uma manifestação da **categoria Espaço (de acordo com a PMEST-TCF)**.

Existe nas relações, um compromisso ontológico para se representar o conhecimento. Esse compromisso está pautado na elaboração de declarações com base em níveis da linguagem e da semântica, crucial para o entendimento de esquemas de conceitos. Considera-se na elaboração de esquemas de relações conceituais um compromisso ontológico do discurso, tem por base e fundamento o significado e a intenção do conceito que é declarado e explicitado, ou seja, a intencionalidade discursiva.

A partir da intensão podemos determinar não somente o que o conceito é, mas também o que significa. É possível defini-lo no contexto, descrevendo, explicitando e definindo, ao nível intensional e não da extensão conceitual. No entendimento de pesquisadores de ontologias:

Um conceito intensivo apela aos atributos implicados pela ideia que se deseja conceituar, busca absorver os princípios que permitem reconhecer que tal coisa cai dentro do conceito, ou não. Apela a propriedades e características que a coisa deve ter para pertencer à categoria implicada pelo conceito. Já um conceito extensivo faz o arrolamento de coisas que “caem” dentro da categoria delimitada pelo conceito. Assim, pode-se tentar definir, um ou outro conceito, usando-se de suas propriedades, ou pode-se arrolar o que pertence ao conceito. (COELHO; BAX; MEIRA JUNIOR, 2012, p.6).

A extensão do conceito é algo a ser pensado em um modelo de análise conceitual, ou o que está subtendido, o que é semelhante, mas não igual. E então voltamos à complexidade do ato de modelização de domínios, pois lidamos com aspectos intensionais e extensionais de um conceito, ou seja, significação e referência, ou ainda de como categorizar conceitos.

Quando falamos em categorizar conceitos nos referimos principalmente, aos relacionamentos entre conceitos, para categorizá-los necessitamos levar em conta seus atributos (propriedades ou qualidades). Não nos referimos aqui, a um sistema universal, mas de uma categorização possível, que refletirá a estrutura de uma linguagem e basear-se-á em uma estrutura cognitiva ou conforme o termo/conceito se apresenta no domínio de conhecimento.

Formalizar conceitos pressupõe um posicionamento filosófico sobre o estudo da essência das coisas. Para o sujeito o que vale é o que é dito sobre as coisas. A linguagem é fundamental, pois torna o dizer realidade a partir da contextualidade, do que é real sob o ponto de vista do domínio de conhecimento. Em outro lado necessitamos de um modelo de análise que leve em consideração os aspectos semânticos da linguagem, sendo assim, a representação é semântica e contextual.

Esta pesquisa se aproxima dos estudos de ontologias de domínios e de representação, e pretende demonstrar que o modelo proposto na tese é importante para tais ferramentas de OC. Passamos a apresentar o modelo de análise.

4 CAMINHOS DA PESQUISA: DA TEORIA AO MODAC

Pensar essa questão é comprometer-me com um vasto caminho de pensamento. Para tanto, precisamos de poucos passos. São passos que, na verdade, não nos levam adiante e sim de volta para onde já somos e estamos. Os passos não são uma sequência de um após o outro, a não ser muito aparentemente. Os passos são bem mais a articulação e a concentração em torno do mesmo, um jogo para trás desse mesmo. A impressão de desvio deve-se ao fato de a vizinhança determinar-se ao se retrair no en-caminhamento. [...] (HEIDEGGER, 2011, p. 164).

Iniciamos este capítulo com uma epígrafe de Heidegger quando em sua obra “a caminho da linguagem” aborda a questão da palavra e trás para o texto o caminho dos modos de dizer, a partir da proximidade da vizinhança, e toma como exemplo poesia e pensamento. Usamos a epígrafe como uma aproximação do que queremos e realizamos nos passos metodológicos para a organização e representação, ou seja, uma aproximação entre os termos e sua vizinhança que declaram um conceito em tempo e espaço. Em um tempo que muda a depender da evolução do contexto em que a palavra se insere.

Além disso, quando medimos a aproximação de palavras em uma rede, de certa forma estamos medindo a vizinhança em um “en-caminhamento” como dito por Heidegger. Estamos trabalhando com espaço e tempo não medido como parâmetro de proximidade dos termos vizinhos, mas sim do “en-contro” de relacionamentos, abertos para outros relacionamentos. A partir dessa ideia passamos então a descrever a metodologia.

A metodologia desta pesquisa está direcionada pela análise do discurso representativo do domínio e análise cognitiva, com o intuito de encontrar e reconhecer o todo discursivo através das “leituras” das redes conceituais. Nessa leitura é necessário deslocar o olhar do analista cognitivo para o olhar “ponto de vista” do leitor/usuário do domínio, para em seguida, seguir o texto pelas macroposições semânticas a partir dos conceitos, buscando traços descritivos que levaram aos argumentos utilizados no texto.

A utilização da análise do discurso e da análise cognitiva tem por objetivo facilitar o trabalho do modelizador de domínios. Através do uso dessas metodologias pode-se elaborar um modelo dinâmico de análise conceitual que permita categorização de conceitos mais coerentes com a realidade do domínio de conhecimento.

O método qualitativo desenvolvido foi alicerçado pela Teoria do conceito, amparado pelos princípios da Análise do Discurso Representativo e da Análise Cognitiva, no processo de “leitura conteudista” do ontologista, para a construção de SOC que tem como alicerce as categorizações conceituais. Esse alicerce permite o refinamento dos significados e suas relações propositivas, fundamentadas a partir dos níveis lógicos; epistemológicos; ontológicos e conceituais.

O método quantitativo esteve presente na análise das redes, pois quantificamos os conteúdos semânticos a partir da incidência e fidelidade com que um par de palavras apareceu na amostra textual com uma frequência de 0,1 por cento do texto. Abaixo desse valor as palavras foram eliminadas da análise por não evidenciarem um peso significativo no discurso do autor do texto e justificasse a análise conceitual e conseqüentemente tornar o conceito um termo “candidato”, aqui entendido como termo sugerido pelo analista cognitivo, a compor a ontologia do domínio e a ser avaliado pela comunidade de especialistas do domínio.

Quanto à natureza, classificamos a pesquisa como pesquisa aplicada, pois objetiva gerar resultado de aplicação prática. Na aplicação incluímos um método híbrido, dedutivo/indutivo de análise conceitual, utilizado na construção de SOC (taxonomias e ontologias de domínios) em um dado “recorte”, que possibilitou traçar o modelo dinâmico de análise.

Quanto à abordagem do problema classificamos o trabalho de tese, como pesquisa qualitativa, onde o conhecimento apreendido é aplicado a partir das bases epistemológicas, teóricas e metodológicas relacionadas à organização do conhecimento, não podendo ser mensurado quantitativamente.

Para alcançarmos os objetivos traçados na tese, norteadores desta pesquisa, apresentados no capítulo introdutório, iniciamos por apresentar teorias que fundamentaram o nosso trabalho, daqui por diante passaremos a construção do modelo dinâmico de organização do conhecimento, mas antes, precisamos apresentar a área de conhecimento e os domínios representados nas redes conceituais.

4.1 MATERIAL E MÉTODO: O DOMÍNIO A SER REPRESENTADO

Escolhemos como corpus a serem trabalhados, para análise das redes semânticas e aplicação do modelo de análise proposto, teses e dissertações, selecionadas no âmbito dos Programas de Pós-graduação do Instituto de Geociências da Universidade Federal da Bahia (UFBA), por entendermos que materiais desta natureza espelham o desenvolvimento científico nas áreas das Ciências da Terra que abarcam domínios inter e multidisciplinar.

Além disso, a autora desta tese trabalhou no Instituto de Geociências, nesta Instituição de Ensino Superior (IES), durante 6 (seis) anos como Bibliotecária e por mais 5 (cinco) anos no domínio da Geoquímica ambiental em outra Instituição Federal de Ensino (IFE), totalizando 11 (onze) anos tratando tecnicamente (catalogando, classificando, indexando) os itens documentais da área das Geociências e de domínios correlatos. Tendo, portanto, conhecimento da terminologia usada por pesquisadores, o que facilita o entendimento dos conceitos e de suas relações com outros conceitos.

Escolhemos a área das Geociências e suas subáreas e especialidades (geologia, geografia e geoquímica) respectivamente. Como a base de ontologias é a taxonomia, dita como a “espinha dorsal” dessa ferramenta, e o conhecimento das áreas e domínios já haviam sido estudados e comprovadamente caracterizados como inter e multidisciplinares, quando realizei o mestrado em CI, preferimos dar continuidade à pesquisa por acreditarmos que o tratamento de conceitos, oriundos desses domínios, possui uma complexidade ampliada no momento da sua categorização. E ainda quando das suas relações ontológicas, pois carregam definições que superam a visão tradicional do conhecimento compartimentado, pois são frutos de articulações e integração de conceitos que enfocam epistemologias, metodologias, procedimentos e dados, a depender da pesquisa realizada.

Nesta pesquisa, no entanto, ampliamos os domínios pesquisados no mestrado, com a intenção de observarmos os discursos das comunidades de práticas acadêmico-científicas e revelarmos como os conceitos, que nascem em domínios interdisciplinares e multidisciplinares se articulam, e podem ser representados por relações ônticas (inerente ao objeto e suas características), ou ontológicas (natureza, realidade e existência) dos objetos. A

organização do conhecimento desses domínios deve respeitar o diálogo entre as diversas disciplinas que se articulam, nas convergências e divergências temáticas.

Como primeiro passo metodológico procuramos entender os domínios a serem representados, através de leituras efetuadas em livros de autores consagrados na área como: Teixeira et al. (2009), Krauskopf (1972), Goldschmidt (1958). Além de pesquisas em sites como o do Instituto de Geociências da UFBA e de seus Programas de pesquisa, para então selecionarmos o corpus a ser trabalhado.

Não realizamos um estudo aprofundado das áreas que abarcam as Ciências da Terra, por não ser objeto de estudo desta tese, mas sim um breve panorama, baseado em leituras, para situar os campos de desenvolvimento das pesquisas do corpus selecionado, objeto de investigação e aplicação do método de análise e modelo dinâmico para organização e representação conceitual.

As Ciências da Terra ou Geociências abrangem o conjunto das ciências que estudam o planeta Terra e seus sistemas complexos, que abrangem: terra sólida, hidrosfera, atmosfera e biosfera. Como as questões ambientais e estudos relacionados às soluções para o meio ambiente, a partir da década de 1960 do século XX foram ampliados, as pesquisas neste domínio evoluíram e receberam aportes das áreas e especialidades da geografia, geofísica, geoquímica ambiental, para dar soluções aos problemas de: contaminação, degradação, recuperação de ambientes continentais, química de atmosfera, gestão ambiental, mudanças globais ambientais, dentre outros.

Entre o século XVII e XVIII, principalmente, com as questões que emergiram no iluminismo, o ser humano passou a perceber que as leis sobrenaturais não explicavam mais os fenômenos ocorridos na natureza, para entendê-los e agir sobre o meio ambiente era necessário observar, pesquisar e investigar. Além disso, com a revolução industrial, iniciada nos meados do século XVIII, ocasionando demanda por matéria prima e recursos de energia oriundos da terra, impulsionando estudos sobre ela, nasceu então à geologia. Os princípios desta nova ciência foram enunciados, segundo Teixeira (2009, p. 308), pelo dinamarquês Nils Stensen.

O nascimento da geoquímica data do século XIX e seu desenvolvimento ao longo do século XX deveu-se como apresentado por Novo (2007, p. 74), “a tendência em tornar a geologia mais quantitativa, incluiu-se a química, física e a matemática [...]”.

Nesta mesma tendência, surgiu a Geofísica, que estuda a estrutura, a composição, as propriedades físicas e os processos dinâmicos da Terra sob a ótica da física, especialmente por estudos da sísmica de reflexão e refração, gravidade, magnetismo, eletricidade, eletromagnetismo e métodos radioativos.

No mesmo século a Geografia, ciência que estuda o planeta terra e sua superfície, foi instituída como disciplina acadêmica. A área se subdivide em Geografia física, Geografia humana e Geografia econômica, abarcando estudos multitemáticos de base territorial e de conhecimento da natureza: física, urbana, rural, social, política, demarcando a dinâmica sócio espacial.

4.2 CORPUS SELECIONADO

Considerando o panorama traçado, o ponto de partida para a delimitação do corpus trabalhado foi identificar teses e dissertações disponibilizadas em formato de arquivo *Portable Document Format* (PDF) no site do IGEO-UFBA, nas áreas e especialidades determinadas para o mapeamento conceitual, posterior análise e aplicação do modelo proposto.

Escolhemos arquivos no formato em PDF porque como iríamos trabalhar com redes semânticas e necessitávamos transformar os textos disponibilizados neste formato em formato TXT, arquivo de texto estruturado em uma sequência de linhas. Esse tipo de formato possibilita o trabalho com programas para geração de redes semânticas. Apresentamos mais adiante, neste capítulo, os procedimentos de construção dessas redes.

Foi necessário um recorte para formação do corpus, pois entendemos que para uma amostra que demonstrasse a abordagem das questões ambientais no âmbito do território baiano e brasileiro na atualidade, pelas subáreas (Geologia, Geografia) e a especialidade (Geoquímica), as dissertações e teses defendidas nos últimos cinco anos revelariam conceitos dinamicamente afetados pelas questões do meio-ambiente como também técnicas e métodos atualizados a partir da interferência e avanço das tecnologias.

Baseados nesta premissa, selecionamos como amostra cinco trabalhos, entre teses e dissertações, de cada domínio, para que pudéssemos efetuar a análise. Esses trabalhos estão referenciados no (Apêndice A). Decidimos por trabalhar primeiro cada domínio separadamente, objetivando conhecer conceitos e suas relações internas. A partir de tais relações observamos se um mesmo conceito utilizado em dois ou mais domínios das subáreas ou especialidades, mantinham a mesma definição ou se sofriam alterações a depender do discurso e das práticas dos diferentes domínios.

Poderíamos ter escolhido trabalhar apenas com as palavras-chave estabelecidas pelos autores dos trabalhos (dissertações e teses) selecionadas, mas isso já havia sido feito no trabalho de mestrado, e embora saibamos que poderíamos trabalhar um modelo de análise a partir dos termos retirados dessas palavras, decidimos por analisar um corpus a partir das redes semânticas estabelecidas com a totalidade dos textos das amostras selecionadas, esperando que a ampliação dos termos facilitasse a análise dos domínios.

Em um segundo momento, com a primeira análise concluída, observamos como se comportava na rede conceitual os termos originados das palavras-chaves, e suas articulações, relações com os conceitos/termos que se associavam, a fim de constatar se as palavras estabelecidas pelos autores dos trabalhos revelavam o conteúdo do texto/contexto e respondiam a expectativa do usuário de encontrar através dessas palavras os fundamentos e resultados da pesquisa realizada pelos autores.

A composição do corpus para análise foi efetuada a partir da construção da Rede Semântica do Domínio e suas sub-redes conceituais. A construção foi realizada a partir de um conjunto de programas para redes semânticas, que apresentamos na sequência desta seção.

4.2.1 Corpus para análise: a construção das redes

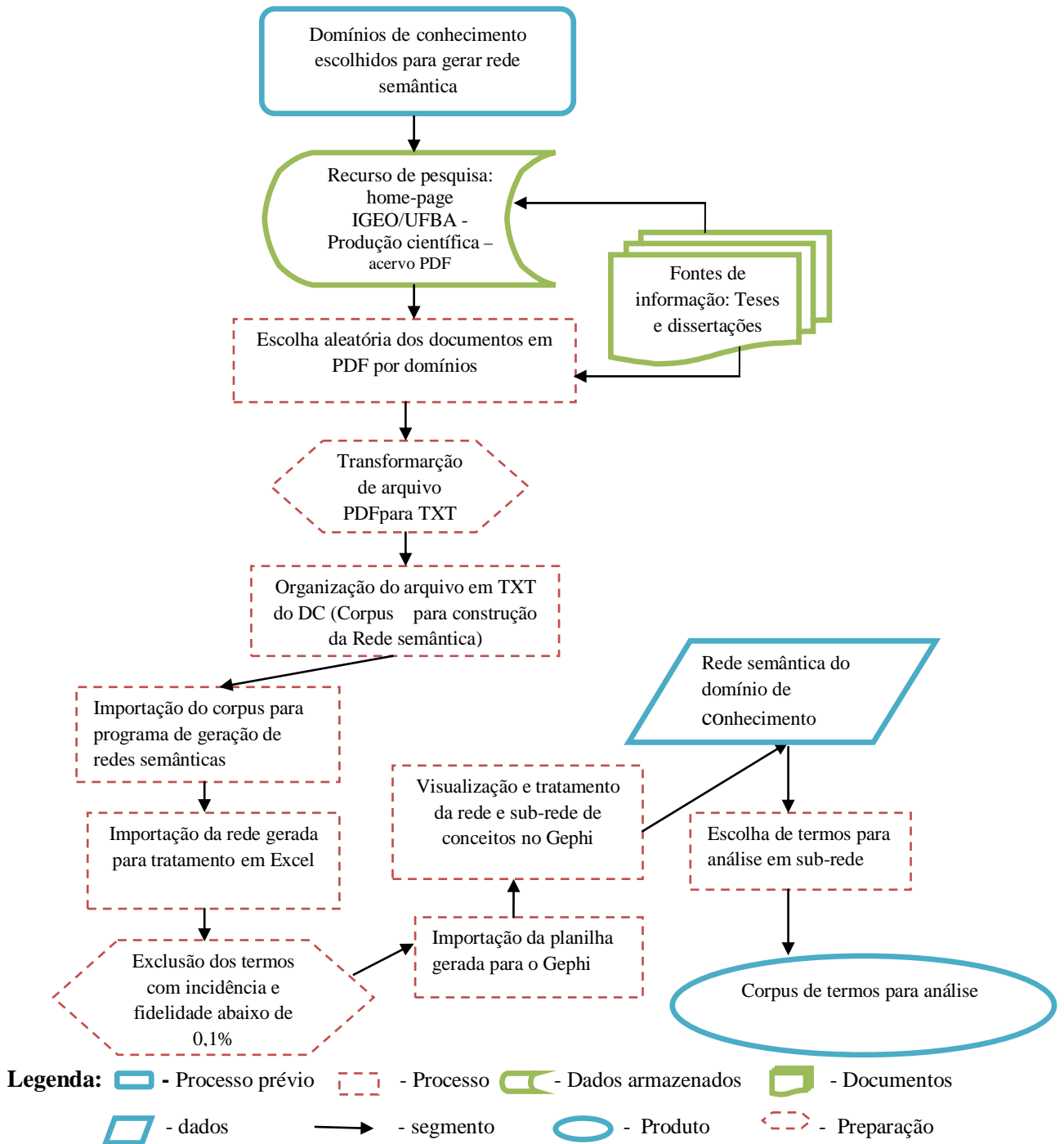
O processo para alcançarmos a Rede Semântica do Domínio (RSD) e consequentemente sub-redes de termos que se interligam e revelam conceitos apreendidos pelos domínios de conhecimento, necessitou do uso de um conjunto de programas, que possibilitaram primeiro, a construção da RSD.

Na sequência, após análise das relações entre os termos dessa rede, construímos sub-redes desses termos. A escolha dos termos foi baseada na incidência e fidelidade das relações

entre eles, observando a fidelidade também ao contexto do domínio, ou quão prudente seria a sua análise conceitual e a sua inclusão em um SOC.

Para a construção da RSD e sub-redes de termos procedimentos foram tomados e se configuraram em etapas dessa construção que evidenciamos passo a passo no diagrama a seguir, figura 6 (seis) e que se apresenta por fluxos de processos.

Figura 6 - Fluxo da construção das redes/sub-redes de termos

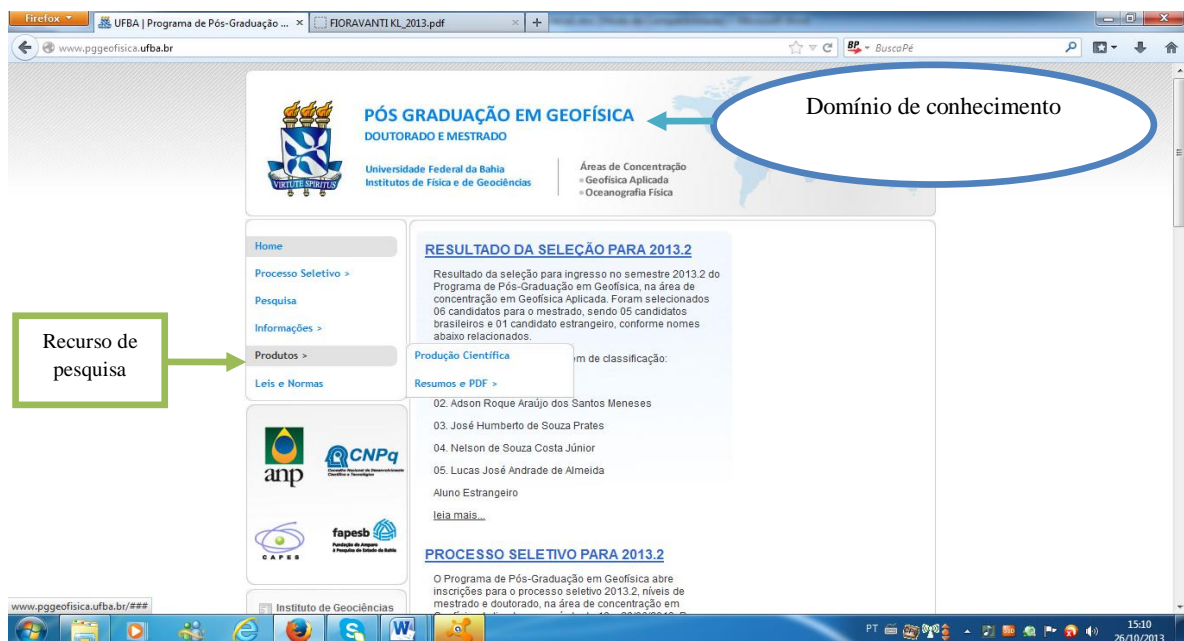


Fonte: construção da autora

O primeiro passo da metodologia para a construção das redes foi à seleção dos domínios e áreas de conhecimento. Escolhemos para apresentarmos nesta seção de demonstração da construção do corpus, duas amostras, a primeira amostra contendo a produção da área de conhecimento (Geografia) e a segunda amostra a produção de um domínio (Geoquímica). Neste primeiro momento, podemos perceber se em uma especialidade ocorria alguma mudança em relação à apropriação de conceitos, ou se esses conceitos se mantinham definidos pela perspectiva da grande área.

Em seguida, o segundo passo foi eleger o corpus, como já citado anteriormente, optamos por investigar o discurso e apropriação dos conceitos em teses e dissertações. Como recurso de pesquisa, para acesso as teses e dissertações dos domínios/áreas utilizamos a “home Page” do IGEO/UFBA, onde se encontram disponibilizados o acervo da produção científica dos programas de pós-graduação em formato PDF, como pode ser visualizado na Figura 7.

Figura 7 – Recurso de pesquisa – Acervo: Produção científica



Fonte: <http://www.twiki.ufba.br/twiki/bin/view/IGeo/WebHome>

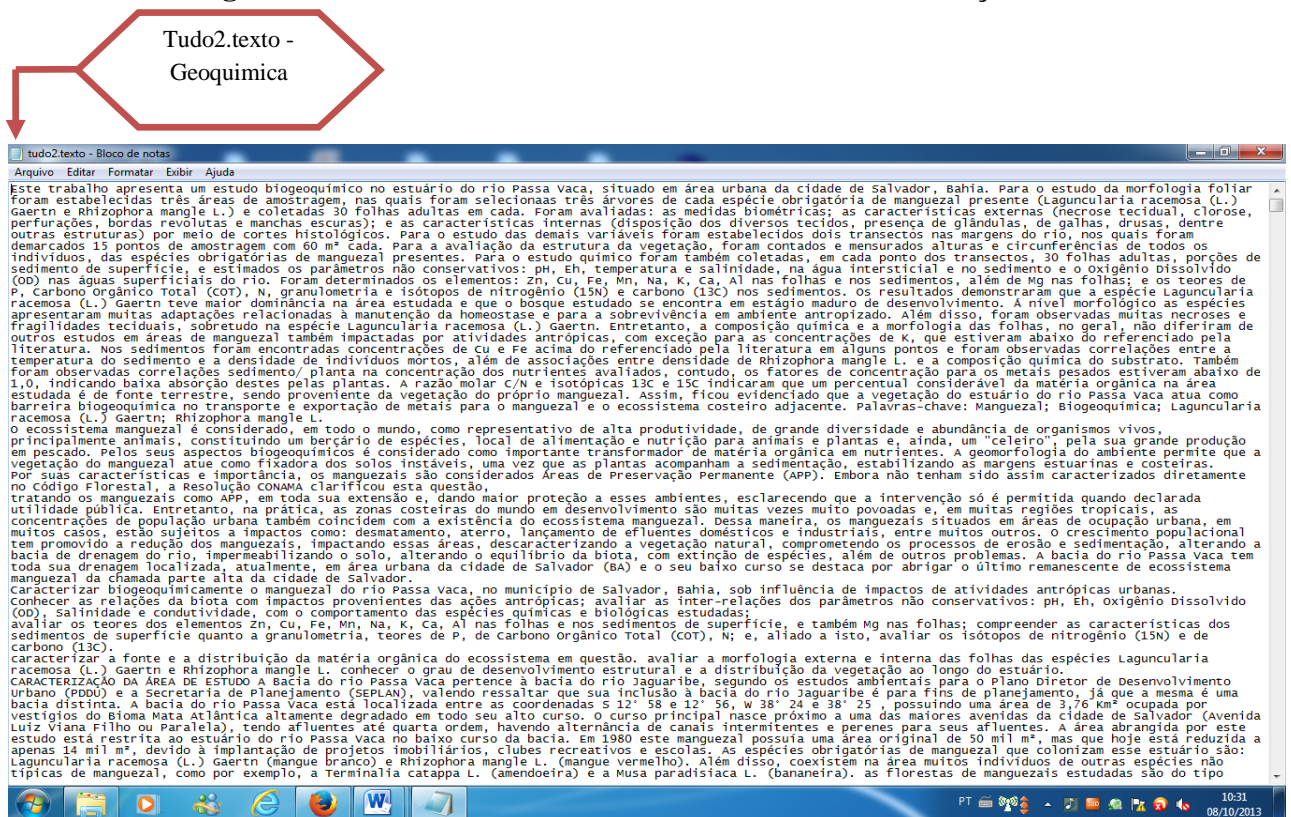
Após a escolha aleatória de cada trabalho, totalizando cinco trabalhos entre dissertações e teses, para cada domínio, disponibilizado em formato PDF nas páginas dos

programas de pós-graduação do IGEO, um arquivo foi aberto para cada área/especialidade. Entendemos que um total de cinco teses por domínio atenderia o propósito da análise, pois o que valida à amostra não é a quantidade de textos, mas o modelo de análise empregado.

Quando afirmamos que a escolha foi aleatória é porque definimos apenas que escolheríamos os trabalhos disponibilizados no repositório de teses e dissertações dos últimos 5 (cinco) anos. Não nos preocupamos, por exemplo, em analisar apenas trabalhos que continham resumos, ou palavras-chave. Ou ainda trabalhos que continham um determinado número de páginas ou palavras. Esses tipos de parâmetros não foram previstos para análise.

Em seguida, cada trabalho, em PDF foi transformado em arquivos de texto com a extensão TXT, (Figura 8). Arquivos dessa natureza facilitam a leitura de alguns programas, como por exemplo, os utilizados para gerar redes semânticas, no nosso caso um conjunto de programas, que geram redes, desenvolvido por Caldeira (2005), e adaptados por Teixeira (2007) e Aguiar (2009).

Figura 8 - Texto transformado em TXT das teses/dissertações



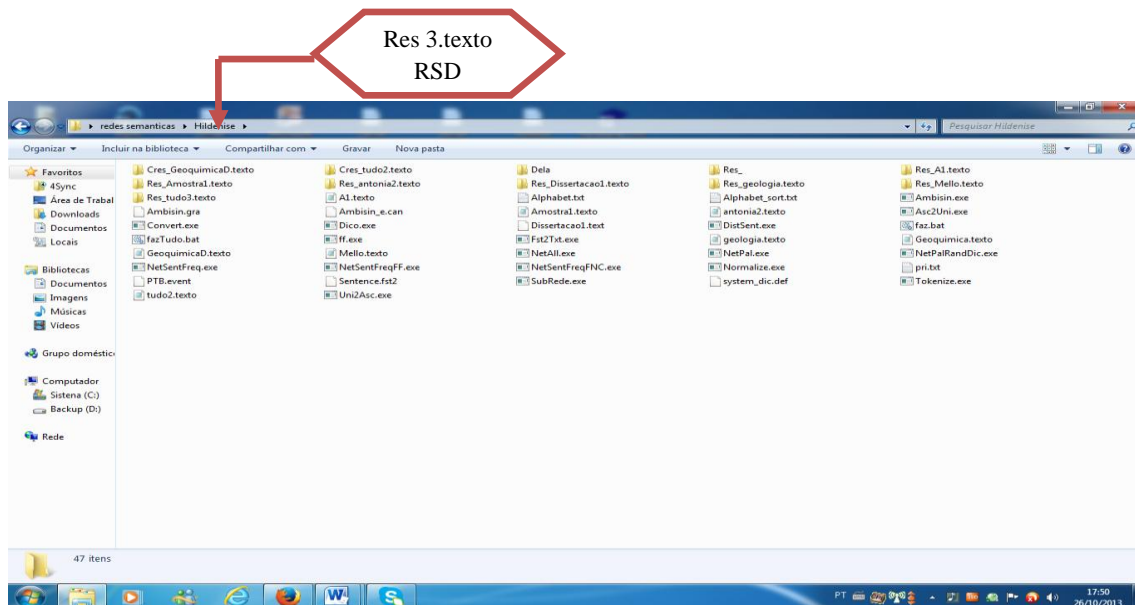
Fonte: construção da autora

É importante evidenciar que os arquivos de texto das teses e dissertações são compostos por 5 (cinco) trabalhos de cada área/domínio, como apresentado na Figura 8, e que compõem o corpus de análise.

Quando transformamos arquivos no formato PDF para o formato TXT, um trabalho manual de limpeza do texto deve ser realizado, eliminando: espaços entre parágrafos, símbolos, tabelas, gráficos e sinais que não serão identificados e reconhecidos pelo conjunto de programas de redes semânticas, além disso, tal processo é necessário para que não ocorra interferência na linearidade textual e conseqüentemente nas conexões das redes. Após este trabalho, o corpus foi importado para o conjunto de programas para redes semânticas. Visualização desta etapa na Figura 9.

Em relação ao corpus da Geofísica, convém esclarecer, quando da transformação do formato PDF para o formato TXT tivemos um problema de ordem técnica. Os arquivos disponibilizados na página quando convertidos para TXT ocorriam erros que nos impediram de realizar uma limpeza eficaz para o reconhecimento do texto pelo conjunto de programas para redes semânticas que utilizamos. Seria necessária a compra de um programa para realizar esse trabalho. Como não havíamos previsto despesas no nosso projeto inicial resolvemos por trabalhar um corpus com três domínios de conhecimento, por entendermos que tal decisão não iria interferir no resultado final e objetivo traçado.

Figura 9 – Conjunto de programas para redes semânticas



Fonte: construção da autora

Como pode ser observado nas Figuras 8 e 9, o arquivo gerado compreende o conteúdo textual das teses/dissertações (Res_tudo) extensão (texto). Após a geração da rede no conjunto de programas, uma nova etapa será necessária no processo de construção da rede, o trabalho de filtragem de termos relevantes para análise das conexões entre os termos. Este processo é realizado numa planilha de Excel, como apresentado na Figura 10.

Figura 10 – Filtragem dos termos e tratamento do corpus no Excel

Tratamento da rede

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	source	target	FF	Type	Corpus																
2	passa	vaca	0.764788	undirecte	Gq																
3	tensão	superficia	0.71935	undirecte	Gl																
4	glicerina	bruto	0.709791	undirecte	Gl																
5	vitória	conquista	0.707462	undirecte	Gg																
6	óleo	diesel	0.546291	undirecte	Gl																
7	rio	vaca	0.451078	undirecte	Gq																
8	rio	passa	0.443855	undirecte	Gq																
9	barra	choça	0.425115	undirecte	Gg																
10	orgânico	matéria	0.32016	undirecte	Gq																
11	antraceno	fenantren	0.310481	undirecte	Gl																
12	estuário	vaca	0.235966	undirecte	Gq																
13	estuário	rio	0.231635	undirecte	Gq																
14	estuário	passa	0.230657	undirecte	Gq																
15	zn	cu	0.21867	undirecte	Gq																
16	docosano	fenantren	0.204526	undirecte	Gl																
17	agitador	orbital	0.184211	undirecte	Gl																
18	fonte	carbono	0.170088	undirecte	Gl																
19	docosano	antraceno	0.170088	undirecte	Gl																
20	cu	fe	0.15572	undirecte	Gq																
21	ser	manguezá	0.152617	undirecte	Gq																
22	área	manguezá	0.151975	undirecte	Gq																
23	conquista	planalto	0.131901	undirecte	Gg																
24	realizar	teste	0.131205	undirecte	Gl																
25	ser	não	0.127563	undirecte	Gg																

Fonte: construção da autora

O principal objetivo desta etapa é o tratamento dos principais elementos que compõem a rede como: determinação dos termos de origem e destino; eliminação de termos com peso 0 (zero), ou seja, que não são relevantes à rede, e a formação de um arquivo com a extensão CSV (separado por vírgula). Assim, a planilha após importada para o Gephi poderá servir ao objetivo deste trabalho de tese, visualização das redes semânticas do domínio de conhecimento para que possamos analisar as relações entre os termos e consequentemente a importância dessas relações no discurso do texto e no domínio de conhecimento. Essa análise facilitará a categorização conceitual, as relações conceituais e consequentemente a construção do SOC.

A RSD gerada totalizou 3346 (três mil trezentos e quarenta e seis) nós e 26990 (vinte e seis mil novecentos e noventa) arestas. Para a visualização dos termos origem e destino de cada domínio, bem como os termos que pertenciam a mais de um domínio efetuamos um trabalho de marcação dos termos por corpus, utilizamos siglas para cada domínio na planilha do Excel (figura 10).

Em seguida, a planilha tratada no Excel seguindo os princípios acima descritos foi importada para o Gephi, versão 0.8.2 beta. Disponível em: <https://gephi.org/>. Conforme descrito no site pode ser definido como uma plataforma interativa de visualização de redes, servindo, portanto, para exploração dos nós e arestas das palavras e conseqüentemente dos conceitos gerados através de filtros dinâmicos.

Após a importação da planilha para o laboratório de dados do Gephi um novo projeto foi criado na plataforma gerando uma tabela de dados de Nós e Arestas da RSD, que poderá ser visualizada e tratada para facilitar a “leitura” e análise discursiva dos textos de origem das redes, como também a análise dos termos e suas relações, reveladores de possíveis conceitos que mapeiam um domínio de conhecimento.

Efetuamos um trabalho de marcação dos termos pertencentes a cada domínio por cores, o que pode ser observado na figura 11 (onze), assim como termos que pertenciam a mais de um domínio de conhecimento, no laboratório de dados da seguinte forma:

- marcamos na coluna corpus a sigla do domínio a que o termo pertence, ou seja: Gg (Geografia); Gl (Geologia); Gq (Geoquímica) e editamos cores para cada domínio;
- marcamos os termos de intercessão, ou seja, que pertencem a dois ou mais domínios por cores distintas das cores determinadas e editadas para o termo origem.

Figura 11 – Trabalho de marcação dos termos por cores no laboratório de dados

The figure consists of two screenshots of the Gephi 0.8.2 software interface, illustrating the process of marking terms by color in a data table.

Top Screenshot: Data Table Configuration

The interface shows the 'Tabela de dados' (Data Table) in 'Configuração' (Configuration) mode. The table has the following columns: Origem, Destino, Tipo, Id, Label, Weight, FF, and Corpus. The 'Filtro' (Filter) is set to 'Origem'. A red arrow points to the 'Origem' filter dropdown menu.

Origem	Destino	Tipo	Id	Label	Weight	FF	Corpus
processo	produto	Não dirigido	39656			1	0,002Gg
processo	clar	Não dirigido	38304			1	0,003Gg
processo	coisa	Não dirigido	39441			1	0,002Gg
processo	mudança	Não dirigido	49997			1	0,003Gg
processo	comunicação	Não dirigido	42427			1	0,001Gg
processo	relevância	Não dirigido	42963			1	0,001Gg
processo	formação	Não dirigido	38919			1	0,002Gg
processo	situação	Não dirigido	55656			1	0,001Gg
processo	mercadoria	Não dirigido	39709			1	0,002Gg
processo	erosivo	Não dirigido	39532			1	0,002Gg
processo	expressar	Não dirigido	42899			1	0,001Gg
processo	influenciar	Não dirigido	51523			1	0,002Gg
processo	participar	Não dirigido	42649			1	0,001Gg
processo	raiz	Não dirigido	48527			1	0,003Gg
petróleo	carbono	Não dirigido	35746			1	0,002Gg
petróleo	oleoso	Não dirigido	34299			1	0,003Gg
petróleo	microbiano	Não dirigido	36764			1	0,001Gg
petróleo	carvalho	Não dirigido	34303			1	0,003Gg
petróleo	et	Não dirigido	32040			1	0,008Gg
petróleo	bruscamente	Não dirigido	33369			1	0,005Gg
petróleo	emitir	Não dirigido	33375			1	0,005Gg
petróleo	principalmente	Não dirigido	34791			1	0,003Gg
petróleo	ramo	Não dirigido	33713			1	0,004Gg
petróleo	nitrogenio	Não dirigido	33748			1	0,004Gg

Bottom Screenshot: Node Property Editing

The interface shows the 'Tabela de dados' in 'Editar' (Edit) mode. The 'Propriedades' (Properties) panel is open, showing the 'Cor' (Color) property set to '[255,0,0]' (red). A red arrow points to the 'Cor' property value. Another red arrow points to the 'Corpus' property value, which is 'Gg'.

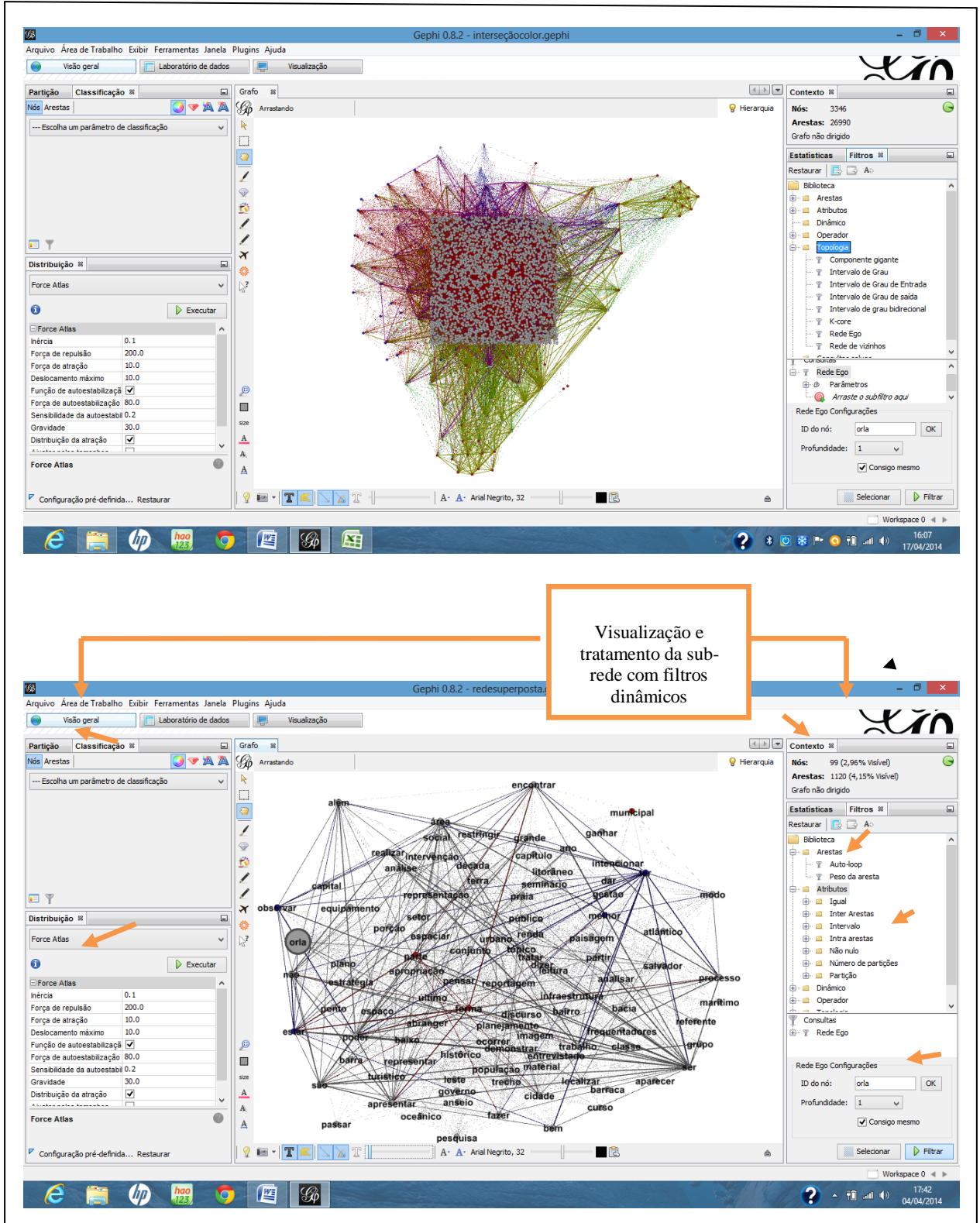
Propriedade	Valor
Cor	[255,0,0]
Id	51523
Label	<valor nulo>
Weight	1,0
FF	0,0021518
Corpus	Gg

Fonte: Construção da autora

No espaço do Gephi denominado de visão geral é possível o trabalho com os nós e arestas das redes. Existe a possibilidade de filtrar o peso das arestas, como também melhor distribuir a distância entre os nós, o que facilita a visualização da rede para análise das relações entre os termos.

O trabalho de visualização e possibilidades de tratamento pode ser observado na Figura 12.

Figura 12– Visualização e tratamento da rede/sub-rede no Gephi



Fonte: construção da autora

Após as etapas da construção da rede, um novo processo deverá ser iniciado para alcançar um corpus de termos que permita aplicar o modelo dinâmico de análise conceitual proposto neste trabalho, para auxiliar o modelizador de domínios na construção de SOC. Esse novo processo inicia um novo passo, constituído por mais algumas etapas como: escolha do termo a partir da análise da RSD; filtragem do termo escolhido para análise através da rede “ego”. A rede ego, visualizada na figura 12 (doze) é um filtro no contexto da topologia da rede, que gera uma sub-rede.

O termo escolhido para exemplificarmos os próximos passos, ou etapas de construção do corpus de termos para aplicação do modelo dinâmico de análise, foi o termo **orla** pertencente à amostra da RSD da Geografia. Esta escolha foi determinada pela percepção de que as relações entre os nós que estão interligados ao termo melhor evidenciariam a percepção da incidência-fidelidade que os pares de termos/palavras evocam no contexto discursivo no domínio.

Convém esclarecer que uma lista dos pares de palavras com seus respectivos valores de incidência-fidelidade são filtrados em uma planilha eletrônica (Excel), antes de sua exportação para o Gephi, de modo a eliminar os pares de palavras com valores de incidência-fidelidade com baixa importância na rede.

Escolhemos um valor limite de 0,001, o que significa uma coocorrência de pares abaixo de 0.1% no texto. A escolha desse valor se deu por um processo empírico, onde percebemos a importância dos pares de palavras evocados no discurso. Dessa forma, avaliamos o quão significativo seria avaliar as palavras como termos tradutores das pesquisas realizadas pelos produtores do conhecimento, pois seriam fiéis ao conteúdo transmitido pelo sujeito.

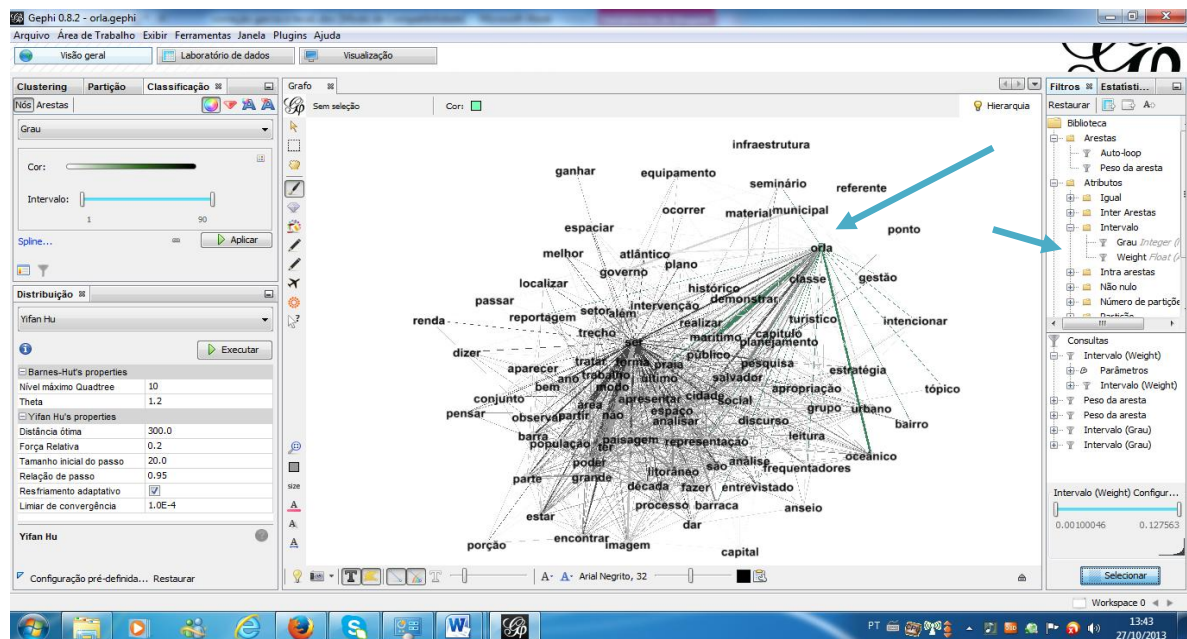
A próxima ação a ser efetuada é a importação da planilha eletrônica para o Gephi, onde o processo de visualização e tratamento da sub-rede deve ser concluído. O valor limite escolhido, de 0,001 pode ser visualizado na figura 11 (onze), na coluna do laboratório de dados.

Após a importação da rede pode-se por fim concluir o tratamento das sub-redes dos termos e visualizar as relações entre os termos, o que facilita a análise conceitual a partir da

percepção do modelizador de domínios sobre o objeto a ser organizado e representado em um SOC, a depender do comportamento, ou das interações do termo na rede conclui-se quão intenso esse termo está para o discurso e conseqüentemente para o domínio.

Na visão geral da sub-rede (figura 13) pode-se observar o intervalo de graus dos nós e o peso mínimo estabelecido para as arestas. Alguns parâmetros foram definidos para a análise da sub-rede: incidência fidelidade e grau do vértice (nó).

Figura 13 – Visão geral da sub-rede – termo orla



Fonte: construção da autora

O intervalo de atributos da flutuação do peso das arestas nos permitiu trabalhar na análise conceitual, apenas com pares de termos que possuam fidelidade no discurso, ou seja, permitindo a percepção de que um termo, bem definido no texto, seja revelador do domínio. Esse termo é denominado de candidato a compor o sistema de representação do conhecimento. Após definido passará a ser categorizado e sistematizado.

A importância da determinação do grau de um vértice, para a análise da rede está na possibilidade de melhor analisar a conexão de um vértice e seus vizinhos, ou seja, de como as ligações mais próximas entre os termos interferem na definição de um termo/conceito (figura 14). Além disso, podem-se avaliar quantas ocorrências do termo estão impressas no texto.

Os termos candidatos são escolhidos pelo analista cognitivo a partir da incidência e fidelidade, ou, a frequência com que um termo é utilizado e consequentemente a fidelidade ao discurso.

Figura 14 – Grau de nós da sub-rede – termo orla

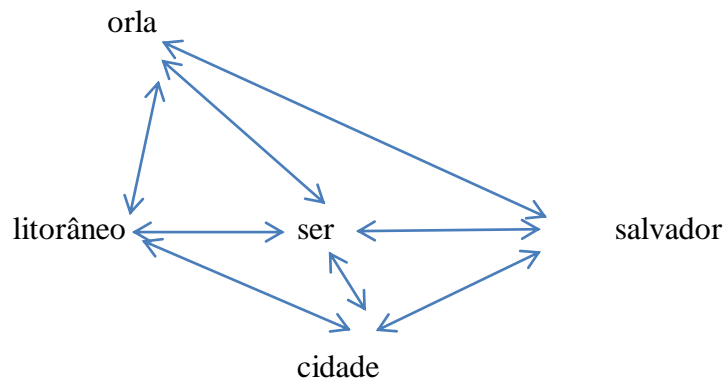
Nodes	Id	Label	Grau
● bairro	bairro	bairro	3
● ano	ano	ano	28
● trecho	trecho	trecho	28
● frequentadores	frequentadores	frequentadores	9
● fazer	fazer	fazer	30
● histórico	histórico	histórico	8
● barraca	barraca	barraca	8
● ser	ser	ser	65
● tratar	tratar	tratar	15
● não	não	não	41
● ocorrer	ocorrer	ocorrer	6
● último	último	último	23
● analisar	analisar	analisar	31
● apropriação	apropriação	apropriação	13
● ter	ter	ter	35
● espaciar	espaciar	espaciar	6
● realizar	realizar	realizar	17
● renda	renda	renda	3
● salvador	salvador	salvador	45
● gestão	gestão	gestão	4

Fonte: construção da autora

A partir das Figuras 13 e 14, analisando sob a ótica da incidência-fidelidade e do grau das arestas podemos afirmar a partir dos pares de palavras que:

- Salvador é uma cidade. Orla é um [espaço] litorâneo. Logo: Salvador é uma cidade litorânea, Figura 15.

Figura 15 – Rede de palavras do texto/discurso



Fonte: construção da autora

Orla - em relação a sua denotação, ou o que o termo revela ou se revela na amostra (Geografia):

- Par de palavras: ser-orla (origem-destino). Frequência: 54 ocorrências e percentual de fidelidade 0,05%
- Par de palavras: Salvador-orla origem-destino. Frequência: 43 ocorrências e percentual de fidelidade 0,17%
- Par de palavras: cidade-orla (origem-destino). Frequência: 40 ocorrências e percentual de fidelidade 0,12%

Nos pares de palavras revelados na rede percebe-se, a “incidência-fidelidade” que existe entre as palavras evocadas no discurso impresso no texto. Essa “incidência-fidelidade” ou “força de interação” pode ser medida pelo peso das arestas. Consequentemente a partir desta análise pode-se afirmar quando uma rede de palavras é classificada como ponderada, ou seja, existe um valor associado às arestas que possibilita a representação de proximidade entre essas arestas, o que propomos neste trabalho é a utilização do peso para filtragem e posterior análise do termo como objeto revelador do contexto/domínio.

Após o trabalho com filtros, das arestas e seus pesos e do intervalo de atributos, além do trabalho estatístico utilizando o grau médio da rede é possível melhor visualizar as sub-redes de termos previamente escolhidos. Observando termos e vizinhos percebem-se as relações entre os termos, o que facilita a análise discursiva sobre os conceitos utilizados pelos autores. Esse trabalho permite uma melhor identificação dos atributos dos termos selecionados, e consequentemente, uma análise conceitual com maior aproximação do domínio de conhecimento facilitando a construção da estrutura das taxonomias e ontologias e o trabalho com modelos de representação.

4.2.2 Aplicação do modelo de raciocínio em redes semânticas

A construção das redes nesta pesquisa, a partir dos textos das teses e dissertações está baseada na hipótese de que palavras que ocorrem juntas em uma mesma sentença teriam sido evocadas de forma associativa na construção da ideia a ser apresentada. Assim podemos construir uma rede, onde as palavras são representadas como os vértices e uma aresta é criada entre pares de palavras que ocorram em uma mesma sentença, conforme apresentado pelos estudos desenvolvidos por Caldeira 2005, Caldeira et. al., 2006 e Teixeira, 2007.

Para Caldeira et al. (2006), as palavras seriam os significantes, as sentenças a menor unidade de significado, e o texto um sistema formado por um grande número de unidades (palavras) interagindo entre si por pequenos grupos (sentenças), através de construções sintáticas (frases). Esta interação é capaz de fazer emergir uma ideia diferente do que cada palavra sozinha significa. Do mesmo modo, sentenças interagem entre si por significantes comuns e também são capazes de fazer emergir ideias no texto, diferente do que cada sentença expressaria se dita sozinha. Tudo depende do contexto em que as palavras estão.

Contudo, a partir deste critério, pares de palavras cuja associação é pouco significativa acabam sendo incluídos na rede e mascaram a estrutura formada pelas associações mais fortes. Desta forma faz-se necessário uma filtragem, de forma a que apenas as associações mais relevantes para o discurso sejam consideradas na construção da rede, como especificado por (Teixeira 2007).

Nesse sentido, Teixeira (2007) define o conceito de Incidência-Fidelidade, como um índice que está associado à frequência com que um determinado par de palavras é evocado em um discurso oral ou escrito. O resultado da análise mostra o quão “fiel” é um par de palavras dentro de um texto.

Antes de apresentarmos o modelo dinâmico de análise se faz necessário entender o objeto de representação, neste caso específico, o conceito.

4.2.3 O conceito como objeto de representação

O objeto de representação é o conceito ou termo que representará o conteúdo das pesquisas desenvolvidas no domínio de conhecimento e que será estruturado em um SOC. O conceito revelará as pesquisas desse domínio e servirá como termo indexador de bases de dados possibilitando resultados de buscas atinentes aos anseios do usuário final que utilizará a base, resultado do SOC construído.

O objeto de representação, conceito foi definido por Dahlberg (1993, p.211) como “unidade mínima de representação” ou “unidades de conhecimento e as muitas possíveis combinações delas em palavras/termos ou declarações”. Denominado pela autora como referente que inserido em um contexto é designado por um nome.

O conceito para ser designado em um contexto dependerá das relações entre conceitos estabelecidas a partir das relações categoriais e que refletem a complexidade da realidade de um dado domínio de conhecimento. A partir de práticas de pesquisas realizadas pelas comunidades e da forma como o conceito é revelado discursivamente no texto, assumirá uma ordenação nas categorias, a depender de propriedades ou características essenciais ou acidentais, como especificado no capítulo 2 deste trabalho.

Além das questões da categorização e seus aspectos metodológicos baseados em teorias, como as do Conceito e TCF, é relevante lembrar a cognoscência do sujeito, ou de como se apreende o objeto ou fenômeno representado. A representação é tarefa executada por um sujeito que percebe e infere sobre o objeto, categorizando-o de acordo com o contexto a ser traduzido pelo termo e atenderá ao que foi sugerido no texto por outro sujeito que expressa suas convicções de pesquisa, mas que, sobretudo atenderá a um terceiro sujeito que busca a informação que lhe dará possibilidades para resoluções de questões, ou pelo menos reduzirá as incertezas.

O termo que declara um objeto ou movimento de objetos deverá ser bem definido, para cada realidade a ser representada, pois o que é esclarecedor para um sujeito pode não ser para outro, existe a subjetividade no entendimento do objeto que sofre influências culturais, políticas, geográficas e de conhecimentos diversos. O problema da subjetividade foi entendido por Foucault (1999, 2008) e demarcado em suas obras como a forma pela qual um sujeito faz a experiência de si mesmo em um jogo de verdade. No caso da análise do objeto para este trabalho, por meios de práticas relacionadas ao conhecimento e as técnicas empregadas nas pesquisas.

Não se pode trabalhar com conceitos com indefinições, pois se corre o risco da imprecisão. Cada definição deverá ter a garantia literária, com base em autores seminais do domínio de conhecimento onde o termo foi cunhado, em um caminho fortuito de acontecimentos, a depender do contexto.

Trata-se de construir representações que identifiquem possibilidades e que funcionem como “verdades” em tempo e espaço. Tal construção dependerá de mecanismos estruturantes, como os de instâncias que permitam distinguir enunciados verdadeiros ou falsos, as regras

estabelecidas segundo o que o sujeito diz como verdade sobre o objeto. Só assim poderemos estabelecer princípios de análise e modelização de domínios. Passamos na próxima seção a evidenciar tais princípios.

4.3 PRINCÍPIOS PARA ANÁLISE CONCEITUAL: UM MODELO BASEADO EM ANÁLISE COGNITIVA

Apresentamos um modelizador de domínios de conhecimento, profissional que conhece e constrói modelos, nesta pesquisa, como interprete cognitivo que infere e pressupõe o que está subentendido no texto, **declarando** o discurso através de uma representação terminológica.

No trabalho de modelização leva-se em conta uma ação comunicativa onde os elementos participantes dessa ação são: texto, contexto, termos/conceitos. Pois o que deve ser representado é muito mais amplo do que a mera representação simbólica da palavra. Aproxima-se do que vivenciamos no mundo fenomenal, portanto do que referenciamos e classificamos a partir dos perceptos, dos espaços mentais, das nossas bases de conhecimento.

Na perspectiva cognitiva do modelizador de domínios, é o processo de dedução e indução que garantirá a coerência na interpretação, organizando o sentido do texto e do conceito que o representará. A partir daí pode-se organizar relações entre conceitos fazendo desses, uma unidade dinâmica de sentidos que evolui e se manifesta de acordo com os significados apresentados no texto.

A prática da inferência permitirá organizar a teia de significados e significantes que o termo/conceito será capaz de traduzir, não de forma aleatória, mas lógica, com base no raciocínio e interpretação que responderá ao objetivo do contexto, satisfazendo a busca do usuário que procura à informação, possibilitando a solução de um problema.

Nesta proposta metodológica cognitiva analisamos as redes conceituais dos textos selecionados para este trabalho a partir de dois aspectos:

- a) nível ampliado de análise: a partir dos espaços referenciais de ligações dos termos, nas redes semânticas, que aqui chamamos de rede conceitual. O objetivo é perceber a construção de sentidos do texto ou de um esquema

discursivo revelado pelas redes de conceitos que caracterizam domínios de conhecimento a partir de subáreas e especialidades.

- b) nível específico de análise: relações conceituais decorrentes das redes. A partir da interpretação e análise discursiva observada nas redes é possível perceber a funcionalidade dos termos/conceitos enquanto possíveis representantes do texto e contexto a ser explicitado em uma categorização.

Após análise dos aspectos citados, o próximo passo foi a aplicação do modelo dinâmico de análise conceitual. Esse modelo está amparado pela análise cognitiva e do discurso representativo, e na Teoria do Conceito. O modelo está baseado em estudos já desenvolvidos e aplicados por autores contemporâneos em ontologias e ampliado com a nossa proposta.

4.4 O MODELO DINÂMICO DE ANÁLISE CONCEITUAL (MODAC) – BASE PARA SOC

Modelo algum pode ser fechado, ou determinado em si mesmo, por este motivo denominamos de modelo dinâmico, pois a nossa pretensão é colaborativa, no sentido de ampliar estudos no âmbito da organização do conhecimento. Acreditamos que se a representação conceitual tem um SOC bem estruturado contribui para a atinência na RI. Segundo Mendonça e Almeida (2012, p.2) “Modelos são entidades fundamentais na prática científica e se prestam a investigar fenômenos complexos do mundo sobre os quais não se tem entendimento completo”.

É certo que a atinência dependerá das relações conceituais estabelecidas e que comporá uma base de dados e ainda da percepção do sujeito que modeliza e do sujeito que vai à busca da informação, mas esperamos que um modelo de análise conceitual possa garantir um mínimo de aproximação das necessidades reais de um usuário.

Chamamos o modelo a ser apresentado de dinâmico porque deverá acompanhar o movimento dos conceitos em seus vários estágios de relação com outros conceitos, podendo ser capaz de se adequar e adaptar-se ao contínuo deslocamento das relações conceituais dependentes de causa, associação e composição (características) do objeto.

Iniciaremos a análise dos termos, a partir da próxima seção. No entanto convém esclarecer que elegemos para este momento 5 (cinco) dissertações da área da geografia, no domínio da geografia física que denominamos de corpus 1 ou **CP1**, nesse corpus estão contidos os trabalhos de: Bittencourt, (2011); DiLauro, (2011); Evangelista, (2010); Novais, (2009) e Souza, (2010). No corpus 3 (três) ou **CP3**, 5 (cinco) dissertações do domínio da geoquímica, contendo os trabalhos de: Andrade, (2011); Caldas, (2011); Ferreira, (2011); Melo, (2011); Moreira, (2011). Os termos da Geologia, corpus CP2, não foram analisados neste capítulo. As referências completas dos trabalhos que compõem os corpus, CP1; CP2 e CP3, encontram-se no **Apêndice A - Referências dos trabalhos que compõem o corpus (CP) selecionados para análise**.

4.4.1 Análise de termo em rede semântica no domínio da Geografia física

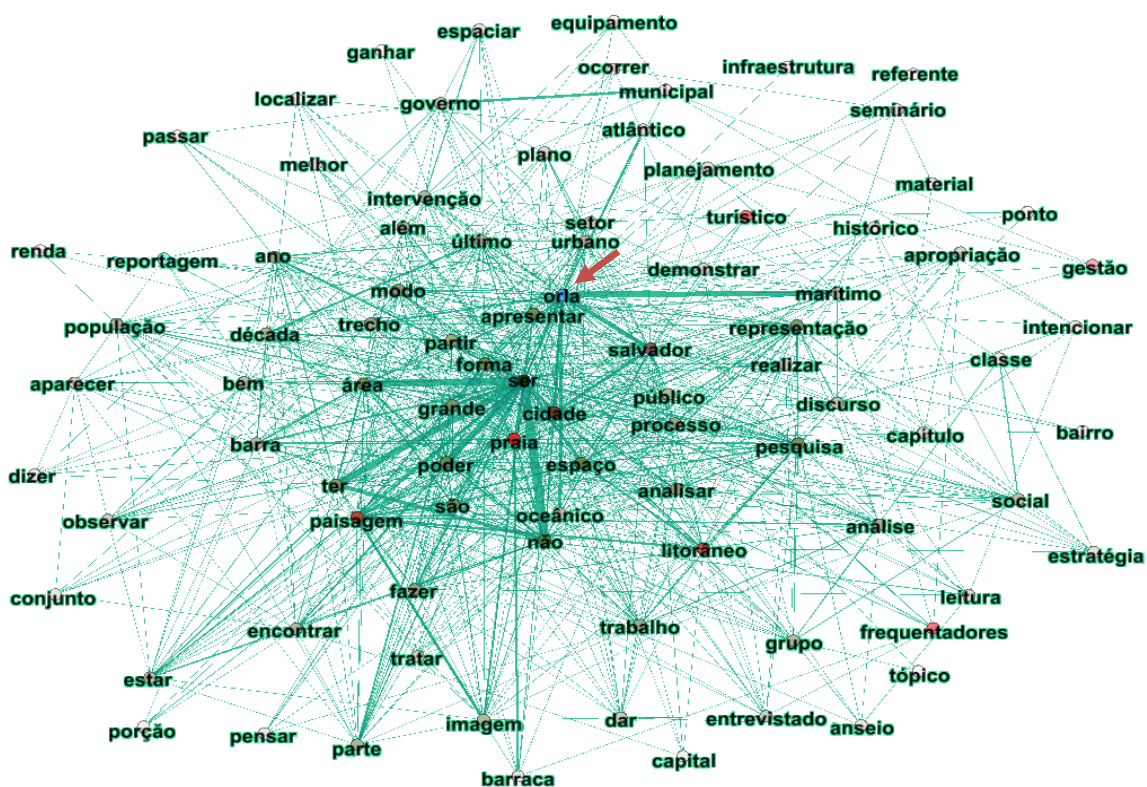
Na primeira leitura de um conceito extraído de uma rede semântica gerada a partir de dissertações de mestrado no âmbito da geografia, destacamos algumas conexões diretas entre conceitos, sem nos preocuparmos em um primeiro momento com as sub-redes deste conceito. Na segunda leitura incluiremos as sub-redes.

As definições conceituais daqui por diante descritas e impressas na análise, partiram de leituras realizadas em fontes de referência como: dicionários, glossários e tesouros de termos publicados em suporte de papel e ou disponibilizados na internet, como também em dissertações e teses dos domínios de conhecimentos analisados.

Utilizamos um processo de análise e síntese sobre o que as fontes descrevem sobre o termo e principalmente procuramos nos aproximar e adequar às definições encontradas ao discurso encontrado no corpus (dissertações e teses) e ao domínio de conhecimento. Essas fontes de referência estão referenciadas no **Apêndice B – Fontes para definição dos termos**.

A partir da construção da sub-rede do termo **orla**, nó localizado no centro da Figura 16, passamos à análise de termos.

Figura 16 – Sub-rede do termo orla



Fonte: construção da autora

O termo **Orla** tem um significado semântico estabelecido na geografia como sendo:

- linha que separa o mar da terra, correspondendo ao limite entre o oceano e o mar. Informalmente designado como margem, ex.: a orla da praia.
- do ponto de vista geográfico, a praia é um acidente que se forma a partir da acumulação de sedimentos não consolidados pela ação do fluxo das ondas, portanto, dinâmico no ponto de vista conceitual, pois pode gerar a depender do contexto abordado relações conceituais complexas, igualmente dinâmicas, a depender do domínio em que o termo seja empregado.

Observando as conexões estabelecidas percebe-se que o termo **orla** foi tratado na tese que compõe a amostra, como um lugar frequentado por pessoas, provavelmente um setor produtivo da cidade, no caso Salvador. O significado que este termo carrega tanto para o trabalho de pesquisa, quanto para o domínio está visivelmente focado nas demarcações da geografia física (enquanto paisagem urbana) bem como na geografia humana (enquanto ambiente social, cultural e político).

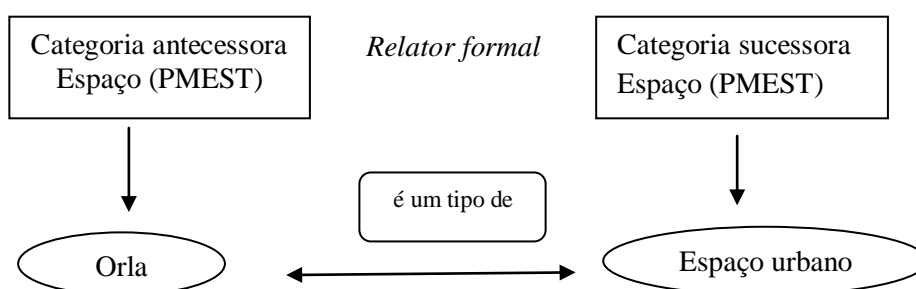
O termo só será entendido a partir da percepção do indivíduo, do conhecimento acerca do significado do termo, e ainda dependente da forma discursiva empregada. Como venha ser abordado na temática, pode mudar a rede pela adição de nós e arestas, a depender da forma de associação individual ou representativa do conhecimento. Assim, para a construção de um modelo dinâmico de relações conceituais é necessário uma linguagem formal, que ao ser processada em uma máquina possa garantir o entendimento do termo para um dado domínio, só assim atingirá o objetivo da difusão do conhecimento, pois o termo poderá ser entendido pelo usuário que busca informação, pelo público em geral.

Em um modelo de relações conceituais, a consistência é garantida a partir das regras estabelecidas para as associações que não podem ser apenas entre categorias de conceitos, mas também se deve levar em conta a forma. Ou seja, o modelo deve ser entendido a partir de três elementos: categoria 1 (antecessora); forma (relação formal); categoria 2 (sucessora), como no modelo triádico apresentado por Sales, (2006), para ontologias de domínios. Segundo Sales, (2008, p.3), “[...] o estabelecimento do modelo triádico é realizado no momento em que se está construindo a parte da Ontologia inteligível por humanos, porém é a partir desta fração da ontologia que se cria a parte computacional de forma coerente.”.

O modelo apresentado por Sales inicia com o levantamento das definições do conceito (etapa 1); seguindo para a interpretação das definições e estabelecimento das categorias conceituais, baseadas na Teoria da Classificação Facetada e suas categorias fundamentais, PMEST e as possíveis combinações entre categorias (etapa 2); e por fim na (etapa 3), o estabelecimento das relações formais entre as categorias e consequentemente entre os conceitos revelados nas categorias.

Exemplificando o modelo triádico entre relações categoriais do conceito do domínio analisado através da rede semântica:

Figura 17 – Exemplificação do modelo triádico de relações categoriais fundamentais



Fonte: construção da autora

Um conceito pode ser definido como um caminho para comunicação de uma ideia, assunto, conteúdo de um texto, contexto de informação. Portanto, encontra-se em um modelo triádico, ou seja, para a categoria (antecessora) que o determina (significante), para a forma verbal que sintetiza as asserções sobre o termo (regra de ação) e para uma segunda categoria (sucessora) que expressa ou revela o seu significado (ou sentido no contexto).

O modelo por si só responde as perguntas:

- O que o conceito expressa?
- O que o conceito explica?
- O que o conceito revela do texto?
- O que o texto propõe?

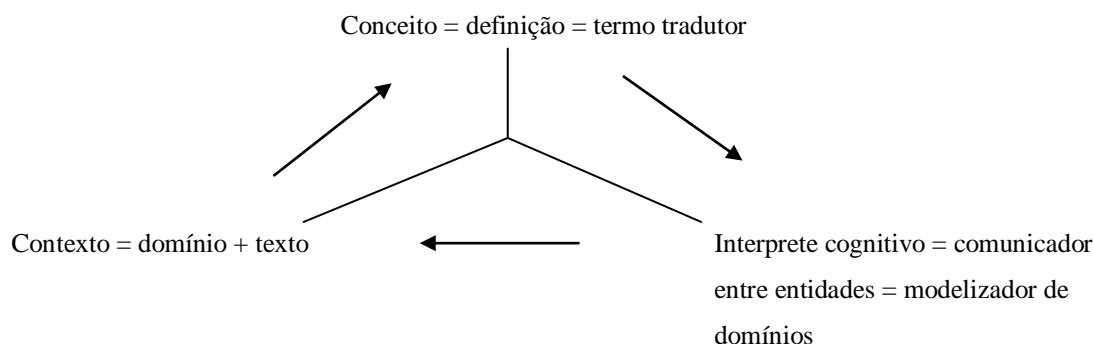
Sabemos que um modelo de análise que se baseia em objetos e atributos (elementos que definem a estrutura de uma classe de conceitos) e as relações entre objetos são dependentes do todo e das partes integrativas do objeto e de como se manifestam nas expressões dos discursos.

Um conceito se mal definido pode interferir nas relações associativas, a depender da tradução dada pelo interpretante do discurso que deverá estar atento em não cometer uma análise equivocada do contexto em que o termo/conceito se manifesta. Apropriando-nos das

considerações de Varela; Thompson; Rosch (2001, p, 47) “[...] hipóteses mais elaboradas são teoricamente reduzidas a possíveis observações por meio de inferências dedutivas.”.

Existe portando uma tríade também na análise do termo e na forma como o termo/conceito seja traduzido para uma linguagem de representação (meta-representação) que comporá um sistema de organização de conhecimento (taxonomias e ontologias), que se caracteriza por (conceito, contexto e interprete cognitivo). Apresentamos na Figura 18 a tríade de análise que propomos:

Figura 18 – Tríade de análise terminológica



Fonte: construção da autora

A organização do conhecimento é dependente de modelos de representação, indicamos a associação de modelos: triádico de categorização e triádico de análise terminológica. Associação desses modelos auxiliará na construção de sistemas meta-representacionais norteadores para SOC.

Defendemos uma prática cognitiva baseada na experiência do interprete, conhecedor do contexto, em uma experiência direta da incorporação humana na prática interpretante do modelizador de domínios. Compreende-se que o modelizador deverá trazer consigo um conhecimento prévio do domínio, utilizando-o para interagir com a análise do texto e das relações expressas nas redes conceituais.

O interprete cognitivo (modelizador de domínios) é o centro de todo o processo de interação entre o conhecimento trazido pelo texto e o conhecimento do usuário que busca

informação. A interpretação é a chave do sentido negociado pelo termo tradutor do contexto/domínio de conhecimento. O processo de interpretação é o primeiro passo para o caminho da comunicação, onde o texto converge às ações da língua, da cognição, da cultura.

Os textos têm bases expositivas, argumentativas (filosofias e teorias), e científicas. Através dessas bases estruturais podem ser conduzidos esquemas que auxiliam na sua interpretação. Os textos expositivos se sustentam em (problema, causas e soluções). Os argumentativos apresentam um ponto de vista que se ajusta através de justificativa, marco teórico que legitimam as circunstâncias e os fatos até a conclusão (tese, argumentos e conclusão). Os de discurso científico estão fundamentados por um problema a ser solucionado (tema, problema, hipótese, metodologia, resultados e conclusão).

Este interprete cognitivo, do ponto de vista do modelizador de domínios, necessariamente precisa estar atento além do discurso/texto, ao contexto da pesquisa em que o termo foi citado. Deverá fazer um deslocamento para o processamento da informação, de como o indexador que usa SOC poderá fazer uso dos termos que comporá tal sistema, e principalmente se deslocar para o lugar do usuário que busca informação. Entender, portanto, todo processo de mediação envolvido neste complexo movimento de sentidos, garantindo um acordo de representação. Partindo do ponto de vista do modelizador retomamos a análise de outro termo elemento da rede, (reportagem), um dos vértices. Exemplificando:

O **termo reportagem** é utilizado na produção de texto jornalístico, escrito ou falado, no caso da sub-rede, originado de um contexto jornalístico que se baseou no testemunho direto de fatos e situações, vivenciadas no ambiente geográfico da orla de Salvador, mas não pertencente ao domínio da Geografia, ou seja, se relaciona ao termo orla por uma questão discursiva e não por traduzir o contexto, portanto não o define. Sendo assim não será um termo candidato a compor a ontologia do domínio.

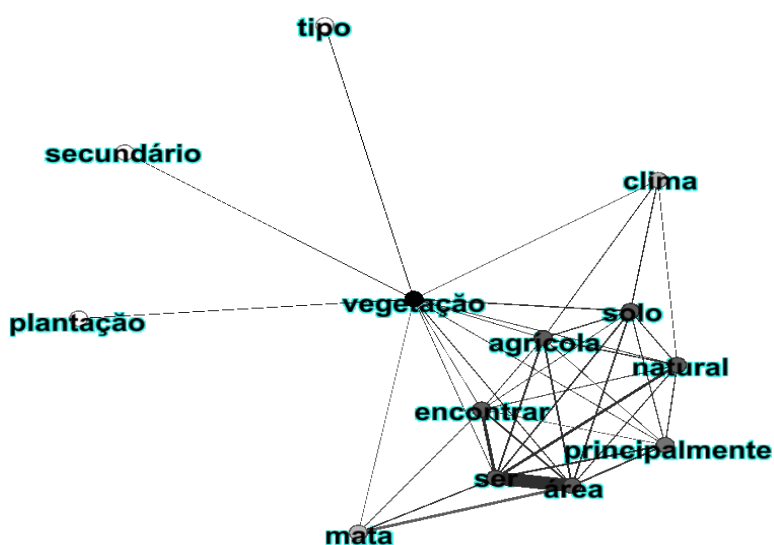
Não se caracteriza ainda, como termo fundamental ao entendimento do termo orla no ponto de vista da geografia física, humana, econômica ou política. Não é um termo fundamental tampouco ao entendimento das pesquisas desenvolvidas na área de meio ambiente, foi utilizado no texto, possivelmente por ter sido objeto de investigação documental sobre questões relacionadas à orla de Salvador, representando um objeto de correlação a outro

objeto, no caso o termo gestor. No entanto, analisando cognitivamente o ponto de vista do usuário do domínio da geografia humana ou física, no contexto de pesquisas sobre planos diretores, por exemplo, o termo gestão ou gestor é um conceito relevante às pesquisas.

4.4.2 Análise de termos em redes semânticas nos domínios da Geografia física e Geoquímica

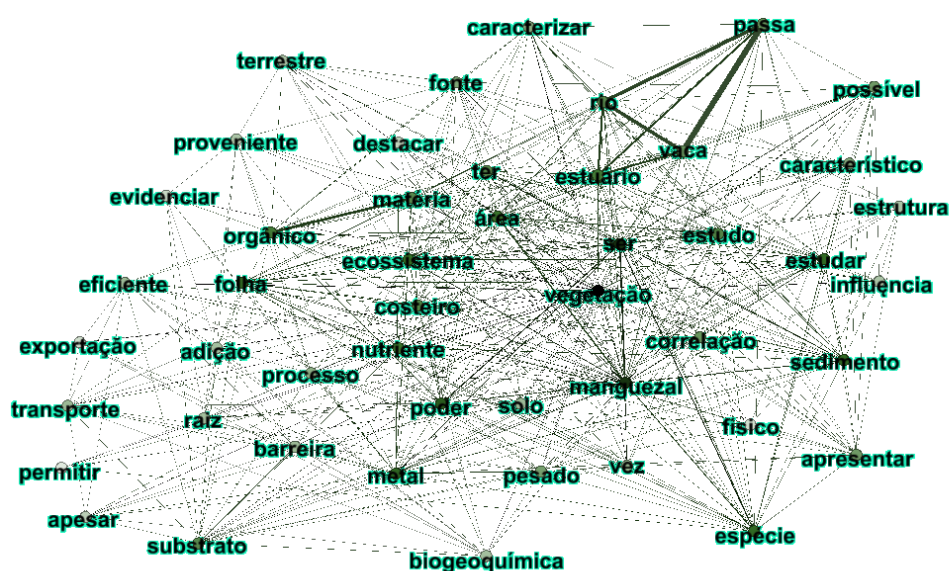
Nesta segunda leitura de um conceito extraído das redes semânticas geradas a partir de dissertações e teses de mestrado e doutorado no âmbito da geografia e da geoquímica observamos as conexões entre conceitos, nas sub-redes do conceito vegetação, Figuras 19 e 20 respectivamente.

Figura 19 – Análise conceitual do termo vegetação no domínio da Geografia



Fonte: construção da autora

Figura 20 – Análise conceitual do termo vegetação no domínio da Geoquímica



Fonte: construção da autora

O conceito de vegetação é atribuído ao grupo dos vegetais existentes em um dado espaço geográfico. Portanto um conceito geral advindo do domínio da botânica (ciência responsável pelo estudo dos vegetais).

No domínio da geografia o conceito **vegetação** tem a mesma apropriação terminológica, sendo usado no contexto geográfico na rede observada (Figura 19), inter-relacionados a conceitos como o de **clima** (conceito advindo do domínio da meteorologia e utilizado por domínios das geociências), e de uma variação do **solo** de determinada região. Solo, no entanto é um conceito multidisciplinar que quando categorizado necessita ser qualificado, como acontece com termos homônimos, ex. Solo (superfície terrestre). Na sub-rede do termo vegetação, percebe-se claramente a inter-relação entre os termos, afinal, solo é um produto do intemperismo, portanto, ligado de forma direta ao clima de um lugar, enquanto o tipo de vegetação dependerá do solo e consequente do clima.

No domínio da geoquímica o termo vegetação na sub-rede analisada (Figura 20) tem uma inter-relação direta com os conceitos: **manguezal** (ecossistema costeiro de transição entre ambientes terrestres e marinhos); **estuário** (ambiente aquático entre rio e mar) e sedimento (detrito resultante de erosão). O **solo** de um manguezal tem características salinas, pobre em oxigênio e rico em nutrientes, com predominância de uma vegetação de halófilos.

Manguezal como ecossistema é um conceito de uso interdisciplinar quando se pensa nos processos biogeoquímicos que o envolvem quanto à produção e acúmulo de energia. Estuário é um conceito de uso multidisciplinar quando utilizado no ponto de vista ecológico pela oceanografia, geologia, geoquímica e biologia que observam em seus estudos determinações de sua geomorfologia, produtividade biológica e aspectos ligados a contaminantes como metais pesados.

Como resultado das análises observa-se que manguezal, clima, estuário, solo e vegetação são termos que constarão em SOC de diferentes domínios, mesmo que com apropriações particulares, dependentes de seus atributos e relações. Alguns conceitos gerais precisam ser especificados como é o caso de vegetação, clima e solo (que tipo de vegetação, clima ou solo?). Quando bem definidos auxiliarão na categorização e conseqüentemente nas relações entre categorias.

Pontuamos então que redes semânticas são úteis para análise dos termos e permite a aplicação da tríade proposta: Tríade da análise terminológica. Permite ao interprete cognitivo (analista cognitivo), através das redes conceituais encontradas a partir de textos, analisar termos sob o ponto de vista do contexto, encontrando suporte para definições que se aproximam do discurso do domínio de conhecimento, garantindo assim uma melhor aplicação da tríade de construção de ontologias de domínios, ou modelo triádico: categoria 1 (antecessora); forma (relação formal); categoria 2 (sucessora).

Reafirmamos que a AnCo, nesta pesquisa, está compreendida no processo de construção do modelo, quando aproxima a análise dos requisitos de informação de forma funcional da seguinte forma:

- reconhecimento da obra: aqui entendida como entidade abstrata de criação intelectual e científica particular;

- reconhecendo a impressão textual: entidade concreta de revelação do conteúdo de pesquisa ou manifestação do domínio de conhecimento na forma da linguagem;
- reconhecendo a pessoa: indivíduo envolvido na criação da obra e indivíduo envolvido na busca de informação contida na obra;
- reconhecendo o conceito: manifestação de uma ideia, unidade de conhecimento;
- e principalmente reconhecendo a liberdade de expressão do indivíduo e de sua percepção sobre o objeto/conceito, pois cada entidade pode ter um ou vários atributos que podem ser percebidos pelo sujeito. A depender da sua percepção, atribuição e uso por parte do sujeito, os atributos do objeto/conceito podem se desdobrar dando sentidos diversos ao conceito ou até revelando novos conceitos.

No contexto do processo, de análise cognitiva, o foco da análise conceitual retrata a ligação entre um conceito e outro, para atender ao usuário em sua investigação no universo do conhecimento retratado por itens documentais que revelam o contexto de domínios, permitindo:

- encontrar manifestações de conhecimento específicas às suas necessidades de informação;
- identificar os conceitos que revelem essas manifestações;
- identificar itens documentais que expressem e retratem o conhecimento;
- selecionar o item que mais se aproxima da sua necessidade de informação;
- obter informação através de um item documental.

A análise que propomos tem por objetivo permitir que o usuário não só encontre o que já sabe ou que complementa o que ele já traz consigo, mas o que leva ao novo ou ao inesperado. Além disso, precisamos ter em mente que existem, usuários que precisam encontrar informações ou obras, mas existem obras que precisam ser encontradas. E isso só

será possível através de uma representação coerente disposta em canais de informação como: catálogos, bases de dados e buscadores on-line, como os disponíveis na internet. Passamos à proposta do modelo na seção que segue.

4.4.3 Proposta de modelo dinâmico de análise para organização do conhecimento

O Modelo Dinâmico de Análise Conceitual (MODAC) que propomos, está baseado na Tríade de Análise Terminológica (TAT), (Figura 18) composta por: conceitos e suas definições; contexto (domínio de conhecimento) e texto; interpretação do modelizador de domínios sobre o texto, contexto e o termo tradutor de domínios (conceito). Apresentamos no quadro 3 (três) a definição dos principais componentes da TAT.

Quadro 3 – Definição formal dos conceitos da TAT

Análise de contexto - análise das redes semânticas (relações entre os termos) dos textos do domínio de conhecimento – base Teoria do Conceito

Definição conceitual - resultado de pesquisa em fontes de referências (dicionários, glossários, tesouros) especializadas nas áreas de conhecimento onde se encontram cunhados os termos pelos domínios em conjunto com a Análise do contexto.

Interpretação cognitiva - resultado da interpretação das “leituras” das relações entre os termos e seus significados semânticos. Baseados em Anco e na Análise de discurso

Fonte: construção da autora

Reafirmamos que o modelo apresentado nesta tese é dinâmico, quando a partir das relações que se dão entre os termos na rede semântica podemos estabelecer relações conceituais mais próximas do domínio e das pesquisas desenvolvidas nesse espaço de pesquisa em constante desenvolvimento.

Acreditamos que em domínios de conhecimento de natureza inter e multidisciplinares essas relações são mais complexas devido ao empréstimo discursivo e conceitual de vários domínios que se integram nas pesquisas. Assim esperamos que o Modac, contribua para a categorização dos conceitos em uma taxonomia e nas relações associativas entre os conceitos em ontologias.

5 APRESENTAÇÃO E APLICAÇÃO DO MODELO DINÂMICO DE ANÁLISE CONCEITUAL - Modac

[...] cada conceito será, pois considerado como o ponto de coincidência, de condensação ou de acumulação de seus próprios componentes. O ponto conceitual não deixa de percorrer seus componentes, de subir e descer neles. Cada componente, neste sentido, é um traço intensivo, uma ordenada intensiva que não deve ser apreendida nem como geral nem como particular, mas como pura e simples singularidade – ‘um’ mundo possível, ‘um’ rosto, ‘certas’ palavras – que se particulariza ou se generaliza, segundo se lhe atribui valores variáveis ou se lhe designa uma função constante. (DELEUZE; GUATARRI, 2010, p.28)

Partindo do pressuposto de “um mundo possível” a partir de um contexto expresso em textos, e do que está revelado como acontecimento conceitual, através das conexões de palavras que se materializam nas entidades, como ato de pensamento, ou ainda como dito por Deleuze e Guatarri (2010, p. 29) “[...] o pensamento operando em velocidade infinita [...]” idealizamos o Modac. De modo que a construção do modelo venha a demonstrar o conjunto de variações existentes na produção de um conceito. Convém esclarecer que utilizamos aqui considerações desses autores para ilustrar o nosso pensamento e não como referencial teórico para o modelo, afinal já apresentamos tal referencial no capítulo 2 (dois).

Para construirmos o modelo dinâmico de análise conceitual proposto necessitávamos de um corpus de análise. Esse corpus foi construído a partir de redes semânticas, que passamos a denominar para o caso desta pesquisa de **rede conceitual**. Caldeira (2005) em dissertação de mestrado sobre caracterização de rede de signos linguísticos preferiu denominá-la de “rede de palavras”. Em 2007, Caldeira em outro trabalho de dissertação onde analisou elementos conceituais esclarece:

À medida que construimos um pensamento a ser transcrito para o papel ou para a tela do computador, escolhemos palavras, que se ligam às anteriormente escolhidas, formando uma rede, como um reflexo aproximado dos significados que queremos transmitir.

A partir dessa premissa, propus outro método de análise de conteúdo dos textos que é a construção das subredes a partir de palavras específicas, com o objetivo de conhecer seus pares, ou seja, as palavras que co-ocorrem com a palavra escolhida nas sentenças. (CALDEIRA, 2007, p.56)

No caso desta pesquisa entendemos rede de palavras, quando a rede não está em funcionamento, ou seja, quando não está sendo analisada ou interpretada por um analista

cognitivo, por exemplo. A partir do momento que a rede produz sentidos ela passa a ser semântica, como denominada no meio acadêmico e na literatura, das áreas que desenvolvem estudos sobre redes como a Ciência da Informação ou Ciência da Computação, dentre outras. Ou ainda rede conceitual, pois explicita significados semânticos através das relações entre termos.

Uma palavra enquanto unidade da linguagem humana possui diferentes significados a depender da ideia que se tenha em mente para designá-la em um determinado espaço e tempo dessa linguagem.

Quando uma palavra representa uma ideia, provocando um sentido por trás do que se quer dizer dar-se o nome de termo, ou ainda conceito. Qualifica-se um conceito como unidade de conhecimento que representa e declara de forma geral ou abstrata um domínio de conhecimento, dando significado ao que a mente concebe sobre uma realidade.

Ao analisarmos um conceito pretendemos dar-lhe sentido em um determinado domínio de conhecimento, para que possamos referenciá-lo, atribuímos predicados, e por fim representá-lo. Quando selecionamos um termo ou palavra e analisamos esse termo a partir de um domínio, determinamos o conceito. Concedemos ao termo características relevantes que o denotam, só assim poderemos representar, ou demonstrar claramente as pesquisas do domínio.

Reafirmamos que a intenção desta pesquisa de doutorado é apresentar um modelo dinâmico de análise conceitual que possa colaborar com SOC, onde os passos de organização incluem os de representação de conhecimento. Neste capítulo apresentamos o modelo, e as etapas do processo de sua construção.

O modelo denominado de Modelo Dinâmico de Análise Conceitual - Modac está simplificado na tríade de análise terminológica TAT apresentado na figura 18, no capítulo 4 (quatro), e a definição formal de seus elementos no quadro 3 (três). Este modelo pretende confirmar a importância da segmentação textual que pode ser auxiliada por uma interpretação da rede semântica/conceitual.

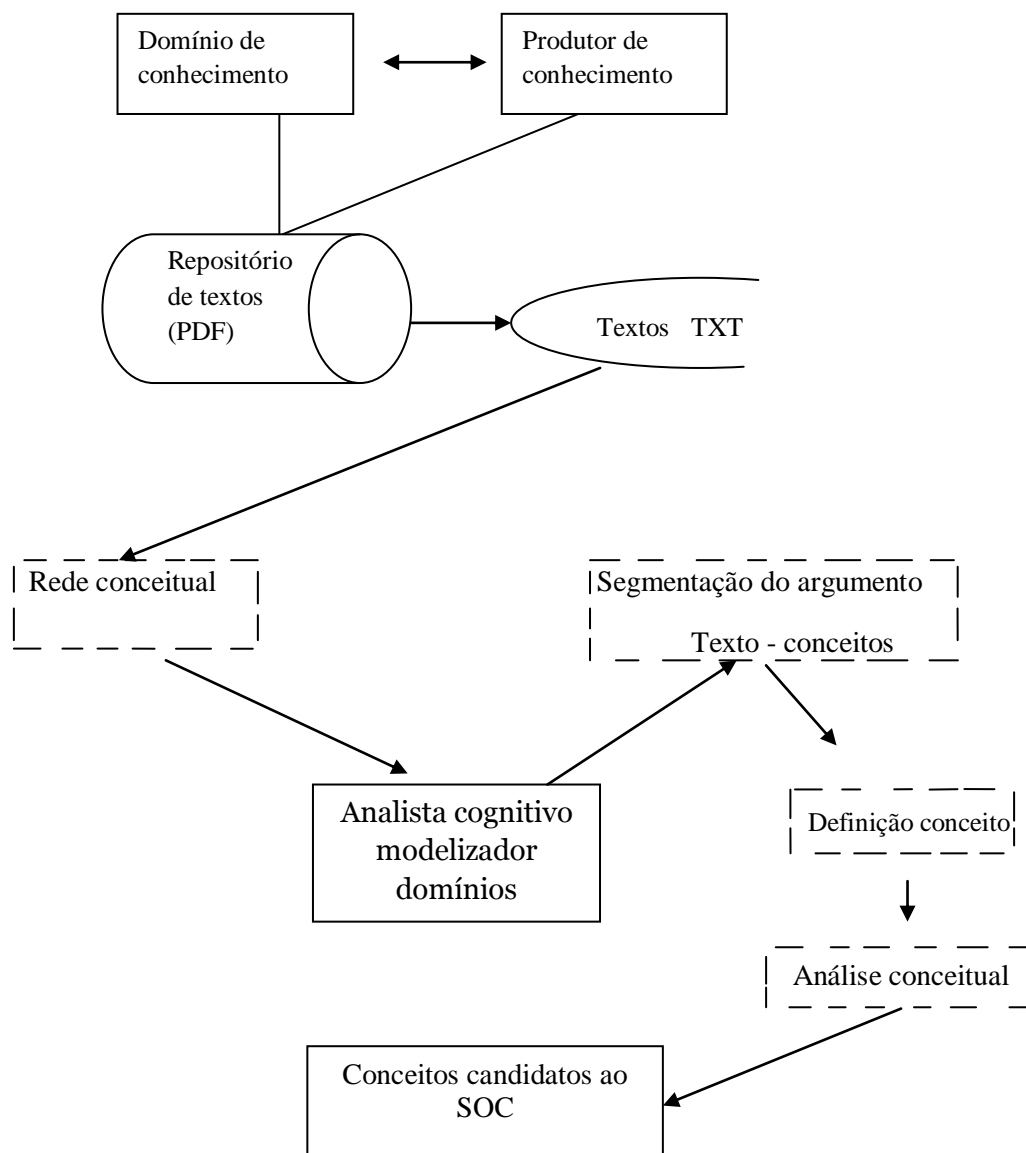
A segmentação só se completa com a definição dos conceitos e análise conceitual realizada por um analista cognitivo que interpreta a rede. Em uma leitura que chamamos de

possível, em blocos fragmentados de palavras que se unem na rede pela coerência da estrutura textual. A cada relação entre as palavras, ou pares de palavras pode-se ter uma leitura diferente a depender da interpretação do analista/modelizador de domínios.






A fragmentação pode resultar em um agregado de pedaços, portanto, há de se ter cuidado. É preciso garantir a coerência na representação, por esse motivo é importante que se tenha um modelo de análise, que inicie com o entendimento do domínio, a partir do olhar do usuário para enfim determinar relações e definições conceituais, que possibilitarão SOC adequado aos usuários que buscam e pretendem encontrar e obter itens documentais.

Na figura 21 que segue, apresentamos o diagrama do fluxo de análise, esse diagrama é composto basicamente de dados e processos.

Figura 21 - Diagrama do fluxo de análise no Modac



Legenda:

-  Armazenamento de dados (aberto)
-  Armazenamento de dados para pesquisa
-  Início e término de processo e fluxo
-  Fluxo
-  Processo

Fonte: construção da autora

Formulação e entendimento do diagrama da figura 21

Procedimentos:

- Para obtermos uma amostra para aplicação do Modac elegemos um corpus de textos identificados e selecionados em bases de dados dos domínios de conhecimentos escolhidos. Cada domínio possui características próprias de pesquisas e produtores de conhecimento e textos. Preferimos dizer que esses textos são dados abertos passíveis de variadas interpretações a depender da percepção do sujeito, por isso a representação de uma elipse aberta.
- Após o entendimento do domínio e de suas propostas de pesquisas, são geradas redes conceituais, a partir da ideia de redes semânticas e conseqüentemente com um conjunto de programas para construção de redes, apresentado no capítulo 4. Nota-se que as redes são fundamentais para que possamos realizar a leitura possível do texto, através dos seguimentos das arestas que ligam os nós (termos) na rede.
- A ligação que se dá entre os termos se aproxima de macroproposições semânticas, ou macroestrutura conceitual, possibilitando ao interprete/analista cognitivo o entendimento do argumento do autor. Embora essa interpretação seja vista como subjetiva ela se fundamenta na Teoria do Conceito, Dahlberg (1978c), (referente, características e forma verbal), com a intenção de fixar os limites do termo em um discurso.
- A definição conceitual é a condição primeira para se categorizar conceitos e entender as relações formais e associativas que se darão em taxonomias e ontologias de domínios - SOC. A definição conceitual permite construir hipóteses dedutivas, e empíricas dedutivas que norteiam as macroproposições semânticas. A partir dessas macroproposições pode-se chegar aos conceitos funcionais operacionais e os funcionais mentais e também compreender os conceitos agenciadores que apresentamos na seção 5.1.

- Na construção do SOC deve-se priorizar o usuário. Entender a multiplicidade de vozes que se expressam através das buscas em bases de dados, por esse motivo um instrumento desenvolvido para esse fim, o de busca da informação deve conter mecanismos de armazenamento que possibilite aos modelizadores de domínios conhecerem as características dos usuários.

Na sequência apresentamos requisitos para o modelo, sintetizados no quadro 4 (quatro) a seguir:

Quadro 4 - Apresentação dos requisitos do modelo (processos)

- Requisito 1 - Construção das redes conceituais
- Requisito 2 - Processamento da Linguagem Natural (LN) dos textos selecionados (dados não estruturados) – palavras/conceitos - análise conceitual
- Requisito 3 - Geração de conceitos – a partir da análise das redes (evidências relevantes de associações conceituais) – Relação entre entidades/ conceitos - base análise discursiva/Teoria do Conceito
- Requisito 4 - Análise dos resultados da amostra – interpretação cognitiva

Fonte: construção da autora

Para implementarmos o modelo de análise necessitávamos de um corpus, elegemos como já descrito 5 (cinco) teses de cada domínio (Geografia, Geologia e Geoquímica), encontradas na home Page do IGEO/UFBA. Em armazenamento desta natureza é natural que se encontre alguns metadados importantes como: autoria e título do item documental, o que não é o bastante para atender às necessidades de busca do usuário.

Observamos que ao acessar a tese (texto completo em PDF) encontrávamos o resumo com palavras-chave em alguns textos. Constatamos que alguns autores não cumprem

exigências do uso de resumo e palavras chaves quando elaboram seus trabalhos de pesquisa. Um resumo mesmo que informativo, como no caso de dissertações e teses contém informações relevantes do conteúdo e resultados alcançados na pesquisa, mas ainda assim, não explicita todo o conteúdo da pesquisa desenvolvida.

A leitura completa do texto por analistas de informação e indexadores, como bibliotecários, por exemplo, é tarefa impossível de ser realizada dado ao grande volume de itens documentais existente na coleção de uma biblioteca ou centro de informação/documentação. Por vezes esses profissionais ficam tolhidos da análise do texto, e impedidos de traçar macroproposições semânticas que descrevam informações significativas e que traduzam o conteúdo dos textos em forma de conceito, optando pelo trabalho com palavras-chave.

Palavras-chave são escolhidas pelos autores de um texto para caracterizar a pesquisa desenvolvida nos trabalhos, mas muitas vezes não expressam o conteúdo de forma atinente. Podemos comparar essas palavras à linguagem natural (LN), pois não passaram por um tratamento artificial da linguagem, como os de linguagem documentária considerada por Dobedei (2002, p.39) como “metarrepresentação”. As LN não se configuram como linguagem controlada, onde o processo de tratamento se dá por conceitos e não por palavras. Lancaster (2004) assim definiu LN:

Considera-se a expressão linguagem natural como sinônimo de ‘discurso comum’, isto é, a linguagem utilizada habitualmente na escrita e na fala, e que é o contrário de ‘vocabulário controlado’. No contexto da recuperação da informação, a expressão normalmente se refere às palavras que ocorrem em textos impressos e por isso, considera-se como seu sinônimo a expressão ‘texto livre’. (LANCASTER, 2004, p. 250).

A Linguagem Natural é carregada de homônimos e sinônimos, próprios das palavras ou grupos de palavras empregadas no processo comunicacional do sujeito com intuito de se expressar constrói frases, sentenças e parágrafos, a depender da forma de comunicação oral ou escrita. Como no caso das redes geradas a partir dos textos das teses e dissertações que ora construímos e apresentamos amostras de suas relações conceituais.

Mas o que nos interessa para representação de domínios é a ideia que um conceito carrega do texto de forma explícita e algumas vezes o que está implícito no discurso.

Interessa-nos entender como se processa o pensar sobre o objeto, a origem de uma ideia. As palavras só existem para revelar uma ideia que se manifesta através do conceito. Esse conceito quando se concretiza na representação leva à informação. A análise do conceito é fundamental e deve tal conceito ser esgotado ontologicamente.

No trabalho de Currás (2010) quanto às ontologias e aspectos ontológicos, encontramos um pensamento que se aproxima do que pretendemos demonstrar nesta tese,

Em alguns tratados de filosofia se lê que a Ontologia é o estudo do que existe e do que admitimos que existe, para conseguir uma descrição coerente da realidade.

A última descrição pode esclarecer sua aceitação por parte dos informáticos. Deseja-se encontrar um paralelismo entre ‘o estudo do que existe’, ou seja, um domínio de conhecimento, com o ‘que admitimos que existe’, quer dizer, a transformação de uma linguagem natural, realidade do domínio escolhido, em uma linguagem codificada; ‘assumimos que existe’, para ‘conseguir uma descrição coerente da realidade’, quer dizer, poder obter daquele domínio, uma ‘resposta’ coerente da realidade. (CURRÁS, 2010, p. 38).

O que admitimos da realidade do domínio, em trabalhos de organização e representação do conhecimento, só é possível na segmentação do texto em macro-proposições semânticas, os conceitos. Além disso, precisamos entender o domínio de conhecimento no qual o conceito tem sua origem e a que uma estrutura conceitual representa.

O esgotamento ontológico a que nos referimos será entendido, não só quando tratamos um conceito como um isolado (presentes nos assuntos antes de serem reunidos em facetas), mas principalmente quando o posicionamos em um universo de assuntos, ou no seu arranjo sistemático em relação a outros conceitos (compreendendo, organizando e classificando-os através de suas características compreensivas), tornando-o inteligível para um determinado domínio de conhecimento. Pois segundo Dahlberg

[...] os conceitos são sínteses rotuladas de enunciados verdadeiros sobre objetos do pensamento: esses enunciados - asserções - levam ao reconhecimento ou à separação das características dos conceitos que também podem ser consideradas como os elementos dos conceitos. (DAHLBERG, 1979, [n.p.]).

Sendo assim podemos analisar um conceito sistematizando-o de forma geral por seus elementos, em qualquer domínio de conhecimento em que venha a se estabelecer,

reconhecendo-o como elemento material de um Sistema de Representação e Organização do Conhecimento e chegar a sua especificidade para o domínio que representa. Para tanto necessitamos transformar dados não estruturados em dados estruturados (informação), que revelem sentidos.

5.1 REDES DE CONCEITOS – DADOS NÃO ESTRUTURADOS

O processo de construção das redes foi discutido no capítulo 4 (quatro) desta tese, mas evidenciamos aqui, o seu principal objetivo. A rede conceitual foi construída a partir dos textos dos domínios de conhecimento das Geociências, que compõem o corpus selecionado. Escolhemos alguns termos para análise conceitual a partir da incidência nos textos e da fidelidade discursiva e que relatavam as pesquisas desenvolvidas nos domínios de conhecimento. Os termos evidenciados nas redes se caracterizam como dados não estruturados.

É necessário estruturar os dados no processo de obtenção da informação. A palavra clima, por exemplo, pode ser configurada como um dado, por ser uma unidade que possui significado próprio. A depender do universo empregado, esse dado poderá apresentar uma estrutura em um domínio de conhecimento e outra estrutura em outro domínio.

O importante é transformar um dado em entidade/conceito, para que possa representar realidade ou abstração, que serão identificadas por sua definição. Transformando-se em objeto que compartilhe características, ou conjunto de atributos, consequência das ocorrências de registros de dados, a depender das relações associativas entre conceitos de um domínio.

A seguir apresentamos uma breve análise de uma entidade e de um conjunto de atributos que ocorrem para determiná-la enquanto conceito:

Quadro 5 - Relações entre entidades

Entidade A	Entidade B	Entidade C	Ocorrências Entidade D	Resultado: Conceito do Domínio (CD)
Clima	Chuva	Maritimidade	Latitude média; latitude alta.	Clima oceânico
Categoria da entidade A	Categoria da entidade B	Categoria da entidade C	Categoria das entidades D	Resultado da relação
Padrão de elemento atmosférico	Fenômeno meteorológico	Fatores climáticos	Coordenadas Geográficas	CD= {Ea, Eb, ... En}

Fonte: construção da autora

Observando a ocorrência de atributos pode-se dizer que existe uma relação associativa e até atributiva, a partir dos atributos, entre os termos analisados, ou seja, o conceito clima oceânico é determinado pela ocorrência de atributos que mantém inter-relacionamentos com atributos próprios.

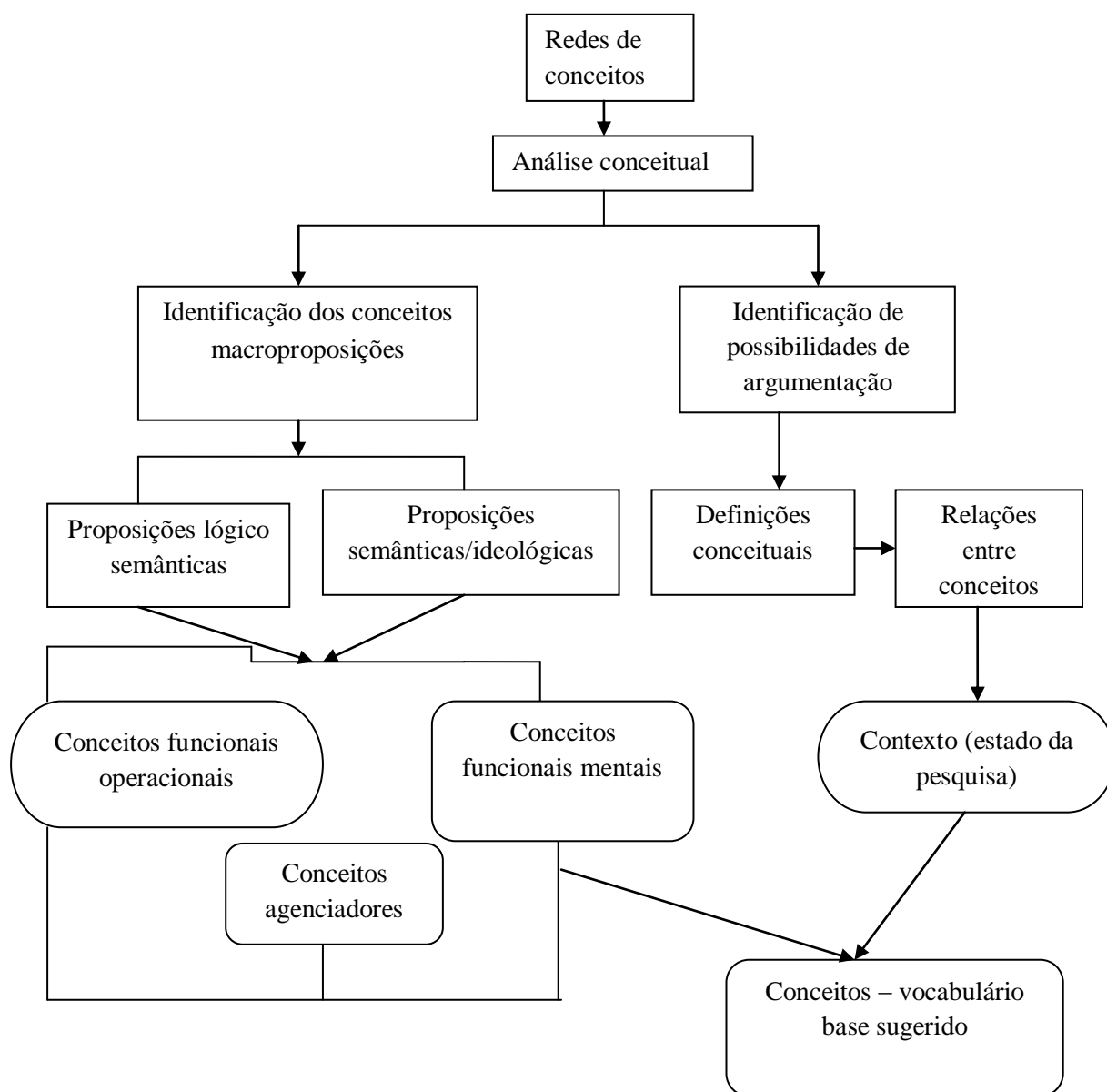
- **Exemplificando:** Um **clima oceânico (conceito)**, ou **clima temperado marítimo**, ocorre em margens ocidentais situadas nas **latitudes médias e altas**. Em regiões que possuem esse tipo de clima, as **chuvas** são constantes. Ou seja, para a definição do conceito existem atributos como (temperado e maritimidade) e conceitos como (latitude média; latitude alta; chuva) que se relacionam ao termo clima para designá-lo. Embora as entidades que se relacionam e constituem o termo pertençam a categorias formais diferentes. Podemos neste exemplo encontrar pelo menos duas categorias formais distintas: **Processos climáticos** e **Métodos de datação** (coordenadas geográficas).
- Uma entidade pode estar relacionada à outra entidade para gerar um conceito em um domínio, ou a muitas entidades entre si. **As relações são de dois tipos:**
- **associativas simples:** $C_D = \{E_1, E_2\}$ – Conceito de um domínio é o resultado da associação da Entidade A e B e seus atributos. Quando as **relações são**

associativas compostas: $C_D = \{E1, E2, \dots, E_n\}$ - Conceito é o resultado da relação de mais de duas entidades e seus atributos.

O que estamos demonstrando é que precisamos estruturar os dados de uma linguagem natural para gerar conceitos e conseqüentemente organizar conhecimento. Tal conhecimento se manifestará através da informação disponibilizada para o usuário e do entendimento que o usuário possui sobre o conceito.

Para selecionarmos os conceitos que configurarão um vocabulário base para o domínio é necessário um processo de segmentação dos conceitos aparentes nas redes em macroproposições, esta tarefa está representada no diagrama, figura 22, que demonstra os caminhos traçados para o referido processo.

Figura 22 - Caminhos da segmentação dos conceitos da rede conceitual



Legenda: Processo de análise Resultado do processo por interpretação cognitiva
 → Seguimento

Fonte: construção da autora

Entendimento da segmentação (figura 22)

No caminho da segmentação e da análise conceitual utilizamos um processo mental esquemático que envolve:

- Leitura da rede de termos: ligação entre os pares de conceitos que englobam ideias ou macroproposições semânticas;
- Análise dos segmentos dos conceitos: ligação entre as sub-redes conceituais;
- Objetivação pragmática: contexto e uso do conceito para identificação dos conceitos: funcionais mentais, funcionais operacionais e agenciadores. Tais conceitos facilitarão a organização das relações: formal categorial, material/paradigmática e funcional sintagmática, propostas na Teoria do Conceito, Dahlberg (1978c), apresentada na seção 2.2. Segundo Cunha:

É através das relações paradigmáticas e sintagmáticas que se torna possível a reunião de unidades léxicas de uma língua em grupos estruturados, de tal modo que cada unidade seja definida pelo lugar que ocupa relativamente à posição das demais. Para tal, a formação do campo semântico possibilita à análise documentária uma melhor organização dos seus procedimentos de análise e a construção de instrumentos e produtos. (CUNHA, 1990, p.64)

O trabalho de Cunha apresenta uma proposta de segmentação textual onde a autora trabalha com macroproposições semânticas, identificando a carga semântica e ideológica por conceitos funcionais e mentais para explicitar informações contidas em diferentes textos. O nosso trabalho difere de Cunha, pois objetivamos a análise dos textos no sentido do fazer pragmático de entender os desdobramentos dos conceitos para organizar e representar domínios de conhecimento.

Neste sentido, quanto aos conceitos:

- os conceitos funcionais mentais revelam: análise, categorias, coordenações, síntese, subordinação e por fim globalizam o objeto/conceito. A principal função é designar o conceito globalmente.

- os conceitos funcionais operacionais revelam: sentido de contradição, extensão, identidade, atribuição e por fim particularizam o objeto/conceito. A principal função é condensar ou resumir o que o conceito é para o domínio.
- os conceitos agenciadores são conceitos possíveis que transitam da designação a determinação. Que não dependem dos limites de ordenação, que, sobretudo atendam as questões inter e multidisciplinares do conhecimento.

O conceito funcional operacional nomeia o objeto; O conceito funcional mental, também nomeia o objeto, como o sujeito processa o objeto de representação; O conceito agenciador é a palavra de “ordem”, a “saída possível”, resultado que pode advir da intuição. Neste aspecto, entendemos que no conceito agenciador encontra-se uma perspectiva social de análise que dá voz ao sujeito, sem a pretensão de individualizar, mas, sobretudo coletivizar.

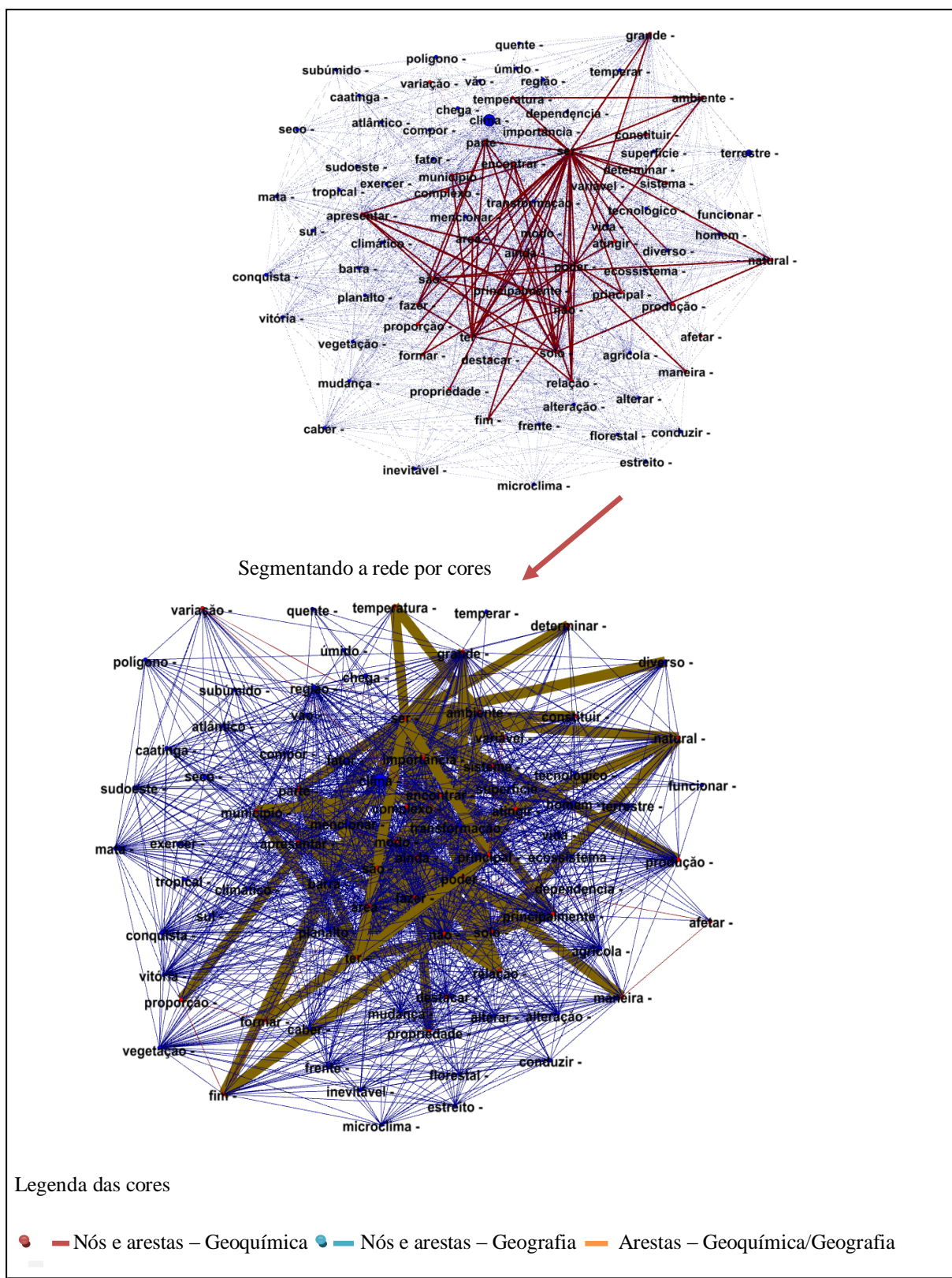
Precisamos de uma representação que ultrapasse as determinações pragmáticas, mas que represente a expressão do enunciado e dê voz ao sujeito que percebe e fala. O conceito agenciador é uma possibilidade do “*continuum imanente*”. O conceito agenciador é um conceito que não limita.

A partir do que está representado na figura 22 (vinte e dois) apresentamos a análise da rede conceitual do termo clima, construída a partir de dissertações e teses dos domínios da Geografia e Geoquímica, com o intuito de elucidar o diagrama.

O termo clima é designado como um padrão de diversos elementos atmosféricos. Portanto, seus estudos estão focados no domínio da Meteorologia, subárea das Geociências, de acordo com a tabela de áreas do conhecimento do CNPq. Sendo apropriado a outros domínios, áreas, subáreas e especialidades, no caso desta análise, da geografia e geoquímica. O conceito é utilizado em investigação de pesquisas que se relacionam ao contexto climático, ou o que o clima possa representar em relação a outros contextos.

Algumas relações podem ser visualizadas na figura 23 apresentada a seguir, na rede sobreposta dos domínios: Geografia e Geoquímica.

Figura 23- Sub-redes de palavras do termo clima – domínios: Geografia e Geoquímica



Fonte: construção da autora

Observando a sub-rede (figura 23) do termo clima nos dois domínios (Geografia e Geoquímica), percebe-se que existem relações (associações) com os termos e suas categorias conceituais, apresentados nos quadros que seguem, onde se inicia a análise conceitual:

Quadro 6 - Termo origem e Categoria formal e fundamental no domínio da Geografia

Termo	Definição do termo origem	Categoria do termo/Formal	Categoria fundamental
Clima	<p>O clima é o estudo médio do tempo para o determinado período ou mês em uma certa localidade. Também, se refere às características da atmosfera inseridas das observações contínuas durante um certo período. O clima abrange maior número de dados e eventos possíveis das condições de tempo para uma determinada localidade ou região. Inclui considerações sobre os desvios em relação às médias, variabilidade climática, condições extremas e frequências de eventos que ocorrem em determinada condição do tempo.</p> <p>http://www.inmet.gov.br/html/informacoes/curiosidade/tempo_clima.html</p>	Elemento atmosférico	Personalidade (PMEST)

Fonte: construção da autora

A categoria formal declara o termo enquanto conceito no domínio de conhecimento, enquanto a categoria fundamental permite pensar o conceito no plano ideacional em uma sequência útil e que traz o que está implícito.

Na sequência nos quadro 7 (sete) e 8 (oito) apresentamos os termos relacionados (destinos ou arestas) ao termo origem (clima) e que podem gerar outras sub-redes para análise.

Quadro 7 - Termos relacionados e categorias no domínio da Geografia

Termos	Definições dos termos destinos	Categorias dos termos/formal	Categoria Fundamental
Quente	Que tem ou produz calor – clima quente	Característica climática	Energia (processo)
Úmido	Impregnado de água ou vapor de água: tempo úmido	Característica Climática	Energia (processo)
Seco	Privado ou desprovido de umidade	Característica	Energia (processo)
Tropical	Relativo aos trópicos: clima tropical, opondo-se um período seco a um período úmido.	Característica	Energia (processo)
Região sudoeste	A Região Sudoeste está localizada entre 13° 02' a 16° 00' de Latitude Sul e 39° a 41°49' de Longitude Oeste (Figura 01), é uma das 15 regiões econômicas do Estado da Bahia propostas pela Superintendência de Estudos Econômicos e Sociais da Bahia – SEI – BA.	Espaço geográfico	Espaço
Vitória da Conquista	Vitória da Conquista é a 3ª maior cidade do estado [Bahia]	Espaço geográfico	Espaço
Planalto	Platô: extensão de terrenos mais ou menos planos, situados em altitudes elevadas em relação aos terrenos subjacentes. O mesmo que planalto.	Espaço geográfico	Espaço
Caatinga	A caatinga é um bioma típico do clima semiárido - quente, com baixa umidade, com chuvas escassas no inverno, e que se estende em grande parte do Nordeste brasileiro .	Ecossistema terrestre	Personalidade

**Quadro 7 - Termos relacionados e categorias no domínio da Geografia
(continuação...)**

Termos	Definições dos termos destinos	Categorias dos termos/formal	Categoria Fundamental
Ecosistema terrestre	O ecossistema é a unidade principal de estudo da ecologia e pode ser definido como um sistema composto pelos seres vivos (meio biótico) e o local onde eles vivem (meio abiótico, onde estão inseridos todos os componentes não vivos do ecossistema como os minerais, as pedras, o clima, a própria luz solar, e etc.) e todas as relações destes com o meio e entre si. Os ecossistemas são classificados de duas formas: em ecossistemas terrestres e ecossistemas aquáticos.	Ecosistema terrestre	Personalidade
Vegetação	Conceito atribuído ao grupo de vegetais existentes num dado espaço geográfico	Vegetal	Personalidade
Agrícola	Referente ao cultivo de plantas	Agricultura	Personalidade
Superfície terrestre	Parte externa da crosta terrestre	Crosta terrestre	Personalidade
Microclima	Corresponde às condições climáticas de uma superfície realmente pequena de duas espécies: natural e de planta	Classificação climática	Energia
Homem	Ser humano de maneira geral	Homo Sapiens Sapiens	Personalidade

Fonte: construção da autora

Quadro 8 - Termos relacionados ao termo origem e Categorias no domínio da Geoquímica

Termo origem	Categoria do termo origem	Termos relacionados	Categoria dos termos relacionados
Clima (ver definição – Geografia)	Elemento atmosférico	Temperatura ambiente	Grandeza física (Métodos e Técnicas)
		Solo natural	Solo (Personalidade)
		Município	Espaço
		Produção	Economia política (Personalidade)
		Ambiente	Meio em que se vive (habitat) – (Espaço)
		Solo	Produto do intemperismo (Personalidade)

Fonte: construção da autora

Como observado nos quadros 7 (sete) e 8 (oito), procuramos demonstrar a associação entre o termo clima (nó na rede) com outros termos (arestas) já categorizados, objetivando facilitar o entendimento da análise conceitual e conseqüentemente à definição de cada termo. Além disso, a análise facilita as relações ontológicas representadas em SOC como as ontologias, que trabalham com princípios triádicos de relações (Categoria 1- relator formal- categoria 2).

Como já descrevemos no capítulo 3 (três) desta tese, a base de uma ontologia é a taxonomia, sendo assim, a presença das relações categoriais nos quadros é fundamental para o entendimento da análise do conceito e da organização categorial.

No entanto é importante dizer que a proposta metodológica desta tese apoia-se em análise discursiva e cognitiva, ou seja, o reconhecimento do todo/discurso, que não se limita as unidades mínimas (conceitos), mas a um esquema mental de “leitura”, da segmentação textual, que se manifesta na rede. Isto significa dizer que procuramos identificar macroestruturas semânticas que nos levam a traços de descrição e argumentação.

Pretendemos uma hipótese de “leitura possível” através de conceitos que revelam a síntese da organização lógico-semântica e ainda dos traços da descrição, ou de como através dos conceitos pode-se perceber o conceito e suas proposições sem que se faça uma leitura total do texto. Consequentemente essa “leitura possível” deverá revelar o pensamento do autor e sua intenção argumentativa.

Para tanto, a partir de uma análise subjetiva da rede, com base no bom senso do analista/interprete cognitivo, conhecedor do domínio, parte-se para uma interpretação de caráter dedutivo e indutivo. Tal interpretação depende da ideia que se tem sobre o domínio e sobre o termo declarado na rede.

Entende-se por ideia a noção do que se deve representar sobre o domínio, representação intelectual de um objeto, do termo que se expressa por palavras, para em seguida atribuir-lhes conceitos-síntese (conceitos funcionais, conceitos agenciadores, conceitos funcionais mentais) “carregados” de sentidos e que darão corpo a um vocabulário base.

A representação do objeto/conceito só é possível através da definição conceitual em um determinado domínio. Tal representação deverá tornar explícito o conceito e explicá-lo no contexto que se revela através dos textos e das proposições dos autores/pesquisadores dos domínios de conhecimento, resultando no estado da pesquisa desses domínios. Isto é possível com a aplicação do método de raciocínio híbrido com princípios de raciocínio dedutivo e indutivo.

O método dedutivo parte de um princípio de raciocínio lógico e permite segundo Campos (2004, p.25) que se elabore “[...] mecanismos de abstração para pensar primeiramente o domínio/contexto, independente de pensar os elementos e suas relações; esta seria uma etapa posterior”. O método dedutivo aplicado nesta tese auxiliou o processo de categorização e o trabalho com meta-níveis conceituais, como os utilizados em taxonomias.

O método indutivo possibilita especificar o objeto de um contexto, a identidade conceitual e a sua dependência ou relação com outros objetos. Observa-se o conceito, e as relações com outros conceitos e por fim generaliza-se, para posteriormente defini-lo em um domínio, especificando-o, levando-se em conta o sujeito e seu pensar heterogêneo e múltiplo sobre o objeto.

A partir da análise discursiva e cognitiva e aplicação dos métodos dedutivo/indutivo é possível processar a linguagem natural, os passos do processamento estão apresentados no quadro 9 (nove).

Quadro 9 - Processamento da linguagem natural – quadro resumo

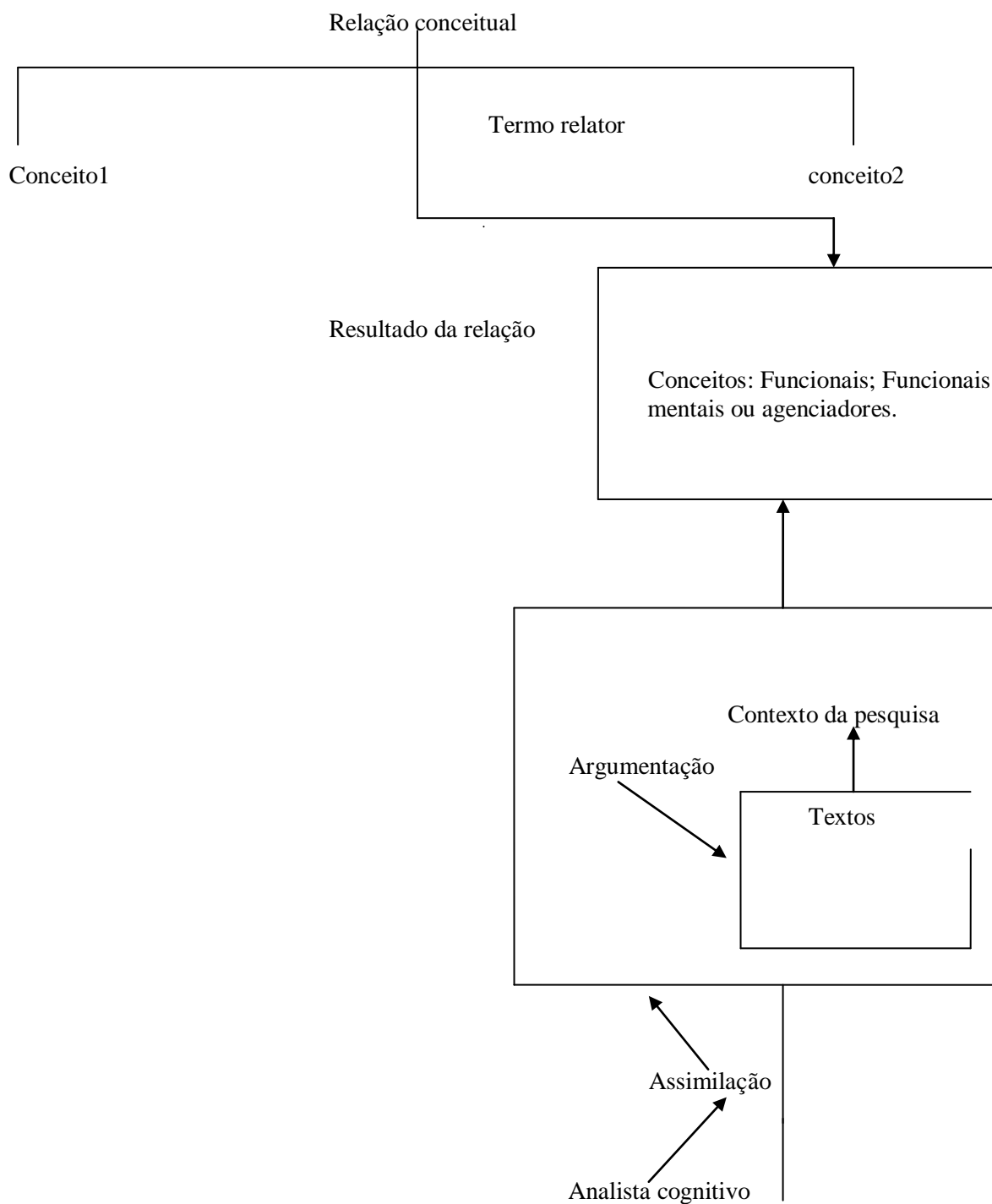
- Interpretar a amostra (rede de conceitos).
- Observar relações entre categorias de conceitos evidentes na rede.
- Identificar proposições: lógico semânticas e semânticas ideológicas.
- Gerar hipóteses discursivas a partir das relações conceituais e suas definições evidenciadas na rede a partir da análise.
- Verificar se tais hipóteses atendem ao contexto e conseqüentemente aos usuários.
- Possibilitar vocabulário base.

Fonte: construção da autora

O processamento da linguagem natural, só é possível se atribuirmos declarações explícitas ao conceito. Para tanto necessitamos analisar a argumentação evidenciada nos textos que revelarão a argumentação utilizada no domínio de conhecimento. Isso só é possível se o analista cognitivo, modelizador de domínios, para o caso desta pesquisa, compreender o contexto da pesquisa e, sobretudo, assimilar o que o usuário do domínio precisa encontrar na declaração conceitual.

O esquema apresentado na figura 24 (vinte e quatro) demonstra como se dá a assimilação conceitual por parte do analista cognitivo.

24 - Esquema das relações conceituais e da assimilação conceitual



Legenda: \longrightarrow seguimento sujeitos a interpretação

Fonte: construção da autora

Entendimento da figura 24

- A tríade das relações conceituais resulta em conceitos que podem ser funcionais, funcionais mentais ou agenciadores.
- A assimilação conceitual só é possível quando se entende a argumentação utilizada nos textos.
- O entendimento da argumentação pode ser percebido através da leitura das redes de termos, construída a partir de textos.
- Através das relações conceituais podem-se declarar os assuntos dos documentos.
- A declaração dos assuntos dos documentos pode ser traduzida através de conceitos.

Ao propormos o Modac, cogitamos a possibilidade de sair da “zona de conforto” dos modelos tradicionais de análise para sistemas de organização e representação de conhecimento e da imposição de hierarquias e relacionamentos rígidos para determinação das relações entre conceitos. A aplicação do modelo pressupõe a adoção da variação contínua de pensamento e conseqüentemente da linguagem representada.

Passamos então ao entendimento da geração dos conceitos a partir das redes e a estruturação dos dados por elas revelados.

5.2 GERAÇÃO DE CONCEITOS A PARTIR DA AMOSTRA DA REDE – DADOS ESTRUTURADOS

A partir de uma amostra de textos foram escolhidas palavras a serem analisadas através da incidência-fidelidade observada nos nós da rede e conseqüentemente com termos relacionados, ou quão fiel essa palavra está para o discurso impresso nos textos. Escolhemos uma coocorrência de pares abaixo de 0,1% do texto, aumentando possibilidades de encontrar maiores relações entre termos e facilitando a interpretação das pesquisas desenvolvidas no domínio, através de conceitos encontrados na amostra.

A rede clima possui um coeficiente de aglomeração de 0,37, um pequeno grupo conectado, mas que revela forte conexão a outros nós o que facilita o entendimento das relações entre termos, possibilitando uma interpretação acerca dos conceitos que tais relações provocam. E ainda uma representação categorial mais aproximada dos espaços lógicos e ontológicos dos conceitos dos domínios.

Após a observação da frequência com que as palavras aparecem nas amostras dos domínios (redes dos domínios), escolhem-se as palavras/termos que serão analisadas. Em seguida aplica-se o modelo de análise terminológica em amostra selecionada (sub-rede) de palavras.

A análise terminologia é necessária para o entendimento de como a palavra se estabelece enquanto termo ou conceito, ou seja, tendo significado para o contexto, traduzindo as pesquisas desenvolvidas no domínio de conhecimento e conseqüentemente os conceitos que os usuários (indexadores) buscam encontrar em um SOC. Essa análise objetiva atender as necessidades de informação dos usuários (pesquisadores) em sistemas de busca.

Usuários indexadores são profissionais que trabalham organizando os conteúdos a serem descritos através de termos contidos em itens informacionais como: documentos, livros, materiais especiais: materiais cartográficos, manuscritos, música impressa, música, gravação de som, filmes cinematográficos, materiais gráficos, recursos eletrônicos, artefatos tridimensionais, realia, microformas e recursos contínuos. Geralmente esses indexadores são Arquivistas, Bibliotecários e Museólogos.

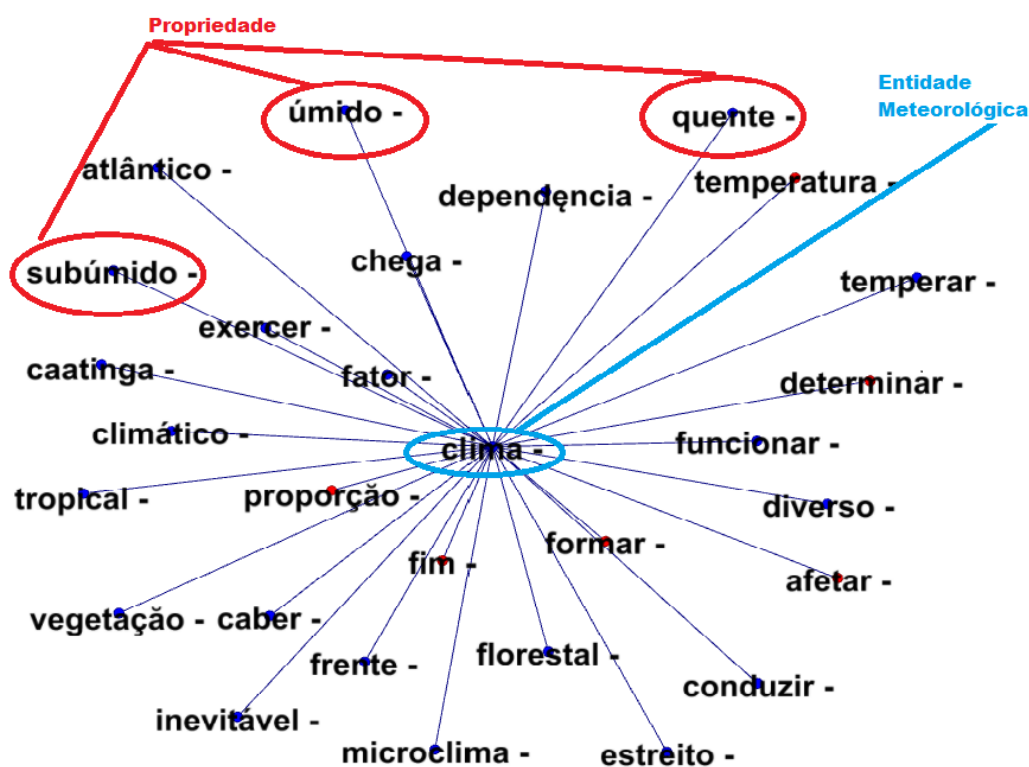
Os usuários pesquisadores a que nos referimos são cidadãos que necessitam de informações que atendam aos seus anseios, ou intenções de pesquisa. Esses usuários tanto podem ser pesquisadores dos domínios analisados, ou pesquisadores de outros domínios de conhecimento, que de uma forma ou de outra necessitam de informações que atendam as suas necessidades de pesquisa.

Analisando o resultado da amostra da **rede clima** (figura 25), pode-se concluir que a interpretação cognitiva a partir da rede, com base nos relacionamentos entre conceitos e suas categorias formais e fundamentais pode levar a uma hipótese discursiva, formando uma

Para o entendimento do conceito é primordial compreender as relações entre os termos, e principalmente às suas categorias formais que se estabelecem a partir do uso do termo em um determinado contexto e conseqüentemente do domínio em que esse termo se integra.

O entendimento do conceito a partir da segmentação da rede conceitual dar-se-á, quando analisarmos o termo (nó) e suas relações com outros termos (arestas) que o determinam na proposta discursiva e no contexto da pesquisa. Observamos o estabelecimento do conceito e sua denotação através da rede e identificamos possibilidades de argumentação, como exemplificado na figura 26.

Figura 26 – sub-rede do termo clima e suas relações



Fonte: Construção da autora

Observamos que os termos: **quente**, **úmido**, **seco** e **subúmido** (arestas) no contexto analisado pertencem à categoria: **Propriedade** que confere ao clima um **estado**. Isto foi possível porque o termo foi analisado na **rede clima**, e no contexto de um texto da **geografia**. O mesmo não aconteceria se analisássemos esses mesmos termos em outros contextos ou domínios.

Exemplificando:

- o clima pode ser quente, úmido, seco e subúmido a depender da localidade e das variações e características atmosféricas de uma região, mas a variação é um estado que ocorre em um determinado processo ou condição.
- mas a temperatura de um indivíduo pode estar quente, ou ainda acima dos padrões normais de temperatura a depender de condições ambientais e de funcionamento do organismo e da saúde desse indivíduo.
- um tecido de uma roupa pode estar seco, úmido ou subúmido a depender do estado e processo que essa roupa tenha passado.
- um alimento pode estar quente a depender do processo de aquecimento pelo qual o alimento tenha passado estando assim em um estado aquecido.

O que não varia é a categoria fundamental do termo. Quente, úmido, seco, subúmido dão particularidades ao clima. Segundo a PMEST o termo clima é uma manifestação da categoria Personalidade. O termo úmido, enquanto propriedade é uma manifestação da categoria Matéria.

As categorias são uteis para identificarmos: as ideias que diferenciam um conceito pela extensão decrescente, ou subordinação em uma categoria, pela relevância desse conceito através da sua coordenação com classes de conceitos e por sua sequência útil e consistente na classe conceitual.

A importância do entendimento das categorias em um trabalho que visa SOC é fundamental, porque a partir delas é que se podem construir relações lógicas e ontológicas sobre o termo como pode ser observado no quadro 10 (dez).

Quadro 10 – Identificação dos conceitos úmido e clima em relação ao domínio Meio Ambiente

Classes de conceitos	Hierarquia	Relacionamento	Função	Instância	Axiomas
Úmido (tipo de estado)	Propriedade	Particularidade	Úmido é um estado	Úmido pode ser...	Úmido é um estado do clima
Clima (estudo do tempo)	Entidade meteorológica	Característica da atmosfera	Clima é um estudo	Clima pode ser ...	Clima é um estudo médio do tempo em uma determinada localidade

Fonte: Construção da autora

As categorias fundamentais e formais não são visualizadas na rede conceitual, mas a partir das relações entre os termos e da análise conceitual que essas relações permitem é possível identificá-las e propor hipóteses discursivas.

Exemplos de hipóteses discursivas e argumentativas a partir da sub-rede clima e da análise conceitual:

- O clima de Vitória da Conquista é tropical de altitude.
- Vitória da Conquista possui região agrícola.
- Vitória da Conquista é um município baiano, localizado no sudoeste do estado.
- O clima depende de características climáticas locais.
- O clima de Vitória da Conquista propicia alta pluviosidade.
- O clima influencia no solo e conseqüentemente na produção agrícola local.
- Vitoria da Conquista é uma região de planalto.

Tais hipóteses podem se configurar como afirmações que podem ser comprovadas quando da leitura dos argumentos dos textos.

Argumentações retiradas do texto de Di Lauro (2011) (CP1 do apêndice A) que comprovam as hipóteses:

- [...] planalto de Vitória da Conquista.
- Vitória da Conquista está situada a sudoeste do Estado da Bahia.
- A região de Vitória da Conquista, compreende os municípios de Barra do Choça, Planalto e Poções.
- Vitória da Conquista um centro polarizador do Território de Identidade.
- [...] Vitória da Conquista a ocupar o lugar de terceira economia do interior do Estado da Bahia.
- Em Vitória da Conquista, a população predomina na zona urbana, contudo, não se pode perder de vista que, os três municípios apresentaram um crescimento populacional considerável, e, possivelmente, deve-se tomar por base, como principal fator de análise, a implantação do café na região.
- Assim como em outras áreas agrícolas, a variável clima, precipitação e temperatura, foi primordial para a introdução da lavoura cafeeira em Vitória da Conquista.

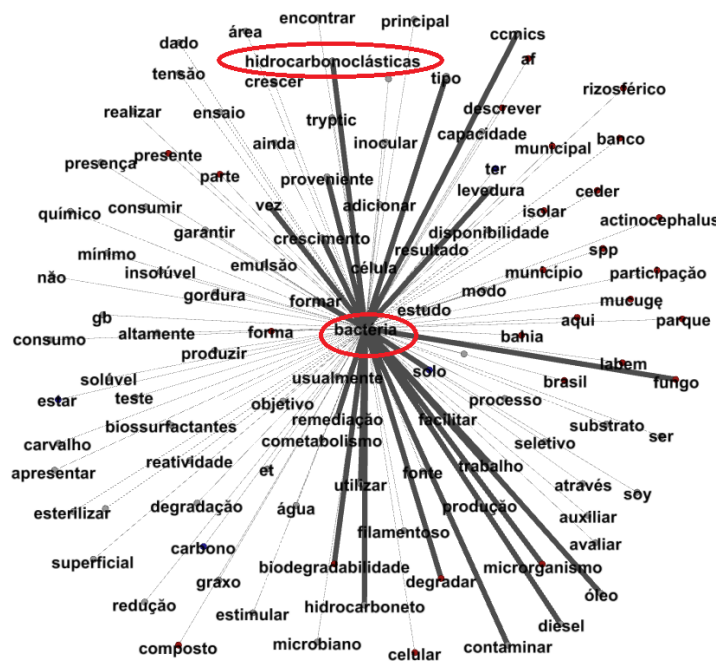
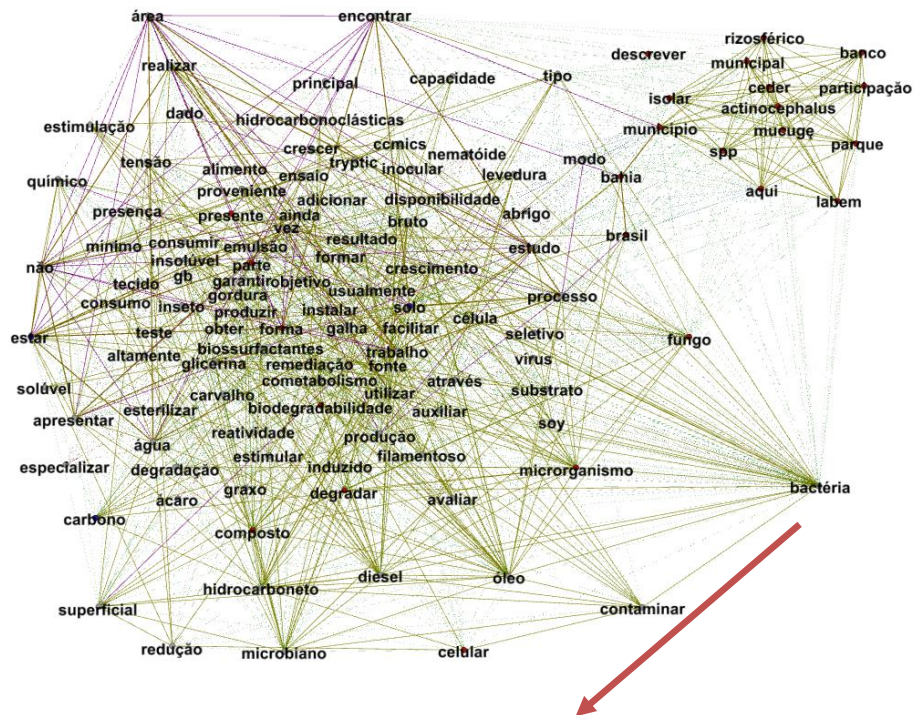
Outras argumentações encontradas no texto, como a que segue, podem ser analisadas por outras sub-redes de termos como, por exemplo, as de Vitória da Conquista, solo e latossolo.

- A oeste de Vitória da Conquista encontra-se uma pequena porção do latossolo vermelho amarelo distrófico.

Na próxima amostra, rede do termo bactéria, (figura 27) nos domínios da Geologia e Geoquímica apresentamos uma análise terminológica com ênfase nos domínios. A análise evidencia questões inter e multidisciplinares que influenciam na formação de um conceito.

Figura 27 – sub-redes do termo bactéria

a) Sub-rede



b) Segmentação da sub-rede

Fonte: construção da autora

Definição terminológica

- **Bactéria** - microrganismos procariontes não apresentam um núcleo definido, estando o seu material genético compactado e enovelado numa região do citoplasma chamada de nucleóide. As bactérias apresentam uma membrana plasmática recoberta por uma parede celular. Diferentes das células eucarióticas, nas bactérias não aparecem organelas delimitadas por membranas.

O termo/conceito nas amostras dos domínios analisados foi utilizado em textos de 02 (dois) domínios (Geologia e Geoquímica) com a mesma apropriação para um tipo de Bactéria (Bactéria hidrocarbonoclástica).

Definição conceitual

- **Bactéria hidrocarbonoclástica** – bactéria que utiliza hidrocarbonetos como fonte de carbono.

Existem termos na amostra, rede analisada, que se relacionam com o conceito como:

no domínio da Geologia

- Termos relacionados: Biosulfactantes, substrato, processo e reatividade.

no domínio da Geoquímica

- Termos relacionados: nematoide, ácaro e abrigo.

Termos da Rede superposta GLGQ (Geologia e Geoquímica)

- Óleo, diesel, hidrocarbonoclástica, remediação, fungo, graxo, degradar, biodegradação, levedura, contaminar e hidrocarboneto.

Analisando o termo por questões e aspectos inter e multidisciplinares pode-se justificar o uso do termo na área da Geologia e no domínio (especialidade) da Geoquímica, por terem relações interdisciplinares em seus estudos, o que está impresso na análise terminológica e conceitual.

Na construção de um conceito, no caso **Bactéria hidrocarbonoclástica** ocorrem à junção de 03 (três) termos: bactéria + hidrocarboneto + clástica. O conceito surge quando do uso da bactéria em processos de **biorremediação** em estudos relacionados à contaminação por óleo diesel ou graxo (compostos derivados do petróleo).

O processo de **Biorremediação** só é possível, quando se aliam estudos: da Geologia e suas práticas de análise de solo, rocha, ambientes aquáticos e terrestres, dentre outros; da Biologia em conjunto com a biotecnologia, em pesquisas com bactérias e seus estudos de fixação e devolução de nitrogênio na atmosfera, e ainda da Química com estudos de carbono.

Um trabalho interdisciplinar dessa natureza pode, por exemplo, ser criado através de processos: biológicos, físicos e químicos. E métodos e técnicas como um método de Datação como os de biodisponibilidade ou ainda indicadores de desenvolvimento sustentável.

A **bactéria hidrocarbonoclástica** é um tipo de bactéria que, por sua vez, é um organismo que se caracteriza como entidade bioquímica, por conter hidrocarbonetos em sua composição. Além de outros componentes como clástica que, em sua constituição, possui fragmentos de rochas que são agregados naturais sólidos compostos de um ou mais minerais.

Existem ainda entidades que por associação podem ser elementos fundamentais a argumentação, exemplo dessas entidades e suas categorias conceituais:

- Biorremediação e Biodegradação – são processos Biológicos
- Contaminação ambiental – processo antropogênico – impacto ambiental
- Óleo diesel (composto formado principalmente por átomos de carbono) - derivado do petróleo – entidade físico-química – substância

E ainda a argumentação pode ser extraída de uma fonte de definição como a de Silva; Garanhan (2006) Apêndice B que utilizamos para descrever os conceitos.

- **Biodegradabilidade**; diferentes dos surfactantes químicos os **biossurfactantes** são facilmente degradáveis na água e no solo;
- O maior mercado para os biossurfactantes é a **indústria petrolífera**, onde são utilizados na produção de petróleo ou incorporados em formulações de óleos lubrificantes.

- Outras aplicações incluem **biorremediação** e dispersão no derramamento de óleos, remoção e mobilização de resíduos de **óleo** em tanques de estocagem e a recuperação melhorada de **petróleo**.
- Os **hidrocarbonetos** naturais são compostos químicos constituídos por átomos de carbono (C) e de hidrogênio (H), aos quais se podem juntar átomos de oxigênio (O), nitrogênio (N) e enxofre (S) dando origem aos compostos orgânicos, classificados de acordo com seus grupos funcionais.

Para o entendimento de um conceito e o seu uso em um determinado domínio é necessário entender suas relações ontológicas o que pode ser evidenciado pelas relações categoriais dos termos relatores/descriptores do domínio.

As relações categoriais de ordem material/paradigmática ordena o conceito por gênero e espécie, assim como em renques de conceitos, particularizando esse conceito. No caso do Conceito Bactéria hidrocarbonoclástica aliam-se a ordem paradigmática um processo funcional sintagmático, que ocorre quando da relação de termos de diferentes entidades para determinar o conceito.

O resultado das relações das entidades para o conceito se dá entre as categorias: **Entidade biológica + Entidade química + Entidade fisico-química = Bactéria hidrocarbonoclástica**. Neste caso conceito do domínio dar-se-á por uma relação composta onde: $CD = \{E1, E2, \dots, En\}$. Este resultado proposto só é possível devido ao Modac, na perspectiva da tríade de análise terminológica que apresentamos no capítulo 4 (quatro), onde trabalhamos com: a análise do contexto (domínio e texto), definição conceitual e interpretação cognitiva.

Importante evidenciar que após os resultados obtidos, os conceitos se configuram como “candidatos” a compor um SOC, que deverão ser analisados e validados por um corpo de especialistas do domínio. Procuramos trabalhar nas amostras com termos que foram anteriormente avaliados por especialistas quando efetuamos o trabalho de mestrado, com o intuito de otimizar tempo para conclusão da tese e melhor análise conceitual.

Na próxima seção passamos a discutir os resultados alcançados e apresentamos considerações a partir das análises efetuadas.

5.3 RESULTADOS E DISCUSSÃO SOBRE O MODELO DINÂMICO DE ANÁLISE CONCEITUAL – MODAC

Observando o resultado das relações conceituais com base no texto e no contexto da pesquisa, nas redes dos domínios de conhecimento é possível declarar alguns conceitos como possibilidades de composição de um vocabulário base, que objetiva facilitar a construção do SOC do domínio. Analisando esse vocabulário o modelizador poderá realizar categorizações conceituais mais aproximadas da realidade. Como também sistematizar as relações entre conceitos, construindo assim uma ontologia de domínios de conhecimento adequada às necessidades do usuário e ao domínio.

Exemplos de análise conceitual e declaração de assunto nos domínios analisados:

Vitória da Conquista (Município Baiano) – conceito declarado por um qualificador, afinal os termos vitória e conquista, mesmo que usados juntos como termo composto, sem um qualificador seria apenas uma junção de termos, contudo, não declara um significado para um domínio de conhecimento. Não possui uma função ou um agenciamento a não ser por um indivíduo pertencente à região ou ao estado a que o conceito está relacionado.

Além disso, é fundamental a identificação da argumentação, com base nos textos e contextos como também a partir de relações conceituais e de definições conceituais:

- Vitória da Conquista é um município – categoria Espaço geográfico;
- Vitória da conquista é um município do estado da Bahia, Brasil. (relação entre conceitos de uma mesma categoria – Espaço geográfico).
- **Vitória da Conquista (Município baiano)** é um conceito cujo resultado se dá por relações do tipo simples, $CD = \{E1, E2\}$.

Um outro conceito importante a ser declarado: **clima tropical de altitude**, afinal esse tipo de conceito foi utilizado na argumentação textual como influenciador da agricultura local.

Clima tropical de altitude – resultado das relações entre conceitos de categorias fundamentais diferentes (Categoria Personalidade e Energia), mas por uma relação entre as

entidades: elemento atmosférico, característica e grandeza física, sendo assim, em uma associação composta, logo: $C_D = \{E_1, E_2, \dots, E_n\}$.

É relevante informar que na dissertação de Di Lauro, 2001, referenciada na CP1, apêndice A, as **palavras chave** indicadas são: **Agricultura tecnológica. Sociedade. Natureza. Globalização. Transformações socioambientais.** Curioso é que Vitória da Conquista é utilizado no título do trabalho e dá todo o sentido à pesquisa. Assim como clima que transforma o espaço nos sentidos: social, cultural, econômico e ambiental. Tanto o local, quanto o clima da região não foram indicados como palavras-chave.

O que reafirma a importância da nossa proposta que é a de trabalhar analisando textos completos a partir da rede de conceitos que são geradas mediante redes semânticas. Essas redes possibilitam a análise conceitual e a segmentação dos argumentos utilizados em textos provenientes de pesquisas como os de teses e dissertações.

Quando analisamos apenas palavras-chave indicadas pelos autores nos limitamos a uma análise sob o ponto de vista do produtor do conhecimento sem levar em conta o usuário que vai à busca da informação em bases de dados e ainda os conceitos que se estabelecem nos domínios através das relações com outros conceitos.

A partir desta premissa apresentamos alguns significados semânticos dos termos analisados para os domínios.

Significado semântico dos termos para os domínios analisados

O clima é um conjunto das condições atmosféricas que caracterizam uma região. No âmbito das pesquisas dos domínios analisados o estudo das variações climáticas é importante para análise do solo e da produção das regiões estudadas.

Sendo assim é importante que se trabalhe com a categorização de todos os tipos de clima para compor o vocabulário base da ontologia, em domínios que efetuam pesquisas em meio ambiente. Para justificar a nossa afirmação apresentamos um argumento utilizado por Carvalho (2009), CP2, referência completa no apêndice A, em um texto de tese da geologia que compõe a amostra apresentada:

“[...] elementos traço parecem estar menos disponíveis em solos altamente intemperizados como aqueles derivados dos sedimentos da Formação Barreiras que em solos derivados de clima temperado.”.

Palavras-chave determinadas pelo autor da tese: Índice geoquímico. Valores orientadores. Teores naturais.

Mais uma vez percebemos que as palavras-chave determinadas pelo autor nem sempre revelam aspectos fundamentais da pesquisa, nem tampouco, para declaração de pesquisas do domínio. Na tese acima citada, ocorreram ainda interpretações de concentrações de metais em solos derivados de sedimentos de formação barreiras.

Observa-se que o tipo de solo ou tipos de solos, onde tais sedimentos foram encontrados e os municípios de pertencimento desses solos não foram citados nas palavras-chave, o que dificulta profundamente a indexação por parte dos bibliotecários, que não podem efetuar uma leitura aprofundada de todo item a ser indexado. Por esse motivo reafirmamos a importância de modelos dinâmicos de análise conceituais para auxiliar a estruturação de SOC.

Sobre o termo bactéria hidrocarbonoclástica podemos notar uma relação associativa de entidades: bactéria + hidrocarbono+clástica. Portanto uma relação onde: $C_D = \{E1, E2, \dots\}$.

Na proposta do Modac, de análise conceitual por proposições, ou identificação de macroproposições semânticas observamos que:

- existe no termo **bactéria** uma proposição semântica mental, ou **conceito funcional mental**, afinal no senso comum entendemos bactéria como um micro-organismo que pode causar enfermidade, ou uma contaminação, dentre outros aspectos e funções imprescindíveis à estudos sobre meio ambiente e vida no planeta .
- No termo **hidrocarbonoclástica** existe a proposição lógico semântica – hidrogênio + carbono = $H + C^2$ do hidrocarboneto. E ainda a junção de outro conceito: clástico (composto de fragmentos de outras rochas). Entendemos que a primeira composição do termo tem origem na química. A segunda

composição do termo é complexa, porque existe no termo clástica muitas variáveis envolvidas quanto a composição de fragmentos de rochas. Mas pode-se afirmar que é um **conceito funcional e que concede uma característica a bactéria**.

- Quando determinamos que uma **bactéria hidrocarbonoclástica** é uma entidade bioquímica a entendemos como um **conceito agenciador** que se contrapõe ao referencial maior e que não se categoriza hierarquicamente pelo conceito primeiro, pois é fruto de um processo conceitual interdisciplinar, é uma categorização possível de um conceito que se estabelece sem estruturas homogêneas, mas sobretudo que surge de um *continuum* do mundo fenomênico.

O que queremos dizer é que se analisarmos o referencial maior do conceito, no caso bactéria, a categorizaríamos enquanto termo pela sua característica biológica, mas em uma perspectiva de análise conceitual, devemos obter uma análise com asserções corretas acerca do item referente (conceito). Isso só é possível quando analisamos as propriedades e características essenciais e acidentais do conceito.

O conceito bactéria hidrocarbonoclástica pode ser categorizado de acordo com os princípios da relação funcional-sintagmática, ou seja entre conceitos de diferentes categorias. Sendo assim, entidade biológica bactéria é suplementada pela entidade química hidrocarbonetos e por uma entidade físico-química do componente clástica. O Componente clástica se deriva de fragmentos de rochas. Sendo assim a classificação do termo se torna complexa.

O que confirma o que foi dito por Dahlberg (1978c) sobre as categorias estruturantes, que fornecem a união do todo, mas que podem ser flexíveis. Para tanto é importante a definição nominal, que fixa o sentido do termo, e a definição real que delimita a intensão conceitual. Aí encontra-se o conhecimento contido em uma linguagem e porque não dizer, é por esse caminho da flexibilidade, que encontra-se o conceito agenciador.

Neste caso o conceito agenciador desta natureza conceitual é entendido pelo usuário do domínio, um conceito advindo do experimento. Mas é importante deixar claro que

existirão outros conceitos agenciadores que surgirão também no senso comum, ou simplesmente do empirismo, que devem ser pensados em construção de SOC, para dar voz aos usuários que merecem ser ouvidos quando vão em busca de informações em bases de dados. Há de se pensar em um modelo híbrido, não só de análise como de estruturas de busca para bases de dados, principalmente as disponibilizadas na internet.

Sabemos que na atualidade existem sites que trabalham com ontologias e folksonomias. Uma forma colaborativa de etiquetagem coletiva, onde os usuários inferem e atribuem rótulos ou metadados aos documentos.

Sugerimos como uma saída possível uma análise de argumentos do usuário que provocaria um conceito agenciador, que pode ser entendido não por uma forma arbórea ou hierárquica, mas sobretudo em um formato endeitado, rizomático. Uma concepção conceitual possível que transite pelos argumentos das macroproposições semânticas às relações conceituais. Uma forma aberta para socializar o conhecimento.

Acreditamos que esta proposta é possível se os modelizadores de domínios, aqui entendidos como profissionais da CI e CC criem nas suas bases de dados um vocabulário base que armazene os conceitos identificados através das macroproposições semânticas e das possibilidades de argumentação advindas das definições e relações conceituais.

Além disso o trabalho com uma base que contemple as questões dos usuários quando da busca à informação e suas características enquanto pesquisador deva ser levado em conta, no processo de encontrar uma forma híbrida de OC. A interseção das bases possibilitará uma linguagem possível e inteligível para todos.

Reafirmamos que taxonomias e ontologias são fundamentais para a organização de domínios de conhecimento, mas que precisam de análises conceituais dinâmicas. Tais análises devem levar em conta os aspectos das entidades nas quais os conceitos descendem e se relacionam para formar outros conceitos.

Por esse motivo o Modac propõe para a organização do conhecimento analisar:

- o domínio ao qual o conceito pertença primitivamente, onde o conceito tem sua origem epistemológica;

- o domínio no qual o conceito esteja sendo empregado;
- os textos e contextos da pesquisa de acordo com argumentos do produtor de conhecimento;

E que a partir dessa análise:

- se definam os conceitos;
- se proceda a análise conceitual com base nas categorias formais e fundamentais as quais os conceitos estejam relacionados.

Por uma agregação desses processos é possível uma análise mais consistente e um modelo flexível e dinâmico que acompanhe o movimento das relações conceituais, pois se as relações se quebram a entidade perde a identidade e passam a ser pedaços desagregados.

Não existe um conceito unitário ou uma só interpretação, como não existe uma ciência unitária, mas sim diferentes abordagens a partir dos níveis de integração que se envolvem na construção do conhecimento. Mas a integração também não garante a perfeição de uma definição conceitual. É preciso o olhar múltiplo sobre o objeto e mais ainda trabalhos interdisciplinares entre a CI, CC e DC para construção de SOC que atendam às questões inter e multidisciplinares de domínios. A nossa crença é que um Modelo dinâmico de análise conceitual possa contribuir para essa construção. Essa é a nossa tese.

Apresentamos na sequência as considerações finais deste trabalho.

6 CONSIDERAÇÕES FINAIS

No início do Doutorado em Difusão do Conhecimento, tínhamos como ideia inicial o estudo de bases teóricas para organização do conhecimento e desenvolvimento de ontologias. A ideia inicial foi amadurecida com os estudos realizados nos componentes curriculares cursados, e principalmente quando encontramos o aporte epistemológico, teórico e metodológico da AnCo.

Entendemos que pesquisas no âmbito da AnCo cresciam em todas as áreas do conhecimento, e nas Ciências da Informação e da Computação isso era um fato que agregava sentido para questões como: análise de domínios de conhecimento, análise de uso e usuários de informação e análise conceitual.

A ideia inicial do estudo de tese foi então ampliada diante desta constatação e decidimos desenvolver um modelo de análise conceitual com o “amparo” da AnCo. Modelos de análise conceitual de certa forma já se encontravam contemplados em pesquisas e estudos desenvolvidos por analistas de informação, analistas de computação e principalmente por profissionais que trabalham com modelagem de dados e modelização de domínios de conhecimento.

Assim optamos por um Modelo Dinâmico de Análise Conceitual (Modac). Chamamos de dinâmico porque intencionamos que a análise proposta possa contemplar: definições nominais que fixem o sentido do conceito, definições reais dos conceitos para o domínio de conhecimento no qual o conceito é revelado e definições enunciadas por usuários.

As questões de enunciação do usuário precisam ser priorizadas em SOC, afinal um sistema desta natureza é criado para atender aos sistemas de recuperação de informação como bases de dados referenciais e buscadores como os utilizados na Internet. Como também auxiliar indexadores como: bibliotecários, arquivistas, documentalistas e museólogos.

Diante desta opção elegemos trabalhar com a Teoria do Conceito, como aporte teórico para análise e entendimento das categorias conceituais e incluímos a Teoria da Classificação Facetada com seus princípios de *continuum*. Quanto ao uso da análise do discurso representativo de Foucault, com o objetivo de analisar a enunciação discursiva de textos. E a

AnCo, principalmente no que concerne, à interpretação dos conceitos com ênfase na percepção do usuário e na intenção de socialização do conhecimento.

A Teoria da Classificação Facetada nos permitiu através do entendimento do plano das idéias e seus postulados a diferenciar um conceito através de sua extensão decrescente que pode ser orientada por uma sucessão de classes de conceitos. Além disso é possível através dos renques e cadeias entender a relevância e modulação do conceito, ou a sua intensidade para o domínio de conhecimento, que varia proporcionalmente a depender da forma como se analisam os assuntos desse domínio. Por esse motivo quando definimos os conceitos da rede no capítulo 5 (cinco), trabalhamos com facetas e com a sequência da PMEST.

O entendimento da Teoria do Conceito nos permitiu a fixação do conteúdo do conceito, quando assumimos que o termo representa o referente (conceito), entendemos as características do referente, desse modo podemos delimitar o seu uso no discurso.

Da análise do discurso apresentada por Foucault recuperamos a ideia de como trabalhar na interpretação das redes conceituais, ou seja, rastreando as palavras pensadas pelo autor do texto para além de um ato de escrita, sobretudo, entendendo-as com a pluralidade das palavras unidas pelo sentido, pela semântica que também podem ser ditas em outros lugares, porque dialogam com outros contextos.

Através dos estudos propostos nos componentes curriculares Análise Cognitiva I e II (Anco I e II) e dos estudos do CAOS, entendemos a AnCo como uma possibilidade de trabalharmos o modelizador de domínios de conhecimento como um analista cognitivo, cujo trabalho de análise conceitual é pensado para o indivíduo que procura, identifica e seleciona informação. Se nos processos de OC encontrarmos um caminho que dê voz ao usuário estamos por fim colocando o sujeito pesquisador e o que busca a informação no interior de todo processo de difusão e socialização do conhecimento.

Além dos aportes encontrados na TCF, TC, Análise do discurso e AnCo, trabalhamos o modelo de análise sempre pensando em taxonomias e ontologias de domínios. Por esse motivo quando analisamos os conceitos procuramos pensá-los em suas categorias fundamentais e formais e nas relações entre esses conceitos ou relações entre entidades, com o

intuito de demonstrar que o modelo que propomos é para auxiliar a construção desses sistemas de organização e representação de conhecimento.

Assim chegamos a uma proposta de análise conceitual que resultou do relacionamento entre entidades e demonstramos que o conceito de um domínio de conhecimento pode ocorrer de forma associativa simples entre duas entidades, $C_D = \{E1, E2\}$, ou de forma associativa composta, quando o conceito se associa a mais de duas entidades, neste caso $C_D = \{E1, E2, \dots, E_n\}$.

Demonstramos que o Modac está fundamentado em uma proposta dinâmica de análise, desde seu primeiro passo de construção quando pensamos em trabalhar com redes conceituais geradas a partir da concepção das redes semânticas. Porque entendemos que a leitura dos documentos gerados nos domínios não poderia ser realizada na sua totalidade. Seria impossível para um analista da informação e um modelizador de domínios efetuarem a leitura integral, dado o volume de textos que é gerado quase que diariamente pelos domínios de conhecimento.

Além disso, por experiência de estudos anteriores de dissertação de mestrado, como também em trabalhos realizados no projeto desenvolvido em AnCo I e II sabíamos que analisar palavras-chave e resumos de textos não revelaria, de forma atinente, o estado da pesquisa de domínios de conhecimento, porque infelizmente os autores de texto de natureza científica negligenciam na confecção dos resumos e na determinação de palavras chave. Na análise com a segmentação textual encontramos a possibilidade de analisar o argumento e encontrar maiores possibilidades de declarações conceituais.

O Modac tem como base para a análise conceitual, as redes conceituais. Uma rede que se auto-representa por apresentar as relações ente pares de termos que são utilizados em sentenças de um texto. Através da representação encontrada nas rede torna-se possível a análise conceitual a partir do que chamamos de segmentação do argumento.

Seguindo a dinâmica do modelo definimos os conceitos e estruturamos esses conceitos através das categorias formais e fundamentais com base nos domínios de conhecimento, para em seguida adequá-los às questões de busca do usuário. O usuário de cada domínio de conhecimento possui características inerentes ao sujeito que percebe e usa as bases de dados,

desse modo, precisamos conhecer o domínio e seus usuários para realizar análises conceituais. Cabe nesse momento evidenciar que o trabalho que realizamos nesta tese contou com a experiência de onze anos trabalhando como Bibliotecária na área das Geociências e conhecendo às necessidades informacionais dos usuários, delimitadas pelo seu perfil de busca.

Quando representamos conceitos concebidos em domínios inter e multidisciplinares a atenção às relações conceituais devem ser redobradas, principalmente se pretendemos construir taxonomias e ontologias de domínios, porque uma entidade pode conter aspectos de duas ou mais entidades a depender do enfoque teórico, metodológico e da aplicação empírica que os estudos e pesquisas carregam. Isto ficou claro quando analisamos os termos e confirmou os pressupostos apresentados na introdução deste trabalho.

O que propomos nesta tese é somarmos perspectivas para atender à diversidade do modo de pensar do indivíduo sobre o conhecimento de um dado domínio. Dessa maneira poderemos representá-lo. Consideramos que o desafio do profissional que trabalha organizando e disponibilizando informações é o de ultrapassar modelos tradicionais e conhecer novos paradigmas, metodologias e teorias.

Quando decidimos por apresentar o Modac como objetivo principal da tese, intencionávamos a elaboração de um modelo de análise dinâmico que viesse a contribuir para a construção de SOC, acreditamos que cumprimos o nosso objetivo.

O Modelo proposto identifica macroproposições conceituais a partir da análise das redes de termos. Essa análise é possível quando interpretamos as ligações que se dão entre os termos, nós visíveis na rede e ainda quando analisamos as sub-redes que se formam através das ligações entre nós e arestas.

Através das arestas podem-se indicar os locais de pertencimento dos conceitos empregados em uma amostra de textos por domínio de conhecimento, ou mesmo quais domínios que utilizam tais conceitos. A partir dessa identificação é possível propor um vocabulário base que servirá para a construção de SOC.

A partir de proposições lógico semânticas e semânticas ideológicas podem-se declarar conceitos que traduzam as pesquisas desenvolvidas nos domínios de conhecimento. Os conceitos declarados podem ser funcionais ou agenciadores. Os conceitos funcionais de uma

forma geral condensam o significado que um conceito tem para um domínio de conhecimento. Enquanto os conceitos agenciadores tem a função de permitir transitar da designação a determinação conceitual. A partir de práticas advindas de pesquisas empregadas nos domínios de conhecimento, é possível dar um “novo” sentido ao conceito, independente da forma categorial tradicional que é usada, como por exemplo em taxonomias de domínios.

Selecionados os conceitos por um índice de incidência e fidelidade no discurso dos textos, que chamamos de corpus para análise. Percebemos que a formação conceitual se dava por relações simples e compostas e que tais relações só são possíveis, quando através da análise conceitual se estabelecem relações entre entidades de conceitos.

As relações entre entidades só são compreendidas a partir do entendimento do contexto em que os conceitos são empregados. Essas relações são possíveis quando os conceitos são definidos com base na sua utilização no domínio de conhecimento. A interpretação das relações das entidades conceituais são fundamentais e são baseadas em um processo de análise cognitiva que é realizado por um intérprete que modeliza domínios de conhecimento.

Passamos então a testar se o modelo de análise por associações entre entidades estava correto quando recorremos aos argumentos utilizados nos textos e percebemos que a construção conceitual revelava o que estava descrito no texto. Sendo assim, o conceito formado por essas associações propostas pelo Modac representavam o texto e o contexto da pesquisa, logo o domínio de conhecimento.

O Modac está alicerçado na identificação dos argumentos textuais, na análise das entidades e na proposição de conceitos que resultam de relações simples e compostas que se dão entre entidades.

Apresentamos teorias e bases epistemológicas que fundamentaram o modelo, como também relacionamentos conceituais com base em argumentos textuais e sob o ponto de vista dos usuários pesquisadores dos domínios. Comprovamos assim que a análise conceitual se configura como metodologia de interpretação cuja base é de análise cognitiva.

Poderíamos dar por encerrada a nossa tarefa, mas acreditamos que trabalhos de pesquisa não se encerram nunca, além disso estudos de modelagem conceitual, vêm sendo

considerado nas últimas décadas como fundamentais, principalmente com o advento da Web social.

Acreditamos que paradigmas de “mundos possíveis” de conhecimentos e de representação precisam ser estudados e saídas para tais possibilidades devam ser encontradas, parafraseando Ranganathan como um “continuum dinâmico”.

Sugerimos estudos das obras de Gilles Deleuze e Félix Guatarri e da concepção rizomática trazida por eles, para pesquisas de organização do conhecimento, pois, envolvem o trabalho com conceitos.

REFERÊNCIAS

AGUIAR, Mádayá dos Santos Figueiredo. **Redes de palavras em textos escritos: uma análise da linguagem verbal utilizando redes complexas**. Salvador, 2009. Dissertação (mestrado em Física) – Instituto de Física, UFBA, 2009.

ALMEIDA, Carlos Cândido de. **Elementos de linguística e semiologia na organização da informação**. São Paulo: Cultura Acadêmica, 2011.

ALMEIDA, M. B. **Interoperabilidade entre fontes heterogêneas: um meta-modelo baseado em ontologias**. 149 f. Dissertação (Mestrado em Ciência da Informação) - Escola da Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2002.

_____. **Um modelo baseado em ontologias para representação da memória organizacional**. 2006. 341f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

ALMEIDA, M. B.; BAX, Marcello P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, Brasília, v. 32, n. 3, p.7-20, set./dez., 2003.

AUTHIER-REVUZ, J. **Palavras incertas: as não-coincidências do dizer**. Campinas: UNICAMP, 1998.

ARDOINO, J. Abordagem multirreferencial (plural) das situações educativas e formativas. In: BARBOSA, J. G. (Coord.). **Multirreferencialidade nas ciências e na Educação**. São Carlos, SP, EdUFSCar, 1988. p. 24-41.

_____. **Para uma pedagogia socialista**. Brasília: Plano, 2013.

BARBOSA, Alice Príncipe. **Teoria e prática dos sistemas de classificação bibliográfica**. Rio de Janeiro: IBBD, 1969.

BARITÉ, Mario et. al. Garantia literária: elementos para uma revisão crítica após um século. **TransInformação**, v. 22, n.2, p. 123-138, maio/ago., 2010.

BLOOM, Benjamin S. et. al. (Eds.). **Taxionomia de objetos educacionais**. Porto Alegre : Globo, 1972. v.1.

BORGES, Jussara. **Participação política, internet e competências infocomunicacionais: estudos com organizações da sociedade civil de Salvador**. Salvador, 2011. 364p. Tese (Doutorado em Comunicação e Cultura Contemporâneas) - FACOM, UFBA, 2011.

BRACHMAN, R. L. On the epistemological status of semantic networks. In: FLINDER, N. V. (ed.). **Associative Neetworks: representation and use of knowledge by computers**. New York: Academic Press, 1979, p. 3-50.

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

CALDEIRA, Sílvia G. M. **Caracterização da rede de signos linguísticos: Um modelo baseado no aparelho psíquico de Freud**. Salvador, 2005. 131 f. (Dissertação de Mestrado Interdisciplinar em Modelagem Computacional) - Centro de Pesquisa e Pós-Graduação da Faculdade Visconde de Cairu (CEPPEV), Salvador, 2005.

_____. **Lendo Bohr ao pé da letra: análise de elementos conceituais em escritos de Niels Bohr**. Salvador, 2007. 134 f. (Dissertação de Mestrado em Ensino, Filosofia e História das Ciências) Universidade Federal da Bahia e Universidade Estadual de Feira de Santana, 2007.

CALDEIRA, Sílvia G. M. et al. The network of concepts in written texts. **European Physical Journal B**, v. 49, p. 523-529, 2006.

CAMPOS, Maria Luiza de Almeida. Integração de ontologias: o domínio da bioinformática. **RECIIS – Revista Eletrônica de Comunicação Informação Inovação em Saúde**, v. 1, n. 1, p. 117-121, jan-jun, 2007.

_____. **Linguagem documentária: teorias que fundamentam sua elaboração**. Niterói: EdUFF, 2001a.

_____. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ciência da Informação**, v.33, n.1, p.22-32, 2004.

_____. A problemática da compatibilização terminológica e a integração de ontologias: o papel das definições conceituais. **Perspectiva em ciência da informação**, v.15, n.1, 2010.

_____. **A organização de unidades de conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para a realização da autoria**. Rio de Janeiro, 2001b. (Tese doutorado em Ciência da Informação) - UFRJ/ECO/IBICT, 2001.

CAMPOS, Maria Luiza de Almeida; CAMPOS, Linair Maria. A organização do conhecimento e suas teorias de representação: a ontologia de fundamentação como um modelo teórico para a representação de domínios. IN: ENCONTRO NACIONAL EM CIÊNCIA DA INFORMAÇÃO, 8, 2012, Rio de Janeiro **Anais...** Rio de Janeiro: ANCIB, 2012.

CAMPOS, Maria Luiza de Almeida; GOMES, Hagar Espanha. Organização da informação e os princípios ranganathianos. **Perspectiva Ciência em Informação**, v. 8, n. 2, p. 150-163, jul./dez. 2003.

CAMPOS, Maria Luiza de Almeida et. al.. O estado da pesquisa do grupo – ontologia e taxonomias: aspectos teóricos e metodológicos. In: SEMINÁRIO DE ONTOLOGIA NO BRASIL, 2009, Rio de Janeiro. **Anais...** Rio de Janeiro, 2009. Disponível em: <<http://ontobra.comp.ime.eb.br/artigos/Gr16.pdf>> acesso em: 17 jan. 2014.

CARLAN, Eliana. **Sistemas de Organização do Conhecimento**: uma reflexão no contexto da Ciência da Informação. Brasília, 2010. Dissertação (mestrado em Ciência da Informação) - UNB, 2010.

COELHO; Eduardo de Mattos Pinto; BAX, Marcello Peixoto; MEIRA JUNIOR, Wagner. As várias naturezas dos compromissos em ontologias aplicadas. In: ENCONTRO NACIONAL EM CIÊNCIA DA INFORMAÇÃO, 13, Rio de Janeiro, 2012. **Anais...** Rio de Janeiro: ANCIB, 2012.

CHIAVEGATTO, Valéria Coelho. Signos entrelaçados: contexto e construção dos sentidos na linguagem. In: HENRIQUES, Claudio Cezar; PEREIRA, Maria Teresa Gonçalves (orgs). **Língua e transdisciplinaridade**: rumos conexões, sentidos. São Paulo: Contexto, 2002.

CUNHA, Isabel Maria Ribeiro Ferin. **Do mito à análise documentária**. São Paulo: Edusp, 1990.

CURRÁS, Emilia. **Ontologias, taxonomia e thesaurus em teoria de sistemas e sistemática**. Brasília: Thesaurus, 2010.

DAHLBERG, I. Conceptual compatibility of ordering systems. **Internacional Classification**. v. 10, n.2, p. 5-8, 1983.

_____. Fundamentos teórico-conceituais da classificação. **Revista de Biblioteconomia**, v.6, n.1, 1978a.

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

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

_____. **Ontical structures and universal classification**. Bangalore: Sarada Ranganathan Endowment, 1978b.

_____. A referent-oriented, analytical concept theory for interconcept. **International Classification**, v. 5, n. 3, p. 142-151, 1978c.

_____. Teoria do conceito. **Ciência da informação**. Rio de Janeiro, v. 7, n. 2, p. 101-107, Jul./dez. 1978d.

_____. Teoria da Classificação Facetada ontem e hoje. In: CONFERÊNCIA BRASILEIRA DE CLASSIFICAÇÃO, 1972, Rio de Janeiro. **Anais...** Rio de Janeiro: IBICT; Brasília: ABDF, 1979. [n.p.], Disponível em <http://www.conexaorio.com/biti/dahlbergteoria/dahlberg_teoriam.htm> acesso em: 4 abril 2013.

_____. Terminological definitions: characteristics and demands. In: TERMIA. **Problèmes de la définition et de la synonymie en terminologie**. Québec: GIRSTERM, 1983.

_____. Towards establishment of compatibility between indexing languages. **Internacional Classification**. 1981; 8(2): 88-91.

DELEUZE; Gilles; GUATARRI, Félix. **O que é a Filosofia?** 3.ed. Rio de Janeiro: Ed. 34, 2010.

DING, Yin; FOO, Schubert. Ontology research and development. Part 1 - a review of ontology mapping and evolving. **Journal of Information Science**, v.28, n.10, p. 123 - 136, 2002.

DING, Yin; FOO, Schubert. Ontology research and development. Part 2 - a review of ontology generation. **Journal of Information Science**, v.28, n.4, p. 375-388, 2002.

DZIEKANIAK, Gisele; PACHECO, Roberto; KERN; Vinicius Medina. Revisitando a Organização do Conhecimento através dos tesouros, folksonomias e ontologias: aportos da Engenharia do Conhecimento. In: SILVA, Fabiano Couto Corrêa da; SALES, Rodrigo de (orgs). **Cenários da Organização do Conhecimento: linguagens documentárias em cena**. Brasília: Thesaurus, 2011. p. 201-238

FADEL, Bárbara et. al. Gestão, mediação e uso da informação. In: VALENTIN, Marta (org.) **Gestão mediação e uso da informação**. São Paulo: Cultura Acadêmica, 2010. p. 13-3.

FADIGAS, Inácio de Sousa et. al. Análise de redes semânticas baseado em artigos de periódicos científicos: o caso dos periódicos de divulgação em educação matemática. **Educação Matemática e Pesquisa**, São Paulo, v. 11, n. 1, pp. 167-193, 2009

FAGUNDES, Norma Carapiá; FRÓES BURNHAM, Teresinha. Transdisciplinaridade, multirreferencialidade e currículo. **Revista da FACED**, n.5, p.39-55, 2001.

FOUCAULT, Michel. **As palavras e as coisas: uma arqueologia das ciências humanas**. São Paulo: Martins Fontes, 1999.

_____. **A arqueologia do saber**. Rio de Janeiro: Forense, 2008.

FRANCELIN, Marivalde Moacir; KOBASHI, Nair Yumiko. Concepções sobre o conceito na organização da informação e do conhecimento. **Ciência da Informação**, v. 40 n. 2, p.207-228, maio/ago., 2011.

FREITAS, Antônio José Gonçalves. Redes semânticas na lógica medieval: o caso de Petrus Hispanus Portucalensis. In: SEMINÁRIO NACIONAL DE HISTÓRIA DA MATEMÁTICA, 10, São Paulo, 2013. **Anais...** São Paulo: Unicamp, 2013.

FRÓES BURNHAM, Teresinha. Análise Cognitiva: Reconhecendo o antes irreconhecido. In: **ANÁLISE Cognitiva e Espaços Multirreferenciais de Aprendizagem: Currículo**,

Educação a Distância e Gestão/Difusão do Conhecimento, Salvador: EDUFBA, 2011, p. 58-75. Palestra proferida no Seminário Mensal da REDPECT / FAGED/UFBA.

_____. Complexidade, multirreferencialidade, subjetividade: três referências polêmicas para a compreensão do currículo escolar. In: BARBOSA, Joaquim Gonçalves (coord.). **Reflexões em Torno da Abordagem Multirreferencial**. São Carlos: EdUFSCar, 1998. p. 35-55

GETTIER, Edmund. Is justified true belief knowledge? **Analysis**, v.23, n. 6, p. 121-123. 1963.

GLERIA, Iram; MATSUSHITA, Raul; SILVA Sergio da. Sistemas complexos, criticalidade e leis de potência. **Revista Brasileira de ensino de Física**, v.20, n.2, p.99-108, 2004.

GOLDSCHMIDT, V. M. **Geochemistry**. London: Oxford, 1958.

GOMES, Hagar Espanha. Tendências da pesquisa em organização do conhecimento. **Pesquisa Brasileira em Ciência Informação**, v.2, n.1, p.60-88, jan./dez. 2009.

GOMES, Hagar Espanha; MOTTA, Dilza Fonseca da; CAMPOS, Maria Luiza de Almeida. **Revisitando Ranganathan**: a classificação na rede. Rio de Janeiro: s.n., 2006. Disponível em: <<http://www.conexaorio.com/bitl>>. acesso em 18 jan. 2014.

GOMES, Henriette Ferreira. A mediação da informação, comunicação e educação na construção do conhecimento. **Datagramazero – Revista de Ciência da Informação**, v.9, n.1, 2008.

GÓMEZ-PÉREZ, A.; FERNÁNDEZ, M.; VICENTE, A. J. Towards a method to conceptualize domain ontologies. In: ECAI WORKSHOP ON ONTOLOGICAL ENGINEERING, 1996, Budapest. **Proceedings...** Disponível em: <<http://citeseer.ist.psu.edu/483876.html>>. Acesso em: 13 maio 2012.

GONÇALVES, Júlia Aparecida; SOUZA, Renato Rocha. Relações e conceitos em ontologias: teorias de farradane e dahlberg. In: SEMINÁRIO DE PESQUISA EM ONTOLOGIAS NO BRASIL, 2008. **Anais...** Niterói: UFF, 2008.

GUARINO, N. Formal ontology, conceptual analysis and knowledge representation. **International Journal of Human and Computer Studies**, v. 43, n, 5/6, p. 625-640, 1995.

_____. Formal ontology and information systems. In: FOIS '98, 1., 1998, Trento, Italy. **Proceedings...** Amsterdam: IOS Press; Tokyo: Omsa, 1998, p. 3-15.

_____. Semantic matching: formal ontological distinction for information organization, extraction, and integration. In: PAZIENZA, M. T. (Ed.) **Information Extraction: a multidisciplinary approach to na emergig information technology**. [s.l] : [s.n], 1997.

GRUBER, T. Toward principles for the design of ontologies used for knowledge sharing. **International Journal of Human-Computer Studies**, v. 43, p. 907-928, 1995.

GRUBER, T. R. A translation approach to portable ontology specifications. **Knowledge Acquisition**. v. 5, p. 199-220, 1993.

HABERMAS, Jürgen. **Legitimation Crisis**. London: Heinemann, 1976.

HEIDEGGER, Martin. **O caminho da linguagem**. 5.ed. Petrópolis: Vozes, 2011.

HENRY, Paul. **A ferramenta imperfeita: linguagem, sujeito e discurso**. Campinas: Editora da Unicamp, 1992.

http://chasqueweb.ufrgs.br/~paul.fisher/aulas/buscas_google/ acesso em: 10 fev. 2014.

http://infojardins.blogspot.com.br/2011_03_01_archive.html acesso em: 10 fev. 2014.

<http://scholar.google.com.br/intl/pt-BR/scholar/about.html> acesso em: 10 fev. 2014.

ISKO – INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION. Disponível em < <http://www.isko.org/> > acesso em: 14 mar. 2013.

JACOB, Elin K. Ontologies and the Semantic Web. **Bulletin of the American Society for Information Science and Technology**, April/May, 2003. p. 19-22.

JORGE, Eduardo Manuel de Freitas. **MOBI – Modelo de Ontologia Baseado em Instâncias**. Salvador, 2012. Tese (Doutorado em Difusão do Conhecimento). UFBA. LNCC, UNEB, UEFS, UFABC, IFBA, SENAI/CIMATEC, 2012.

KOBASHI, Nair Yumiko. Análise documentária e representação da informação. **Ciência da Informação**, v.2, n.2, p.5-27, 1996.

KRAUSKOPF, Kourand B. **Introdução à geoquímica**. São Paulo: Polígono, 1972.

KUMAR, Krishan. **Theory of classification**. 2.nd revised. New Delhi: Vikas Publishing House, 1981.

LACAN, Jacques. A instância da letra no inconsciente ou a razão desde Freud. In: **Escritos**. Rio de Janeiro: Jorge Zahar, 1998.

LAGE, Ana Lúcia; FRÓES BURNHAM, Teresinha; MICHINEL, José Luis. Abordagens epistemológicas da cognição: a análise cognitiva na investigação da construção de conhecimento. In: FRÓES BURNHAM, Teresinha (Ed.). **Análise Cognitiva e Espaços Multirreferenciais de Aprendizagem: Currículo, Educação a Distância e Gestão/Difusão do Conhecimento**. Salvador: EDUFBA, 2012, p. 76-97.

Le Coadic, Yves-François. **A ciência da informação**. 2.e.d. Brasília: Briquet de Lemos, 2004.

THE LÉON Manifesto. 2007. Disponível em: < <http://www.iskoi.org/ilc/leon.htm> > Acesso em: 20 set. 2013.

LIMA, Gercina Ângela Borém. Mapa conceitual como ferramenta para organização do conhecimento em sistema de hipertextos e seus aspectos cognitivos. **Perspectiva em ciência da Informação**, v.9 n.2, p. 134-145, jul./dez. 2004

LIMA, Vânia Mara Alves; LARA, Marilda Lopes Ginez de. A terminologia e a legitimação do conhecimento nos artigos das revistas científicas. In: POBLACIÓN, Dinah Aguiar [et. al.] **Revistas científicas: dos processos tradicionais às perspectivas alternativas de comunicação**. Cotia, SP: Ateliê Editorial, 2011.

LIMA-MARQUES, Mamede. **Ontologias: da filosofia à representação do conhecimento**. Brasília: Thesaurus, 2006.

MARCONDES, Carlos Henrique. Organização e representação do conhecimento em ambiente web: do formato textual linear aos artigos semânticos. **PontodeAcesso**, Salvador, V.7, n.1 ,p. 7-41, abr 2013. Disponível em: <www.pontodeacesso.ici.ufba.br>. acesso em 14 jul. 2013.

_____. O papel das relações semânticas na organização e representação do conhecimento em ambientes digitais. In: SILVA, Fabiano Corrêa da; SALES, Rodrigo de (orgs). **Cenários da organização do conhecimento: linguagens documentárias em cena**. Brasília: Thesaurus, 2011. p. 129-168.

MATURANA, Humberto. **Cognição, ciência e vida cotidiana**. Belo Horizonte: UFMG, 2001.

_____. **Emoções e linguagem na educação e na política**. Belo Horizonte: UFMG, 2002.

MENDONÇA, Fabrício Martins; ALMEIDA, Maurício Barcelos. Modelos e teorias para representação: uma teoria ontológica para o sangue humano. In: ENANCIB, 13, Rio de Janeiro, 2012. **Anais...** Rio de Janeiro: ANCIB, 2012.

MERLEAU-PONTY, Maurice. **A prosa do mundo**. São Paulo: Cosac & Naify, 2002.

MONTEIRO, Silvana Drumond. As múltiplas sintaxes dos mecanismos de busca no ciberespaço. **Informação & Informação**, v.14, n. esp., p. 68-102, 2009.

MOREIRA, Walter; LARA, Marilda Lopes Ginez de. Relações conceituais e categorias filosóficas: aportes das ontologias e da terminologia para representação do conhecimento. In Encontro Nacional de Pesquisa em Ciência da Informação – ENCONTRO NACIONAL EM CIÊNCIA DA INFORMAÇÃO, 12, 2011, Brasília. **Anais...** Brasília, 2011.

MOREIRO GONZÁLEZ, José Antonio. Evolução ontológica das linguagens documentárias: relato de uma experiência de curso organizado conjuntamente para o DT/SIBi-USP e o PPGCI/ECA. **Ciência Informação e Documentação**, v. 2, n. 1, p. 143-164, jan./jun. 2011a

_____. **Linguagens documentárias e vocabulários semânticos para a Web**. Salvador: EDUFBA, 2011b.

MORIN, Edgar. **A cabeça bem-feita: repensando a reforma, reformar o pensamento**. 2.ed. Rio de Janeiro: Bertrand Brasil, 2000. 128p.

_____. **Introdução ao pensamento complexo**. 3.ed. Porto Alegre: Sulina, 2007. 120p.

NAESS, Arne; CHRISTOPHERSEN, Jens A.; KVALO, Kjell. **Democracy, ideology, and objectivity, studies in the semantics and cognitive analysis of ideological controversy**. Oslo: Norwegian Council for Science and the Humanities; Oslo University Press, 1956.

NONATO, Rafael dos Santos. **Teoria do conceito e hipertextos: uma proposta para determinação de relacionamentos em links conceituais**. Belo Horizonte, 2009. dissertação (Mestrado em Ciência da Informação) - UFMG, 2009.

NONATO, Rafael dos Santos; LIMA, Gercina Ângela Borém de Oliveira. A Teoria do Conceito aplicada a interpretação de links hipertextuais: consideração sobre modelagem conceitual. In ENCONTRO NACIONAL EM CIÊNCIA DA INFORMAÇÃO, 8, Salvador, 2007. **Poster**. Rio de Janeiro: ANCIB, 2007

NOVO, Hildenise Ferreira. **A elaboração de taxonomias: princípios classificatórios para domínios interdisciplinares**. Niterói, 2007. Dissertação (Mestrado em Ciência da Informação) - UFF/IBICT, 2007.

PEREIRA, Ana Lúcia Lage. **A rede como espaço multirreferencial de aprendizagem: construção do conhecimento na produção de inovação em TIC em um Instituto de Ciência e Tecnologia brasileiro**. Salvador, 2013. Tese (Doutorado em Difusão do Conhecimento) - DMMDC, 2013.

PLATÃO. **Teeteto**. Tradução NUNES, Carlos Alberto. Disponível em: <http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=2299>. Acesso em: 09 fev. 2014.

RANGANATHAN, S. R. **Philosophy of library classification**. New Delhi: Ejnar Munksgaard, 1951.

_____. **Colon Classification**. 6.ed. Bombay: Ásia Publishing House, 1963.126p. Reimpressão.

_____. **The five laws of library Science**. Bombay: Ásia Publishing House, 1963. 449p. Reimpressão.

_____. **Prolegomena Io library classification**. Bombay: Ásia Publishing House, 1967.640p.

RIOS, Jocelma Almeida; FRÓES BURNHAM, Teresinha. A emergência da análise cognitiva: entrevista. **Poiésis**, v.5, n.9, p. 173-195, 2012.

ROSADOS, Helen Beatriz Frota. A Ciência da Informação em sua aproximação com as Ciências Cognitivas. **Em Questão**, v.9, n.1, p. 79-94, 2003.

SALES, Luana Farias. Modelo triádico de relações para ontologias. In: SEMINÁRIO DE PESQUISA EM ONTOLOGIA NO BRASIL, 2008, **Anais...** Rio de Janeiro, UFF. 2008. Disponível em: <<http://www.uff.br/ontologia/artigos/13.pdf>> acesso em 18 de maio 2013.

_____. **Ontologias de domínios: estudos das relações conceituais e suas aplicações.** Nitéroí, 2006. Dissertação (Mestrado em Ciência da Informação) - UFF/IBICT, 2006.

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

SALES, Luana Farias; SAYÃO, Luís Fernando. Um modelo triádico de relações: um protótipo de modelagem conceitual para a área nuclear. In: ENCONTRO NACIONAL EM CIÊNCIA DA INFORMAÇÃO, 13, Rio de Janeiro, 2012. **Anais...** Rio de Janeiro: ANCIB, 2012.

SANTAELLA, Lucia. **Semiótica aplicada.** São Paulo: Cengage Learning, 2012.

SANTOS, Maria Carolina Alves dos. A demarcação Platônica de novas fronteiras epistêmicas para o discurso filosófico: um estudo sobre sofista. **TransFormação**, v.24, p. 273-299, 2001.

SCHOPENHAUER, Arthur. **O mundo como vontade e representação.** Rio de Janeiro: Contraponto, 2001.

SILVA, Daniela Lucas da; SOUZA, Renato Rocha; ALMEIDA, Maurício Barcellos. Ontologias e vocabulários controlados: comparação de metodologias para construção. **Ciência da Informação**, v. 37, n. 3, p. 60-75, set./dez. 2008.

SIMÕES, Darcilia. Comunicação em tempos de crise: signos em intercâmbio. In: HENRIQUES, Claudio Cezar; PEREIRA, Maria Teresa Gonçalves. **Língua e transdisciplinaridade: rumos conexões, sentidos.** São Paulo: Contexto, 2002.

SMITH, B. Ontology. In: FLORIDI, L. **Blackwell guide to the philosophy of computing and information.** Oxford: Blackwell, 2003. p.155-166.

SOUZA, Hélio José dos Santos. **O problema da motivação moral em Kant.** São Paulo: Cultura Acadêmica, 2009.

SOUZA JUNIOR; Mario bastos; CAFÉ, Ligia. Ontologias: abordagens em teses e dissertações das universidades brasileiras. **Informação & Sociedade**, v.22, n.2, p. 81-98, maio/ago. 2012.

SOWA, John F. Relating diagrams to logic in conceptual graphs for knowledge representation, Lectures Notes. **Computer Science**, vol. 699, p. 1-35, 1993.

_____. **Semantic networks**. [s.l.] : [s.n.], [1987]. Disponível em:
<<http://www.jfsowa.com/pubs/semnet.htm>> acesso em: 13 de set. 2013.

TEIXEIRA, Gesiane Miranda. **Redes semânticas em discursos orais**: uma proposta baseada na psicologia cognitiva utilizando redes complexas. Salvador, 2007. Dissertação (mestrado Interdisciplinar em Modelagem Computacional) - Fundação Visconde de Cairú, 2007.

TEIXEIRA, Gesiane Miranda et al. Complex semantic networks. **International Journal of Modern Physics C**. v.21, n.3, 2010.

TEIXEIRA, Wilson et al. (org). **Decifrando a terra**. São Paulo: Companhia Editora Nacional, 2009.

VAN DER WAL, Thomas. **Folksonomy coinage and definition in off the top**. 2007. Disponível em <<http://www.vanderwal.net/folksonomy.html>> . acesso em 11 nov. 2013.

VARELA, Aida. **Informação e autonomia**: a mediação segundo Feuerstein. São Paulo: Senac, 2007.

VARELA, Franciso J.; THOMPSON, Evan; ROSCH, Eleanor. **A Mente Incorporada**: Ciências Cognitivas e Experiência Humana. Porto Alegre: Artmed, 2001.

VITAL, Luciane Paula; CAFÉ, Ligia Maria Arruda. Ontologias e Taxonomias: diferenças. **Perspectivas em Ciência da Informação**, v.16, n.2, p.115-130, abr./jun. 2011.

VICKERY. B. C. Ontologies. **Journal of Information Science**, v. 23, n. 277, p. 277-286, 1997.

WÜSTER, E. Begriffs-und Themaklassifikationen: unterschiede in ihrem Wesen und ihrer Anwendung. **Nachrichten fuer dokumentation**, v.22, n.3, p.98-104, 1971.

L'Etude scientifique générale de la Terminologie, zone frontalière entre la Linguistique, la Logique, l'Ontologie, l'Informatique et les Sciences des Choses. In: RONDEAU, G. ; FELBER, H. (org.). **Textes choisis de Terminologie. I. Fondements théoriques de la terminologie**. Québec: GIRSTERM, 1981. p. 57-114.

APENDICE A – REFERÊNCIAS DOS TRABALHOS QUE COMPÕEM O CORPUS (CP) SELECIONADOS PARA ANÁLISE NA QUALIFICAÇÃO DO DOUTORADO

CP 1 - GEOGRAFIA

BITTENCOURT, Adriana Santana. **A paisagem urbana de Cachoeira-BA: diferentes olhares e interfaces.** Salvador, 2011, 157f. Dissertação (mestrado em Geografia) - Curso de Pós-graduação em Geografia. Instituto de Geociências. Universidade Federal da Bahia, 2011.

DI LAURO, Aluztane. **Análise do uso e ocupação da terra em Vitória da Conquista, Barra do Choça e planalto: estudo das transformações socioambientais a partir da implantação da lavoura cafeeira (1970 a 2008).** Salvador, 2011, 129f. Dissertação (mestrado em Geografia) - Curso de Pós-graduação em Geografia. Instituto de Geociências. Universidade Federal da Bahia, 2011.

EVANGELISTA, Antonia dos Reis Salustiano. **O processo de ocupação do bioma caatinga e suas repercussões socioambientais na Sisalândia, Bahia.** Salvador, 2010, 199f. Dissertação (mestrado em Geografia) - Curso de Pós-graduação em Geografia. Instituto de Geociências. Universidade Federal da Bahia, 2010.

NOVAIS, Marcos Paulo Souza. **Aspectos da apropriação e ocupação do espaço na microbacia Rio do Ouro, Jacobina – Bahia e suas repercussões socioambientais.** Salvador, 2009. 159f. Dissertação (mestrado em Geografia) - Curso de Pós-graduação em Geografia. Instituto de Geociências. Universidade Federal da Bahia, 2009.

SOUZA, André Nunes. **Orla oceânica de Salvador: um mar de representações.** Salvador, 2010. 158f. Dissertação (mestrado em Geografia) - Curso de Pós-graduação em Geografia. Instituto de Geociências. Universidade Federal da Bahia, 2010.

CP2 – GEOLOGIA

CARVALHO, Sérgio Roberto Lemos de. **Referência geoquímica regional para a interpretação das concentrações de metais pesados em solos derivados de sedimentos da formação barreiras.** Salvador, 2009. 202f. Tese (doutorado em Geologia) - Instituto de Geociências. Universidade Federal da Bahia, 2009.

LIMA, Geraldo MarceloPereira. **Análise do espaço de acomodação, das taxas de sedimentação e das condições de preenchimento do estuário Jacuípe, Litoral Norte da Bahia, Brasil.** Salvador, 2010. 112f. Tese (doutorado em Geologia) - Instituto de Geociências. Universidade Federal da Bahia, 2010.

MAIA, Paulo Henrique Prates. **Um novo método para avaliar a vulnerabilidade dos aquíferos.** Salvador, 2011. 133f. Tese (doutorado em Geologia) - Curso de Pós-graduação em Geociências. Instituto de Geociências. Universidade Federal da Bahia, 2011.

NASCIMENTO, Sérgio Augusto de Moraes. **Diagnóstico hidrogeológico, hidroquímico e da qualidade da água do aquífero freático do alto cristalino de Salvador – Bahia**. Salvador, 2008. 214f. Tese (doutorado em Geologia) - Geologia. Curso de Pós-graduação em Geociências. Instituto de Geociências. Universidade Federal da Bahia, 2008.

NUNES, Alina Sá. **A utilização da Geologia na identificação dos habitats mais adequados para o estabelecimento de áreas marinhas protegidas na Costa do Dendê, Bahia, Brasil**. Salvador, 2009. 178f. Tese (doutorado em Geologia) - Geologia. Curso de Pós-graduação em Geociências. Instituto de Geociências. Universidade Federal da Bahia, 2009.

CP3 - GEOQUÍMICA

ANDRADE, Consuelo Lima Navarro de. **Biogeoquímica do ambiente estuário do rio Passa Vaca em área urbana de Salvador, BA**. Salvador, 2011, 117f. Dissertação (mestrado em Geoquímica) - Curso de Pós-graduação em Geoquímica, Petróleo e Meio Ambiente. Instituto de Geociências. Universidade Federal da Bahia, 2011.

CALDAS, Alex Soares. **Aproveitamento do resíduo de coco verde (cocos nucifera L.) como sorvente em derramamento de petróleo em ambiente marinho**. Salvador, 2011, 55f. Dissertação (mestrado em Geoquímica) - Curso de Pós-graduação em Geoquímica, Petróleo e Meio Ambiente. Instituto de Geociências. Universidade Federal da Bahia, 2011.

FERREIRA, Aline Nogueira. **Avaliação do impacto da dragagem sobre a associação fitoplanctônica do Porto de Aratu, Baía de Todos os Santos, Bahia**. Salvador, 2011, 73f. Dissertação (mestrado em Geoquímica) - Curso de Pós-graduação em Geoquímica, Petróleo e Meio Ambiente. Instituto de Geociências. Universidade Federal da Bahia, 2011.

MELO, Eduardo Gomes Vieira de. **Avaliação da glicerina bruta na estimulação de bactérias hidrocarbonoclásticas para remediação de áreas contaminadas por hidrocarbonetos**. Salvador, 2011, 68f. Dissertação (mestrado em Geoquímica) - Curso de Pós-graduação em Geoquímica, Petróleo e Meio Ambiente. Instituto de Geociências. Universidade Federal da Bahia, 2011.

MOREIRA, Ícaro Thiago Andrade. **Avaliação da eficiência de modelos de remediação aplicados em sedimentos de manguezal impactados por atividades petrolíferas**. Salvador, 2011, 221f. Dissertação (mestrado em Geoquímica) - Curso de Pós-graduação em Geoquímica, Petróleo e Meio Ambiente. Instituto de Geociências. Universidade Federal da Bahia, 2011.

APÊNDICE B – FONTES PARA DEFINIÇÃO DOS TERMOS

REFERÊNCIAS

ANTAS, Luiz Mendes. **Glossário de termos técnicos inglês-português**. 6.ed. São Paulo: Traço, 2009.

CPRM. **Glossário Geológico**. Disponível em:
<http://www.cprm.gov.br/Aparados/glossario_geologico.htm> acesso em: 2 set. 2013.

GRISI, Breno Machado. **Glossário de ecologia e ciências ambientais**. 3.ed. João Pessoa: UFPB, 2007.

GUERRA, Antonio Teixeira; GUERRA, Antonio José Teixeira. **Novo dicionário geológico-geomorfológico**. Rio de Janeiro: Bertrand Brasil,

[http://ambientes.ambientebrasil.com.br/biotecnologia/artigos_de_biotecnologia/bacterias_\(geral_e_variedades\).html](http://ambientes.ambientebrasil.com.br/biotecnologia/artigos_de_biotecnologia/bacterias_(geral_e_variedades).html)

<http://bahia.com.br/destinos/caminhos-do-sudoeste/>

http://www.cprm.gov.br/geocoturismo/pedra_gavea/glossario.html

<http://www.dicio.com.br/quente/>

<http://www.infoescola.com/biologia/ecossistema/>

http://www.inmet.gov.br/html/informacoes/curiosidade/tempo_clima.html

<http://www.rc.unesp.br/museudpm/rochas/introducao.html>

<http://www.significados.com.br/caatinga/>

<http://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Uva/UvaAmericanaHibridaClimaTemperado/clima.htm>

SILVA, Lorival Hoffmann da; GARANHANI, Thiago. **Biosurfactantes**. Florianópolis: UFSC, 2006. Disponível em: <www.enq.ufsc.br/labs/probio/disc_eng_bioq/.../biosurfactantes> acesso em: 2 jan. 2014.