

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO

MURILLO LIMA MODESTO

**UM ESTUDO DE CASO SOBRE AQUISIÇÃO DO CONHECIMENTO EM
ONTOLOGIAS: VALIDAÇÃO DE CONHECIMENTO ESPECIALIZADO A PARTIR
DE LINGUAGENS NATURAIS CONTROLADAS**

Belo Horizonte

2021

MURILLO LIMA MODESTO

**UM ESTUDO DE CASO SOBRE AQUISIÇÃO DO CONHECIMENTO EM
ONTOLOGIAS: VALIDAÇÃO DE CONHECIMENTO ESPECIALIZADO A PARTIR
DE LINGUAGENS NATURAIS CONTROLADAS**

Dissertação apresentada ao Programa de Pós-Graduação em Gestão & Organização do Conhecimento da Escola de Ciência da Informação da Universidade Federal de Minas Gerais para obtenção do grau de Mestre em Gestão e Organização do Conhecimento.

Área de Concentração: Ciência da Informação

Linha de Pesquisa: Gestão & Tecnologia da Informação e Comunicação

Orientador: Maurício Barcellos Almeida

Co-orientadora: Livia Marangon Duffles
Teixeira

BELO HORIZONTE

2021

M691e

Modesto, Murillo Lima.

Um estudo de caso sobre aquisição do conhecimento em ontologias [recurso eletrônico]: validação de conhecimento especializado a partir de linguagens naturais controladas / Murillo Lima Modesto . - 2021.

1 recurso eletrônico (117 f. : il., color): pdf.

Orientador: Mauricio Barcellos Almeida.
Coorientadora: Livia Marangon Duffles Teixeira
Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

Referências: f. 110-117.

Exigências do sistema: Adobe Acrobat Reader.

1. Ciência da Informação – Teses. 2. Ontologias (recuperação da informação) – Teses. 3. Aquisição de conhecimento (sistemas especialistas) – Teses. 4. Representação do conhecimento (sistemas especialistas) – Teses. I. Título. II. Almeida, Mauricio Barcellos. III. Teixeira, Livia Marangon Duffles. IV. Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

CDU:025.4.03



FOLHA DE APROVAÇÃO

Um Estudo de Caso em Ontologias: Validação de conhecimento especializado a partir de linguagens controladas

MURILLO LIMA MODESTO

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, como requisito para obtenção do grau de Mestre em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, área de concentração CIÊNCIA DA INFORMAÇÃO, linha de pesquisa Gestão e Tecnologia da Informação e Comunicação.

Aprovada em 24 de fevereiro de 2021, todos por videoconferência, pela banca constituída pelos membros:

Prof.(a). Mauricio Barcellos Almeida (orientador)
ECI-UFMG

Prof.(a). Livia Marangon Duffles Teixeira (co-orientadora)
Fundação CPQD

Prof.(a). Benildes Coura Moreira dos Santos Maculan
ECI-UFMG

Prof.(a). Fabricio Martins Mendonça
DCC-UFJF

Belo Horizonte, 24 de fevereiro de 2021.



UNIVERSIDADE FEDERAL DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E ORGANIZAÇÃO DO CO-
NEHECIMENTO



ATA DA DEFESA DA DISSERTAÇÃO DO ALUNO
MURILLO LIMA MODESTO

Realizou-se, no dia 24 de fevereiro de 2021, às 09:30 horas, todos por videoconferência, pela Universidade Federal de Minas Gerais, a defesa de dissertação, intitulada *Um Estudo de Caso em Ontologias: Validação de conhecimento especializado a partir de linguagens controladas*, apresentada por MURILLO LIMA MODESTO, número de registro 2019663516, graduado no curso de EDUCAÇÃO FÍSICA, como requisito parcial para a obtenção do grau de Mestre em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, à seguinte Comissão Examinadora: Prof(a). Mauricio Barcellos Almeida - ECI/UFMG (Orientador), Prof(a). Livia Marangon Duffles Teixeira - Fundação CPqD (co-orientadora), Prof(a). Benildes Coura Moreira dos Santos Maculan - ECI/UFMG, Prof(a). Fabrício Martins Mendonça – UFJF.

A Comissão considerou a dissertação:

(x) Aprovada

() Reprovada

Finalizados os trabalhos, lavrei a presente ata que, lida e aprovada, vai assinada por mim e pelos membros da Comissão.

Belo Horizonte, 24 de fevereiro de 2021.

Prof.(a). Mauricio Barcellos Almeida (orientador)

Prof(a). Livia Marangon Duffles Teixeira (co-orientadora)

Prof(a). Benildes Coura Moreira dos Santos Maculan

Prof(a). Fabrício Martins Mendonça
DCC-UFJF

DEDICATÓRIA

Aos meus pais, Ildo e Solange, por todo esforço e apoio incondicional a todos estes anos de estudo.

Dedico às Ciências de modo geral, que por muitas vezes, ao longo dos anos, vem sendo subjugada.

AGRADECIMENTOS

Ao meu orientador Prof. Dr. Mauricio Barcellos Almeida; a co-orientadora Dr^a Livia Marangon Duffles Teixeira: Sem a ajuda, a paciência e a dedicação de ambos os percalços seriam muito maiores. Aprendi muito com vocês;

A CEMIG e em especial ao Engenheiro Henrique Diniz por toda a ajuda no processo de aquisição do conhecimento e validação da ontologia.

Aos professores, Dr^a Benildes Maculan e Dr. Fabrício Mendonça pelas importantes contribuições na banca de qualificação;

Aos meus colegas de pesquisa e professores do ReCOL;

Aos colegas e professores da ECI, em especial a Yluska, Rafael e Prof. Dr. Max Cirino de Mattos por todas as dicas e ensinamentos;

Aos meus grandes amigos da vida: Caíque, Anne, Ana Paula, Aline, Tomás, Leonardo, Cabral (O Velho Lobo do Mar), Jéssica Caetano e Carla. As amizades de todos vocês são essenciais durante toda a minha vida.

A Ana Kaizer, por todo o carinho, compreensão, paciência e amor que nestes últimos 4 meses foram muito importantes, incentivando e acalmando nos momentos de dúvidas e medos, lendo todo o conteúdo desta dissertação me ajudando nas correções.

*“...Tempo de chorar, e tempo de rir; tempo de
prantear, e tempo de dançar;
Tempo de espalhar pedras, e tempo de ajuntar
pedras; tempo de abraçar, e tempo de afastar-
se de abraçar...”*

Eclesiastes 3:4,5

*“A mind full of questions, and a teacher in my
soul, and so it goes”*

“Uma mente cheia de perguntas, e um
professor em minha alma, e assim vai”

Guaranteed – Edward Louis Severson III

RESUMO

A Aquisição do Conhecimento é fase importante, porém pouco estudada, do processo de construção de ontologias. Em função disso, constitui-se em uma lacuna na compreensão por especialistas de domínio e mesmo por profissionais da informação. A linguagem natural, utilizada na comunicação é informal e de fácil compreensão por humanos. Uma linguagem formal consiste na expressão da linguagem natural por meio uma sintaxe também formal, bem definida no sentido de ser restrita, com semântica livre de ambiguidades e suporte aos métodos formais de raciocínio. As linguagens formais são adotadas no processo de construção de ontologias, artefatos formais de Representação do Conhecimento. Na construção desses artefatos, após as etapas de Aquisição do Conhecimento e de Conceitualização é necessário validar toda a representação com especialistas de domínio, de forma a confirmar que o modelo realmente representar o conhecimento especializado. Entretanto, a linguagem lógica usada nas ontologias nem sempre é de fácil entendimento pelas pessoas. Nesse contexto, identificou-se o problema de pesquisa inserido na lacuna entre a linguagem do artefato ontológico e a linguagem natural do especialista, a partir da dificuldade e até mesmo impossibilidade da validação adequada da ontologia. Busca-se, assim, encontrar respostas para a seguinte questão: como planejar e realizar a validação da representação formal do conhecimento em ontologias pelos especialistas de domínio usando uma interface baseada em linguagem natural controlada? O **objetivo geral** da presente dissertação é criar um método para validar a aquisição do conhecimento confrontando resultados daquela etapa com os axiomas da fase de formalização, a partir do uso de linguagens controladas. A pesquisa se **justifica** pela pouca literatura encontrada tanto sobre método de aquisição do conhecimento quanto sobre validação da representação formal, mesmo sendo ambos relevantes para a pesquisa de ontologias em Ciência da Informação (CI). Para tal atingir esse objetivo foi feito um recorte da ontologia do Manual de Controle Patrimonial do Setor Elétrico, desenvolvido pela Agência Nacional de Energia Elétrica. Os axiomas referentes aos termos escolhidos foram traduzidos utilizando-se da linguagem amigável para representação de ontologias, e após a tradução, validados por um especialista de domínio. Compreender como o conhecimento é extraído e registrado pelos especialistas de domínio, consolidado pelos especialistas em CI e, por fim, validado pelos especialistas de domínio impacta na qualidade dos artefatos ontológicos produzidos e em sua reutilização, se configurando na principal contribuição da presente pesquisa.

Palavras-chave: Aquisição do Conhecimento. Ontologia. Conhecimento especializado.

ABSTRACT

Knowledge Acquisition is an important, but not studied, stage in the process of constructing ontologies. It constitutes a gap in the understanding by domain experts and even information professionals. Natural language is used in communication, it is informal and easily understood by humans. A formal language consists of the expression of natural language through a formal syntax, well defined in the sense of being restricted, with semantics free of ambiguities and supporting the formal methods of reasoning. Formal languages are adopted in the process of constructing ontologies, which are formal artifacts of Knowledge Representation. In the construction of these artifacts, after the Knowledge Acquisition and Conceptualization stages, one needs to validate the entire representation with domain experts, in order to confirm that the model really represents the specialized knowledge. However, the logical language employed in ontologies is not always easy for humans. In this context, a research issue, within the gap between the formal language of the ontological artifact and the natural language of the specialist, was identified based on the difficulty and even impossibility of proper validation of the ontology. Thus, we seek to find answers to the following question: how to plan and carry out the validation of formal representation of specialized knowledge in ontologies by domain experts by using an interface based on controlled natural language? Our general goal is to create a method to validate the acquisition of knowledge in comparing the results of that stage with the axioms of the formalization phase using controlled languages. The research is justified by the little literature found both on the methods of knowledge acquisition and on the validation of formal representations, even though both are relevant to the research of ontologies in Information Science (IC). To achieve this goal, a fragment of a domain ontology was taken from the Patrimonial Control Manual for the Electric Sector, developed by Brazilian National Electric Energy Agency. The axioms related to the chosen terms were translated using friendly language to represent ontologies, and after the translation, validated by a domain experts. Understanding how knowledge is extracted and registered by domain experts, consolidated by CI specialists and, finally, validated by domain experts impacts on the quality of the ontological artifacts produced and in their reuse, configuring the main contribution of this research.

Keywords: Knowledge Acquisition. Ontology. Expertise.

LISTA DE FIGURAS

Figura 1 – <i>Protégé 5.5 e Plug in ACEView</i>	74
Figura 2 – Fragmento de tabela para registro das entrevistas	81
Figura 3 – Imagem tela <i>Protégé Frames 5.5</i>	85
Figura 4 – Fragmento da ontologia para validação com especialista	87
Figura 5 – Questionário validação especialista de domínio	87
Figura 6 – Exemplo de erro de conteúdo na AC	90
Figura 7 – Exemplo de erro de atribuição de operador lógico	91
Figura 8 – Modelo de formulário para validação	94
Figura 9 – Modelo de questionário para validação	95
Figura 10 – Linguagem formal <i>Manchester</i> – tela do <i>Protégé 5.5</i>	97
Figura 11 – <i>Plug in ACEView</i>	98
Figura 12 – Planilha de registro - validação ontologia	99

LISTA DE TABELAS

Tabela 1 – <i>String</i> de busca por publicações em língua portuguesa	48
Tabela 2 – <i>String</i> de busca por publicações em língua inglesa.....	48
Tabela 3 – Tabela Verdade Conjunção.....	67
Tabela 4 – Tabela Verdade Disjunção	68
Tabela 5 – Quantificadores	68
Tabela 6 – Constructos em linguagem <i>Manchester</i>	69
Tabela 7 – Validação ontologia – entrevista 1	90
Tabela 8 – Validação ontologia - Entrevista 2	91

LISTA DE QUADROS

Quadro 1 – Quadrado Ontológico	30
Quadro 2 – Descritivo das etapas das metodologias para construção de ontologias	37
Quadro 3 – Origem e propósito das ontologias	39
Quadro 4 – “Armadilhas” na construção de ontologia	41
Quadro 5 – Representação esquemática de técnicas de Aquisição do Conhecimento	57
Quadro 6 – Constructos lógicos	67
Quadro 7 – Termos e Axiomas – Equipamentos de Manufatura para Serviços.....	85
Quadro 8 – Roteiro de questionário estruturado e possíveis resultados	88

LISTA DE ABREVIATURAS

AC	—	AQUISIÇÃO DO CONHECIMENTO
ACE	—	ATTEMPTO CONTROLLED ENGLISH
ANEEL	—	AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA
APE	—	ATTEMPTO PARSING ENGINE
BFO	—	BASIC FORMAL ONTOLOGY
CI	—	CIÊNCIA DA INFORMAÇÃO
DRS	—	DISCOURSE REPRESENTATION STRUCTURES
GC	—	GESTÃO DO CONHECIMENTO
GF0	—	GENERAL FORMAL ONTOLOGY
LD	—	LÓGICA DESCRITIVA
LNC	—	LINGUAGEM NATURAL CONTROLADA
LPO	—	LÓGICA DE PRIMEIRA ORDEM
MCPSE	—	MANUAL DE CONTROLE PATRIMONIAL DO SETOR ELÉTRICO
MCSE	—	MANUAL DE CONTABILIDADE DO SETOR ELÉTRICO
NASA	—	NATIONAL AERONAUTIC AND SPACE ADMINISTRATION
ND	—	NORMAS DE DISTRIBUIÇÃO
ONS	—	OPERADOR NACIONAL DO SISTEMA ELÉTRICO
OWL	—	WEB ONTOLOGY LANGUAGE
P&D	—	PESQUISA E DESENVOLVIMENTO
PRODIST	—	PROCEDIMENTOS DE DISTRIBUIÇÃO
RDF	—	RESOURCE FOR DESCRIBING FRAMEWORK

TI	_	TECNOLOGIA DA INFORMAÇÃO
TUC	_	TIPO DE UNIDADE DE CADASTRO
UFO	_	UNIFIED FOUNDATION ONTOLOGY
WIKI	_	WIKIPEDIA
XML	_	EXTENSIBLE MARKUP LANGUAGE

SUMÁRIO

1 INTRODUÇÃO	18
2 BACKGROUND	22
2.1 ONTOLOGIA: VISÃO GERAL	22
2.2 ONTOLOGIA: CARACTERIZAÇÃO FILOSÓFICA.....	28
2.3 ONTOLOGIA: METODOLOGIAS DE CONSTRUÇÃO.....	31
2.3.1 Methontology	33
2.3.2 Metodologia NeON	34
2.3.3 OntoForInfoScience.....	35
2.3.4 Comparativo de metodologias	36
2.3.5 Armadilhas na construção de ontologias	40
2.4 AQUISIÇÃO DE CONHECIMENTO	46
2.4.1 Especialistas e conhecimento especializado.....	49
2.4.2 Extração de conhecimento especializado	51
2.4.2.1 Geração de protocolos	52
2.4.2.2 Análise de protocolo	53
2.4.2.3 Baseadas em matriz.....	53
2.4.2.4 Classificação de cartões (Cardsorting).....	54
2.4.2.5 Limitações de informação e tarefa de restringir processamentos	54
2.4.2.6 Diagrama	55
2.4.2.7 Técnicas de Grupo	55
2.4.2.8 Técnicas automatizadas.....	55
3. LINGUAGENS E REPRESENTAÇÃO.....	61
3.1 LINGUAGENS DE REPRESENTAÇÃO DO CONHECIMENTO	61
3.1.1 Linguagem natural.....	61
3.1.2 Linguagem Lógica	63
3.2.3 Linguagens naturais controladas.....	69

4 METODOLOGIA DE PESQUISA.....	75
4.1 CONTEXTO DA PESQUISA	75
4.1.1 A empresa do setor de energia	76
4.1.2 Construção da ontologia.....	79
4.2 PASSOS METODOLÓGICOS.....	84
4.2.1 Passo 1: selecionar de um recorte da ontologia.....	84
4.2.2 Passo 2: identificar axiomas (Manchester) na ontologia	84
4.2.3 Passo 3: traduzir da lógica para Linguagem Natural Controlada	86
4.2.4 Passo 4: estruturar roteiro de perguntas	86
5 RESULTADOS.....	89
5.1 VALIDAÇÃO DA ONTOLOGIA.....	89
5.2 VALIDAÇÃO DA LINGUAGEM NATURAL CONTROLADA (ACEVIEW)	92
5.3 MÉTODO DE VALIDAÇÃO	93
6 DISCUSSÃO	96
6.1 VALIDAÇÃO UTILIZANDO <i>PLUG IN ACEVIEW</i>	96
6.2 LINGUAGEM NATURAL CONTROLADA EM PORTUGUÊS	99
6.3 LITERATURA SOBRE A AQUISIÇÃO DO CONHECIMENTO.....	100
6.4 TÉCNICAS DE AQUISIÇÃO DO CONHECIMENTO.....	101
6.5 COMPLEXIDADE DO DOMÍNIO E A BUSCA PELO ESPECIALISTA.....	101
6.6 PESQUISADORES E A AQUISIÇÃO DO CONHECIMENTO	102
6.7 PERDA DE INFORMAÇÃO	103
6.8 COMPLEXIDADE DAS ONTOLOGIAS	103
7 CONSIDERAÇÕES FINAIS E CONCLUSÃO.....	106
REFERÊNCIAS.....	110

1 Introdução

A busca pelo saber surge a partir do momento que o indivíduo reconhece em si uma lacuna de conhecimento relacionado a um assunto. Todo ser humano possui a necessidade de busca por conhecimento como uma característica vital e inerente a ele, denominada de necessidade informacional, tornando-o um indivíduo com a capacidade de adaptar-se às condições externas de sua existência. Essa necessidade conduz o usuário na busca pela informação em diferentes fontes (como livros, sítios da *internet*, ou por meio de um especialista) e representada de diferentes formas. Assim compete ao usuário a tarefa de extrair o conteúdo relevante para o contexto em que se pretende utilizá-la (GONZALEZ, 2003), tornando os processos de aquisição e troca de conhecimento dinâmicos. A partir do momento que o conhecimento é adquirido ele deve ser registrado provendo um esquema global de nominação à informação e melhorando a reutilização de conceitos (SALEENA; SRIVATSA, 2010).

Choo (2006, p. 29) afirma que o indivíduo faz uso da informação quando ele seleciona e processa uma informação ou uma mensagem levando a uma mudança na capacidade do indivíduo de fazer senso de uma experiência e agir ou responder a um novo entendimento. Por isso, registrar o conhecimento adquirido em forma de manuais, textos, procedimentos, métodos ou outras formas de registro é importante para a transformação do conhecimento tácito em explícito para o compartilhamento com outros usuários. Esse registro é realizado por meio de linguagens, tais como a linguagem natural, formal e natural controlada, que são tipos de linguagens compreendidas pelos seres humanos, máquinas e por ambos, podendo ser utilizadas em diversos equipamentos e programas de comunicação.

Como a linguagem natural é raramente escrita ou falada corretamente, compreendê-la não é apenas entender as palavras, mas compreender também os conceitos e como eles estão ligados entre si para criação de significados (BARBOSA et al., 2017). No âmbito da Ciência da Computação e da Ciência da Informação, a utilização da linguagem natural levanta algumas questões, tais como a ambiguidade, a imprecisão e a potencial inconsistência, que prejudicam a recuperação da informação. Consequentemente, para representar conhecimento em computadores é necessário utilizar da Linguagem Formal. Segundo Barbosa et al

(2017) apesar da facilidade que temos, como humanos, de aprender a linguagem natural, a ambiguidade é o que torna a compreensão um problema difícil para computadores.

No contexto da Ciência da Informação (CI), a linguagem é utilizada para a representação do conhecimento seja por meio das ontologias, das taxonomias, dos esquemas de classificação ou tesauros. As ontologias são importantes aplicações que podem ser utilizadas como Sistemas de Organização do Conhecimento (SOC). O termo “ontologia” deriva dos estudos aristotélicos envolve as “categorias” utilizadas em classificação e caracterização de entidades em um mesmo domínio do conhecimento.

O processo de construção de uma ontologia consiste em três etapas que são: *i) aquisição do conhecimento; ii) desenvolvimento e documentação* e *iii) avaliação da ontologia*. A Aquisição do Conhecimento (AC) consiste em uma etapa onde o conhecimento é extraído das mais variadas fontes (YEH et al., 2003) processado e então validado por especialistas.

O conhecimento pode estar registrado e estruturado em normas, em procedimentos, em manuais, em, livros e artigos, na literatura acadêmica e também na mente dos especialistas, e nesses casos de forma não estruturada. O que se observa, entretanto, é uma grande quantidade de conhecimento não estruturado disponível, de difícil transformação em um modelo significativo e útil para sistemas de informação (ZAHARUDIN; NOAH; NOOR, 2009).

O uso das ontologias contribui para esses objetivos ao consolidar o conhecimento de um domínio por meio das entidades representadas por termos e respectivas definições, por meio das relações entre essas entidades e através dos axiomas que traduzem ambos em linguagem formal e livre de ambiguidades.

Uma Linguagem Formal consiste na expressão da linguagem natural por meio de uma sintaxe bem definida, no sentido de que é restrita, com semântica livre de ambiguidades e suportando os métodos formais de raciocínio. A linguagem formal se preocupa com as estruturas de significado, onde a sintaxe irá espelhar diretamente todos os princípios de construção de estruturas (formais) que podem ser encontrados no mundo (SMITH; MULLIGAN, 1983).

Uma linguagem assim não é compreendida por especialistas de domínio que não possuem familiaridade com esse tipo de representação no seu dia a dia.

Por outro lado, trata-se de um mecanismo favorável às aplicações computacionais, uma vez que existem algumas delas que são decidíveis, ou seja, produzem algoritmos que retornam resultados válidos e que podem ser interpretados computacionalmente (MENDONÇA, 2015).

A Linguagem Natural Controlada (LNC) é um tipo de linguagem formal, ou seja, possui sintaxe e semânticas bem definidas, porém utiliza-se sentenças muito parecidas com a linguagem natural, tornando a LNC mais acessível, essa é principal vantagem desse tipo de linguagem: proporciona maior entendimento e estimula a representação do conhecimento. São um tipo de linguagem que se enquadram em dois tipos principais: aquelas que melhoram a legibilidade para leitores humanos e aquelas que permitem a análise semântica automática confiável da linguagem (FUCHS; KALJURAND; KUHN, 2008).

Um exemplo de LNC é a *Attempto Controlled English* (ACE), linguagem utilizada na construção de ontologias. Possui semântica e sintaxe controladas por meio de regras simples e de fácil compreensão por indivíduos com pouco ou nenhum treinamento em linguagens formais (lógica). Devido à plataforma da ACE possuir uma interface voltada para linguagem natural utilizando-se de regras de sintaxe e semântica bem definidas, é considerada uma LNC que possui o benefício de ser compreendida por máquinas, na qual podem ser inseridos dados em uma interface, por indivíduos sem treinamento em linguagem lógica.

No processo de construção de ontologias, após realizadas as etapas de AC e Conceitualização, é necessário validar toda a representação realizada, com os especialistas de domínio. O **trabalho tem como objeto o estudo do problema** existente entre a lacuna que se identifica entre a linguagem do artefato e a linguagem natural do especialista, dificultando e até mesmo impossibilitando a execução da etapa de validação. A LNC é um mecanismo facilitador, reduzindo a lacuna, tanto para o profissional da informação como o especialista de domínio. Além disso, facilita o intercâmbio, entre a linguagem lógica e a linguagem natural. Pretende-se, assim responder a seguinte questão de pesquisa: que aspectos a LNC pode auxiliar na validação da representação formal do conhecimento em ontologias, por especialistas do domínio?

O **objetivo geral** da presente dissertação consiste em criar um método para validar processos de aquisição do conhecimento (AC) na construção de

ontologias, em um contexto de conhecimento especializado, confrontando resultados daquela etapa com os axiomas da fase de formalização a partir do uso de linguagens naturais controladas. Os **objetivos específicos** do trabalho envolvem:

- Identificar e sistematizar os tipos de falhas no processo de AC;
- Desenvolver método para profissionais em CI, que atuando na AC, possam se valer da LNC para validação;
- Desenvolver mecanismos para a efetiva participação de especialistas de domínio na representação do conhecimento, a partir da LNC, conhecida como ACE.

A pesquisa foi realizada no âmbito do setor elétrico brasileiro, o qual possui um rico aparato linguístico e terminológico. Além disso, a vivência e o conhecimento adquirido por seus profissionais ao longo dos anos, gerou um grande volume de conhecimento registrado em diferentes documentos, apresentando vocabulários diversos e em diferentes estruturas classificatórias. Nesse sentido, a busca por soluções relacionadas ao conhecimento em um setor tão complexo, pode resultar na qualidade dos insumos durante o processo de tomada de decisões, trazendo melhoria de competitividade e vantagens financeiras.

A pesquisa **se justifica** pela escassa literatura encontrada, tanto sobre o método de AC em desenvolvimento de ontologias quanto sobre a validação da representação formal, mesmo sendo ambos relevantes para a pesquisa em ontologias para a Ciência da Informação. Compreender como o conhecimento é extraído e registrado pelos especialistas de domínio, consolidado pelos especialistas em CI e, por fim, validado pelos especialistas de domínio impacta na qualidade dos artefatos ontológicos produzidos e em sua reutilização.

O restante do presente trabalho é constituído por sete capítulos. O capítulo 2 apresenta uma visão geral do tema ontologia e das metodologias de construção; o capítulo 3 discorre sobre a AC e Representação do Conhecimento, abordando formas de extração do conhecimento proveniente do especialista e as linguagens lógicas; o capítulo 4 trata da metodologia utilizada na presente pesquisa; o capítulo 5 apresenta os resultados obtidos da validação com especialistas de domínio; o capítulo 6 traz uma discussão sobre os resultados e, finalmente, o capítulo 7 traz as conclusões e considerações finais.

2 Background

O presente capítulo apresenta *background* com assuntos relevantes para o tema principal desta pesquisa. Na Seção 2.1 é apresentada uma visão geral de ontologias. A seção 2.2 caracteriza as ontologias e apresenta suas raízes Aristotélicas. A Seção 2.3 é voltada para a construção de ontologias e apresenta as metodologias utilizadas e respectivas etapas. A seção 2.4 apresenta a fase de AC e sua importância na construção de ontologias. Por fim, a seção 2.5 apresenta trabalhos relacionados a AC e a CI na utilização de ontologias.

Para tal, foram realizadas buscas por artigos científicos em periódicos, trabalhos apresentados em congressos, teses e dissertações nas diversas fontes como as bases de dados mais importantes na CI como também no portal Capes, além de busca por conhecimento em alguns livros, no período de maio a dezembro de 2020. A estratégia de busca foi por estudos que abordassem as áreas correlatas à CI, por serem áreas de principal interesse para o presente trabalho. A pesquisa no Portal de Periódico CAPES não se limitou a uma base de dados específica como a *Web of Science*, a escolha aconteceu a partir da seleção de áreas, como as Ciências Sociais Aplicadas e a Engenharia.

2.1 Ontologia: visão geral

A CI tem muito a contribuir no processo de organização do conhecimento, para que o uso da informação seja efetivo e eficiente, por seus usuários. Para tanto, o campo de pesquisa desenvolveu ao longo de sua trajetória como ciência, tanto teórica, quanto metodológica para possibilitar o melhor uso da informação. Esse repertório inclui instrumentos terminológicos para organização da informação, que incluem os vocabulários controlados, tesauros, terminologias e, a partir da década de 90 também as ontologias (MENDONÇA, 2015).

Pode-se caracterizar as ontologias como áreas de pesquisa pertencentes a três grandes domínios do conhecimento, Filosofia, Ciência da Computação, CI e Linguística. Na Filosofia o termo relaciona-se ao estudo do ser, ao estudo daquilo que existe. Na Ciência da Computação, está relacionado a artefatos de *softwares* em linguagens formais e à Inteligência artificial, a partir dos anos 70, e a modelagem

de sistemas de informação durante os anos 80. Na CI, as ontologias são definidas como um artefato para representação do conhecimento por meio de uma estrutura taxonômica que combina as categorias, as classes e as relações definidas entre ambos (TEIXEIRA, 2019).

Campos como a Linguística também começaram a desenvolver e fazer uso de ontologias, estimulados pelos avanços em áreas como inteligência artificial e *web semântica* (SCHALLEY, 2019). Na Ciência da Computação, alguns enxergam ontologias como um meio de estruturar uma base de conhecimento, outros concebem uma ontologia para ser usada como parte de uma base de conhecimento (USCHOLD; KING, 1995).

As ontologias são utilizadas para descrever o conhecimento. Mais especificamente, as ontologias possuem três pode refletir elementos básicos da teoria dos conjuntos aplicada às linguagens lógicas, presentes em qualquer sistema de representação do conhecimento: conjuntos, membros de conjuntos e relações entre esses conjuntos e membros (WARREN, 2017, p. 9).

As ontologias vêm se caracterizando ao longo dos últimos anos, como uma aplicação utilizada como artefatos em sistemas diversos, como Medicina, Engenharia, Gestão do Conhecimento (GC), plataformas de comércio eletrônico, dentre outras. As ontologias são também utilizadas na padronização dos vocabulários como uma alternativa às terminologias de um determinado domínio, trazendo maior poder de expressividade na representação do conhecimento do domínio a partir de seus formalismos lógicos e também possibilitando uma representação destes domínios mais próximas da realidade (MENDONÇA, 2015). Outro estudo afirma que as ontologias podem ser utilizadas em diversas aplicações, como se fossem um sistema multiagente, fazendo com que os agentes troquem informações entre si, ou seja, fazendo o que podemos também chamar de interoperabilidade de sistemas (PONTES JUNIOR; CARVALHO; AZEVEDO, 2013).

Segundo Coelho (2012, p. 56) na “Ciência da informação, a temática é abordada como forma de representação e organização do conhecimento, propondo meios de estruturá-lo e torná-lo disponível para uso das pessoas ou por máquinas”. A utilização desta representação formal é observada no setor Biomédico, nas Engenharias e no Setor Elétrico. É também possível destacar a importância das ontologias, reconhecidas em áreas como Engenharia do Conhecimento, Integração do Conhecimento e GC, além da criação de modelos (AN; PARK, 2018).

Atualmente, as potencialidades de uso das ontologias em CI e nas demais áreas correlacionadas estão reforçadas pelo grande volume e diversidade de dados manipulados no contexto da *web*, além do visível avanço no uso de tecnologias semânticas, as quais fazem parte do que se convencionou chamar de *web* semântica (MENDONÇA, 2015). Para Farinelli (2017), as ontologias são estudadas com foco na Representação do Conhecimento, permitindo o raciocínio computacional e a integração da informação, promovendo o processo de aplicação de linguagens naturais e assegurando o significado semântico.

O tema é estudado há muito tempo pela Filosofia, que a considera como um ramo da metafísica, que se preocupa com as coisas que existem. As ontologias possuem suas raízes em Aristóteles, este termo possui o significado de “categoria”, que é utilizada para designar o ato de classificar e caracterizar alguma coisa. Recentemente, em função da crescente demanda, está associada a recursos de *web* amplamente utilizados na Computação (COELHO, 2012; MARCONDES, 2015).

Segundo Almeida (2006) o termo ontologia tem origem no grego *ontos* (ser) e *logos* (estudo), sendo um termo relativamente novo na história da Filosofia, originalmente introduzido com o objetivo de diferenciar o estudo do ser humano em sua essência, do estudo dos vários tipos de outros seres das Ciências Naturais.

Segundo Mendonça (2015, p. 51), para a Filosofia, as ontologias “podem ser compreendidas como uma descrição básica das coisas que verdadeiramente existem no mundo”. Por outro lado, a visão da CI e da Computação as compreendem como um artefato de Engenharia e um componente importante, diferenciando-as da visão filosófica.

Apesar do termo ontologia remontar aos tempos Aristotélicos, no mundo moderno ela somente passou a ser mais utilizada na década de 90. Porém, vários estudos e teorias ligados à CI precedem os estudos direcionados à sua construção desde a década de 30. Segundo Mendonça (2015) “a partir da década de 1930 surgiram teorias sobre terminologias mais conhecidas em CI: i) a *teoria da classificação facetada*; ii) a *teoria geral da terminologia*; e iii) a *teoria do conceito*”. Já na década de 40 foi desenvolvido o modelo matemático para explicar a comunicação humana entre os polos emissor e receptor. O autor também afirma que na década de 60 as pesquisas relacionadas aos bancos de dados apresentaram três tipos principais de modelos de dados, que foram denominados como: modelo hierárquico,

modelo em rede e modelo relacional. Já nos anos 70 foram produzidos os primeiros modelos semânticos usados em modelagem conceitual a serem utilizados para padronização dos sistemas de gerenciamento de dados.

Durante a década de 80 surgiram os primeiros estudos sobre as ontologias como alternativa aos demais instrumentos usados para compreender e modelar o conhecimento de diferentes domínios, apresentando uma natural proximidade com as técnicas metodológicas usadas pela CI para organização do conhecimento humano (MENDONÇA, 2015). A partir dos anos 90, o autor afirma, que, apesar da proximidade com o assunto, apenas nesta década surgiu o interesse da CI, através dos trabalhos de Soergel (1997) e Vickery (1997), considerados precursores nesta área ao pesquisarem a sua aplicação no campo da biblioteconomia.

Modelos baseados em ontologias equivalem a uma teoria usada para representar os principais fatos e regras que governam certa parte da realidade, e com a criação das ontologias passa a se avaliar fatos reais, identificando a estrutura sobre a qual se organizam, realizando generalizações e abstrações (MENDONÇA, 2015). Problemas relacionados à representação, reuso, compartilhamento, aquisição e integração do conhecimento; processos de linguagem natural; tradução automática; interoperabilidade entre sistemas, agentes companhias ou pessoas, podem ser resolvidos também através da utilização das ontologias (PONTES JUNIOR; CARVALHO; AZEVEDO, 2013). Atualmente, busca-se pela utilização das ontologias nos processos inteligentes de *web* e *web* semântica, onde é possível realizar a padronização das terminologias por serem formadas por termos, definições e relações.(SALES; CAMPOS; GOMES, 2008; MENDONÇA, 2015).

No campo da Linguística as ontologias são entendidas como uma nova representação ou estrutura de modelagem para a pesquisa linguística, com as capacidades computacionais resultantes, permitindo novas formas de coleta de conhecimento e desenvolvimento de hipóteses por linguistas (SCHALLEY, 2019). O autor ainda afirma que no campo da IA e na representação do conhecimento, estamos lidando com um domínio, por exemplo, uma linguagem específica, um fenômeno linguístico específico como a evidencialidade ou uma subdisciplina linguística específica como a Fonologia.

Gruber (1993) define uma ontologia como sendo uma especificação explícita de uma conceitualização, onde que esta conceitualização é definida por uma visão abstrata e simplificada do mundo que desejamos representar para algum propósito. Segundo Arp, Smith e Spear (2015) uma definição básica afirma que ontologias são artefatos representacionais que compreendem uma taxonomia como parte apropriada, onde suas representações visam designar alguma combinação de universais, classes definidas e certas relações entre elas, onde são consideradas um instrumento de representação com maior expressividade que as terminologias, devido ao seu uso por computadores, possibilitando uma representação do conhecimento do mais próxima da realidade.

É possível também definir uma ontologia como uma descrição formal de entidades e suas propriedades, relacionamentos, restrições e comportamentos, obtida pelo processo de construção compreendida por etapas, tais como, a especificação, conceitualização formalização, integração, implementação e avaliação ontológica. As ontologias são um instrumento para a organização e representação do conhecimento, capazes de capturar, mapear e compartilhar o conhecimento de diversos tipos de domínio, com o intuito de estabelecer a verdade sobre a realidade, encontrando respostas para a questão 'do que existe' (GRUNINGER; FOX, 1995; COELHO; ALMEIDA, 2012a; MENDONÇA, 2015).

As ontologias podem ainda ser definidas como artefatos de representação, o qual possui uma taxonomia central, onde as unidades representacionais designam uma combinação de universais, classes definidas e relações definidas entre os universais e classes (ANDRADE, 2013). O mesmo autor afirma que as ontologias devem ter uma relação direta com a realidade sob alguma perspectiva definida e compartilhada, desta forma, excluindo informações probabilísticas, contextuais e epistemológicas. Ontologias são expressas em formalismos baseados em lógica, de modo que motores de inferência e algoritmos baseados em lógica possam manipular as definições de acordo com a estrutura da ontologia. Já An e Park (2018) as definem como uma linguagem utilizada na construção de modelos que suportam integração conceitual dos dados distribuídos de um domínio e a inferência de relações entre os conceitos como resultado de uma atividade, tal como análises de conceitos e modelagem de domínio utilizando-se de metodologias padrão.

Segundo Mendonça (2015) as ontologias são tipificadas em três apresentações: i) *ontologias de alto nível*; ii) *ontologias de domínio* e; iii) *ontologias de tarefa*. *Ontologias de domínio* descrevem o vocabulário de um domínio, por exemplo, Medicina, setor Automotivo ou Elétrico, foram desenvolvidas com o objetivo de realizar uma descrição de um vocabulário relacionado a um domínio, como: o setor Elétrico, a Medicina, o Direito entre outros, expressando conceitualizações específicas para um determinado domínio do conhecimento. Toda ontologia de domínio deve utilizar como referência algum tipo de ontologia de nível superior, onde os principais objetivos para o desenvolvimento de uma ontologia de domínio estão ligados ao compartilhamento das informações, reuso de elementos de domínio, separação do conhecimento de um domínio de conhecimento operacional e análise do conhecimento do domínio (COELHO, 2012; MENDONÇA, 2015).

Já as *ontologias de alto nível* descrevem conceitos gerais como espaço, tempo, matéria, objeto, evento, ação, dentre outros, que são independentes do problema ou domínio. As ontologias de alto nível, segundo Coelho (2012) foram desenvolvidas para descreverem conceitos gerais como espaço, tempo, matéria ou um objeto não dependendo de um domínio. São conhecidas como ontologias básicas ou de fundamentação por serem concebidas como ontologias dos constituintes básicos da realidade, sendo assim, não é papel deste tipo de ontologia definir classes ou entidades funcionais. Ontologias de alto nível foram concebidas especificamente para serem pequenas para que possam ser capazes de representar as categorias de alto nível comuns às ontologias de domínio. As *ontologias de tarefa* descrevem uma tarefa ou atividade como, por exemplo, diagnóstico ou compras, através da inserção de termos especializados.

Um exemplo de uma ontologia de alto nível é a *Basic Formal Ontology* (BFO), utilizada como base ou ponto de partida para a construção da ontologia de domínios específicos. É uma ontologia de formação realista, que fornece suporte aos desenvolvedores das ontologias de domínio, orientando-se em suas tomadas de decisões ontológicas. Sua primeira publicação aconteceu no ano de 2004 na versão 1.0. A BFO é uma ontologia que usa tanto *Instâncias* quanto *Universais* (MENDONÇA, 2015). Seu propósito é promover os níveis mais genéricos para a classificação de entidades do mundo, especialmente no domínio Biomédico, reconhecendo apenas as entidades que existem na realidade, rejeitando as

entidades negativas (que indicam faltas, ausências, inexistências, possibilidades e similares), incluindo um conjunto de entidades (denominadas universais e instâncias) e a relação entre essas (TEIXEIRA, 2019).

Um segundo exemplo é a *Unified Foundational Ontology* (UFO). Assim como a BFO, a UFO, também é uma ontologia de alto nível, desenvolvida em 2009 como um projeto unificado da *General Formal Ontology* (GFO) e da ontologia de alto nível subjacente a *OntoClean*. Possui como seu propósito principal promover uma teoria para resolver os problemas clássicos da modelagem conceitual. Inclui teoria de partes e todos os tipos, estruturas taxonômicas, relacionamentos, atributos e atributos de espaço, papéis e qualidades individuais, além de outras coisas. Do ponto de vista filosófico, a UFO é uma ontologia de *universais* e não de *particulares*. Assim como outras ontologias de alto nível sua distinção básica é entre *endurantes* ou entidades *perdurantes* (eventos) (MENDONÇA, 2015).

As ontologias também podem ser caracterizadas por seus componentes básicos: i) *classes*: organização dos conceitos de um domínio, dispostos através de uma taxonomia; ii) *relações*: representam todo tipo de interação entre as classes de domínio; iii) *axiomas*: regras formais usadas para restringir a interpretação e o uso dos termos envolvidos, geralmente, especificados em uma linguagem formal da lógica; iv) *instâncias*: exemplos ou casos particulares das classes, usadas para representar objetos específicos. Além das quatro características básicas apresentadas é importante destacar as *questões de competência*. As questões de competência são utilizadas para especificar as tarefas e os problemas que uma ontologia pode solucionar, antes mesmo da sua construção, ou seja, uma forma de auxiliar na definição do escopo e das características da ontologia.

Segundo Uschold e King (1995) as questões de competência podem ser vistas como relacionadas ao propósito de uma ontologia em termos muito específicos, podendo desempenhar o papel de um documento de referência em uma metodologia abrangente.

2.2 Ontologia: caracterização filosófica

Como já mencionado, o campo das ontologias vem sendo estudado por filósofos desde a Grécia antiga, por filósofos como Aristóteles e Platão, que pensam

que não há obstáculo em princípio para que saibamos pelo menos algo sobre a realidade como ela é em si (LOWE, 2006).

As teorias filosóficas são classificadas de acordo com duas posições, realistas e nominalistas, onde o termo realista apresenta a noção de que existe um mundo físico independente da mente e o nominalismo é mostrado como a realidade é fundamentalmente mental, construída mentalmente ou imaterial, como foi descrito por Kant (ALMEIDA, 2013). Ainda segundo o autor, realismo e nominalismo são duas filosofias diferentes que são estudadas para construir ontologias. O realista acredita que existem categorias universais, por outro lado, os nominalistas não acreditam na existência dos universais, utilizando apenas os particulares.

Aristóteles utilizou a linguagem como uma forma de descrição das categorias ontológicas, já Kant utilizou alguns conceitos como forma de abordar categorias de objetos de cognições possíveis. Ambos os filósofos tinham como objetivo descrever um sistema de categorização que o mundo possui de acordo com os pensamentos e a linguagem humana (ALMEIDA, 2013).

Tanto no nominalismo quanto no realismo, a representação das categorias ontológicas por categorias lógico-gramaticais de uma teoria é de forma lógica e direta. Já as representações no realismo lógico sejam talvez ainda mais diretas e simples do que o nominalismo (COCCHIARELLA, 1996).

A visão Aristotélica demonstra pontos importantes que também categorizam uma ontologia dentro de um esquema chamado de quadrado ontológico, onde é possível representar a realidade e suas entidades. O quadrado ontológico (Quadro 1) divide as entidades entre *universais* e *particulares*. As entidades *universais* referem-se a tipos de coisas que existem no mundo real e que possuem características em comum, que são instanciais ou exemplificadas, por mais de um particular (MENDONÇA, 2015). Já as entidades *particulares* são chamadas de *instâncias*, *tokens* ou *individuais* e referem-se a um objeto específico ou um indivíduo do mundo real. O mesmo autor demonstra uma importante diferença entre tais tipos de entidades e refere-se à predicação em uma declaração lógica onde particulares não podem ser predicados de um sujeito, já *universais* podem aparecer tanto como sujeito ou predicado.

Quadro 1 – Quadrado Ontológico

	Substanciais	Acidentais (Não Substanciais)
Universais	Tipos (por exemplo, Ser Humano)	Atributos (por exemplo, Sabedoria)
Particulares (individuais)	Substâncias (por exemplo, Sócrates)	Modos ou acidentes (por exemplo, Sabedoria de Sócrates)

Fonte: Mendonça, 2015

A visão de Kant demonstra que não é possível afirmar saber nada sobre a realidade "como ela é em si mesma", de modo que a ontologia pode ser coerentemente concebida apenas como a ciência do nosso pensamento sobre o ser, ao invés da ciência do ser como tal (LOWE, 2006).

Ainda na visão de Kant, os julgamentos são cognições conscientes complexas que se referem a objetos diretamente, por meio de intuições ou indiretamente por meio de conceitos. As proposições analíticas contêm o conceito de predicado a priori no conceito de sujeito, enquanto proposições sintéticas não (ALMEIDA, 2013).

O termo nominalismo apareceu pela primeira vez durante o século 12, e as suas definições geram controvérsias, onde o conceito afirma que não há universais que possam ser predicados de outras coisas além da expressão predicativa da linguagem. Alguns consideram que os universais são invocados erroneamente por realistas para explicar as características semânticas da linguagem que podem ser explicadas de forma bastante adequada de outras maneiras (BURGESS, 1983; COCCHIARELLA, 1996; BURKHARDT; SMITH, 2001; LOWE, 2006).

Filósofos nominalistas adotam uma filosofia da ciência instrumentalista, onde a ciência é uma mitologia útil, e não há nenhum tipo de aproximação ou idealização da verdade (BURGESS, 1983).

Segundo Andrade (2013) o objetivo mais comum das ontologias em Sistemas de Informação é compartilhar bases de conhecimento; permitir comunicação entre agentes; integrar conjuntos de dados heterogêneos; apoiar suporte à decisão; proporcionar arcabouços semânticos para sistemas de informação organizacionais; representar vocabulários em linguagem natural; representar a semântica de serviços e aplicações de *softwares*; e apoiar a busca por conhecimento, proporcionando aprimoramento e melhorias.

O processo de construção de uma ontologia é antecipado por um momento importante, que é a escolha da metodologia correta para a sua construção de acordo com as necessidades do projeto. Diversas metodologias estão disponíveis, tanto para *ontologias de alto nível* quanto para *ontologias de domínio*. Na próxima subseção serão apresentadas brevemente as metodologias utilizadas para a construção de uma ontologia e as etapas necessárias para a sua construção.

2.3 Ontologia: metodologias de construção

O auge do desenvolvimento das metodologias para a construção de ontologias ocorreu durante os anos 90 e se estendeu até o início dos anos 2000, ocorrendo uma proliferação de diversas metodologias e métodos (MENDONÇA, 2015). As metodologias atuais não são utilizadas de forma individualizada, na maioria das vezes elas são usadas em conjunto, para que possam suprir as limitações apresentadas em cada uma delas. As metodologias possuem diferentes detalhamentos, algumas são mais bem descritas, em determinadas etapas em relação a outras e, por isso, o uso em conjunto de várias metodologias é uma solução possível durante o processo de construção de uma ontologia.

É necessário fazer uma diferenciação entre metodologia e método conforme apontado por Mendonça (2015). Para o autor, **metodologia** é uma série integrada de técnicas ou métodos criando uma teoria geral de sistemas referentes a uma classe de pensamentos a serem executados. **Método**, por sua vez, é um conjunto de processos ou procedimentos ordenados usados na engenharia de um produto ou na realização de um serviço, onde processos são compostos de atividades e tarefas atribuídas a um ou mais membros do projeto.

“Metodologia refere-se a conhecimento sobre métodos, isto é, determina “como” e “quando” uma dada atividade pode ser realizada. Assim, uma metodologia é composta de métodos que possui em suas próprias técnicas” (SILVA; SOUZA; ALMEIDA, 2008, p. 63).

Uma metodologia abrangente deve incidir sob as circunstâncias que as questões de competência devem ser usadas para esclarecer o propósito da captura da ontologia e os estágios de codificação, bem como fornecer informações para a fase de avaliação (USCHOLD; KING, 1995). Segundo Uschold e Gruninger (1996)

em geral as metodologias se limitam a estudos de caso do desenvolvimento de uma única ontologia, ou um projeto em particular.

Cada metodologia segue um conjunto de princípios, critérios de projeto e fases no processo de desenvolvimento, entretanto, a ausência de diretrizes e métodos articulados dificulta o desenvolvimento na criação de novas ontologias, onde é possível afirmar que a fonte dos problemas durante o processo de criação é a ausência de um modelo conceitual explícito e totalmente documentado sobre o qual a ontologia é construída (GÓMEZ-PÉREZ; FERNÁNDEZ; DE VICENTE, 1996).

Uma importante consideração a ser feita neste momento consiste na não existência de uma metodologia padrão para a construção de uma ontologia. Cada metodologia possui as suas características e são desenvolvidas de acordo com a necessidade de cada projeto, porém uma ontologia de domínio deve sempre estar ligada a uma ontologia de alto nível para que os direcionamentos para a construção sejam corretos (MENDONÇA, 2015).

Silva, Souza e Almeida (2008) também apontam em seu trabalho problemas relacionados à falta de um padrão para a construção de ontologias e a falta de explicações de como e quais limites podem ser utilizados nas abordagens teóricas dentro do processo de elaboração de uma ontologia.

Uma metodologia é composta por ciclos evolutivos de implementação e *workflows* que descrevem uma série de operações. O número de ciclos de implementação varia de acordo com a necessidade dos desenvolvedores de associar novos elementos de contexto na filtragem de informações, e uma nova fonte de informação (banco de dados) inserida no processo de filtragem (MARAN et al., 2018).

Quando uma nova ontologia é construída, três questões básicas surgem relacionadas às metodologias de ferramentas e linguagens a serem utilizadas no processo de desenvolvimento: i) quais métodos e metodologias podem ser usados para a criação de uma ontologia; ii) que ferramentas que darão suporte no processo de desenvolvimento; e iii) qual linguagem deve ser usada para se implementar uma ontologia (CORCHO; FERNÁNDEZ-LÓPEZ; GÓMEZ-PÉREZ, 2003).

A escolha de uma metodologia abrangente deixa claro o que se aplica à construção de ontologias, em oposição às bases de conhecimento em geral,

esclarecendo também sob quais circunstâncias os estágios de captura e codificação podem ser mesclados (USCHOLD; KING, 1995).

Ainda segundo Uschold e King (1995, p. 6) “O fornecimento de orientações e ferramentas na integração de ontologias podem ser um dos maiores desafios no desenvolvimento de uma metodologia abrangente.”

É possível encontrar diversas metodologias já estudadas, como a *Metodologia de Uschold e King*, *Methontology*, *Método Kactus*, *Método 101*, *Método CYC*, *Metodologia NeOn*, dentre outras. Na sequência serão apresentadas breves descrições de algumas destas metodologias, como a *Methontology*, Metodologia NeON e a Metodologia *OntoForInfoScience*, desenvolvida especificamente para o campo da CI.

2.3.1 Methontology

É uma metodologia bem estruturada para construção de ontologias, tendo seu desenvolvimento entre os anos de 1996 e 1997, por meio da experiência de seus autores Gómez-Pérez, Fernández e De Vicente na construção de uma ontologia no domínio bioquímico (MENDONÇA, 2015). Segundo Gómez-Pérez, Fernández e De Vincente (1996a), a engenharia ontológica requer a definição e padronização de um ciclo de vida que vai das definições de requisitos à manutenção do produto acabado, bem como metodologias e técnicas que impulsionam o seu desenvolvimento.

A *Methontology* contempla um conjunto de estágios de desenvolvimento dentro de um ciclo de vida baseado na evolução de protótipos e técnicas para realizar as atividades de planejamento, desenvolvimento e suporte (SILVA; SOUZA; ALMEIDA, 2008).

Este ciclo de vida de desenvolvimento ontológico inclui seis atividades principais: i) AC; ii) construção do documento de especificação de requisitos; iii) conceitualização da ontologia; iv) implementação da ontologia; v) avaliação durante cada fase do ciclo de desenvolvimento; e vi) documentação após cada fase. Como a grande maioria das metodologias, a *Methontology* passou a ser utilizada no desenvolvimento de outras metodologias em diferentes domínios do conhecimento,

sendo bem aceita até hoje (GÓMEZ-PÉREZ; FERNÁNDEZ; DE VICENTE, 1996; MENDONÇA, 2015).

2.3.2 Metodologia NeON

Baseada em três abordagens (Trabalho Metodológico de Engenharia de *Software*, Metodologias e Métodos para Construção de Ontologias e Práticas e Experiências Prévias), a metodologia NeOn caracteriza-se por ser uma metodologia para a construção de redes ontológicas baseado no desenvolvimento colaborativo e argumentativo de ontologias.

No contexto do fim dos anos 90 não existiam metodologias que ajudassem os desenvolvedores de ontologias a construir grandes ontologias de maneiras diferentes, reutilizando e possivelmente realizando reengenharia de recursos do conhecimento, usando alinhamentos, e tendo em vista a contínua evolução das ontologias, além do uso de ontologias embutidas em redes construídas de forma colaborativa por equipes (GÓMEZ-PÉREZ; SUÁREZ-FIGUEROA, 2009).

A metodologia NeON tem sido amplamente utilizada no desenvolvimento de ontologias direcionadas à área de engenharia de *softwares* por apresentar atividades importantes no ciclo de vida de um *software*. Seus desenvolvedores traçaram um panorama de um conjunto de nove cenários possíveis na construção de ontologias e redes ontológicas. São estes: i) construção de redes ontológicas heterogêneas sem o reuso de recursos ontológicos existentes; ii) construção de redes ontológicas pelo reuso e engenharia de recursos não ontológicos; iii) construção de redes ontológicas pelo reuso de recursos ontológicos; iv) construção de redes ontológicas pelo reuso e reengenharia de recursos não ontológicos; v) construção de redes ontológicas pelo reuso e fusão (*merging*) de recursos ontológicos; vi) construção de redes ontológicas pelo reuso, fusão (*merging*) e reengenharia de recursos ontológicos; vii) construção de redes ontológicas pelo reuso de padrões de desenvolvimento ontológicos (*ontology design patterns* – ODPs); viii) construção de redes ontológicas pela reestruturação de recursos ontológicos; e ix) construção de redes ontológicas pela localização de recursos ontológicos (SUÁREZ-FIGUEROA, 2010; MENDONÇA, 2015).

Essa estrutura de nove etapas é baseada em quatro pilares: i) um glossário de processos e atividades; ii) um conjunto de nove cenários para construir ontologias e redes de ontologias; iii) dois modos de organizar desenvolvimentos de ontologias, chamados modelos de ciclo de vida de ontologias; e iv) um conjunto de diretrizes metodológicas precisas para a execução de processos específicos e atividades (SUÁREZ-FIGUEROA; GÓMEZ-PÉREZ; FERNÁNDEZ-LÓPEZ, 2015).

Cada um dos cenários é decomposto em diferentes processos e atividades daqueles incluídos no Glossário NeOn. Para cada processo ou atividade, são fornecidas diretrizes metodológicas detalhadas. Nesse sentido, a metodologia é escrita com um processo ou uma abordagem centrada na atividade e de maneira prescritiva e não descritiva (SUÁREZ-FIGUEROA, 2010).

2.3.3 *OntoForInfoScience*

Iniciativas na criação de metodologias próprias para a construção de ontologias, reutilizando etapas de metodologias já existentes, realizando as adaptações necessárias e têm surgido em diferentes pesquisas relacionadas ao campo de estudo referente às ontologias, sendo bem aceitas pela comunidade científica que investiga as ontologias. Neste sentido a metodologia *OntoForInfoScience*, vem com o intuito de tornar o processo de construção de ontologias mais palatáveis aos profissionais da CI (MENDONÇA; ALMEIDA, 2016).

Desenvolvida a partir da experiência de construção de uma ontologia sobre componentes do sangue humano – HEMONTO -, localizada em um domínio biomédico e caracterizada pelos componentes do sangue humano usados para fins terapêuticos, os quais são chamados de hemocomponentes e hemoderivados do sangue. Com a intenção de que a metodologia seja um instrumento útil e de suporte aos desenvolvedores de ontologias, principalmente aqueles ligados à CI, através de uma linguagem simples, foram desenvolvidas etapas essenciais no ciclo do desenvolvimento ontológico, que não estão bem explicadas ou explícitas nas metodologias de construção atuais (MENDONÇA, 2015).

A *OntoForInfoScience* utilizou etapas de metodologias conhecidas como *NeOn*, *Methontology* e o *Método 101*, detalhando as etapas e fazendo adaptações necessárias de cada uma em uma linguagem mais adequada aos desenvolvedores

de ontologia, incluindo os cientistas da informação (MENDONÇA, 2015). Assim, a *OntoForInfoScience* possui nove etapas: i) avaliação da necessidade da ontologia; ii) especificação da ontologia; iii) aquisição e extração do conhecimento; iv) Conceitualização; v) fundamentação ontológica; vi) formalização da ontologia; vii) avaliação da ontologia; viii) documentação da ontologia; e ix) disponibilização da ontologia (MENDONÇA, 2015). A intenção da criação desta metodologia foi criar um modelo que fosse padronizado, com etapas e processos bem definidos, baseado em metodologias já conhecidas.

2.3.4 Comparativo de metodologias

É importante realizar uma comparação das metodologias existentes mais conhecidas para fins de análise. Cada metodologia possui seus passos e processos definidos, mostrando a inexistência de um padrão e a multiplicidade de etapas conforme apresentado no quadro 2. Ainda visando análise comparativa das metodologias é importante conhecer a origem de cada metodologia e seus domínios de aplicação, conforme apresentado no quadro 3. Os dois quadros são apresentados na sequência:

Quadro 2 – Descritivo das etapas das metodologias para construção de ontologias

Metodologia	Etapas (processos e atividades)
<i>Methontology</i> (1996)	1. Especificação: identificar o propósito da ontologia, incluindo os usuários pretendidos, cenários de uso, o grau de formalidade requerido, etc., e o escopo da ontologia, incluindo o conjunto de termos a ser representados. 2. Aquisição do Conhecimento: etapa realizada em paralelo com a etapa (1). Esta etapa é não-prescritiva e, assim qualquer método pode ser usado, embora seja mais comum as entrevistas com especialistas e análises de textos do domínio tratado. 3. Conceitualização: identificação de termos do domínio, tais como conceitos, instancias, relações verbais e propriedades, com cada termo tendo uma representação intermediária. 4. Integração: deve-se analisar termos de outras ontologias, tal como o padrão Ontolingua, que possam ser reutilizados na ontologia em construção. 5. Implementação: a ontologia deve ser representada formalmente através de uma linguagem formal, tais como a Ontolingua e a lógica descritiva. 6. Avaliação: são usadas técnicas baseadas nos métodos de validação e verificação dos sistemas de base de conhecimento. Há também um conjunto de diretrizes para avaliar incompletudes, inconsistências e redundância. 7. Documentação: recomenda-se a especificação de uma documentação em linguagem natural ao final de cada fase do ciclo de vida do desenvolvimento da ontologia.

NeON (2010)	1. Gerenciamento de processos e atividades: inclui controle e medida da qualidade da ontologia. 2. Desenvolvimento orientado de processos e atividades: a) Pre-desenvolvimento de processos e atividades: (I) estudo do ambiente; (II) estudo de viabilidade; (III) reuso de ontologias; (IV) reengenharia ontológica; (V) reuso de recursos não-ontológicos; e (VI) reengenharia de recursos não-ontológicos. b) Desenvolvimento de processos e atividades: (I) especificação de requisitos da ontologia; (II) conceitualização da ontologia; (III) formalização; (IV) implementação da ontologia; (V) integração com outras ontologias; (VI) modularização da ontologia; (VII) reestruturação da ontologia; (VIII) merge da ontologia; (IX) alinhamento da ontologia; (X) atualização da ontologia; (XI) modificação da ontologia; (XII) localização da ontologia; (XIII) tradução da ontologia; (XIV) anotação da ontologia; (XV) customização. c) Pós-desenvolvimento de processos e atividades: (I) atualização da ontologia; (II) versionamento da ontologia; (III) evolução da ontologia. 3. Suporte aos processos e atividades: engloba as seguintes etapas; (I) aquisição de conhecimento; (II) avaliação da ontologia; (III) documentação da ontologia; (IV) resumo da ontologia; (V) avaliação da ontologia; e (VI) gerenciamento da configuração da ontologia.
OntoForInfoScience (2015)	As 9 etapas para construção da <i>OntoForInfoScience</i>: 1. Avaliação da necessidade da ontologia: realizar avaliação prévia sobre a real necessidade de criação de uma ontologia antes de se iniciar o projeto. 2. Especificação da Ontologia: preencher <i>template</i> de especificações com informações sobre o domínio/escopo geral, propósito geral; classes de usuários; uso pretendido; tipo de ontologias; graus de formalidade e delimitação do escopo de cobertura. 3. Aquisição e extração do conhecimento: adotar métodos de aquisição e extração do conhecimento, tais como a seleção de materiais de referência do domínio. 4. Conceitualização: construir dicionário de conceitos; elaborar a tabela de conceitos e valores; elaborar tabela de conceitos e propriedades; construir o dicionário de verbos e representar modelos conceituais. 5. Fundamentação ontológica: selecionar ontologia(s) de fundamentação para uso e aplicar ontologia selecionada no desenvolvimento. 6. Formalização da ontologia: construir taxonomia geral; definir propriedades descritivas; criar definições formais das classes; criar instâncias das classes; especificar relações ontológicas e definir propriedades das relações. 7. Avaliação da ontologia: definir parâmetros de avaliação; estabelecer critérios de validação e estabelecer

	critérios de verificação. 8. Documentação da ontologia: produzir o documento final com o conteúdo da ontologia. 9. Disponibilização da ontologia: gerar representação da ontologia em linguagem formal e apresentar a ontologia em formato e meio eletrônico.
--	--

Fonte: adaptado de (MENDONÇA, 2015)

Quadro 3 – Origem e propósito das ontologias

Metodologia	Origem e Propósito Principal	Aplicação
<i>Menthontology</i> Gómez-Perez, Fernandez-Lopes Vicente (1996)	Metodologia desenvolvida no Laboratório de Inteligência Artificial da Universidade Politécnica de Madri entre 1996 e 1997, que possibilita a construção de uma ontologia por reengenharia sobre outra ontologia, utilizando-se do conhecimento do domínio tratado. Tal metodologia pode ser usada, segundo seus autores, em quaisquer domínios do conhecimento, embora seu uso mais conhecido e citado ocorreu na criação de uma ontologia no domínio da química.	Diversos
Metodologia <i>NeON</i> Suarez-Figueroa (2010)	Metodologia para construção de redes ontológicas baseado em um desenvolvimento colaborativo e argumentativos de ontologias. Tal metodologia foi desenvolvida em uma abordagem híbrida que combina o trabalho metodológico da área de Engenharia de <i>Software</i> e algumas metodologias para construção de ontologias, especificamente, a <i>Menthontology</i> , a <i>On-To-Knowledge</i> , a <i>DILIGENT</i> e outros métodos ontológicos, como o de Grüninger e Fox. A <i>NeON Mentontology</i> inclui: (i) o Glossário <i>NeOn</i> de Processos e Atividades (ao todo tem-se 59 processos e atividades definidos), o qual identifica e define os processos e atividades potencialmente envolvidos quando redes ontológicas são construídas colaborativamente ; e (ii) Duas Redes Ontológicas sobre os modelos de ciclo de vida.	Diversos
<i>OntoForInfoScience</i> Mendonça (2015)	Desenvolvida de uma a partir de da ontologia HEMONTO, localizada em um domínio biomédico e caracterizado pelos componentes do sangue humano usados para fins terapêuticos, os quais são chamados de hemocomponentes e hemoderivados do sangue. A <i>OntoForInfoScience</i> utilizou de algumas etapas de metodologias conhecidas como <i>NeOn</i> , <i>Menthontology</i> e o <i>Método 101</i> , detalhando as etapas e fazendo adaptações necessárias de cada uma em uma linguagem mais adequada aos desenvolvedores de ontologia, incluindo os cientistas da informação	Biomédico

Fonte: adaptado de (MENDONÇA, 2015)

2.3.5 Armadilhas na construção de ontologias

No âmbito da AC, cabe ainda citar questões que envolvem a ontologia como um todo, referentes a anomalias ou a falta de boas práticas durante o processo. A pesquisa e a inspeção manual de ontologias já implementadas resultaram na identificação de “armadilhas” comuns que ocorrem mesmo que sejam seguidas diretrizes de modelagem usuais e fundamentos ontológicos (POVEDA-VILLALÓN; SUÁREZ-FIGUEROA; GÓMEZ-PÉREZ, 2010). Um catálogo dessas armadilhas envolve cinco tipos principais que fazem parte da etapa de AC e são denominados de várias formas (GÓMEZ-PÉREZ, FERNÁNDEZ-LÓPEZ e CORCHO, 2004): i) falhas de entendimento humano; ii) consistência lógica; iii) representação do mundo real; iv) modelagem computacional; e, v) especificação da linguagem de representação, conforme apresentado no quadro 4.

Quadro 4 – “Armadilhas” na construção de ontologia

Item	Tipo	Nome do Erro	Descrição
1	Entendimento Humano	Criação de Elementos Polissêmicos	Um elemento de ontologia cujo nome tem significados diferentes é incluído na ontologia para representar mais de uma ideia conceitual.
2	Entendimento Humano	Criação de Classes Sinônimas	Várias classes cujos identificadores são sinônimos são criadas e definidas como equivalentes.
3	Problemas de Modelagem	Criação da <i>Relação “is”</i> ou <i>“rdf:type”</i> ou <i>“owl: same as”</i>	o relacionamento <i>“is”</i> é criado na ontologia ao invés de usar primitivas OWL para representar o relacionamento da subclasse (<i>“subclassOf”</i>), a associação a uma classe (<i>“instanceOf”</i>) ou a igualdade entre as instâncias (<i>“sameAs”</i>).
4	Problemas de Modelagem	Criação de elementos ontológicos não-conectados	Os elementos da ontologia (classes, relacionamentos ou atributos) são criados sem relação com o resto da ontologia.
5	Consistência Lógica / Problemas de Modelagem / Representação do Mundo	Definindo relações inversas erradas	Duas relações são definidas como relações inversas quando, na verdade, não são
6	Consistência Lógica/ Problemas de Modelagem	Inclusão de Ciclos na Hierarquia	Um ciclo entre duas classes na hierarquia está incluído na ontologia, embora não se pretenda ter tais classes como equivalentes. Ou seja, alguma classe A tem uma subclasse B e ao mesmo tempo B é uma superclasse de A.
7	Problemas de Modelagem	Mesclar diferentes Conceitos em uma Mesma Classe	É criada uma classe cujo identificador se refere a dois ou mais conceitos diferentes.
8	Entendimento Humano	Anotações Perdidas	Os termos da ontologia não possuem propriedades de anotação. Este tipo de propriedades melhora o entendimento da ontologia e a usabilidade.

9	Representação do Mundo	Ausência de Informações Básicas	As informações necessárias não são incluídas na ontologia. Às vezes, essa armadilha está relacionada aos requisitos do ORSD que não são cobertos pela ontologia. Outras vezes, está relacionado com conhecimentos que poderiam ser agregados à ontologia para torná-la mais completa.
10	Problemas de Modelagem / Representação do Mundo	Disjunções Perdidas	A ontologia carece de axiomas disjuntos entre classes ou entre propriedades que deveriam ser definidas como disjuntas.
11	Entendimento Humano / Problemas de Modelagem	Ausência do Domínio ou do Alcance das Propriedades	Relacionamentos e / ou atributos sem domínio ou intervalo (ou nenhum deles) são incluídos na ontologia. Existem situações em que a relação é muito geral e o alcance deve ser o conceito mais geral <i>"Things"</i> . No entanto, em outros casos, as relações são mais específicas e é uma boa prática especificar seu domínio e / ou extensão.
12	Entendimento Humano / Problemas de Modelagem	Ausência da Propriedade Equivalente	Quando uma ontologia é importada para outra, as classes que são duplicadas em ambas as ontologias são normalmente definidas como classes equivalentes. No entanto, o desenvolvedor da ontologia perde a definição de propriedades equivalentes nos casos de relacionamentos e atributos duplicados.
13	Problemas de Modelagem	Ausência da Relação Inversa	Existem dois relacionamentos na ontologia que devem ser definidos como relações inversas.
14	Consistência Lógica / Problemas de Modelagem	Uso Incorreto de <i>"owl:allValuesFrom"</i>	Essa armadilha aparece de duas formas: i) a anomalia é usar a restrição universal (<i>"allValuesFrom"</i>) como o qualificador padrão em vez de usar a restrição existencial (<i>"someValuesFrom"</i>). ii) erro é incluir <i>"allValuesFrom"</i> para excluir outras adições para uma determinada propriedade.
15	Consistência Lógica / Problemas de Modelagem	Uso indevido de <i>"not some"</i> e <i>"some not"</i>	Confundir a representação de <i>"not some"</i> com <i>"some not"</i> , ou o contrário.

16	Consistência Lógica / Problemas de Modelagem	Uso indevido de classes primitivas e definidas	Deixar de fazer a definição "completa" em vez de "parcial" (ou "necessário e suficiente" em vez de apenas "necessário"). É fundamental entender que, em geral, nada será inferido para ser incluído em uma classe primitiva pelo classificador. Esta armadilha implica que o desenvolvedor não entende a suposição de mundo aberto
17	Problemas de Modelagem	Hierarquia Muito Especializada	A hierarquia na ontologia é especializada de tal forma que as folhas finais não podem ter instâncias, porque na verdade elas são instâncias e deveriam ter sido criadas dessa forma em vez de serem criadas como classes.
18	Consistência Lógica / Problemas de Modelagem	Grande Especificação do Alcance ou Domínio	Não para encontrar um domínio ou intervalo que seja em geral o suficiente.
19	Entendimento Humano / Consistência Lógica / Problemas de Modelagem	Troca da Intercessão ou União	Os intervalos e / ou domínios das propriedades (relacionamentos e atributos) são definidos pela intersecção de várias classes nos casos em que os intervalos e / ou domínios devem ser a união dessas classes.
20	Entendimento Humano	Troca de etiqueta e comentário	O conteúdo das propriedades de anotação etiqueta e comentário são trocados.
21	Problemas de Modelagem	Uso de uma classe diversa	Criar em uma hierarquia de uma classe que contém as instâncias que não pertencem às classes irmãs, em vez de classificar essas instâncias como instâncias da classe no nível superior. A classe, em geral, é "Outros" ou "Diversos".
22	Entendimento Humano	Uso de Critérios Diferentes de Nomes na Ontologia	Nenhuma convenção de nomenclatura é usada nos identificadores dos elementos da ontologia.
23	Problemas de Modelagem	Uso Incorreto de Elementos da Ontologia	Um elemento de ontologia (classe, relacionamento ou atributo) é usado para modelar uma parte da ontologia que deve ser modelada com um elemento diferente. Um caso particular dessa armadilha em relação ao uso indevido de classes e valores de propriedade.
24	Problemas de Modelagem	Uso de Definições Recursivas	Um elemento de ontologia é usado em sua própria definição.

25	Problemas de Modelagem	Definindo um relacionamento inverso consigo mesmo	Um relacionamento é definido como o inverso de si mesmo.
26	Problemas de Modelagem	Definindo relações inversas em uma relação simétrica	Uma propriedade de objeto simétrico (<i>owl: SymmetricProperty</i>) é definido como inverso de outra propriedade do objeto (usando <i>owl: inverseOf</i>).
27	Consistência Lógica / Problemas de Modelagem / Representação do Mundo	Definindo Relação Equivalente Errada	Duas propriedades de objeto ou duas propriedades de tipo de dados são definidas como equivalentes, usando <i>owl: equalProperty</i> , mesmo que eles não tenham a mesma semântica.
28	Consistência Lógica / Problemas de Modelagem / Representação do Mundo	Definindo Relação Simétrica Errada	Um relacionamento é definido como simétrico, usando <i>owl: SymmetricProperty</i> , quando ele não é necessariamente simétrico.
29	Representação do Mundo .	Definindo Relação Transitiva Errada	Um relacionamento é definido como transitivo, usando <i>owl: TransitiveProperty</i> , quando o relacionamento não é necessariamente transitivo.
30	Problemas de Modelagem	Ausência de Classes Equivalentes	Esta armadilha consiste em perder a definição de classes equivalentes (<i>owl: classe equivalente</i>) em caso de conceitos duplicados. Quando uma ontologia reutiliza termos de outras ontologias, classes que tenham o mesmo significado devem ser definidos como equivalentes de forma a beneficiar a interoperabilidade
31	Consistência Lógica	Definindo Relação Equivalente Errada	Duas classes são definidas como equivalentes, usando <i>owl: equivalentClass</i> , quando não são necessariamente equivalentes.
32	Entendimento Humano / Problemas de Modelagem	Diferentes Classes com o Mesmo Rótulo	Duas ou mais classes têm o mesmo conteúdo para anotações e nomenclatura, por exemplo <i>rdfs:label annotation</i> . Pode envolver falta de precisão ao definir termos.
33	Consistência Lógica / Problemas de Modelagem	Criação de uma cadeia de propriedade com apenas uma propriedade	O construto OWL 2 <i>owl: propertyChainAxiom</i> permite um propriedade a ser definida como a composição de várias propriedades

34	Especificação da Linguagem Ontológica	Classes Sem Tipos	Um elemento de ontologia é usado como uma classe sem ter sido explicitamente declarado como tal usando as primitivas <i>owl: Class</i> ou <i>rdfs: Class</i> .
35	Especificação da Linguagem Ontológica	Propriedades Sem Tipos	Um elemento de ontologia é usado como uma propriedade sem ter sido declarado como tal, usando as primitivas <i>rdf:Property</i> , <i>owl:ObjectProperty</i> ou <i>owl:DatatypeProperty</i> .

Fonte: adaptado de (Gómez-Pérez; Fernández-López; Corcho, 2004; Poveda-Villaon; Suárez-Figueróa; Gómez-Pérez, 2010; “OOPS! – OntoOlogy Pitfall Scanner! – Pitfall Catalogue”,2020)

2.4 Aquisição de Conhecimento

Os estudos relacionados ao comportamento das pessoas e como elas buscam por uma informação datam do meio do século XX, época de considerável desenvolvimento tecnológico. Porém, a busca e o uso de informações teve seu maior desenvolvimento quando agências governamentais, no final dos anos 40, começaram a apoiar e a garantir fundos a um grande número de grupos de pesquisa, técnicos-científicos, ligados às agências de defesa americanas e a *National Aeronautic and Space Administration* - NASA (CHOO, 2006).

A busca por informação é um processo dinâmico e contínuo, constituído tanto por ações e necessidades de um indivíduo ou da sociedade, quanto pelo ambiente físico onde o indivíduo coleta e usa as informações. A necessidade pela busca da informação leva um indivíduo, ou seu grupo social, a realizar a busca e a coleta em um ambiente onde esta informação pode ser registrada, compreendida e transmitida para outros indivíduos ou grupos. Tais ações são importantes em muitas atividades humanas e nos sistemas sociais, onde o uso e a análise de uma informação têm se tornado um componente importante dentro das disciplinas de pesquisa (CHOO, 2006).

As atividades que envolvem a AC são fundamentais. Entretanto, estas atividades, carecem de metodologias de execução confiáveis, tanto para a extração do conhecimento quanto para a representação cognitiva pelos agentes envolvidos (sejam os especialistas de domínio ou os profissionais da informação), ou seja, o registro do conhecimento é uma barreira existente nesse processo (COELHO; ALMEIDA, 2012, 2015).

A AC é um método utilizado desde os anos 80, quando surgiram os primeiros estudos com sistemas computacionais, referindo-se como o conhecimento produzido por especialistas de um domínio pode ser representado usando sistemas computacionais (COELHO; ALMEIDA, 2015). Já nos anos 90, esse tipo de atividade estava ligado às práticas de GC, sendo integrada a um grupo de estratégias pioneiras que buscaram capturar, o conhecimento individual, e converter em conhecimento organizacional. Muitos exemplos históricos podem ilustrar a marcante evolução dos sistemas, técnicas e/ou máquinas utilizados para a recuperação da

informação, desde as mais primordiais até as mais modernas formas, transformando a recuperação da informação em um processo altamente interativo, das bases documentais para as bases de conhecimento em diversas formas (SARACEVIC, 1996).

Assim, diferentes definições, teorias e técnicas podem ser encontradas na literatura sobre a AC, porém, o tema recai sobre diversos campos de pesquisa, a Psicologia, a Semiótica e a Linguística (MENDONÇA et al., 2012; COELHO; ALMEIDA, 2015). Dehghani e Akhavan (2017) definem a AC como processo de extrair, estruturar e organizar o conhecimento de diversas fontes, tais como especialistas, ou ainda, o processo de acessar e absorver o conhecimento de forma direta ou indireta por meio de interação com a fonte conhecimento. Já Postolache (2016) aponta que a “AC é uma metodologia sistemática, que envolve a apresentação de processos, procedimentos e passos que surgem através de uma proposta de AC”.

Em ontologias, a etapa de AC é pouca estudada, o que pode ser comprovado pelo baixo número de publicações, ainda que o assunto seja de grande relevância (COELHO; ALMEIDA, 2012; SHU; QIAN; XIE, 2019). Quando se busca por trabalhos sobre o tema nas bases de dados, encontra-se muitos trabalhos voltados para AC relacionada à aprendizagem escolar, ou um conhecimento geral sobre um determinado assunto.

Após a realização de uma busca nas bases de dados relevantes para a CI, além do portal CAPES, por trabalhos publicados nos últimos cinco anos, foram retornados 63 trabalhos em língua portuguesa ou inglesa sobre o tema AC, conforme Tabelas 1 e 2. Destes, após realizada uma análise dos títulos e resumos, foram excluídos 58, por não estarem relacionados a AC e a construção das ontologias. Os cinco trabalhos restantes foram incluídos nesta seção.

Tabela 1 – *String* de busca por publicações em língua portuguesa

Combinação 1	("Aquisição do Conhecimento") OR ("Elicitação do Conhecimento") OR ("Representação do Conhecimento") OR ("Análise de Domínio") AND ("Ontologia")					
Base/Ano	2015	2016	2017	2018	2019	Total por Base
Web of Science	0	0	0	0	0	0
Google Academics	0	0	0	0	0	0
Scopus	0	0	0	0	0	0
CAPES	4	2	2	2	2	12
JASIST	0	0	0	0	0	0
Total por Ano	4	2	2	2	2	12

Referência: autor

Tabela 2 – *String* de busca por publicações em língua inglesa

Combinação 2	("Knowledge Acquisition") OR ("Knowledge Elicitation") OR ("Knowledge Representation") OR ("Domain Analysis") AND ("Ontology")					
Base/Ano	2015	2016	2017	2018	2019	Total por Base
Web of Science	7	10	3	9	5	34
Google Academics	0	0	0	0	0	0
Scopus	0	1	0	0	0	1
CAPES	2	3	6	3	2	16
JASIST	0	0	0	0	0	0
Total por Ano	9	14	9	12	7	51

Referência: autor

Segundo Shu, Qian e Xie (2019) a AC é o processo de extrair conhecimento útil de um conjunto de dados (conhecimento especializado) para obter conhecimento relevante a um domínio para consequente construção de um sistema baseado em conhecimento.

Uma definição voltada para a criação de ontologias é indicada por Dehghani e Akhavan (2017, p. 494), ao afirmarem que “é o processo de reunir conhecimento sobre um domínio, geralmente de um especialista, e incorporá-lo a um programa de computador”. Este processo é referente a uma das etapas do processo

de construção das ontologias, quando o conhecimento é extraído dos especialistas de domínio. É importante que especialistas com grande conhecimento do domínio estejam dispostos a contribuir na construção do conhecimento de forma clara e coesa. Assim por meio do seu profundo conhecimento, é possível coletar, registrar, analisar e validar as informações levantadas (DEHGHANI; AKHAVAN, 2017).

Por essa característica, a AC é considerada por muitos pesquisadores uma importante ferramenta, uma vez que os especialistas possuem diferentes tipos de conhecimento que podem ser extraídos por diferentes técnicas.

2.4.1 Especialistas e conhecimento especializado

Os especialistas de domínio são essenciais ao longo processo de construção de ontologias. Como dito na seção anterior, é por meio do seu profundo entendimento em um determinado domínio que se é possível extrair o conhecimento desejado, tendo em vista que, para Wang et al (2006), as ontologias são o conhecimento básico estruturado de um domínio gerado através do conhecimento adquirido do especialista.

A busca por uma informação com especialistas de domínio pode ser considerada uma fonte informal muito importante, pois o seu conhecimento tácito, acumulado ao longo da sua experiência com o domínio, pode ser transformado em conhecimento explícito através de entrevistas dentro do processo de AC.

Segundo Akhavan, Shahabipour e Hosnavi (2018, p. 414), “um dos maiores desafios no desenvolvimento de Sistemas de GC é a AC através de especialistas de domínio e a representação do conhecimento na forma mais apropriada e aplicável”

Enquanto a informação possui uma manifestação física como um documento ou registro, o seu contexto e significado são novamente criados cada vez que ela é utilizada por um usuário. A informação é fabricada pelos indivíduos, que utilizam das suas experiências passadas para construir novas, de acordo com as exigências de situações particulares em que a informação é melhor utilizada (CHOO, 2006).

A criação do conhecimento é alcançado através do reconhecimento da relação de sinergia entre os conhecimentos tácitos e explícitos e através do *design*

de processos sociais que criam um novo conhecimento pela conversão do conhecimento tácito em explícito (CHOO, 1996).

Ainda de acordo com o autor, o conhecimento tácito é um conhecimento pessoal que é difícil de se formalizar ou de transmitir para outras pessoas, consistindo de um conhecimento subjetivo, intuições que vem de uma pessoa que está imersa em uma atividade por um período extenso de tempo. Tal definição de conhecimento tácito se aplica ao especialista de domínio que possui um amplo conhecimento através de sua experiência profissional ou acadêmica, entretanto, o conhecimento explícito é um conhecimento formal, facilmente disseminado entre os indivíduos de um grupo, frequentemente articulado por fórmulas matemáticas, regras, especificações e demais formas de documentação (CHOO, 1996).

Tal conhecimento possui fácil disseminação, pois é encontrado em livros, artigos científicos e bibliotecas, porém não conta com a experiência profissional. Desta forma, o autor também aponta que as organizações devem tornar-se habilitadas em converter esse conhecimento pessoal e tácito, em um conhecimento documentado e explícito, o que pode estimular a inovação.

É possível considerar como especialista, aquele indivíduo que adquire um determinado nível de conhecimento sobre um domínio específico ao longo do tempo. Tal conhecimento, é adquirido através de um processo de aprendizagem e experiência (COELHO, 2012) onde os indivíduos chamados especialistas são aqueles que, por critérios objetivos e ao longo do tempo, demonstram um desempenho superior em atividades típicas de um domínio. É possível afirmar que um indivíduo é especialista quando é analisada a sua experiência, tanto profissional quanto acadêmica, já o desempenho de um especialista é demonstrado e comparado, na maioria das vezes, com seus pares que possuem menor experiência em um determinado domínio (COELHO, 2012).

O conhecimento especializado pode ser adquirido pela comunicação oral, encontros ou até mesmo por conversas feitas por telefone, principalmente em conversas de pouca duração (CHOO, 2006). O grande trabalho do especialista de domínio neste processo é conseguir fazer com que o conhecimento seja transmitido para o profissional da informação com o menor número de dúvidas possível, sendo necessário, portanto, empreender técnicas específicas para tal. Tais técnicas devem ser devidamente estudadas para se encontrar a que melhor se encaixa no perfil do

estudo pesquisador e especialista. A complexidade das tarefas depende do conhecimento, ferramentas e técnicas utilizadas para transformação dos resultados da informação em resultados organizacionais (CHOO, 2006).

2.4.2 Extração de conhecimento especializado

O processo de AC é de grande importância, porém não é um processo fácil de se realizar por diversos motivos. Dentre eles pode-se considerar a dificuldade na disponibilidade de especialistas na transmissão do conhecimento de forma compreensível. Por este motivo é necessário que todo conhecimento adquirido pelo pesquisador informacional seja validado.

Como dito na seção anterior, a escolha de um especialista de domínio é baseada no tempo de experiência e em critérios específicos relacionados ao domínio e tipo de pesquisa. O papel desta figura é indicar se o caminho buscado pelos profissionais da CI está correto, apontando os erros e ajudando na correção do estudo. Isto só é possível quando se encontra a melhor técnica de AC dentro das inúmeras existentes.

A AC requer diversos passos, devido à existência de um grande volume de dados do mundo real. Para se extrair estes dados são necessárias técnicas para reduzir a quantidade (ou dimensão) dos dados de entrada para torná-lo gerenciável por algoritmos de processamento, como agrupamento e métodos estatísticos (GANZ; BARNAGHI; CARREZ, 2016).

A pesquisa aqui explora as diversas técnicas de AC existentes numa tentativa de identificar aquelas que pudessem se adequar ao processo de criação de ontologias. É relevante considerar que se propõe identificar técnicas aplicáveis a cada uma das fases e que possibilitem interação com o especialista de domínio. É importante destacar que existem diversas propostas de classificação baseadas em diversos autores, que explicam como cada técnica é utilizada (COELHO, 2012).

As técnicas para realização da AC podem ser divididas em três propostas de qualificação, a saber i) sob o ponto de vista da aplicação; ii) pelo tipo de interação; e iii) em relação ao conhecimento obtido (COELHO, 2012). A autora define que *sob o ponto de vista da avaliação* as técnicas podem ser classificadas como técnicas manuais ou técnicas baseadas em computador, onde a primeira

possui raiz na psicologia. Já as técnicas baseadas em computador fazem uso de ferramentas semiautomáticas, conhecidas como interativas, e ferramentas do processamento da linguagem natural, em geral identificadas pela expressão aprendizagem de máquinas.

As técnicas classificadas *pelo tipo de interação* se dividem em diretas ou indiretas, de acordo com o tipo de interação tida com o especialista. As técnicas diretas são usadas para eliciar conhecimento por meio de questionamentos diretos aos especialistas, de modo que este possa expressar sua perícia. Contudo, as técnicas indiretas são usadas para viabilizar questionamentos indiretos em situações em que o conhecimento especializado não é facilmente expressado pelo especialista ou encontra-se registrado em outras fontes (COELHO, 2012). A classificação *em relação ao conhecimento obtido* demonstra que a partir de diferentes métodos de eliciação, derivam-se de diferentes tipos de conhecimento, sendo conhecida na literatura científica como a hipótese do acesso diferenciado.

A importância de se conhecer as técnicas existentes consiste na capacidade de adaptar a melhor forma de se extrair o conhecimento do especialista de acordo com o projeto e as pessoas que fazem parte dele. Desta forma, é possível reduzir algumas barreiras como distância, tempo e compressão sobre o domínio. É possível exemplificar algumas técnicas como Entrevistas Estruturadas, *Card Sorting*, Mapas Conceituais e Protocolos de Análise (DEHGHANI; AKHAVAN, 2017).

A seguir serão apresentadas algumas das técnicas no processo de AC com uma breve descrição de como elas são aplicadas em um processo que envolve o engenheiro informacional e o especialista.

2.4.2.1 Geração de protocolos

A geração de protocolos corresponde a um conjunto de técnicas que inclui entrevistas, que podem ser não estruturadas, semiestruturadas ou estruturadas, além de técnicas de comunicação, como o auto relato, e de técnicas de observação. O objetivo dessas técnicas é produzir um protocolo que seja um registro de comportamentos, que podem ser documentados em áudio, vídeo ou qualquer outro tipo de mídia eletrônica (COELHO, 2012). As entrevistas são um

processo que o engenheiro informacional interroga o especialista sobre o domínio de seu conhecimento. A observação é um processo que um observador determina a informação útil e observa uma tarefa a ser completada pelo especialista. Ainda segundo a autora o sucesso de uma entrevista depende das perguntas a serem feitas e da capacidade do especialista em articular aquilo que conhece.

Boose (1990) demonstra a existência de sete tipos diferentes de entrevista: i) entrevista não estruturada; ii) entrevista semiestruturada; iii) entrevista estruturada; iv) entrevista de demonstração; v) programação neolinguística; vi) entrevista tutorial; e vii) entrevista de processo. Já dentre as técnicas de observação, o mesmo autor cita: i) a de observador (*Observation Participant*), o profissional da informação torna-se um observador ou um participante no processo de resolução problema do especialista de domínio, e ii) a de análise de protocolo, o profissional da informação registra ou transcreve o que o especialista de domínio diz (BOOSE, 1990).

2.4.2.2 Análise de protocolo

A análise de protocolo é usada na transcrição de entrevistas ou de outros registros textuais onde é realizada a análise do registro gerado pelo especialista ou para identificar os diferentes tipos de conhecimento. Pode também atuar como uma ponte entre o uso das técnicas de protocolo e as técnicas de modelagem do conhecimento, além de envolver a identificação dos objetos do conhecimento no âmbito de um protocolo (COELHO, 2012).

2.4.2.3 Baseadas em matriz

Essa técnica envolve a construção de grades, onde se é possível realizar indicações de problemas e possíveis soluções. Neste caso, as matrizes são utilizadas para representar as propriedades de conceitos e o *Repertory Grid* para elicitar, calcular, analisar e categorizar as propriedades dos conceitos. O *Repertory Grid* é útil na resolução de problemas analíticos pelo uso da classificação heurística, que é o processo que envolve questões e respostas do especialista de domínio (WOLF; DELUGACH, 1996). Esse método também faz o uso da psicologia de

construção pessoal e métodos relacionados para obter e analisar o conhecimento, assim como a construção e preenchimento de uma matriz de duas dimensões, como por exemplo, conceitos versus propriedades, problemas versus soluções, hipóteses versus técnicas de diagnóstico, dentre outras (BOOSE; GAINES, 1989; COELHO, 2012).

2.4.2.4 Classificação de cartões (*Cardsorting*)

A Classificação de Cartões é um método de *design* centrado no usuário frequentemente usado para entender as relações de conceito, onde os participantes (usuários representativos) recebem uma lista de conceitos (cartões) para classificar em grupos significativos (REESE et al., 2018).

As entidades de um domínio são classificadas para determinar como o especialista classifica o seu conhecimento, sendo utilizada com o intuito de capturar o modo como especialistas comparam e ordenam os conceitos. A forma mais conhecida e mais simples é a classificação de cartões (*cardsorting*), onde são utilizados cartões com nomes de determinados conceitos.

O especialista tem a tarefa de classificar os cartões em grupos que tenham algo em comum, visando encontrar e organizar termos ou conceitos e seus relacionamentos em um domínio particular, tais como a classificação dos objetos, as hierarquias e outras descrições estáticas dos objetos do domínio de acordo com a visão do especialista (BOOSE, 1990; BURGE, 1998; COELHO, 2012).

2.4.2.5 Limitações de informação e tarefa de restringir processamentos

De acordo com Coelho (2012) as técnicas são utilizadas quando se pretende limitar o tempo e a informação a disposição do especialista, quando este desempenha uma tarefa que normalmente exigiria recursos para desempenhá-la.

Ao especialista é permitido fazer perguntas ao profissional da informação, que responde sim ou não. As perguntas e a ordem na qual elas são proferidas sugerem aspectos importantes e que devem ser considerados para fins de aquisição.

2.4.2.6 Diagrama

Diz respeito à geração e ao uso de redes de diagramas, como mapas conceituais, redes de transição de estado e mapa de processos. O engenheiro informacional suscita ideias no especialista, através de referência em um diagrama de papel ou computador. Tal referência também é empregada quando se utiliza de outros métodos. Mapas conceituais são muito empregados quando se pretende adquirir conhecimento para o desenvolvimento de sistemas orientados ao objeto.

Sob o ponto de vista do emprego dessa técnica, é apresentada uma metodologia e um *plug-in* do *Protégé* para a eliciação do conhecimento a partir de documentos, que utilizam de outras técnicas de AC. O principal objetivo do trabalho é integrar as técnicas de eliciação do conhecimento ao processo de construção de ontologias, de modo a tornar esse processo mais eficiente e menos propenso a erros. A proposta é baseada em princípios de desenvolvimento centrados no usuário, permitindo capturar requisitos para a construção de ontologias.

2.4.2.7 Técnicas de Grupo

Diversas técnicas de AC podem ser realizadas em forma de dinâmicas de grupo. O *brainstorming* é uma técnica muito utilizada na geração rápida de ideias (BOOSE, 1990).

Consiste de uma grande geração de ideias em um curto espaço de tempo, onde é possível divergir em diferentes cenários ou formas de se entender um determinado assunto.

2.4.2.8 Técnicas automatizadas

No campo das ontologias, os estudos referentes a AC automática têm se ampliado e desenvolvido em diversos domínios, o que aumenta a necessidade de compartilhamento e reutilização de ontologias disponíveis (DHINGRA; BHATIA, 2015).

As técnicas automatizadas geralmente utilizam algoritmos para realizar generalizações ou indução do conhecimento a partir de exemplos. Diversas ferramentas combinam métodos interativos com métodos de aprendizagem de

máquinas, onde tipicamente a informação é elicitada dos especialistas de domínio e analisadas (BOOSE, 1990). O principal foco dessas ferramentas de análise é a identificação, categorização das proposições e a organização semântica dos segmentos de texto. Algumas dessas ferramentas são desenvolvidas dentro de um contexto científico, onde seu desempenho é aproximado das técnicas manuais (MASTELLA, 2005).

É necessário desenvolver ferramentas de AC que possam processar os dados disponíveis e transformá-los em uma representação do conhecimento. Em domínios onde estes dados estão presentes em uma forma estruturada de fácil processamento, isso pode parecer uma tarefa fácil, porém muitos sistemas de armazenamento de textos em linguagem natural são complexo, de tal forma que extrair conhecimento destes textos está distante de ser trivial (MÉSZÁROS; KISS, 2018).

Ganz, Barnaghi e Carrez (2016) apresentam em seu trabalho uma abordagem que automaticamente gera uma ontologia semântica a partir de dados brutos que representam informações e novos insights coletados por meio de um processo automatizado de aquisição de conhecimento.

No quadro 5 é apresentada uma representação esquemática de técnicas de AC:

Quadro 5 – Representação esquemática de técnicas de Aquisição do Conhecimento

Categoria	Subcategoria	Objetivo	Instrumento	Interação	Resultado
Entrevista	Estruturada, não estruturada e semiestruturada	Visão Geral do domínio; Elicitar conhecimento básico (facilmente verbalizado); Detalhar o conhecimento sobre conceitos, suas propriedades e relações	Manual	Direta	Varia de acordo com a pergunta. Elicitar principais termos e conceitos representativos do domínio.
	Questionário	Questões que podem ser aplicadas a partir de uma amostragem significativa e representativa de especialistas	Manual	Direta	Varia de acordo com o objetivo da tarefa.
Classificação de Conceitos	Classificação de cartões (<i>Card Sorting</i>)	Revelar o conhecimento em relação a classes, propriedades e prioridades. Organizar termos ou conceitos, seus relacionamentos num dado domínio.	Manual	Indireta	Classificação de entidades, objeto, hierarquias e outras descrições estáticas dos objetos do domínio.
Elicitação de constructos	<i>Repertory Grid</i> (Grades de Repertório)	Obter informações sobre como o especialista discrimina as entidades no domínio	Manual	Indireta	Atributos (entidades fornecidas) Mapeamento conceitual do domínio
	<i>Laddering</i>	Classificar relações entre os constructos elicitados pela técnica de <i>repertory grid</i> e organiza-los em relações hierárquicas. Tipos de <i>ladders</i> : Conceitos : mostrar conceitos e instancias, classes e subclasses às quais pertencem. Composicional : mostrar a maneira como um objeto é composto de suas partes; Decisões : mostrar caminhos de ação alternativos para uma tomada de decisão. Apresentar	Manual	Indireta	Mapa hierárquico das tarefas do domínio.

		conhecimentos procedimental. Atributos: mostrar atributos e seus valores como sub-nós. Mostrar propriedades associadas a um conceito. Processo: Mostrar processos no domínio e seus subprocessos, também representa conhecimento procedimental.			
Observação	Observação Focada	Fornecer impressão global do domínio.	Manual	Direta	Identificação de conhecimento implícito dos especialistas
	Participação Ativa		Manual	Direta	Conhecimento e habilidades necessários para desempenhar uma tarefa
Feedback	<i>Teachback</i>	Obter <i>feedback</i> dos especialistas a partir da descrição de uma parte do conhecimento adquirido com sessões anteriores	Manual	Direta	Correção de equívoco conceitual Esclarecimento de mal entendidos
Mapeamento	Mapeamento conceitual	Ajudar especialistas explorar os relacionamentos entre conceitos	Interativo	Indireta	Mapa conceitual
Análise de Documentos	Imersão na Literatura	Levantar os conceitos básicos o domínio em questão; Compreender minimamente o domínio do conhecimento	Manual	Indireta	Obtenção de principais conceitos do domínio para conhecimento. Compreender um mínimo do domínio, para que as entrevistas possam ser mais produtivas.
	Análise de Assunto	Extrair conceitos que traduzem a essência do conhecimento registrado em um documento	Manual	Indireta	Conceitos que melhor expressam o assunto do documento
	Análise Informal	Estudar principais conceitos nos domínios, obtidos em livros, manuais, publicações, entre outros	Manual	Indireta	Varia dependendo dos documentos disponíveis e se haverá ou não interação com os especialistas

Análise de Domínio	Bibliometria	Mapear campos do conhecimento a partir do domínio	Manual	Indireta	O que é o domínio em estudo
		Definir o domínio do conhecimento	Automático	Indireta	Quem são os especialistas e a literatura Identificar os especialistas a partir das publicações O que é o domínio em estudo
Técnicas de Grupo	<i>Brainstorming</i>	Elicitar conhecimento baseado em consenso	Manual	Direta	Geração de consenso em grupo de especialistas
	<i>JAD – Joint Application Design</i>		Manual	Direta	Discussões em grupo, onde decisões são baseadas em consenso.
Análise de Cenários	Narrativas	Elicitar conhecimento por meio de narrativas	Manual	Direta	Dados sobre as necessidades e atividades dos funcionários das empresas
Formação de consenso	Ferramentas Wiki	Obter consenso	Interativa	Indireta	Obtenção de conhecimento especializado em recursos centralizados Permite aos usuários anexarem comentários

Fonte: adaptado de Coelho (2012)

Devido ao grande número de técnicas para se utilizar durante o processo de AC, o grande desafio, é encontrar a que mais se adequa às necessidades dos profissionais da informação e dos especialistas, de forma a extrair o conhecimento de maneira compreensível e sem ambiguidades. A literatura aponta para o tipo personalização durante o processo de AC e nas técnicas de avaliação (DEHGHANI; AKHAVAN, 2017).

Após a extração do conhecimento, problemas podem ocorrer durante a fase de validação das análises feitas pelos especialistas de domínio e os profissionais da informação. Especialistas tendem a utilizar expressões inadequadas, criando dificuldades para que os profissionais da informação possam compreender e registrar de forma correta o que foi transmitidos pelos especialistas (COELHO; ALMEIDA, 2015).

É importante criar uma relação de troca entre o especialista de domínio e os profissionais da informação envolvidos, para que o conhecimento seja transmitido de forma objetiva e livre de ambiguidades. A disponibilidade ofertada pelos especialistas é importante, para quando surgir os momentos de dúvida, essas possam ser sanadas da melhor forma possível.

Um projeto relacionado à AC sempre estará sujeito a riscos que envolvem o indivíduo que possui determinado conhecimento, podendo levar a uma imprecisão da informação e barreiras de compartilhamento (AKHAVAN; SHAHABIPOUR; HOSNAVI, 2018).

3. Linguagens e representação

O capítulo três descreve linguagens e representações do conhecimento: na seção 3.1, abordam-se as linguagens de representação do conhecimento – enquanto na 3.1.1 aborda-se brevemente a linguagem natural utilizada pelos seres humanos na comunicação, na 3.1.2 traz a linguagem lógica compreendida por máquinas – e, finalmente, na subseção 3.2 a LNC, linguagem livre de ambiguidades e de sintaxe controlada, compreendida tanto por humanos quanto por máquinas.

3.1 Linguagens de representação do conhecimento

As linguagens de representação do conhecimento são definidas como um conjunto de códigos que surgiram em função de um esforço de formalização dos procedimentos de decisão de línguas regulares. É possível definir linguagens formais como um tipo de código de tentativas infinitas, ou seja, como árvores de profundidade infinita se ramificando sobre o alfabeto (TRAYTEL, 2017).

3.1.1 *Linguagem natural*

Linguagem natural é a linguagem utilizada por seres humanos no processo de comunicação, de fácil compreensão, porém com diversas ambiguidades, o que impossibilita a compreensão deste formato de linguagem por máquinas.

A linguagem é um dispositivo natural, um componente da mente humana, representado fisicamente no cérebro e parte do patrimônio biológico da espécie. É a natureza da linguagem humana em suas muitas formas e a maneira como atende às nossas necessidades comunicativas. Em outras palavras, a linguística visa a uma compreensão mais profunda das restrições e forças que moldam as línguas humanas (CHOMSKY, 2002; BOLHUIS; EVERAERT, 2013).

Segundo Chantree (2004), a interpretação semântica do texto em linguagem natural constitui o “Santo Graal” para a linguística computacional, onde interpretar uma frase não significa apenas capturar seu conteúdo proposicional,

mas também requer axiomas ontológicos que apoiem a inferência de implicações adicionais de significado que a frase tem além do que é explicitamente expresso.

Segundo Maculan, “toda linguagem é um sistema complexo e heterogêneo de subsistemas inter-relacionados, e o que difere a linguagem natural da linguagem de especialidade são os seguintes princípios: (1) a linguagem natural é um conjunto de regras, unidades e restrições, sendo de conhecimento da maior parte dos falantes de uma língua; (2) as unidades linguísticas da linguagem natural são utilizadas em situações que podem ser classificadas como não-marcadas; (3) a linguagem de especialidade é parcialmente coincidente com o subcódigo da linguagem natural; (4) as linguagens especializadas possuem características específicas, tais como interlocutores, temática, contexto de comunicação, entre outros, que as tornam linguagens marcadas. Nesse sentido, uma unidade lexical pode ser um termo, desde que aplicado a uma determinada situação de comunicação, estando, portanto, sujeito aos fenômenos típicos dos sistemas linguísticos, tais como as variações” (MACULAN, 2015, p. 74).

. De fato, quando falamos, selecionamos livremente uma estrutura gerada por nosso procedimento recursivo e que está de acordo com nossas intenções comunicativas, uma seleção particular em uma situação de discurso específica é um ato livre de liberdade condicional, mas o procedimento subjacente que especifica os possíveis "padrões regulares" é estritamente governado por regras (CHOMSKY, 2002).

Warren (2017) afirma que linguagem natural oferece a possibilidade de prover a representação do conhecimento em um formato padronizado para a compreensão do ser humano, porém este formato de linguagem gera diversas ambiguidades. Para as máquinas, no entanto, a tarefa de interpretar a linguagem natural, especialmente compreender o significado além do conteúdo literal, tem se mostrado extremamente difícil e requer um grande conhecimento prévio (CHANTREE, 2004).

"Com a semântica estruturalista, pode-se entender que uma linguagem artificial de especialidade é parte integrante da linguagem natural, sendo, de forma similar, complexa (ambiguidade e polissemia)"(MACULAN, 2015, p. 63)

Para humanos interpretarem a linguagem natural, e especialmente compreender o significado além do conteúdo literal, provou ser extremamente difícil e requer uma grande quantidade de conhecimento prévio, em muitos casos, o significado entendido de uma sentença ou expressão em linguagem natural

contém muitas peças implícitas que precisam ser trazidas à tona. (CIMIANO; UNGER; MCCRAE, 2014).

O que permite que os humanos entendam o conteúdo implícito e inferido de sentenças e enunciados é o fato de possuímos muito conhecimento prévio sobre o mundo em que vivemos. Esse conhecimento nos permite traçar todo tipo de inferências semânticas e pragmáticas, que guiam nosso processo de interpretação (CIMIANO; UNGER; MCCRAE, 2014).

A abordagem da interpretação da linguagem natural baseada em ontologia a coloca no centro do processo de interpretação, a fim de abranger e guiar esse processo de uma forma baseada em princípios. Nessa linha, o nível de granularidade representacional em que o significado da linguagem natural é capturado, não é impulsionado pela linguagem, mas pelas distinções semânticas feitas em uma ontologia e, portanto, por distinções que são relevantes no contexto de um domínio particular (CHANTREE, 2004).

As gramáticas que são usadas para interpretar a linguagem natural no contexto de uma determinada ontologia são em grande medida geradas automaticamente a partir do léxico da ontologia. Essas gramáticas contam com um vocabulário alinhado à ontologia e contêm informações sintáticas e semânticas que é crucial para o processo composicional de construção do significado de uma frase particular com respeito àquela ontologia dada (CHANTREE, 2004).

É através da transformação das definições de linguagem natural para os termos contidos na ontologia em um formato, de linguagem lógica, que se é possível realizar a formalização lógica em uma ontologia (ARP; SMITH; SPEAR, 2015).

3.1.2 Linguagem Lógica

Um dos grandes obstáculos para se proporcionar suporte avançado para os usuários *web* é o fato de grande parte dos conteúdos encontrados não serem possíveis de serem acessados por máquinas. Atualmente existe um grande número de ferramentas capazes de coletar textos, fragmentar, verificar ortografia e contar palavras, entretanto, tais ferramentas ainda possuem

limitações que impedem que máquinas possam ser capazes de compreender o que está registrado (MENDES FERREIRA; DE BARROS VIDAL, 2016). Para que seja possível compreender as informações que se encontram na *web*, humanos e máquinas devem compartilhar do mesmo entendimento do mundo real, onde ambos devem compreender as informações contidas no mundo virtual.

A linguagem utilizada na representação de uma ontologia é a linguagem formal, baseada em lógica, e compreensível por máquinas. Com o aumento dos estudos sobre as ontologias nos últimos anos, diversos tipos de linguagens e *softwares* foram desenvolvidos voltados, especificamente, para este campo de pesquisa. As linguagens mais utilizadas e que serão descritas em ordem cronológica nesta seção, conforme aceitação do *World Wide Web Consortium* (W3C) são: a *Resource for Describing Framework* (RDF), *Web Ontology Language* (OWL) e 2 *Web Ontology Language* (OWL2).

No ano de 2004 foi aceito pela W3C a linguagem RDF como uma estrutura para descrição de recursos da *web*. Entretanto, a RDF não inclui constructos específicos para a compreensão das ontologias. A RDFS, um esquema de linguagem RDF, fornece uma estrutura para descrição de uma aplicação específica de classes e propriedades (HEIYANTHUDUWAGE; SCHWITTER; ORGUN, 2016).

A OWL é uma linguagem de *web* semântica desenvolvida para representar o conhecimento sobre coisas, grupo de coisas e o relacionamento entre elas, sendo usada na representação de ontologias como linguagem lógica (linguagem formal). Considerada o padrão oficial para a representação de ontologias, foi aceita em 2004 pelo W3C como substituto das linguagens RDF/RDFS. Por meio da OWL é possível representar diversos aspectos do mundo para fins computacionais através da Lógica Descritiva (LD) - *Description Logics* – um fragmento da Lógica de Primeira Ordem (LPO) – nome dado a uma família de formalismos da Representação do Conhecimento que representa conceitos de um domínio, as propriedade desse conceito e os indivíduos que os instanciam (MENDONÇA, 2015; MENDES FERREIRA; DE BARROS VIDAL, 2016; HEIYANTHUDUWAGE; SCHWITTER; ORGUN, 2016).

O OWL é dividido em três segmentos principais: *axiomas*, *entidades* e *expressões*. *Axiomas* são as expressões básicas de declarações que podem ser

verdadeiras ou falsas; *entidades* são elementos usados para se referir a objetos do mundo real; *expressões* são combinações de entidades de descrições complexas de elementos básicos (MENDES FERREIRA; DE BARROS VIDAL, 2016). Em todas as suas versões – *lite*, *DL* e *full* – a OWL possui sintaxes normativas baseadas em RDF e XML, ambas inerentemente difíceis de ler e escrever (KALJURAND; FUCHS, 2006).

Segundo Kaljurand e Fuchs (2006) é necessário que o usuário possua um amplo conhecimento de LD para se trabalhar em OWL. Com o intuito de possibilitar a adoção de OWL na *web*, os usuários são encorajados ao uso de ferramentas de *front-end*¹, como *Protégé* e *SWOOP*. Ambos são editores gráficos amigáveis, porém, na utilização de descrições de classes complexas, remetem ao uso da LD e falham em esconder as complexidades da OWL (KALJURAND; FUCHS, 2006).

Usuários da linguagem OWL-DL enfrentam problemas lógicos que geram dificuldades, entre elas: i) uma tendência em assumir que os quantificadores existenciais “*only*” e “*some*”, os quais são usados da mesma forma; ii) confusão no uso dos operadores “*and*” e “*or*”; iii) a combinação dos fatores anteriores, onde o uso inadequado do “*and*” entre classes disjuntas gera o constructo *owl:nothing*; e iv) confusão entre *P some* (não X) e *not (P some X)* (WARREN, 2017).

Segundo Heiyanthuduwege; Schwitter e Orgun (2016) desde de 2009, a segunda versão da OWL, a OWL2, tem sido usada como recomendação da W3C como uma linguagem para ontologias para *web* semântica, por ser uma versão mais expressiva da OWL, em que aprimorou principalmente as relações e as expressividades de tipos de dados da linguagem. A OWL2 possui um conjunto de constructos que foram desenhados pensando em aplicações particulares e tarefas de raciocínio.

O perfil OWL 2-EL é adequado para ontologias com um grande número de conceitos ou propriedades para as quais as tarefas básicas de raciocínio exigem tempo polinomial em termos do tamanho da ontologia. Já o perfil OWL 2-QL é recomendado para aplicativos que trabalham com grandes

¹ Interface Visual: Parte visual de um site, programa de computador, aquilo que o usuário final consegue interagir.

volumes de dados, onde as tarefas de resposta a consultas, requerem espaço logarítmico em termos do tamanho dos dados. O perfil OWL 2 RL foi proposto para domínios que exigem raciocínio escalável, porém que apresentam muita energia expressiva em comparação com a OWL 2, podendo usar mecanismos de raciocínio baseados em regras. A resposta de consulta por ontologias da OWL 2 RL exige tempo polinomial em termos do tamanho da ontologia.

Como ontologias são operadas através da utilização de linguagens lógicas é importante compreender como é construído o pensamento lógico e suas funcionalidades. O ser humano utiliza a lógica através do seu raciocínio, seus conjuntos de pensamentos sobre um determinado assunto, domínio, causa ou problema. A lógica é o estudo do que se apresenta (do que vale, *counts*) como uma boa razão, para o quê e por que. Pensando desta maneira, se faz necessário entender esta afirmação, em que a lógica é composta de inferências (PRIEST, 2017).

A lógica se manifesta em um conjunto de inferências, que são um conjunto composto por premissas e conclusões. Premissas apresentam-se antes de um constructo, e possuem a função de dar razão, já as conclusões são as partes da inferência que após o constructo, possuem a função de compreensão que deve ser razão para algo. A compreensão dessas inferências, pode indicar se elas são verdadeiras ou falsas, de acordo com as premissas e conclusões (PRIEST, 2017).

Segundo Goldstein et al (2007) um argumento é válido quando a conclusão segue-se logicamente das premissas, onde um argumento legítimo é aquele que é não apenas válido, mas também possui premissas que sejam verdadeiras, desta maneira a conclusão de um argumento legítimo deve ser verdadeira. Um argumento válido com ao menos uma premissa falsa pode ter uma conclusão falsa, porém um argumento válido pode ter uma premissa falsa, entretanto, com uma conclusão verdadeira (GOLDSTEIN et al., 2007).

A lógica está diretamente ligada ao raciocínio humano, porém ela não é uma investigação empírica acerca dos processos de raciocínio das pessoas, nem dos produtos de tais processos, onde a lógica fornece embasamento às normas do raciocínio certo, ou seja, é uma disciplina normativa (GOLDSTEIN et al., 2007). O autor ainda afirma que o intuito do raciocínio lógico não é tentar

dizer como as coisas são, mas sim como as pessoas devem efetivamente raciocinar.

Utilizar a linguagem é uma das atividades próprias aos seres humanos enquanto agentes normais. Um outro ato básico realizado pelos humanos consiste em agrupar unidades de informação através das conjunções, ou seja, a compilação de unidades de informação. A conjunção é representada por um conector lógico em português: “e” e em inglês “*AND*” (GOLDSTEIN et al., 2007; PRIEST, 2017). Já a disjunção é um tipo de sentença onde é possível verificar uma ou mais suposições. A disjunção é representada pelo conector lógico em português: “ou” e em inglês pelo “*OR*” (GOLDSTEIN et al., 2007; PRIEST, 2017). Abaixo no quadro 6 são representados os conectores lógicos para conjunção e disjunção:

Quadro 6 – Constructos lógicos

Tipo	Conector Simbólico	Conector
Conjunção	&	<i>AND</i>
Disjunção	V	<i>OR</i>

Fonte: autor

As tabelas verdade são usadas para indicar a correspondência das condições de verdade de uma sentença lógica. Com elas é possível verificar se uma inferência é verdadeira (V) ou falsa (F) através das combinações de premissas e conclusões (PRIEST, 2017). Nas tabelas abaixo, 3 e 4, é possível visualizar as tabelas verdade para conjunção e disjunção.

Tabela 3 – Tabela Verdade Conjunção

a	b	a & b
V	V	V
V	F	F
F	V	F
F	F	F

Fonte: adaptado de Priest (2017)

Tabela 4 – Tabela Verdade Disjunção

a	b	a & b
V	V	V
V	F	V
F	V	V
F	F	F

Fonte: adaptado de Priest (2017)

Segundo Priest (2017), quantificadores são diferenciados de nomes próprios. Mesmo que quantificadores e nomes possam ser sujeitos gramaticais de sentenças, eles devem funcionar de maneiras bem diferentes. Existem três tipos diferentes de quantificadores conforme apontado na tabela 5.

Tabela 5 – Quantificadores

Quantificador	Conector
Quantificador Particular	<i>SOME</i>
Quantificador Universal	<i>EVERY</i>
Quantificador de Negação	<i>NO</i>

Fonte: adaptado de Priest (2017)

A linguagem OWL é uma sintaxe amigável para as descrições de OWL2, mas também pode ser usada para escrever ontologias inteiras do OWL2. A versão original da sintaxe OWL foi criada para o OWL e atualizada para OWL2. É utilizada no *Protégé 4* para inserir e exibir descrições associadas às classes (“OWL 2 Web Ontology Language Manchester Syntax (Second Edition)”, 2012). Ela reúne informações sobre nomes de maneira semelhante a *frames*, em oposição ao RDF / XML.

É, portanto, mais próximo da sintaxe abstrata para OWL do que as sintaxes para OWL2, no entanto, analisar a sintaxe *Manchester* OWL na especificação estrutural OWL2. Como a sintaxe de *Manchester* é baseada em *frames*, não pode lidar diretamente com todas as ontologias da OWL2. No entanto, há uma transformação simples que leva qualquer ontologia da OWL2

que não sobrecarrega entre propriedades de objeto, dados e anotação ou entre classes e tipos de dados em um formulário que pode ser gravado na sintaxe de *Manchester* (“OWL 2 Web Ontology Language Manchester Syntax (Second Edition)”, 2012). Abaixo a tabela 6 demonstra os conectores para constructos em Linguagem *Manchester*.

Tabela 6 – Constructos em linguagem *Manchester*

Tipo	Conectores
Coordenação (<i>Coordination</i>)	<i>AND; OR</i>
Quantificação (<i>Quantification</i>)	Quantificador Existencial: <i>SOME</i> Quantificador Universal: <i>EVERY; ONLY</i> Restrição de Número (Quantidade): <i>EXACTLY; MAX; MIN; VALUE</i>
Negação (<i>Negation</i>)	Quantificador Existencial de Negação: <i>NO</i> Quantificador Universal de Negação: <i>NOT</i> Negação VP: <i>DOES NOT</i> Sentença na Negação: <i>IT IS FALSE; AND THAT</i>
Subordinação (<i>Subordination</i>)	Frases relativas; Sentenças condicionais; Sentenças de Subordinação; Modalidade

Fonte: autor

3.2.3 Linguagens naturais controladas

As LNCs foram desenvolvidas com o objetivo de facilitar para o usuário, que possui pouco ou nenhum conhecimento em linguagens formais, a trabalhar com a representação do conhecimento. Diferentes tipos de LNC foram propostas para utilização na representação do conhecimento, especificamente para *web Semântica*, com o intuito de superar o problema das linguagens formais, que são geralmente difíceis de se compreender por pessoas que não são familiarizadas com notações formais (KUHN, 2012).

Segundo Kuhn (2012) as LNCs são definições artificiais de um subconjunto de linguagens naturais com vocabulários, sintaxes e semânticas

restritas, objetivando a criação uma linguagem que possa parecer natural, porém sem ambiguidades.

O principal benefício do uso das LNCs é que elas são processáveis por máquinas e podem ser traduzidas para outras representações sem problemas de ambiguidades (MÉSZÁROS; KISS, 2018). A complexidade de uma linguagem natural justifica o uso das linguagens formais para realizar a interação das sentenças ontológicas formais de uma maneira fácil e rápida para os usuários não familiarizados com os métodos e notações formais (KUHN, 2012). Segundo Fuchs, Kaljurand e Schneider (2006), a LNC pode ser considerada como um subconjunto de uma determinada linguagem natural, que é utilizada para documentação, especificação ou uma linguagem de representação do conhecimento.

É possível encontrar LNCs como: *Simplified Technical English*, *Caterpillar Technical English*, *IBM's Easy English*, *Common Logic Controlled Englis* e *Boeing's Computer Processable Language*, porém todas desenvolvidas para a área industrial e na língua inglesa (FUCHS; KALJURAND; KUHN, 2008)

A ACE é uma linguagem controlado da língua inglesa, precisamente definida, com um subconjunto tratável em inglês, com sintaxe restrita, que pode automaticamente e sem ambiguidades ser traduzido em LPO, compreensível tanto por humanos quanto por máquinas (FUCHS; KALJURAND; KUHN, 2008; IVANOV; VARFOLOMEYEV, 2017). Segundo Kuhn (2012) a linguagem ACE é mais genérica e mais expressiva que o OWL.

O processo de criação da ACE teve seu início no ano de 1995 através das pesquisas realizadas por Norbert E. Fuchs e Rolf Schwitter, ambos pesquisadores da University of Zurich (KUHN, 2010). Inicialmente a ACE teve seu desenvolvimento voltado para uma linguagem específica, porém na geração de modelos. Em 2004 ocorreu uma mudança de foco voltando as pesquisas para representação do conhecimento.

Considerada uma LNC, possui semelhança com a linguagem natural, porém é uma linguagem formal definida por uma gramática abstrata e uma gramática concreta, livre de ambiguidades por possuir um conjunto de regras de sintaxe e além de um conjunto de interpretações (FUCHS; KALJURAND; SCHNEIDER, 2006; KUHN, 2008a, 2008b, 2008c).

Suas regras gramaticais são intuitivas e devidamente conhecidas por pessoas que conhecem a língua. Através de experimentos foi possível demonstrar que a ACE pode ser aprendida em poucos dias (KALJURAND; FUCHS, 2006). A interface utilizada pela ACE assemelha-se a uma linguagem natural, contudo é definida como uma linguagem formal através de um conjunto de regras de interpretação. Desta maneira, a ACE pode ser compreendida tanto por máquinas quanto por qualquer pessoa que saiba a língua inglesa (FUCHS; KALJURAND; KUHN, 2008).

A sintaxe utilizada nos processos da ACE é definida de forma generalizada e informal por um número de regras utilizadas em palavras, frases com substantivos, frases com verbos, cláusulas relativas (frase com verbo arbitrário), coordenação (e/ou), quantificadores (todo/toda), negação, modalidade, frases condicionais, frases completas, pronomes anafóricos e variedades (FUCHS; KALJURAND; SCHNEIDER, 2006; KUHN, 2008a, 2008b, 2008c, 2010).

As regras para a semântica são definidas por um mapeamento utilizado em representação lógica, sendo que tais restrições eliminam a maioria das ambiguidades existentes. A utilização desta linguagem facilita o processo de AC por possuir diversas ferramentas que facilitam na tradução de linguagem natural para a linguagem formal e de formal para natural, além da *web* semântica.

Uma destas ferramentas é o *Attempto Parsing Engine* (APE), que incorpora a construção e regras de interpretação da ACE, traduzindo textos que não possuem ambiguidade em estruturas de representação de discursos (*Discourse representation structures* – DRS), uma variante da LPO que pode ser traduzida facilmente em qualquer linguagem formal equivalente a um subconjunto da LPO (KALJURAND; FUCHS, 2006; FUCHS; KALJURAND; KUHN, 2008). Segundo Ivanovs e Varfolomeyev (2017), os textos utilizando ACE podem ser traduzidos em DRS e correlacionados com fórmulas da LPO, portanto a DRS pode ser traduzida em qualquer linguagem formal que é equivalente a um subconjunto da LPO.

A ACE utiliza-se do APE, um *software open source*, que é utilizado no processo de implementação da ACE. É através da utilização do APE que é possível acessar a ACE para que os textos possam ser traduzidos da LNC em linguagem lógica, atribuindo semântica formal (FUCHS; KALJURAND;

SCHNEIDER, 2006; KUHN, 2008a, 2008b, 2008c, 2010). Para que a ACE obtenha sucesso, a OWL ACE deve ser de fácil utilização para os usuários, ou seja, os usuários devem aprender rapidamente a resolver os erros sintáticos e semânticos durante a entrada dos textos na OWL ACE (KALJURAND; FUCHS, 2006).

A ACE oferece constructos de linguagem como substantivos contáveis e de massa: plurais, coletivos e distributivos; quantificadores generalizados; pronomes indefinidos; negações, conjunções e disjunções de frases substantivas; frases verbais, sentenças e referências anafóricas a frases substantivas por meio de nomes próprios; frases de nomes nítidos; pronomes e variáveis (KALJURAND; FUCHS, 2006). A gramática utilizada na ACE define e limita a forma e o significado tanto das sentenças como nos textos, sendo que é expressada como uma pequena construção de regras e o significado dos textos em ACE é definido por um pequeno conjunto de regras de interpretação (IVANOV; VARFOLOMEYEV, 2017).

Segundo Fuchs; Kaljurand e Kuhn (2008) aplicações para *web* semântica, tais como o desenvolvimento de traduções da ACE para OWL/SWRL ou de OWL/SWRL para ACE são possíveis de realizar com a utilização da ACE. Também possui um *plug-in* para o editor de ontologias *Protégé-Frames* e *Protégé 5.5* (FUCHS; KALJURAND; KUHN, 2008). O *Protégé* é um dos editores de ontologia mais populares com o suporte de muitas tarefas de engenharia ontológica, possibilitando ao usuário criar ontologias por definição de conceitos, especificações, relacionamento, anotações e outras informações (WANG et al., 2006).

Para um conhecimento baseado no *Extensible Markup Language* (XML), um programa pode ser traduzido tanto diretamente da XML em OWL ou gerar afirmações em ACE de textos XML (IVANOV; VARFOLOMEYEV, 2017). Um texto em ACE, quando gerado a partir de um XML, forma uma base de conhecimento específica que representa informações extensivas sobre um documento (IVANOV; VARFOLOMEYEV, 2017).

Durante o processo de transformação de OWL para ACE podem ocorrer algumas restrições referentes a palavras, frases e tipos de sentenças, que são facilmente explicáveis por não existir suporte para verbos no intransitivo

e ditransitivo, frases preposicionais, advérbios, adjetivos intransitivos e a maioria das formas do plural (KALJURAND; FUCHS, 2006).

Ainda segundo os autores, outras restrições podem ocorrer nas estruturas sintáticas das sentenças, sendo necessário realizar implicações, como cada sentença que faz referência anafórica apenas para nomes próprios e outros objetos de nível superior.

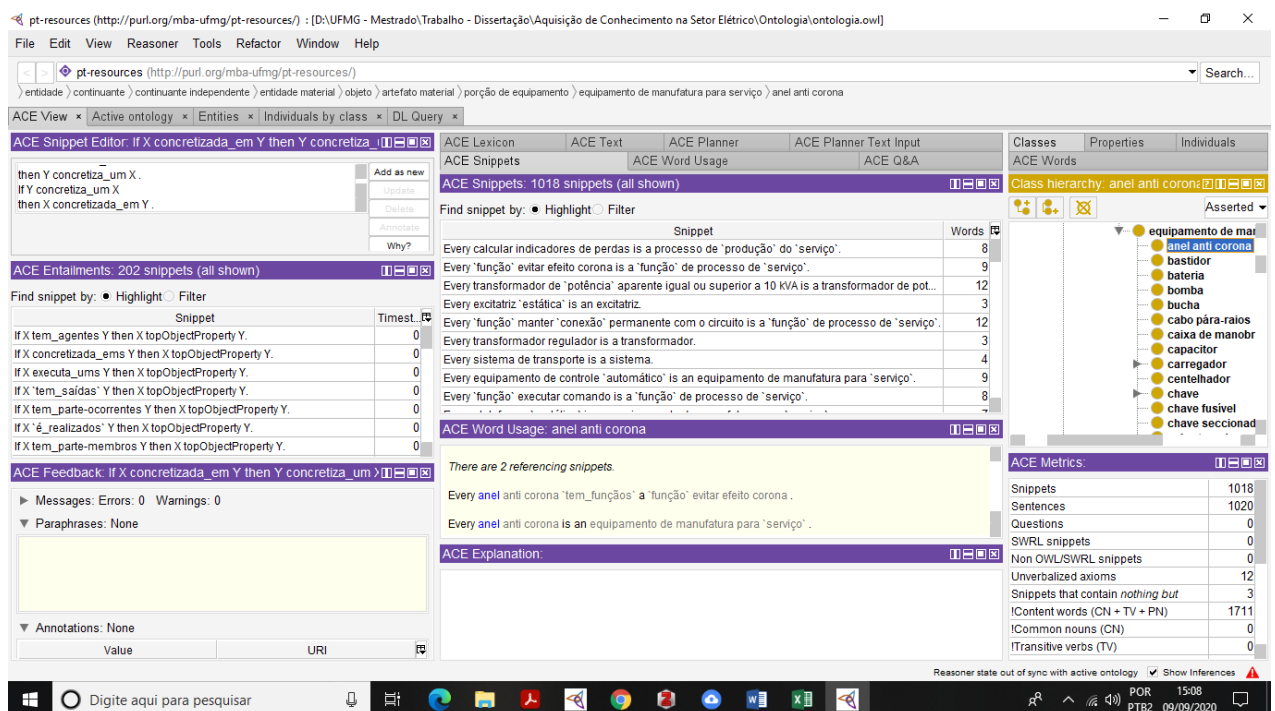
A ACE ainda possui uma *ACEwiki* que utiliza linguagens naturais controladas para criação de *wikipedia* (*wiki*) semânticas, permitindo a usuários comuns, que não são familiarizados com as regras da linguagem lógica e ontologias, a entenderem, modificar e complementar o conteúdo formal da *wiki*. A facilidade de uso faz com que usuários comuns também possam criar ontologias de forma colaborativa e intuitiva (KUHN, 2008a, 2008b, 2008c, 2010).

O foco principal da *ACEwiki* é aprimorar a agregação da representação e do conhecimento, tornando mais fácil sua utilização e entendimento do que outras *wiki* semânticas, que utilizam a *OWL reasoner Pallet* como motor de inferência para transformação das sentenças da *wiki* em OWL possuindo suporte de alto nível de expressividade.

A *ACEwiki* fornece um editor de predicativos que mostra passo-a-passo as palavras que são sintaticamente possíveis de serem utilizadas em uma sentença, a interface da *ACEwiki* não utiliza termos técnicos como “elementos ontológicos”, “propriedades” ou “subclasses”, porém, utiliza termos como “palavras”, “verbos transitivos” ou “hierarquias”, que são familiares às pessoas com pouco ou nenhum conhecimento em lógica (KUHN, 2008a). Todas sentenças adicionadas na *ACEwiki* são automaticamente verificadas após a criação, para que seja possível ser adicionada de forma correta na ontologia.

Uma outra ferramenta importante para é o *plug-in ACEView* para o *software* editor de ontologias *Protégé*. Utilizados para as versões 5.5 do *Protégé* (figura 1), o *ACEView* é um *plug in* que utiliza-se da linguagem controlada ACE para visualizar e editar ontologias em linguagens OWL2 e SWRL, sendo possível abrir as ontologias nestas linguagens e visualizá-las em textos feitos através do *ACEview* (KALJURAND, 2008).

Figura 1 – *Protégé* 5.5 e *Plug in ACEView*



Fonte: autor

No capítulo seguinte será apresentada a metodologia e os passos realizados durante o processo de pesquisa.

4 Metodologia de Pesquisa

A metodologia é o um conjunto de planos que o pesquisador utiliza durante o processo de desenvolvimento da pesquisa, abrangendo tanto o momento mais amplo da pesquisa até o mais detalhado, coletando, analisando e interpretando dados através de métodos minuciosos que se transformam em um projeto de pesquisa. Essa planificação envolve uma série de decisões que são tomadas, tendo como objetivo que a pesquisa seja compreensiva para qualquer leitor (CRESWELL, 2013).

O autor também aponta que “a seleção da metodologia é também baseada na natureza do problema de pesquisa ou nas adversidades encontradas, experiência profissional e o público alvo” (CRESWELL, 2013, p.13).

O desenvolvimento metodológico deste trabalho, de cunho pragmático, de acordo com (GIL, 2008), pode ser classificado como:

- Baseada na natureza da pesquisa aplicada com o objetivo de gerar conhecimentos para aplicação prática no processo de AC;
- Quanto à abordagem do problema, pesquisa qualitativa, visto que o conhecimento apreendido, no âmbito da atividade da AC para representação em ontologias, não pode ser mensurado quantitativamente;
- Do ponto de vista de seus objetivos, exploratória, pois busca maior familiaridade com o problema, com o objetivo maior de explicá-lo;
- Quanto aos procedimentos técnicos, é um estudo de caso, ao utilizarmos a ACE para transformação da ontologia em LNC, os termos serão validados novamente por especialista.

As demais seções estão organizadas da seguinte forma: a seção 4.1 apresenta um contexto geral da pesquisa e a seção 4.2 apresenta os procedimentos metodológicos do trabalho.

4.1 Contexto da pesquisa

Com o processo de privatização do setor elétrico a partir de 1996 criou-se o Operador Nacional do Sistema Elétrico (ONS), com a finalidade de

realizar operações interligada dos sistemas elétricos nacionais. Além disso várias outras mudanças marcaram o setor buscando maior controle do Estado.

Assim a partir de 2004 foi implantado um novo modelo de investimento onde o Estado retomou a responsabilidade pelo planejamento e desenvolvimento do setor. Buscando uma maior transparência para o segmento elétrico, o governo propôs e aprovou a criação da Agência Nacional de Energia Elétrica (ANEEL).

Entre as principais atribuições da ANEEL se destacam: i) a iniciativa de criação e monitoramento dos requisitos técnicos que deverão ser observados com o objetivo de garantir a qualidade dos serviços oferecidos aos usuários; ii) solicitar, quando achar necessário, a abertura de processos licitatórios para as concessões de geração, transmissão e distribuição de energia elétrica; iii) garantir que todas as concessionárias tenham as mesmas condições de participação no mercado, ou seja, não ocorra o privilégio de uma em detrimento de outra; iv) estabelecer os critérios adotados para a gestão de custo de transmissão; v) realização de maneira periódica da avaliação e determinação das tarifas que deverão ser praticadas no varejo (FERREIRA, 2000).

Várias iniciativas em direção à padronização técnica e administrativa do setor têm sido implementadas pela ANEEL, como o Manual de Contabilidade do Setor Elétrico (MCSE), Manual de Controle Patrimonial do Setor Elétrico (MCPSE) e os Procedimentos de Distribuição (PRODIST). O conjunto de padronização tem o objetivo de melhorar a fiscalização dos serviços prestados pelas concessionárias e garantir a auditabilidade dos processos regulados pelo órgão, como os contábeis. Entretanto, tais iniciativas, além de apresentarem divergências terminológicas entre si, também apresentam características impeditivas ao estabelecimento da interoperabilidade semântica entre os sistemas de uma empresa de energia.

4.1.1 A empresa do setor de energia

Com sua sede localizada em Belo Horizonte, a CEMIG – Companhia Energética de Minas Gerais, é uma das principais concessionárias de energia elétrica do Brasil e a maior integradora do setor, aglutinando as concessões de

geração, transmissão e distribuição de energia, com unidades espalhadas por todo o estado de Minas Gerais.

A empresa foi fundada em 1952, durante o governo do presidente Juscelino Kubitscheck, com o nome de “Centrais Elétricas de Minas Gerais”. Em 1984, teve seu nome alterado para Campainha Energética de Minas Gerais, quando a organização decidiu atuar de forma mais abrangente no mercado de fonte de energia.

Atualmente a área de concessão da CEMIG cobre 96% do Estado de Minas Gerais, atendendo a cerca de 774 municípios, com o total de 20 milhões de habitantes atendidos. Possui a maior rede distribuição da América Latina com mais de quinhentos e trinta seis mil quilômetros de extensão (CEMIG, 2020). O principal negócio da CEMIG é a comercialização de energia elétrica, que é distribuída em formas contratuais que levam em conta a disponibilidade sazonal (horária, semanal, mensal e anual) e quantitativa (por demanda), condições de interruptibilidade, pontos de medição modalidades de geração, tensão de fornecimento e uso, além da prestação de diversos serviços de consultoria, planejamento e execução (ALMEIDA, 2006).

A atuação da organização está expressa em sua visão, sua missão e seus valores. A visão da CEMIG estabelece estar entre os três melhores grupos integrados de energia elétrica do Brasil em governança, saúde financeira, desempenho de ativos e satisfação de clientes. A empresa tem como missão prover soluções integradas de energia limpa e acessível à sociedade, de maneira inovadora, sustentável e competitiva. Os valores da organização são: respeito à vida, integridade, geração de valor, sustentabilidade e responsabilidade social, comprometimento e inovação (CEMIG, 2020). As diretrizes estratégicas, a visão, a missão e os valores da CEMIG são orientados para conduzir e operar com sucesso a organização da Empresa sendo concebido de forma a melhorar continuamente seu desempenho. A implantação, a manutenção e o aprimoramento do sistema de gestão têm como premissas as diretrizes estratégicas expressas na visão e missão e no Planejamento Estratégico da Companhia (CEMIG, 2020).

A presente pesquisa se desenvolve como parte de um projeto de Pesquisa e Desenvolvimento (P&D) na CEMIG, financiado pela ANEEL. O proceder metodológico do Projeto é baseado em preceitos estudados pela CI.

Teixeira e Ávila (2019) explicam que, para melhor contextualização do problema para o qual se propõe uma solução, é necessário compreender universo de gestão de ativos pela Norma ISO 5500, que normatiza a disciplina. Um ativo consiste em um item (ou conjunto de itens) que possui valor real (ou potencial) em uma organização.

Este valor pode ser tangível ou intangível, financeiro ou não financeiro, positivo ou negativo (em diferentes estágios da vida do ativo) e ainda inclui as considerações de riscos e passivos. Os ativos físicos (que no setor elétrico correspondem a equipamentos e instalações) são o oposto de ativos intangíveis, que são ativos não físicos (como contratos, marcas, ativos digitais, direitos de uso, licenças, direitos de propriedade intelectual, reputação ou acordos).

De acordo com a Norma, a Gestão de Ativos compreende um conjunto de procedimentos coordenados para estabelecer maior controle e maior valor sobre um ativo (como já dito, no setor elétrico corresponde a um equipamento ou instalação) e obter melhores resultados do seu uso, equilibrando seus custos, riscos e desempenho. Já um Sistema de Gestão para a Gestão de Ativos tem a função de estabelecer a política e os objetivos da Gestão de Ativos. Por consequência, a Gestão de Ativos deve ser observada sob os aspectos estratégico e operacional.

O nível estratégico está vinculado à administração da organização e o nível operacional remete ao gerenciamento do ativo, ou seja, às etapas do seu ciclo de vida: desde a concepção do projeto, sua especificação técnica, seus processos de aquisição, implantação, operação, manutenção e monitoramento até o seu descomissionamento e descarte. E é exatamente neste nível operacional onde se localizam as entradas de informações relacionadas à vida dos ativos em diferentes aplicações de uma infraestrutura corporativa de Tecnologia da Informação (TI). Entretanto, tanto na CEMIG quanto em diferentes organizações não só do setor elétrico, a arquitetura de TI não favorece a gestão e tampouco o gerenciamento, uma vez que as informações se encontram dispersas,

em silos, fazendo uso de diferentes padrões terminológicos e sem integração semântica entre as diferentes bases.

Assim, com o propósito de otimizar as iniciativas vigentes na empresa em função da implantação de um Sistema de Gestão para a Gestão de Ativos, assim como melhorar a obtenção de suas receitas obtidas tanto por meio da manutenção dos seus ativos quanto pela remuneração de seus investimentos a eles relacionados, foi proposto o desenvolvimento de um artefato ontológico.

4.1.2 Construção da ontologia

Este estudo utilizou uma LNC como base para a validação de partes de uma ontologia por um especialista de domínio. A ontologia é uma ontologia experimental, ondem sua construção será descrita nas próximas seções.

4.1.2.1 Aquisição do conhecimento pelo profissional da informação

A etapa de AC consiste, basicamente no desenvolvimento das seguintes etapas:

- i) Identificar e analisar material de referência – mapear fontes de informação e conhecimento de referência que permitam um melhor entendimento de um domínio, assim como trabalhos correlatos;
- ii) Identificar os especialistas de domínio – identificar os especialistas do domínio que participarão dos processos de extração do conhecimento e validação do conhecimento registrado pelos profissionais da CI;
- iii) Aplicar técnicas colaborativas – selecionar, personalizar e aplicar o método de extração do conhecimento adequado e personalizado para o contexto;
- iv) Registrar o conhecimento extraído em formato ontológico.

4.1.2.2 Resultados obtidos

A primeira etapa, referente ao levantamento e análise de material de referência, iniciou-se com entrevistas não estruturadas com colaboradores e

gestores de áreas (departamentos) relacionadas aos assuntos “gerenciamento e gestão de ativos” na CEMIG. Esses levantamentos preliminares propiciaram uma visão macro do contexto, maior elucidação do problema e o direcionamento para o levantamento do material de referência.

A partir desse momento foram mapeadas terminologias, ontologias, modelo de informação, textos acadêmicos, documentos normativos e regulatórios relacionados à gestão e operação. Os materiais foram estudados e selecionados como fontes de informação e conhecimento e como corpus candidatos ao processo de extração terminológica.

Teixeira e Ávila (2019), apresentam todo o processo realizado, que culminou também na confirmação e o escopo do artefato ontológico: o MCPSE. Uma vez que o mesmo fora definido como escopo do Projeto e considerando a amplitude terminológica do mesmo e o prazo de desenvolvimento do Projeto, delimitou-se um recorte da Distribuição que contemple as entidades mais recorrentes em remuneração para a CEMIG. Cada entidade é denominada Tipo de Unidade de Cadastro (TUC) sendo que no MCPSE existem aproximadamente 105 TUCs. No recorte foram delimitadas 20 TUCs que resultaram em aproximadamente 400 termos para representação na ontologia. Importante destacar que, no momento de extração dos termos do MCPSE, algumas informações ou definições relativas aos mesmos foram identificadas previamente nos próprios materiais de referência.

A segunda etapa, identificação de especialistas de domínio, consistiu em grande desafio. Em sua concepção, o projeto não contemplava o desenvolvimento de um artefato ontológico e os recursos, tanto humanos quanto financeiros, não eram adequados para tal. Nesse sentido, três especialistas de domínio (um da CEMIG e dois da executora) foram chamados a contribuir. Entendeu-se que, em função da disponibilidade dos mesmos e da localização geográfica distribuída, a objetividade para a extração do conhecimento era um requisito essencial.

A terceira etapa, aplicação de técnicas colaborativas, foi planejada para ser executada por meio da plataforma *wiki*. Entretanto, em função da locação de mão de obra, o recurso não foi disponibilizado a tempo. Assim, optou-se pela realização de entrevistas aos três diferentes especialistas. Três

pesquisadores em CI realizaram 13 entrevistas do tipo estruturadas. Cada seção se desenvolveu em aproximadamente 2:30h de duração e foram realizadas tanto presencialmente quanto por videoconferência. Foram apresentados aos especialistas os termos das TUCs, de forma que pudessem se dedicar de acordo com a suas experiências e vivências. Para direcionar o momento e registrar o conhecimento extraído, gerou-se um formulário. Cada um dos especialistas definiu os termos contidos em cada TUC, atentando-se para que cada definição fosse realizada de acordo com o contexto da mesma e não pela definição global do termo.

Desta maneira, foi possível avaliar visões diferentes, complementares e iguais de cada um deles. Tal abordagem foi importante para que ocorresse o mínimo possível de ambiguidade e para que os termos fossem definidos com maior exatidão na etapa após a AC. A figura 2 apresenta fragmento do formulário por meio do qual estabeleceram-se os registros, sendo considerando que KA1 corresponde às informações identificadas previamente (na primeira etapa) e KA2 às informações obtidas nas entrevistas.

Figura 2 – Fragmento de tabela para registro das entrevistas

TUC 160: CHAVE		
EQUIPAMENTOS (OBJETOS CONSTITUINTES)	KA1	KA2
câmara de extinção a vácuo	Câmara de extinção à vácuo é um equipamento industrial que suporta a função de interrupção da corrente elétrica (em um disjuntor ou chave), onde o processo de proteção de equipamentos e circuitos elétricos realiza a função de de interrupção da corrente elétrica e é um processo industrial. Source: 1-Glossary of Terms For the Electric Utility Industry, North Carolina State University 2- https://slideplayer.com.br/slide/1734084/	[cemig] componente de um dispositivo de manobra onde ocorre a separação dos contatos com a consequente geração de arco elétrico e onde esse arco é extinto
chave	Chave é um equipamento industrial que suporta a função de ativar, desativar ou alternar o fornecimento de energia elétrica, onde o processo de proteção de equipamentos e circuitos elétricos realiza a função de ativar, desativar ou alternar o fornecimento de energia elétrica e é um processo industrial. Source: 1-Glossary of Terms For the Electric Utility Industry, North Carolina State University 2- http://www.ngeletrica.com.br/chaves	[cemig] Chave composta de um contato fixo e outro móvel utilizada para promover seccionamento visível no circuito elétrico. [cpqd] dispositivo de manobra ou seccionamento do circuito elétrico

Fonte: Elaborado pelo grupo de pesquisa

Observou-se no decorrer do método que o setor elétrico possui diversos glossários e vocabulários sobre o setor elétrico. Os sites da ANEEL, da Eletrobrás e das distribuidoras de energia possuem uma área onde é possível encontrar informações relevantes aos termos, organizados de forma alfabética e, em alguns casos, com motores de busca. Entretanto, muitos desses recursos estão voltados para os consumidores e para os investidores. Assim, apresentam descrições e abreviaturas voltadas para o trabalho realizado pelo órgão regulador e pelas próprias distribuidoras, e não termos voltados para instalações, equipamentos e sistemas, que é o foco do projeto.

Além disso, a extração de conhecimento por meio das entrevistas, além de toda a riqueza proporcionada, evidenciou muitas variações (sinônimas) e inconsistências terminológicas do setor. Alguns termos apresentaram divergências significativas de entendimento e outros foram questionados pelos entrevistados sob diferentes perspectivas, tais como:

i) em relação à obsolescência do equipamento, ou seja, equipamentos em desuso nas instalações das distribuidoras: impressora (contido na TUC 305);

ii) em relação ao desconhecimento sobre qual equipamento se refere na realidade: transformador de linha (contido na TUC 565);

iii) equipamento não utilizado na distribuição de energia, mas apresentado no recorte como tal: banco de capacitores inseridos em série ou paralelo (contido na TUC 130);

iv) desconhecimento da função no contexto: *smartphone* (contido na TUC 485); e

v) inconsistência significativa entre os registros dos especialistas: sistema de telealarme (contido na TUC 485) dentre outros exemplos.

O processo de AC e suas respectivas etapas, planejado e organizado sistematicamente, evidenciou a carência de conhecimento registrado no setor. Apesar do mapeamento de fontes de informação e conhecimento e de pontuais consultas ao material de referência, todo o conteúdo de todos os termos cobertos pelo escopo foi extraído por meio das entrevistas aos especialistas. Esses momentos foram extremamente ricos de conteúdo e sem que os especialistas consultassem outros materiais. Muitas vezes as reflexões culminaram na

identificação de diferenças entre o descrito no MCPSE e o encontrado na realidade e diferenças terminológicas regionais.

4.2.2.3 Pré-conceitualização

Uma análise preliminar do resultado obtido indicou a não uniformidade terminológica em relação aos processos em que as funções identificadas atuavam. Esta lacuna impactou diretamente na construção das definições dos termos pelos especialistas da CI. Nesse sentido, procedeu-se à busca pelo site da ANEEL, da CEMIG e de empresas correlatas. Na indisponibilidade de respostas, foi necessário entrevistar especialista para se obter um rol de processos padronizados. Tal busca não alcançou os resultados pretendidos, uma vez que essa padronização de processos de negócios não existe para a geração, para a transmissão e para a distribuição.

Nesse sentido, partiu-se em busca de insumos que pudessem subsidiar esse levantamento. Justificado por seus conteúdos e objetivos, optou-se pelo uso dos 11 Procedimentos de Distribuição (PRODIST²) da ANEEL e das 26 Normas de Distribuição (ND³) da CEMIG, uma vez que ambos padronizam as atividades técnicas da distribuição. Após estudo e leitura dos documentos, foi gerada uma tabela que subsidiou o desenvolvimento das definições. Foram identificados 10 macroprocessos e 51 processos relacionados.

Sete pesquisadores foram treinados para gerar as definições em padrão ontológico. Cada pesquisador recebeu aproximadamente 50 termos para serem definidos, além dos insumos para desenvolverem seus trabalhos, que são o conhecimento registrado dos especialistas de domínio e a tabela de processos desenvolvida. Todas as quase 400 definições geradas foram revisadas, sendo adequadas ou modificadas, segundo o padrão ontológico estabelecido. Das definições foram identificadas 296 funções que, após extração daquelas sinônimas foram também padronizadas. Das 221 funções identificadas dos equipamentos, extraíram-se 93 verbos. Nesse sentido, obteve-se nesta etapa tanto o glossário de termos definidos quanto o glossário de verbos.

² Disponível em: <<https://www.aneel.gov.br/prodist>>

³ Disponível em: <http://www.cemig.com.br/pt-br/atendimento/Clientes/Paginas/norma_tecnica.aspx>

4.2 Passos Metodológicos

O objetivo deste trabalho consiste em criar um método para validar a AC confrontando resultados desta etapa com os axiomas da fase de formalização, a partir do uso da LNC. A apresentação dos passos que antecedem ao que se propõe pesquisar se justifica pela visão metodológica e sequencial do processo de construção de uma ontologia.

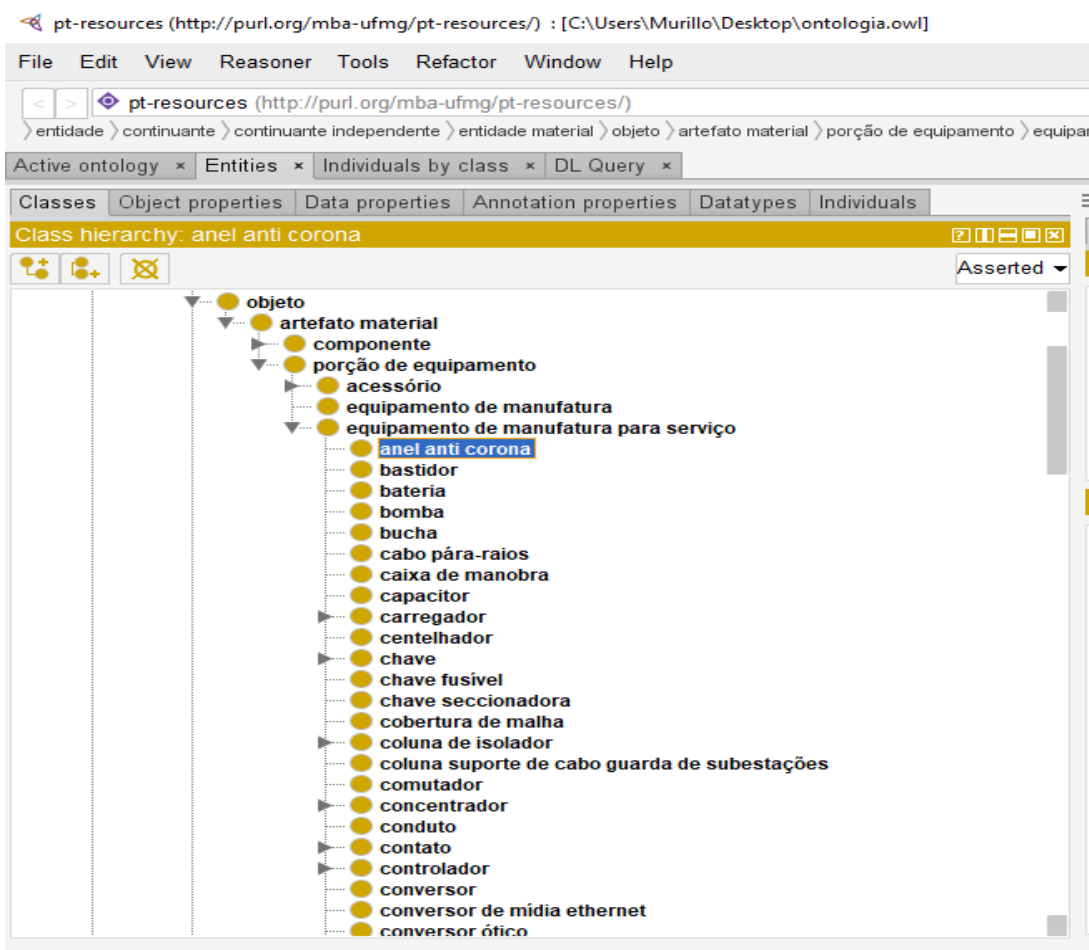
4.2.1 Passo 1: *selecionar de um recorte da ontologia*

Foi realizado um recorte que representa uma parte da ontologia, onde foram selecionados determinados termos que seriam validados com um especialista de domínio. Tal escolha se deu pela importância dos termos estudados dentro do MCPSE, desenvolvido pela ANEEL.

O manual é dividido em TUC onde estão contidos todos os equipamentos (termos) de acordo com o que é padronizado para a construção de estruturas de geração, transmissão e distribuição de energia elétrica. Foram selecionadas as TUC's mais importantes ou que tiveram algum tipo de inconsistência durante o período de AC.

4.2.2 Passo 2: *identificar axiomas (Manchester) na ontologia*

Para se encontrar os axiomas estudados na ontologia foi necessário utilizar o programa *Protégé* versão 5.5. O caminho para chegar nos axiomas seguiu a coluna onde se encontram as partes da ontologia e foi seguido o seguinte caminho, baseado na ontologia de alto nível BFO; Equipamento de Manufatura para Serviço, selecionando o termo (equipamento) conforme figura 3 e quadro 7.

Figura 3 – Imagem tela *Protégé Frames 5.5*

Fonte: o autor

Quadro 7 – Termos e Axiomas – Equipamentos de Manufatura para Serviços

Termo	Axiomas de Equipamento de Manufatura para Serviço
Anel Anti Corona	Tem_função SOME 'função de evitar efeito corona'
Bastidor	Tem_função SOME 'função abrigar dispositivo'
Bateria	Tem_função SOME 'função alimentar sistemas críticos'
Bomba	Tem_função SOME 'função drenar líquidos'
Bucha	Tem_função SOME 'função realizar transição de condutor em potencial elevado'
Capacitor	Tem_função SOME 'tem função ajustar fator de potência'

Fonte: o autor

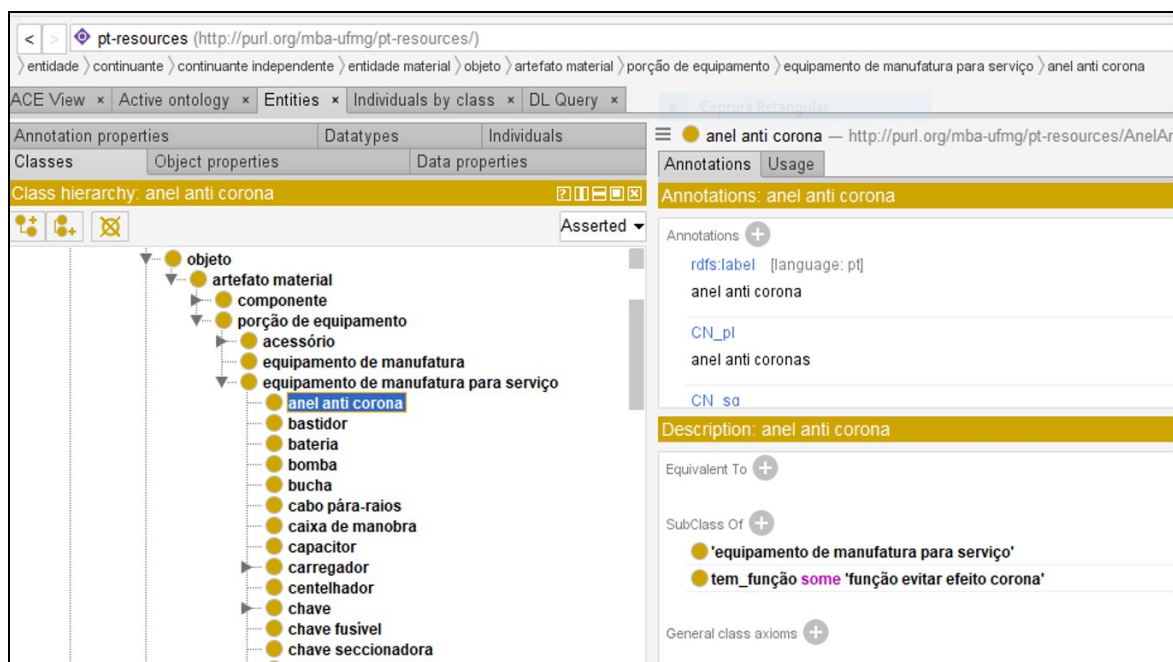
4.2.3 Passo 3: traduzir da lógica para Linguagem Natural Controlada

Foi realizada a tradução dos axiomas referentes em linguagem *Manchester* para uma LNC utilizando a ACE. A tradução foi realizada de forma automática através do *plug in ACEView* instalado no *software Protégé*. A tradução dos axiomas estão parte em português e parte em inglês, devido ao fato da ACE ter como base a língua inglesa. Para que a tradução tenha melhor compreensão pelo especialista de domínio, foi realizado um ajuste, traduzindo todas as partes que permaneceram em inglês para o português.

4.2.4 Passo 4: estruturar roteiro de perguntas

A validação da tradução dos termos com o especialista ocorreu via videoconferência na CEMIG, quando foi realizada uma entrevista com a apresentação dos termos, figuras 4 e 5 e quadro 8, com a tradução para a LNC, a validação das atribuições lógicas dos axiomas e demais “armadilhas” que podem ser enfrentadas durante o processo de construção de uma ontologia.

Figura 4 – Fragmento da ontologia para validação com especialista



Fonte: o autor

Figura 5 – Questionário validação especialista de domínio

• "anel anti-corona" tem_função "evitar efeito corona"

• 0- isso está certo?

• SIM

• NAO

• 1- algum outro equipamento serve para evitar o efeito o corona?

• SIM

• NAO

• 2- Tudo que "evitar efeito corona" é um "anel anticorona"?

• SIM

• NAO

The screenshot shows the same ontology editor interface as in Figure 4. The class hierarchy on the left is identical. The right pane shows the annotations for 'anel anti corona', including 'rdfs:label [language: pt] anel anti corona', 'CN_pl anel anti coronas', and 'CN_sq'. The 'Description: anel anti corona' section shows 'Equivalent To' and 'SubClass Of' relationships. The 'SubClass Of' section lists 'equipamento de manufatura para serviço' and 'tem_função some "função evitar efeito corona"'. The 'General class axioms' section is also visible.

Fonte: o autor

Quadro 8 – Roteiro de questionário estruturado e possíveis resultados

	Perguntas
1	Isto está certo? SIM- vá para 1 NAO - tem um erro de conteúdo na ontologia
2	Algum outro equipamento serve para.... SIM - indicativo de some, vá para 2 NAO - indicativo de <i>ONLY</i> , acabou
3	Tudo que é um equipamento X é um... SIM - ele está equivocado NAO - é de fato <i>SOME</i>

Fonte: o autor

5 Resultados

O capítulo contém os resultados da pesquisa. A seção 5.1 traz os resultados da validação da ontologia apresentada e a seção 5.2 expõe a validação da LNC utilizando o *plug in ACEView* para o *software Protégé 5.5*.

5.1 Validação da Ontologia

A proliferação de ontologias tem resultado na qualidade dos dados e das informações, porém não elimina totalmente erros provenientes das dificuldades e complexidades inerentes ao processo, os quais abrangem diferentes variáveis como: métodos de AC variados, especialistas de domínio com diferentes perspectivas e responsáveis pelo desenvolvimento de ontologias que não conhecem e compreendem bem o domínio. O domínio da energia elétrica é complexo e apresenta peculiaridades.

Segundo Poveda-Villalón, Suárez-Figueroa e Gómez-Perez (2010, p. 1), a “qualidade não é totalmente garantida porque os desenvolvedores devem enfrentar uma ampla gama de dificuldades e desvantagens ao modelar ontologias.” Ainda segundo os autores tais dificuldades podem implicar no aparecimento de anomalias ou a não utilização de boas práticas no processo de construção de ontologias, sendo importante avaliar as metodologias de construção antes de usá-las ou reutilizá-las em outras ontologias ou aplicações semânticas.

Todas as quase 400 definições geradas foram revisadas, sendo adequadas ou modificadas, segundo o padrão ontológico estabelecido. A validação analisou 50 termos – em duas etapas de 22 e 28 termos – contidos na classe “Equipamentos de Manufatura para Serviço”, onde se encontrava o maior número de representações de ativos. Dos 22 termos analisados na primeira etapa foram encontrados 11 erros de atribuição e 11 atribuições corretas apresentadas na tabela 7.

Tabela 7 – Validação ontologia – entrevista 1

Entrevista 1 (09/10/20)	
Atribuição corretas (SOME correto)	11
Erros de atribuição (SOME incorreto)	7
Erros de conteúdo	4
Total de termos	22

Fonte: o autor

Observou-se que esses erros ocorreram por necessidade de especialistas durante a modelagem. Confrontado com a situação, o responsável pela modelagem informou que, na dúvida ou por simples desconhecimento do domínio, diversos operadores existenciais foram atribuídos sem confirmação. Operadores existenciais são mais comuns e sabidamente menos problemáticos em ontologias, de forma que o procedimento que não se configura com uma má prática em versões iniciais (“OWL 2 Web Ontology Language Manchester Syntax (Second Edition)”, 2012).

Também foram encontrados 04 erros de conteúdo (Figuras 6 e 7), onde as definições atribuídas estavam de fato erradas. Tal tipo de erro é proveniente das entrevistas com especialistas.

Figura 6 – Exemplo de erro de conteúdo na AC

Termo 14	Relé Principal				
tem_função some 'função executar principal função'					
Perguntas:					
0	Sim		Não	X	Observações: Erro. Relé primário ou Secundário. Titular ou reserva. Principal e Backup
1	Sim		Não		
2	Sim		Não		
Resultado	Erro na ontologia				

Fonte: o autor

Figura 7 – Exemplo de erro de atribuição de operador lógico

Termo 11	Seguidor Solar									
tem_função some 'função movimentar placas fotovoltaicas'										
Perguntas:										
0	Sim	X	Não		Observações:					
1	Sim		Não	X						
2	Sim		Não							
Resultado	ONLY									

Fonte: o autor

Já na segunda etapa, dos 28 termos analisados, foram encontrados 06 erros de atribuição e 12 atribuições corretas, apresentadas na tabela 8. Foram encontrados ainda 10 erros de conteúdo. O resultado revela cerca de 25% de erros de atribuição do operador lógico e 30% de erros de conteúdo.

Tabela 8 – Validação ontologia - Entrevista 2

Entrevista 2 (03/11/20)	
Atribuição corretas (SOME incorreto)	12
Erros de atribuição (SOME incorreto)	6
Erros de conteúdo	10
Total de termos	28

Fonte: o autor

Como foi apresentado na subseção 2.3.5, é possível verificar que, uma ampla lista de armadilhas identificadas na literatura apresentada por Gómez-Pérez; Fernández-López, Chorcho (2004) e por Poveda-Villaon; Suárez-Figueróa; Gómez-Pérez (2010), onde a ontologia desenvolvida neste projeto de P&D continha ao menos 03 erros derivados.

O primeiro caso – “anotação perdida” – correspondia ao fato de que a informação registrada ainda na fase de AC em comentários, junto as classes candidatas, foi perdida em parte ou em sua totalidade.

O segundo caso – “ausência de equivalências” – correspondia ao fato de que nem todas as equivalências foram listadas devido ao grande número de possibilidades em um domínio complexo como o abordado.

O terceiro caso – “ausência de instanciação” – derivou do fato de que a instanciação dependia de dados reais da empresa, os quais não foram disponibilizados ao longo do projeto e da pesquisa. Esse fato ilustra que os erros mencionados nem sempre são causados por modelagem, mas também por contingências comuns nas organizações.

5.2 Validação da Linguagem Natural Controlada (ACEView)

Durante o processo de validação da ontologia com o especialista de domínio também foi avaliado a tradução da linguagem formal (*Machester* OWL) para a LNC através do *plug in* ACEView instalado no editor de ontologia *Protégé* 5.5.

Foi possível verificar que a tradução feita pelo ACEView tem maior funcionalidade quando a ontologia é feita em língua inglesa. Entretanto, após a realização de ajustes de tradução, feitos para que a leitura em língua portuguesa se tornasse mais amigável para o especialista, foi possível verificar que é viável utilizar a LNC como uma forma mais fácil de compreensão da linguagem utilizada pelas ontologias, auxiliando a compreensão pelo especialista de domínio e ajudando na validação do artefato ontológico.

Dos 50 termos estudados nas duas rodadas de entrevistas, todos foram de fácil compreensão para o especialista, contribuindo para a descoberta de erros de conteúdo, assim que a sentença foi lida e compreendida pelo especialista, principalmente dos termos que possuíam axiomas mais complexos.

Segundo Kaljurand (2008) axiomas de estrutura muito simples, tem a sua verbalização em linguagem natural, não trazendo significativa melhoria em relação a uma sintaxe OWL tradicional, entretanto, as apresentações de todas as

implicações como uma única lista de frases, apresentadas pelo *plug in ACEView*, em LNC podem fornecer uma boa visão geral e de fácil leitura.

A LNC transforma a leitura das definições dos termos, inseridos na ontologia em linguagem formal, em uma leitura amigável e natural para qualquer indivíduo, com treinamento mínimo, facilitando a compreensão das definições e a validação dos termos da ontologia.

5.3 Método de Validação

O objetivo geral da presente dissertação é a criação de um método de validação do processo de AC que possa ser aplicado por outros pesquisadores e profissionais da informação. O resultado da criação deste método será apresentado nesta seção.

Os pesquisadores interessados em validar uma ontologia devem seguir quatro passos que já foram descritos no capítulo da metodologia desta dissertação.

- i) **Passo 1:** consiste na busca pelos termos inseridos na ontologia e que necessitam serem validados por um especialista de domínio. Os termos devem ser inseridos em um formulário que será utilizado na entrevista de validação, conforme figura 8. Este formulário deve conter um campo com o termo a ser validado; um campo para adição do axioma a ser validado; um campo das perguntas para validação; um campo para observações e um campo o resultado da validação com relação ao quantificador lógico;
- ii) **Passo 2:** após a identificação dos termos, é necessário encontrar o axioma referente ao termo na ontologia que está sendo validada. O axioma deve ser inserido no formulário;
- iii) **Passo 3:** os axiomas, que são registados na linguagem lógica dentro da ontologia, deve ser traduzido utilizando o *plug in ACEView* instalado no editor de ontologias *Protégé 5.5* ([HTTP://ATTEMPTO.IFI.UZH.CH/ACEVIEW/](http://attempto.ifi.uzh.ch/aceview/), 2018). Esta tradução deve passar por ajustes de tradução, pois a ACE é uma LNC desenvolvida utilizando a língua inglesa. Por este motivo parte da

tradução da linguagem lógica ficará em inglês e outra parte em português. Sendo necessário fazer ajustes de tradução da língua. A tradução final da linguagem lógica deve ser registrada em um arquivo de *PowerPoint* com as perguntas utilizadas no processo de validação;

- iv) **Passo 4:** criação de um roteiro de perguntas visando a validação dos quantificadores lógicos, conforme demonstrado na figura 9. As perguntas têm o intuito de validar tanto o processo de AC, por conter o conteúdo adquirido neste processo e o quantificador lógico.

Figura 8 – Modelo de formulário para validação

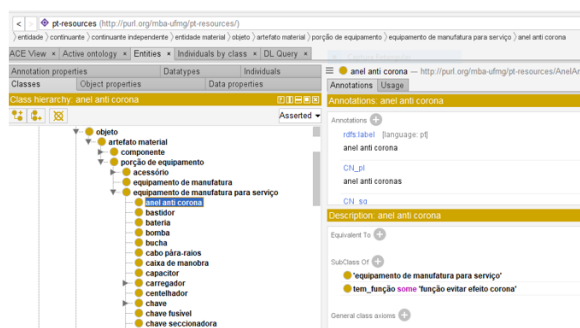
Termo nº X	Inserir termo neste campo			
Axioma:				
Perguntas:				
0	Sim		Não	Observação:
1	Sim		Não	
2	Sim		Não	
Resultado	Inserir resultado da validação neste campo			

Fonte: o autor

Este método contempla a validação dos quantificadores lógicos, que é umas das “armadilhas” que mais confundem os pesquisadores durante o processo de construção de uma ontologia. Entretanto, este método também contribui para a validação do processo de AC, pois é através da tradução da linguagem lógica, pela LNC utilizada, que o especialista de domínio irá validar o quantificador lógico, e ocorrendo erros de conteúdo, eles são fáceis de serem identificados.

Figura 9 – Modelo de questionário para validação

- "anel anti-corona" tem_função "evitar efeito corona"
- 0- isso está certo?
- SIM
- NAO
- 1- algum outro equipamento serve para evitar o efeito o corona?
- SIM
- NAO
- 2- Tudo que "evitar efeito corona" é um "anel anticorona"?
- SIM
- NAO



Fonte: o autor

6 Discussão

A discussão sobre um assunto tão complexo, como o que envolve a construção de ontologias e a fase de AC, pode levar a considerações diversas. Nos capítulos iniciais deste trabalho, observa-se assuntos que envolvem questões complexas de difícil compreensão.

O estudo da AC é de grande importância para a CI quando se trata do estudo da construção de artefatos de representação do conhecimento através das ontologias. A fase de AC deve ser tratada com cuidado e bem planejada, onde as técnicas escolhidas sejam as ideais, tanto para o trabalho do profissional da informação, quanto para a compreensão do que é transferido de conhecimento pelo especialista para o profissional da informação.

É possível verificar que foram encontrados pontos positivos e negativos relacionados ao processo de construção de ontologias, à fase de AC, à utilização de uma LNC e ao tipo de ontologia que estava sendo desenvolvida no projeto de P&D dentro da CEMIG. Esses serão discutidos ao longo das próximas subseções.

6.1 Validação utilizando *plug in ACEView*

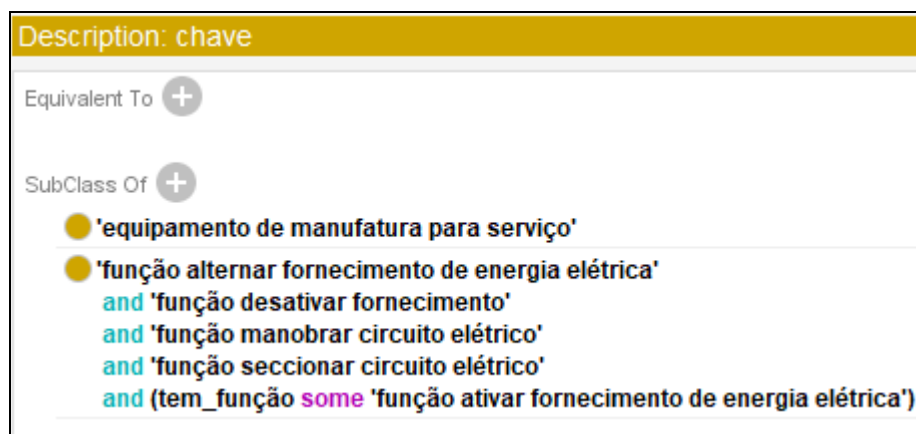
A utilização do *ACEView* dentro do *software Protégé 5.5* facilitou o entendimento da linguagem lógica, traduzida da Linguagem *Manchester* para a ACE, tanto pelo profissional da informação quanto pelo especialista de domínio.

O objetivo do *ACEView* é simplificar a exploração e edição de ontologias OWL 2 expressivas e sintaticamente complexas e conjuntos de regras SWRL baseando a interface do usuário na ACE (FUCHS; KALJURAND; KUHN, 2008, p. 18).

Isso torna o *ACEView* radicalmente diferente dos editores OWL / SWRL atuais, que são baseados em sintaxes lógicas formais, que muitas vezes são vistos como complexos na sua utilização por especialistas sem experiência prévia em métodos formais (FUCHS; KALJURAND; KUHN, 2008). Kaljurand (2008) demonstra os benefícios da utilização do *ACEView* no processos de convenções de nomenclatura e consistência, na complexidade estrutural das expressões de classe e no retorno para sintaxes e métodos existentes.

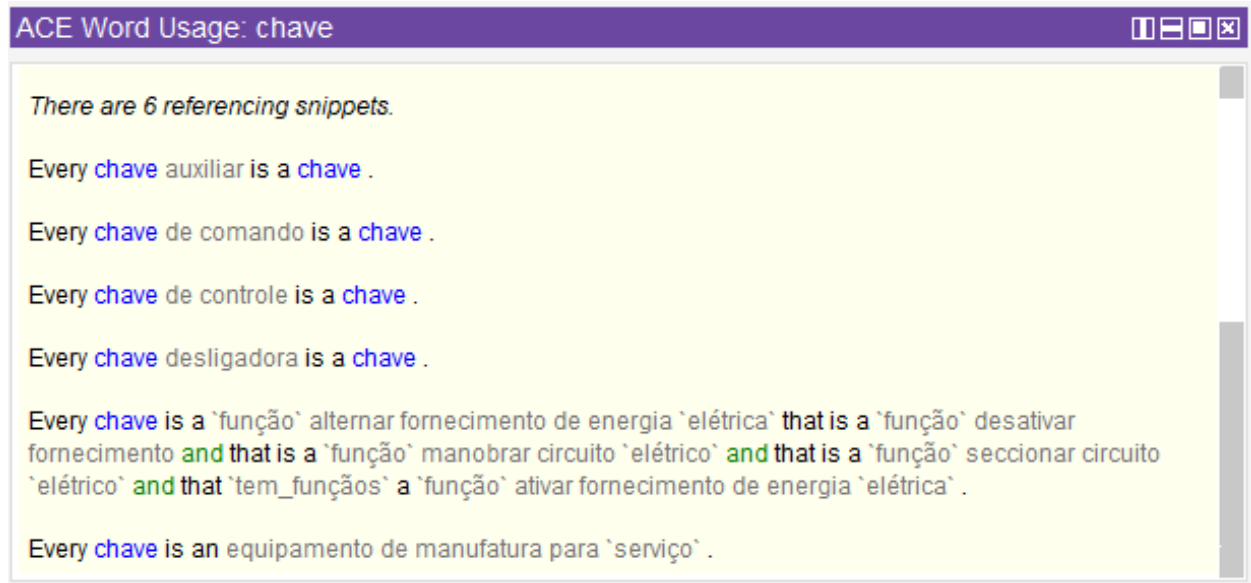
Quando se analisa o resultado gerado pelo *plug in*, de forma abrangente, no formato de uma lista completa com a tradução de todos os axiomas contidos na ontologia estudada, é possível compreender o significado dos termos de uma forma mais acessível, facilitando o processo de validação da ontologia. O processo de validação é dificultado quando se analisa somente a linguagem formal em *Manchester*, conforme demonstrado nas figuras 10 e 11.

Figura 10 – Linguagem formal *Manchester* – tela do *Protégé* 5.5



Fonte: o autor

Figura 11 – Plug in ACEView



Fonte: o autor

Esses axiomas possuem uma estrutura simples, no entanto, a apresentação de todas as implicações como uma única lista de frases em linguagem natural, pode fornecer uma visão geral boa e de fácil leitura (KALJURAND, 2008).

Ao longo da última década diversas LNCs foram projetadas e usadas para escrever as especificações de *softwares* com o objetivo de apoiar o processo de Aquisição e Representação do Conhecimento, entre elas a ACE (LUZ, 2013).

Entretanto a ACE é uma LNC desenvolvida na língua inglesa, sendo assim a tradução de uma ontologia de domínio desenvolvida em português. Utilizando-se de uma ontologia de alto nível, também na língua inglesa, a tradução necessitou de ajustes devido aos atributos lógicos estarem em inglês e as definições em português. Tal fato não alterou o sentido final das traduções, porém os ajustes foram necessários para que a frase apresentada ao especialista ficasse melhor compreendida. Não demonstrando uma tradução incompleta ou sem sentido, caso o especialista não tivesse conhecimento da língua inglesa.

A metodologia utilizada na validação tornou o processo fácil de ser realizado e registrado. Durante as entrevistas de validação com o especialista os termos foram apresentados em um arquivo *PowerPoint*TM e registrados em uma

planilha de *Excel*, apresentada na figura 11. Sendo este o local onde foram registradas as traduções realizadas pelo *ACEView*, atribuições adequadas e se os conteúdos das definições estavam corretos ou não, além das adequações necessárias de definição.

Figura 12 – Planilha de registro - validação ontologia

Termo 16	Transformador de força				
tem_função some 'função manter a potência elétrica constante'					
Perguntas:					
0	Sim	X	Não		
				Observação: Mofica a tensão de corrente na entrada para a saída mantendo ela constante.	
1	Sim		Não		X
2	Sim		Não		
Resultado	ONLY				

Fonte: o autor

O processo de aprendizagem para a utilização do *ACEView* é rápido e fácil de ser compreendido e utilizado pelo pesquisador/profissional da informação. A utilização do *ACEView* demanda de menor tempo para ser aprendido e é mais compreensível por quem o utiliza (KUHN, 2012)

Ainda segundo Kuhn (2012), a LNC pode potencialmente levar a interfaces de usuário que são mais fáceis e rápidos de entender do que interfaces baseadas em linguagens de ontologia comuns.

6.2 Linguagem natural controlada em português

Como apresentado na sessão anterior, o *ACEView* é um *plug in* derivado da ACE, o que levou a tradução da ontologia estudada a uma mistura de palavras entre a língua portuguesa e a língua inglesa.

O principal benefício de usar LNCs é que elas são facilmente processáveis por analisadores e podem ser traduzidas para outras representações, sem problemas de ambiguidade (MÉSZÁROS; KISS, 2018).

Existe também a necessidade de ser realizado o desenvolvimento de uma LNC na língua portuguesa. Como um artefato de representação do conhecimento muito importante, as ontologias estão sendo cada vez mais utilizadas por desenvolvedores que utilizam da língua portuguesa, porém ainda

não existe uma LNC desenvolvida para este público. O que encontramos atualmente é um conjunto de palavras, glossários, dicionários ou tesouros em língua portuguesa, que não possui uma semântica e sintaxe controladas, além de uma grande quantidade de ambiguidades em suas definições.

Segundo Luz (2013, p. 7) “o que chamamos de interoperabilidade entre ontologia e linguagem natural é o fato dos elementos da ontologia serem escritos de maneira ‘clara’, com termos bem definidos em linguagem natural”.

Utilizar o português controlado tem como objetivo facilitar o desenvolvimento de ontologias pelo usuário, uma vez que a construção de sentenças utilizando um subconjunto na língua portuguesa, é mais fácil do que aprender uma nova língua (LUZ, 2013).

6.3 Literatura sobre a aquisição do conhecimento

Como apresentado no capítulo de revisão de literatura o tema AC possui pouca literatura sobre o assunto. Durante a realização das pesquisas nas bases de dados sobre o assunto, muitos temas sinônimos a AC são retornados como resultados, porém o cerne principal dos trabalhos estão direcionados para a AC relacionada à aprendizagem escolar, entretanto, quando se busca por AC relacionada a um domínio específico ou para o processo de construção de ontologias, poucos artigos são apresentados nas bases de dados, mesmo sendo um tema de grande importância para a CI e para a construção de ontologias.

Os resultados retornados referentes a buscas por Representação do Conhecimento combinado com a AC, apresentam artigos relacionados aos artefatos de representação do conhecimento, como a ontologia, porém não apresentam suas fases de construção, onde a AC está inserida.

A fase de AC tem importância no processo de construção de ontologias, pois é através das técnicas existentes que se extrai o conhecimento de um especialista de domínio. A AC é o processo de extrair conhecimento útil de o domínio, de modo a construir uma base de conhecimento (SHU; QIAN; XIE, 2019).

Um número maior de publicações sobre o assunto referente a uma fase muito importante no processo de construção de ontologias poderá ajudar na

criação de metodologias melhores, diminuindo o número de erros na eliciação do conhecimento.

6.4 Técnicas de aquisição do conhecimento

A existência de diversas técnicas de AC proporciona um grande leque de opções ao profissional da informação para as utilizarem durante o processo de construção de uma ontologia.

Como é possível observar na seção *background*, a maioria das técnicas apresentadas são manuais, o que faz com que o profissional informacional, tenha durante o processo de construção, um contato direto com os diferentes especialistas de domínio e demais profissionais ligados ao setor, buscando diversas informações relevantes.

Durante o presente trabalho foram utilizadas as técnicas de análise de documentos, o que trouxe uma grande gama de informações, para quando o processo evoluísse para entrevistas com os especialistas, a compreensão do que estava sendo abordado fosse mais fácil para o profissional da informação.

As técnicas utilizadas nesta dissertação foram todas manuais, o que garantiu uma grande proximidade de todos os envolvidos no projeto, ajudando na resolução de dúvidas com maior facilidade, porém nos últimos anos tem se realizado muitos estudos na busca da AC de forma automatizada através do emprego de diversos bancos de dados.

6.5 Complexidade do domínio e a busca pelo especialista

O Setor Elétrico é um domínio complexo de ser compreendido por profissionais que não fazem parte do dia a dia do setor. Diversos profissionais compõem uma cadeia de ações complexas voltadas para o planejamento, a implementação de equipamentos e a manutenção dos mesmos.

Dentro do trabalho realizado foi necessário, além do conhecimento obtido pela análise de documentos, compreender o que o especialista de domínio informava durante o processo de AC. É necessário ter os especialistas para julgar se os termos extraídos fazem sentido no domínio. O vocabulário também tem uma função relevante, onde definir harmonicamente o significado dos termos

usados na prática e na pesquisa, é possível através da AC diretamente de especialistas (COELHO; ALMEIDA, 2015).

A busca pelo especialista de domínio esbarra em algumas barreiras, que vão além dos critérios necessários para sua escolha, como a experiência e o conhecimento sobre o domínio e a disponibilidade para realizar o processo longo de AC. Sendo assim, pode ser necessária a realização de diversas entrevistas com um tempo longo de duração.

A principal dificuldade para o especialista de domínio, em participar de uma atividade de AC, de modo geral, é o pouco tempo disponível por acumular outras atividades dentro do seu escopo de trabalho (COELHO; ALMEIDA, 2012).

O processo de AC em geral é uma experiência que consome uma grande demanda de tempo tanto pelo especialista quanto para o profissional da informação, além de ser um processo que pode vir a ser caro, o que explica o seu negligenciamento em muitos casos (MENDONÇA et al., 2012).

Também é importante que o especialista acredite que o processo que está sendo realizado é importante para o seu campo de atuação e que a sua dedicação a essa atividade trará melhorias para o setor no futuro.

Segundo Coelho e Almeida (2012) a discussão sobre a atividade de AC em domínios especializados, torna o trabalho mais eficiente e, no longo prazo, auxilia o especialista do domínio.

6.6 Pesquisadores e a Aquisição do Conhecimento

Como apontado na metodologia, sete pesquisadores participaram do processo de AC, compondo a construção de uma ontologia experimental para a CEMIG.

A diversidade de pesquisadores é importante para que visões diferentes sobre um determinado assunto possam trazer benefícios para o processo de AC. Entretanto, a técnica a ser utilizada tem que ser a mesma para todos, para que informações importantes não se percam, levando a erros quando estas forem inseridas na ontologia pelo desenvolvedor. Reduzindo assim, a quantidade de erros de conteúdo que podem surgir.

Essa diversidade de pesquisadores pode ser também benéfica quando observado que a compreensão de um domínio complexo pode levar a dificuldade aos pesquisadores em compreender o que o especialista tem em mente. As ontologias são desenvolvidas de forma colaborativa, sendo essencial considerar a forma como as pessoas enxergam o mundo e compreendem os processos sociais que levam ao desenvolvimento (MENDONÇA et al., 2015).

A visão do especialista, muitas vezes prática, pode conduzir a incompreensões sobre o assunto, devido a facilidade prática de lidar com o dia a dia do setor, para o profissional da informação, que está registrando o que está sendo dito, pode não estar claro.

6.7 Perda de Informação

Diversos fatores podem levar a perda de informações importantes durante o processo de AC, o que pode justificar o número de erros de conteúdos encontrados no trabalho.

Compreender o que é dito pelo especialista de domínio é importante para que tais erros possam ser evitados, além do registro e manipulações corretas das informações que estão sendo elicitadas pelo especialista.

O especialista de domínio deve garantir que o conhecimento transmitido a todos os envolvidos esteja claro, para que durante todo o processo, inclusive o de validação da ontologia, os erros identificados possam ser sanados com sucesso.

6.8 Complexidade das Ontologias

O domínio de conhecimento onde se encontra o estudo das ontologias possui uma grande complexidade, sendo um processo longo que envolve diversas metodologias e a necessidade do conhecimento de um desenvolvedor.

A construção de ontologias é um processo difícil, que requer a presença de um especialista de domínio, além da dificuldade de se utilizar ontologias em larga escala (DE FREITAS; VIEIRA, 2008). As ontologias são compreendidas como um instrumento para a organização e representação da

informação, capazes de capturar, mapear e compartilhar o conhecimento de diferentes domínios do conhecimento (MENDONÇA; ZAIDAN, 2019).

Outro fator, é que existem diversas metodologias, cada uma com suas propriedades e passos para construção. A inexistência de uma metodologia padronizada para a construção de uma ontologia pode levar a erros. A *OntoForInfoScience* é uma metodologia desenvolvida para ser utilizada por profissionais da CI, buscando uma padronização das etapas. É uma metodologia que utiliza uma linguagem comum e simples, facilitando a compreensão das atividades, realizadas na construção de ontologias, mantendo a alta expressividade das linguagens lógicas (MENDONÇA; ALMEIDA, 2016).

Como demonstrado por Gomez-Pérez; Fernández e Corcho (2004) existem 35 diferentes “armadilhas” que são imperceptíveis para o desenvolvedor. Tais “armadilhas”, podem levar a erros, que podem impossibilitar a utilização de uma ontologia. Mesmo com toda a complexidade existente no processo de construção de uma ontologia, o artefato gerado por ela torna a compressão de um domínio e a organização das informações referentes a ele mais fáceis de acessar e utilizar.

Segundo de Freitas e Vieira (2008, p. 8), “as ontologias representam o ponto mais elevado já atingido em termos de representação, compartilhamento e reutilização do conhecimento.” No processo de desenvolvimento de produtos, as ontologias ajudam a tornar explícito o conhecimento adquirido de um certo domínio, permitindo representar melhor as informações na transição entre o projeto e o desenvolvimento do produto (MENDONÇA; ZAIDAN, 2019).

No cenário atual, as ontologias de alto nível são todas realizadas utilizando-se de ontologias desenvolvidas na língua inglesa. Assim como é necessário a criação de uma LNC em português, também se faz necessário a criação de uma ontologia de alto nível totalmente construída na língua portuguesa.

Segundo de Freitas e Vieira (2008), pesquisas relacionadas a automatização da tarefa de construção de ontologias, estão sendo realizadas levando em conta bases textuais como fonte de conhecimento, desta forma, ferramentas e métodos dependem diretamente do idioma, já que as informações coletadas encontram-se escritas em determinada língua. Ainda segundo os

autores existe um esforço para a construção de ontologias em língua portuguesa, sendo esses ainda esforços pequenos e pontuais.

Existem ainda poucas propostas na área para o idioma português, além da constante dificuldade no compartilhamento das ontologias existentes em língua portuguesa (DE FREITAS; VIEIRA, 2008).

A ontologia entregue à CEMIG, foi concebida na forma experimental, por este motivo alguns erros foram encontrados. Alguns erros necessitam ser corrigidos para que ela possa ser operacionalizada e transformada em um projeto real de implementação.

Tomando como base o estudo realizado por Gomez-Pérez; Fernández e Corcho (2004), os erros encontrados nesta ontologia experimental estão relacionados a problemas de modelagem e entendimento humano, indicando erros de ausência da propriedade equivalente e hierarquias muito especializadas conforme descrito nos resultados.

7 Considerações finais e conclusão

O presente trabalho teve por objetivo criar um método para validar um processo de AC confrontando resultados desta etapa com os axiomas da fase de formalização, a partir da utilização da LNC. Para tanto, utilizou-se a LNC ACE através do seu *plug in ACEView* para o *software* editor de ontologias *Protégé 5.5*.

Na seção 1 apresentou-se a pesquisa, seus objetivos e questões. A seção 2 apresentou um *background* com noções básicas sobre as ontologias e a AC. A seção 3 descreveu as linguagens e a representação do conhecimento apresentando conceitos básicos sobre a linguagem natural, linguagem lógica e LNC. A seção 4 apresentou o contexto onde a pesquisa está inserida e a metodologia utilizada com os quatro passos descritos na validação da LNC. A seção 5 demonstrou os resultados encontrados na validação da ontologia e da LNC através do *ACEView*. A seção 6 discutiu os resultados obtidos. Finalmente, o restante da presente seção, oferece considerações finais e conclusões, bem como perspectiva de trabalhos futuros.

A falta de publicações relacionadas ao tema da AC é um ponto importante a se abordar, sendo um fator complicador dentro do processo de construção e de estudo sobre o tema, que deve ser aprofundado, tanto por esta pesquisa quanto por trabalhos futuros. Desta forma, foi importante delimitar ainda no início do trabalho, o escopo do problema ao nível da validação da fase de AC com o especialista, utilizando-se de uma LNC, porém, analisando os outros processos adequados e as dificuldades na criação de uma ontologia como um todo.

A fase de AC, inserida em todas as metodologias para a construção de uma ontologia, é contemplada por um vasto número de técnicas de elicitação do conhecimento. A escolha da técnica ideal leva o trabalho a um dinamismo necessário para que as informações possam ser elencadas, catalogadas e organizadas de forma a facilitar e reduzir erros no processo de construção de uma ontologia.

Para que a fase de AC seja realizada, é necessário que entrevistas sejam feitas por profissionais da informação com o especialista de domínio.

Durante o estudo dos termos, o papel do especialista de domínio provou ser importante, pois além de definir os termos de uma forma que fosse de fácil compreensão, por muitas vezes foi necessário utilizar-se de desenhos e de explicações mais aprofundadas, garantindo a qualidade da informação fornecida.

A experiência disponível na literatura sobre a AC em diversos campos menciona dificuldades de comunicação entre especialistas de domínio e profissionais que lidam com informação (MENDONÇA et al., 2012). Segundo Coelho e Almeida (2012) o vocabulário especializado dificulta a compreensão das terminologias e repassa conceitos complexos para quem não pertence ao domínio estudado e é um fator que pode levar a dificuldades no processo de AC.

A utilização do ACE e o seu *plug in* ACEView no processo de tradução da linguagem formal para linguagem lógica trouxe grandes benefícios para o processo de validação pela forma abrangente como o ACEView apresenta a tradução dos axiomas contidos na ontologia.

Estudos demonstram a existências de diversas técnicas manuais para a realização da AC. Tais técnicas devem ser utilizadas de acordo com o tipo de estudo que está sendo realizado e a que se molde melhor ao processo de pesquisa utilizado. Entretanto, as pesquisas relacionadas ao tema nos últimos cinco anos, apontam para a maior utilização de técnicas automatizadas de AC, através da utilização de algoritmos de mineração de dados.

Com estas afirmações é possível concluir que os objetivos da presente dissertação são pertinentes ao estudo dentro do campo da CI.

Objetivo 1: foi possível sistematizar as possíveis falhas no processo de AC. Tais falhas englobam o a fase de AC, mas também os demais processos de construção de uma ontologia. É possível identificar erros de compressão do que o especialista transfere de conhecimento para os profissionais informacionais por diversos motivos, porém, erros relacionados ao objetivo da ontologia construída e experiência do desenvolvedor foram identificados, demonstrando como é complexo a criação de uma ontologia e também a sua importância como artefato de representação do conhecimento, quando essas possíveis falhas são identificadas e corrigidas.

É importante manter as boas práticas no processo de construção de ontologias e durante a fase de AC, reduzindo a possibilidade de ocorrências de

erros ou anomalias, mesmo assim, é um processo complexo e sujeito a erros. Uma ampla lista de “armadilhas” é apresentada na seção de resultados, o que pode ajudar desenvolvedores a manterem as boas práticas em trabalhos futuros.

Objetivo 2: Foi desenvolvido um método, composto de quatro passos para que profissionais da CI possam realizar de forma mais amigável, utilizando-se de uma LNC. A linguagem lógica é de difícil compreensão, tanto por profissionais da informação, quanto por especialistas de domínio. A criação deste método torna o processo mais rápido e preciso, podendo identificar com facilidade erros decorrentes da fase de AC.

Objetivo 3: A utilização de uma LNC na validação de uma linguagem lógica provou-se possível, mesmo que a linguagem utilizada ACE, através de seu *Plug in ACEView*, tenha as suas limitações, já que as traduções de axiomas simples tenham ficado próximo da estrutura da linguagem *Manchester*. Entretanto, a linguagem se tornou acessível e de fácil compreensão pelo especialista de domínio, facilitando a sua validação durante as entrevistas realizadas.

É importante a realização futura de novos trabalhos relacionados a um tema tão importante como a AC, a CI e às ontologias, que ao longo do tempo vem sendo negligenciado por diversos motivos. O estímulo ao estudo e a novas publicações sobre o assunto ajudariam os profissionais da informação e o campo da CI no desenvolvimento de novas técnicas de eliciação do conhecimento, mais precisas e provavelmente diminuindo os erros de conteúdo relacionados a ontologias.

Mais estudos podem ser desenvolvidos utilizando a ACE, pelas suas facilidades de instalação e utilização dentro da interface do *Protégé* 5.5. É importante também que sejam desenvolvidos estudos relacionados à criação de ontologias de alto nível e uma LNC na língua portuguesa. O ACE foi desenvolvido utilizando-se a língua inglesa, o que trouxe a necessidade de realizar pequenos ajustes na tradução da linguagem lógica de uma antologia desenvolvida em língua portuguesa. Tais ajustes não trouxeram prejuízos para os resultados obtidos, porém demonstra a necessidade do desenvolvimento de uma LNC, e ontologias de alto nível, em língua portuguesa, devido ao grande número do desenvolvimento de tais artefatos de representação do conhecimento também em língua portuguesa., sendo que a própria

ACE possa ser adaptada para o português, por seu *plug in* ser um *software open source*.

Por fim, é pertinente neste momento indicar as percepções da realização da pesquisa em todo o seu percurso. O trabalho contemplou o amplo escopo da pesquisa visando alcançar os objetivos, geral e específicos, desta pesquisa dentro contexto da AC, da relação especialista de domínio e profissional da informação e a construção de ontologias.

O domínio referente às ontologias é muito complexo, que demanda muita dedicação ao estudo constante para sua compreensão, principalmente para aqueles pesquisadores iniciantes que possuem pouco ou nenhum conhecimento sobre o domínio. Compreender as diferentes linhas de pesquisa relacionadas ao tema no campo da Filosofia, da Ciência da Computação, da CI e da Linguística é muito importante para o que o pesquisador possa compreender a sua utilização, sendo necessário uma base de estudo filosófico bem ampla, o que também facilita a compreensão do tema.

Uma outra característica que demandou bastante dedicação foram as entrevistas com os especialistas de domínio. A busca pelo o especialista ideal para um projeto de AC deve levar em conta a busca por um profissional experiente, com vasto conhecimento sobre o domínio estudado. Compreender o domínio é importante para que a construção de uma ontologia ocorra com menos dores, sendo importante uma sinergia tanto da equipe de profissionais da informação, atuantes nos processos de construção, como entre a equipe de profissionais da informação e o especialista de domínio.

Para que a participação do especialista de domínio seja realmente eficiente, ele deve acreditar na importância do projeto de construção de ontologias e que dentro deste processo ele terá que transmitir informações específicas do domínio para pessoas que não conhecem a fundo sobre o assunto. Tal reflexão é importante para que a informação transmitida seja clara e compreensível. Essa compreensão leva o especialista de domínio a olhar para o seu setor de atuação com um olhar diferente, onde possivelmente terá a oportunidade de aprender um pouco mais sobre ele.

Referências

- AKHAVAN, P.; SHAHABIPOUR, A.; HOSNAVI, R. A Model for Assessment of Uncertainty in Tacit Knowledge Acquisition. **Journal of Knowledge Management**, v. 22, n. 2, p. 413–431, 12 mar. 2018.
- ALMEIDA, M. B. **Um modelo baseado em ontologias para representação da memória organizacional**. 2006. Universidade Federal de Minas Gerais, 2006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362006000300013&lng=pt&tlng=pt>. Acesso em: 3 jun. 2020.
- ALMEIDA, M. B. Revisiting Ontologies: A Necessary Clarification. **Journal of the American Society for Information Science and Technology**, v. 64, n. 8, p. 1682–1693, ago. 2013.
- AN, J.; PARK, Y. B. Methodology for Automatic Ontology Generation Using Database Schema Information. **Mobile Information Systems**, v. 2018, p. 1–13, 2018.
- ANDRADE, A. Q. D. **A LINGUAGEM MÉDICA UTILIZADA EM PRONTUÁRIOS E SUAS REPRESENTAÇÕES EM SISTEMAS DE INFORMAÇÃO: AS ONTOLOGIAS E OS MODELOS DE INFORMAÇÃO**. 2013. Universidade Federal de Minas Gerais, Belo Horizonte, 2013.
- ARP, R.; SMITH, B.; SPEAR, A. D. **Building ontologies with Basic Formal Ontology**. Cambridge, Massachusetts: Massachusetts Institute of Technology, 2015.
- BARBOSA, J. L. N. et al. Introdução ao Processamento de Linguagem Natural usando Python. **Livro Anais - Artigos e Manuscritos**, v. 1, n. 1, p. 25, 2017.
- BOLHUIS, J. J.; EVERAERT, M. (ed.). **Birdsong, speech, and language: exploring the evolution of mind and brain**. Cambridge, Mass: MIT Press, 2013.
- BOOSE, J. H. Knowledge Acquisition Tools, Methods, and Mediating Representations. **Knowledge Acquisition for Knowledge-Based Systems (Motoda, H. et al., Eds)**, IOS Press, Ohinsha Ltd., Tokyo, p. 123–168, 1990.
- BOOSE, J. H.; GAINES, B. R. Knowledge Acquisition for Knowledge-Based Systems: Notes on the State-of-the-Art. **Machine Learning**, v. 4, p. 337–394, 1989.
- BURGE, J. E. **Knowledge Elicitation for Design Task Sequencing Knowledge**. 1998. Worcester Polytechnic Learning, US, 1998.
- BURGESS, J. P. Why I am not a nominalist. **Notre Dame Journal of Formal Logic**, v. 24, n. 1, p. 93–105, 1983.
- BURKHARDT, H.; SMITH, B. (ed.). **Handbook of Metaphysics and Ontology: 2-Volume-Set Students Edition**. [s.l.] Philosophia Verlag, 2001.

CEMIG. Disponível em: <https://www.cemig.com.br/pt-br/a_cemig/nossos_negocios/Paginas/distribuicao.aspx>. Acesso em: 17 jun. 2020.

CHANTREE, F. Ambiguity management in natural language generation. In: 7th Ann. CLUK Research Colloquium, **Anais...**2004.

CHOMSKY, N. **On Nature and Language**. Cambridge: Cambridge University Press, 2002.

CHOO, C. W. The Knowing Organization: How Organizations Use Information to Construct Meaning, Create Knowledge and Make Decisions. **International Journal of Information Management**, v. 16, n. 5, p. 329–340, 1996.

CHOO, C. W. **The knowing organization: how organizations use information to construct meaning, create knowledge, and make decision**. 2. ed ed. New York, NY: Oxford Univ. Press, 2006.

CIMIANO, P.; UNGER, C.; MCCRAE, J. **Ontology-based interpretation of natural language**. [s.l.] Morgan & Claypool Publishers, 2014.

COCCHIARELLA, N. B. Conceptual Realism as a Formal Ontology. In: **Formal Ontology**. Nijhoff International Philosophy Series. [s.l.] Roberto Poli & Peter Simons, 1996. 53p. 27–60.

COELHO, K. C. **Aquisição de conhecimento especializado para construção de ontologias: um estudo no domínio das ciências da vida**. 2012. Universidade Federal de Minas Gerais, 2012. Disponível em: <<https://repositorio.ufmg.br/handle/1843/ECID-8XLQS4>>. Acesso em: 8 dez. 2019.

COELHO, K. C.; ALMEIDA, M. B. Aquisição de conhecimento para construção de ontologias: uma proposta de roteiro metodológico aplicado ao contexto da hematologia. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 17, n. 35, 18 dez. 2012a. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/19175>>. Acesso em: 9 jul. 2019.

COELHO, K. C.; ALMEIDA, M. B. Aquisição de conhecimento para construção de ontologias: uma proposta de roteiro metodológico aplicado ao contexto da hematologia. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 17, n. 35, 18 dez. 2012b. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/19175>>. Acesso em: 2 mar. 2020.

COELHO, K. C.; ALMEIDA, M. B. Representation of Biomedical Expertise in Ontologies: A Case Study about Knowledge Acquisition on HTLV Viruses and Their Clinical Manifestations. **Proceedings of the 15th World Congress on Health and Biomedical Informatics**, p. 5, 2015a.

COELHO, K. C.; ALMEIDA, M. B. Representation of Biomedical Expertise in Ontologies: A Case Study about Knowledge Acquisition on HTLV Viruses and Their Clinical Manifestations. **Proceedings of the 15th World Congress on Health and Biomedical Informatics**, p. 5, 2015b.

CORCHO, O.; FERNÁNDEZ-LÓPEZ, M.; GÓMEZ-PÉREZ, A. Methodologies, Tools and Languages for Building Ontologies. Where Is Their Meeting Point? **Data & Knowledge Engineering**, v. 46, n. 1, p. 41–64, jul. 2003.

CRESWELL, J. W. **Research design: qualitative, quantitative, and mixed methods approaches**. [s.l.: s.n.]

DE FREITAS, L. A.; VIEIRA, R. Ontologias e língua portuguesa. **Anais do CELSUL**, v. 1, n. 2, p. 5, 2008.

DEHGHANI, M.; AKHAVAN, P. An Experimental Investigation of Knowledge Acquisition Techniques. **Journal of Management Development**, v. 36, n. 4, p. 493–514, 8 maio 2017.

DHINGRA, V.; BHATIA, K. K. Development of Ontology in Laptop Domain for Knowledge Representation. **Procedia Computer Science**, v. 46, p. 249–256, 2015.

FARINELLI, F. **Realismo Ontológico Aplicado a Interoperabilidade Semântica Entre Sistemas de Informação: Um estudo de caso do domínio obstétrico e neonatal**. 2017. UNIVERSIDADE FEDERAL DE MINAS GERAIS ESCOLA DE CIÊNCIA DA INFORMAÇÃO, Belo Horizonte, 2017.

FERREIRA, C. K. L. Privatização do setor elétrico no Brasil. 2000.

FUCHS, N. E.; KALJURAND, K.; KUHN, T. Attempto Controlled English for Knowledge Representation. In: BAROGLIO, C. et al. (Ed.). **Reasoning Web**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. 5224p. 104–124.

FUCHS, N. E.; KALJURAND, K.; SCHNEIDER, G. Attempto Controlled English Meets the Challenges of Knowledge Representation, Reasoning, Interoperability and User Interfaces. **FLAIRS'2006**, 2006. Disponível em: <<https://www.zora.uzh.ch/id/eprint/33148>>. Acesso em: 9 jan. 2020.

GANZ, F.; BARNAGHI, P.; CARREZ, F. Automated Semantic Knowledge Acquisition From Sensor Data. **IEEE Systems Journal**, v. 10, n. 3, p. 1214–1225, set. 2016.

GIL, A. C. **Métodos e técnicas de pesquisa social**. [s.l.] 6. ed. Editora Atlas S.A., 2008.

GOLDSTEIN, L. et al. **Lógica**. 1 edition ed. [s.l.] Artmed, 2007.

GÓMEZ-PÉREZ, A.; FERNÁNDEZ, M.; DE VICENTE, A. Towards a method to conceptualize domain ontologies. In: ECAI96 Workshop on Ontological Engineering, Budapest, **Anais...**1996a.

GÓMEZ-PÉREZ, A.; FERNÁNDEZ, M.; DE VICENTE, A. J. Towards a Method to Conceptualize Domain Ontologies. **ECAI96 Workshop on Ontological Engineering, Budapest**, p. 12, 1996b.

GÓMEZ-PÉREZ, A.; FERNÁNDEZ-LÓPEZ, M.; CORCHO, O. **Ontological engineering: with examples from the areas of knowledge management, e-commerce and the Semantic Web**. London ; New York: Springer, 2004.

GÓMEZ-PÉREZ, A.; SUÁREZ-FIGUEROA, M. C. Scenarios for building ontology networks within the NeOn methodology. In: Proceedings of the fifth international conference on Knowledge capture - K-CAP '09, Redondo Beach, California, USA. **Anais...** In: THE FIFTH INTERNATIONAL CONFERENCE. Redondo Beach, California, USA: ACM Press, 2009. Disponível em: <<http://portal.acm.org/citation.cfm?doid=1597735.1597773>>. Acesso em: 24 jan. 2021.

GONZALEZ, M. Recuperação de Informação e Processamento da Linguagem Natural. p. 49, 2003.

GRUBER, T. R. A Translation Approach to Portable Ontology Specifications. **Knowledge Acquisition**, v. 5, n. 2, p. 199–220, jun. 1993.

GRUNINGER, M.; FOX, M. S. Methodology for the Design and Evaluation of Ontologies. p. 10, 1995.

HEIYANTHUDUWAGE, S. R.; SCHWITTER, R.; ORGUN, M. A. OWL 2 Learn Profile: An Ontology Sublanguage for the Learning Domain. **SpringerPlus**, v. 5, n. 1, dez. 2016. Disponível em: <<http://www.springerplus.com/content/5/1/291>>. Acesso em: 24 abr. 2020.

[HTTP://ATTEMPTO.IFI.UZH.CH/ACEVIEW/](http://attempto.ifi.uzh.ch/aceview/). **ACE View — a natural language interface to knowledge engineering**. Disponível em: <<http://attempto.ifi.uzh.ch/aceview/>>. Acesso em: 10 jan. 2020.

IVANOV, A.; VARFOLOMEYEV, A. Service-Oriented Architecture of Intelligent Environment for Historical Records Studies. **Procedia Computer Science**, v. 104, p. 57–64, 2017.

KALJURAND, K. ACE View --- an Ontology and Rule Editor Based on Attempto Controlled English. **s.n.**, 2008. Disponível em: <<https://www.zora.uzh.ch/id/eprint/8822>>. Acesso em: 10 jan. 2020.

KALJURAND, K.; FUCHS, N. E. Bidirectional Mapping Between OWL DL and Attempto Controlled English. In: ALFERES, J. J. et al. (Ed.). **Principles and Practice of Semantic Web Reasoning**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. 4187p. 179–189.

KUHN, T. AceWiki: A Natural and Expressive Semantic Wiki. p. 8, 2008a.

KUHN, T. AceWiki: Collaborative Ontology Management in Controlled Natural Language. p. 5, 2008b.

KUHN, T. Combining Semantic Wikis and Controlled Natural Language. p. 2, 2008c.

KUHN, T. **Controlled English for knowledge representation**. 2010. 2010. Disponível em: <<https://www.zora.uzh.ch/id/eprint/33191>>. Acesso em: 9 jan. 2020.

KUHN, T. The Understandability of OWL Statements in Controlled English. **IOS Press**, n. 1, p. 1–15, 2012.

LOWE, E. J. Ontological Categories and Categorial Schemes. In: **The Four-Category Ontology A Metaphysical Foundation for Natural Science**. [s.l.] Oxford University Press, 2006.

LUZ, F. F. **Consulta a ontologias utilizando linguagem natural controlada**. 2013. Universidade de São Paulo, 2013. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/45/45134/tde-02012014-224412/>>. Acesso em: 3 dez. 2020.

MACULAN, B. C. M. dos S. Estudo e aplicação de metodologia para reengenharia de tesouro: remodelagem do THESAGRO. 24 jun. 2015. Disponível em: <<https://repositorio.ufmg.br/handle/1843/BUBD-9ZKMUV>>. Acesso em: 16 jun. 2020.

MARAN, V. et al. UPCaD: A Methodology of Integration Between Ontology-Based Context-Awareness Modeling and Relational Domain Data. **Information**, v. 9, n. 2, p. 30, 30 jan. 2018.

MARCONDES, C. H. Análise ontológica de definições de informação: em busca da sua essência. **Transinformação**, v. 27, n. 2, p. 105–122, ago. 2015.

MASTELLA, L. S. **Um modelo de conhecimento baseado em eventos para aquisição e representação de sequencias temporais**. 2005. Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005.

MENDES FERREIRA, D.; DE BARROS VIDAL, F. Towards a Temporal Extension to OWL 2: A Study based on tOWL Language: In: Proceedings of the 12th International Conference on Web Information Systems and Technologies, Rome, Italy. **Anais...** In: 12TH INTERNATIONAL CONFERENCE ON WEB INFORMATION SYSTEMS AND TECHNOLOGIES. Rome, Italy: SCITEPRESS - Science and Technology Publications, 2016. Disponível em: <<http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0005857502270234>>. Acesso em: 20 fev. 2020.

MENDONÇA, F. M. et al. Knowledge Acquisition in the Construction of Ontologies: A Case Study in the Domain of Hematology. **International Conference of Biomedical Ontologies**, p. 5, 2012.

MENDONÇA, F. M. et al. From a Consensual Conceptual Level to a Formal Ontological Level. **SEMAPRO**, p. 6, 2015.

MENDONÇA, F. M. **Ontoforinfoscience: metodologia para construção de ontologias pelos cientistas da informação - Uma aplicação prática no desenvolvimento da ontologia sobre componentes do sangue humano (HEMONTA)**. 2015. 2015. Disponível em: <<https://repositorio.ufmg.br/handle/1843/BUBD-A35H3K>>. Acesso em: 29 nov. 2019.

MENDONÇA, F. M.; ALMEIDA, M. B. ONTOFORINFOSCIENCE: A DETAILED METHODOLOGY FOR CONSTRUCTION OF ONTOLOGIES AND ITS APPLICATION IN THE BLOOD DOMAIN. **Brazilian Journal of Information Science**, v. 10, n. 1, 1 mar. 2016. Disponível em: <<http://revistas.marilia.unesp.br/index.php/bjis/article/view/5426>>. Acesso em: 11 set. 2020.

MENDONÇA, F. M.; ZAIDAN, F. H. Ontologias para organização da informação em processos de transformação digital. **Em Questão**, v. 25, n. 1, p. 295, 1 jan. 2019.

MÉSZÁROS, T.; KISS, M. Knowledge Acquisition from Critical Annotations. **Information**, v. 9, n. 7, p. 179, 20 jul. 2018.

OWL 2 Web Ontology Language Manchester Syntax (Second Edition). Disponível em: <<https://www.w3.org/TR/owl2-manchester-syntax/>>. Acesso em: 17 jun. 2020.

PONTES JUNIOR, J. de; CARVALHO, R. A. de; AZEVEDO, A. W. Da recuperação da informação à recuperação do conhecimento: reflexões e propostas. **Perspectivas em Ciência da Informação**, v. 18, n. 4, p. 2–17, dez. 2013.

POSTOLACHE, F. ONTOLOGY TOOL FOR KNOWLEDGE ACQUISITION IN A VIRTUALISED ICT INFRASTRUCTURE. "**Scientific Bulletin**" **Mircea Cel Batran**" **Naval Academy**", v. 19, n. 1, p. 484–489, 2016.

POVEDA-VILLALÓN, M.; SUÁREZ-FIGUEROA, M. C.; GÓMEZ-PÉREZ, A. A double classification of common pitfalls in ontologies. In: Workshop on Ontology Quality (OntoQual 2010), Co-located with EKAW, **Anais...**2010.

PRIEST, G. **Logic - A Very Short Introduction**. 2ª ed. New York, NY: Oxford University Press, 2017. v. 1

REESE, T. et al. Patient Information Organization in the Intensive Care Setting: Expert Knowledge Elicitation with Card Sorting Methods. **Journal of the American Medical Informatics Association**, v. 25, n. 8, p. 1026–1035, 1 ago. 2018.

SALEENA, B.; SRIVATSA, D. A Novel Approach to Develop a Self-Organized Domain Specific Search Mechanism for Knowledge Acquisition using Ontology. **International Journal of Computer Applications**, v. 12, n. 4, p. 1–4, 2010.

SALES, L. F.; CAMPOS, M. L. de A.; GOMES, H. E. Ontologias de domínio: um estudo das relações conceituais. **Perspectivas em Ciência da Informação**, v. 13, n. 2, p. 62–76, ago. 2008.

SARACEVIC, T. Ciência da Informação: Origem, evolução e relações. **Perspectivas em Ciência da Informação**, v. 1, n. 1, p. 41–62, jun. 1996.

SCHALLEY, A. C. Ontologies and ontological methods in linguistics. **Language and Linguistics Compass**, v. 13, n. 11, p. e12356, 2019.

SHU, W.; QIAN, W.; XIE, Y. Knowledge Acquisition Approach Based on Incremental Objects From Data With Missing Values. **IEEE Access**, v. 7, p. 54863–54878, 2019.

SILVA, D. L. da; SOUZA, R. R.; ALMEIDA, M. B. Ontologias e vocabulários controlados: comparação de metodologias para construção. **Ciência da Informação**, v. 37, n. 3, p. 60–75, dez. 2008.

SMITH, B.; MULLIGAN, K. Framework for Formal Ontology. **Topoi**, v. 2, n. 1, p. 73–85, jun. 1983.

SOERGEL, D. Functions of a thesaurus/classification/ontological knowledge base. 1997.

SUÁREZ-FIGUEROA, M. C. **NeOn Methodology for building ontology networks: specification, scheduling and reuse**. 2010. Informatica, 2010.

SUÁREZ-FIGUEROA, M. C.; GÓMEZ-PÉREZ, A.; FERNÁNDEZ-LÓPEZ, M. The NeOn Methodology Framework: A Scenario-Based Methodology for Ontology Development. **Applied Ontology**, v. 10, n. 2, p. 107–145, 15 set. 2015.

TEIXEIRA, L. M. D. **Princípios Ontológicos Aplicados à Classificação Internacional de Doenças: Alternativas para a busca por interoperabilidade semântica entre sistemas de prontuários eletrônicos de pacientes**. 2019. Universidade Federal de Minas Gerais, Belo Horizonte, 2019.

TEIXEIRA, L. M. D.; ÁVILA, I. M. A. A semântica nos sistemas de informação de energia: demanda regulatória. ENANCIB. **Proceedings XX ENANCIB**, set. 2019.

TRAYTEL, D. Formal Languages, Formally and Coinductively. **Logical Methods in Computer Science ; Volume 13**, p. Issue 3; 18605974, 2017.

USCHOLD, M.; GRUNINGER, M. Ontologies: Principles, Methods and Applications. **Knowledge Engineering Review**, v. 11, p. 93–136, 1996.

USCHOLD, M.; KING, M. Towards a Methodology for Building Ontologies. **Workshop on Basic Ontological Issues in Knowledge Sharing**, p. 15, 1995.

VICKERY, B. C. Ontologies. **Journal of information science**, v. 23, n. 4, p. 277–286, 1997.

WANG, Y. et al. Knowledge Elicitation Plug-In for Protégé: Card Sorting and Laddering. **The Semantic Web – ASWC 2006**, v. 4185, p. 552–565, 2006.

WARREN, P. **Human Reasoning and Description Logics: Applying Psychological Theory to Understand and Improve the Usability of Description Logics**. 2017. The Open University, 2017.

WOLF, R.; DELUGACH, H. S. Knowledge Acquisition via Tracked Repertory Grids. **Computer Science Dept. Univ. Alabama in Huntsville**, p. 1–19, 31 jul. 1996.

YEH, I. et al. Knowledge acquisition, consistency checking and concurrency control for Gene Ontology (GO). **Computer Applications in the Biosciences - CABIOS**, v. 19, n. 2, p. 241–248, 2003.

ZAHARUDIN, I.; NOAH, S. A.; NOOR, M. M. Knowledge acquisition from textual documents for the construction of medicinal herbs domain ontology. **Journal of Applied Sciences (Faisalabad, Pakistan)**, v. 9, n. 4, p. 794–798, 2009.