

Guilherme Francis de Noronha

**Um Método Para Transcrição de Termos
Biomédicos e um Protótipo de Ferramenta
Didática para Ensino de Lógica em Ontologia
Aplicada**

Belo Horizonte/MG, Brasil

Junho de 2017

Guilherme Francis de Noronha

**Um Método Para Transcrição de Termos Biomédicos e
um Protótipo de Ferramenta Didática para Ensino de
Lógica em Ontologia Aplicada**

Dissertação de Mestrado apresentada à coordenação do PPGGOC/UFMG com o objetivo de obtenção de título de Mestre em Gestão e Organização do Conhecimento.

Universidade Federal de Minas Gerais – UFMG

Escola de Ciência da Informação

Programa de Pós-Graduação em Gestão e Organização do Conhecimento

Orientador: Maurício Barcellos Almeida

Belo Horizonte/MG, Brasil

Junho de 2017

N852m

Noronha, Guilherme Francis de.

Um método para transcrição de termos biomédicos e um protótipo de ferramenta didática para ensino de lógica em ontologia aplicada [manuscrito] / Guilherme Francis de Noronha.– 2017.

145 f., enc. : il.

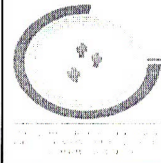
Orientador: Maurício Barcellos Almeida.

Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

Referências: f. 138-145.

1. Ciência da informação – Teses. 2. Ontologias (Recuperação da informação) – Teses. 3. Inferências (Lógica) – Teses. I. Título. II. Almeida, Maurício Barcellos. III. Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

CDU: 025.4.03



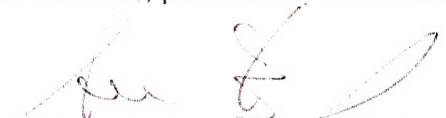
FOLHA DE APROVAÇÃO


Um Método Para Transcrição de Termos Biomédicos e um Protótipo de Ferramenta Didática para Ensino de Lógica em Ontologia Aplicada


GUILHERME FRANCIS DE NORONHA


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 REPRESENTAÇÃO DO CONHECIMENTO, linha de pesquisa Gestão e Tecnologia.

Aprovada em 22 de junho de 2017, pela banca constituída pelos membros:


Prof(a). Mauricio Barcellos Almeida - Orientador
ECI/UFMG


Prof(a). Renato Rocha Souza
FGV/RJ


Prof(a). Manoel Palhares Moreira
PUC/MG


Prof(a). Christiano Pereira Pessanha
UNILESTE


Prof(a). Renata Maria Abrantes Baracho Porto
ECI/UFMG

Belo Horizonte, 22 de junho de 2017.



ATA DA DEFESA DA DISSERTAÇÃO DO ALUNO GUILHERME FRANCIS DE NORONHA

Realizou-se, no dia 22 de junho de 2017, às 14:00 horas, Sala 1000 ECI/UFMG, da Universidade Federal de Minas Gerais, a defesa de dissertação, intitulada *Um Método Para Transcrição de Termos Biomédicos e um Protótipo de Ferramenta Didática para Ensino de Lógica em Ontologia Aplicada*, apresentada por GUILHERME FRANCIS DE NORONHA, número de registro 2015666421, graduado no curso de CIÊNCIA DA COMPUTAÇÃO, 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 - Orientador (ECI/UFMG), Prof(a). Renato Rocha Souza (FGV/RJ), Prof(a). Manoel Palhares Moreira (PUC/MG), Prof(a). Christiano Pereira Pessanha (UNILESTE), Prof(a). Renata Maria Abrantes Baracho Porto (ECI/UFMG).

A Comissão considerou a dissertação:

Aprovada

Reprovada

Finalizados os trabalhos, lavrei a presente ata que, lida e aprovada, vai assinada por mim e pelos membros da Comissão.
Belo Horizonte, 22 de junho de 2017.

Prof(a). Mauricio Barcellos Almeida

Prof(a). Renato Rocha Souza

Prof(a). Manoel Palhares Moreira

Prof(a). Christiano Pereira Pessanha

Prof(a). Renata Maria Abrantes Baracho Porto

Aos meus pais, pois em nenhuma universidade do mundo eu aprenderia o que eles me ensinaram durante toda a vida. Muito obrigado.

Agradecimentos

Gostaria de agradecer primeiramente ao professor Maurício pela oportunidade dada, por acreditar em mim, pela paciência e pelos ensinamentos. Espero que este trabalho retribua uma parte de toda a confiança.

Ao professor Palhares que me acompanhou durante toda a minha graduação, que despertou em mim o gosto pela Ciência da Informação e suas áreas. Por ter acreditado em mim e me ajudado quando precisei na minha vida acadêmica.

À banca da qualificação e da defesa pela sua presença, por seu tempo gasto para ler e avaliar este trabalho. Pelas críticas sempre construtivas e pelas dicas valiosas que tanto agregaram ao trabalho e a minha pessoa.

Às meninas da secretaria do PPG-GOC e também do PPG-CI pelos serviços prestados, pela atenção, cordialidade e disponibilidade em resolver meus "pepinos". Adiantaram minha vida e sobrou mais tempo para me preocupar com o que realmente importava, este trabalho.

Ao médico Eduardo Ferreira Lobato por me acompanhar e auxiliar, ajudando-me a recobrar das enfermidades que venho sofrido. Pelas sugestões dadas e pelo empenho demonstrado em mim. Parte desta vitória também é sua.

E ao grande amor da minha vida, Natanna Carvalho. Obrigado pelo carinho, paciência, amizade, companhia e até mesmo pelas críticas e cobranças. Sem você essa jornada sequer teria terminado. Muito obrigado mesmo, de todo o meu coração.

"Algumas regras para a vida:

- Questione a autoridade.

- Nenhuma ideia é verdadeira só porque alguém diz que é, incluindo eu.

- Pense por si próprio, questione a si próprio.

- Não acredite em algo só porque quer acreditar, acreditar em algo não o torna verdadeiro.

- Teste ideias pelas evidências adquiridas, pela observação e experimentação.

- Se uma ideia prevalecente falhar num teste bem desenvolvido, está errada. Supere.

- Siga as evidências onde quer que elas levem, se não houver evidências, evite julgamento.

E talvez a regra mais importante de todas:

- Lembre-se, você pode estar errado."

(Neil deGrasse Tyson)

"Ausência de evidência não é evidência de ausência, alegações extraordinárias exigem evidências extraordinárias."

(Carl Sagan)

Resumo

O campo de ontologia aplicada está em ascensão devido à modernização dos sistemas de informação e à necessidade de atribuição semântica tanto de domínios especializados quanto da web. É então feito um grande esforço para a formalização do conteúdo para que sejam passíveis de leitura por máquinas computacionais e que informação implícita se transforme em informação explícita. A construção de ontologias envolve um esforço que exige interdisciplinaridade dos envolvidos por tratar de assunto especializado. Além do conhecimento de ontologia aplicada, o profissional da ciência da informação deve ter noções de sistemas de informação e tecnologia, lógica descritiva e apoio de um especialista do domínio. Com o suporte destas três bases, se torna possível construir uma ontologia bem fundamentada que atende aos requisitos da representação da informação. Contudo, notou-se que profissionais da ciência da informação nem sempre cursam disciplinas que envolvam a parte lógica da construção de ontologias. A lógica pertence ao núcleo da construção de uma ontologia, mas ela não é usada de maneira adequada para aproveitar seu potencial. Este trabalho é uma tentativa de esclarecer o uso da lógica em ontologias. Para isto, foi criado um método de transcrição de definições formais para lógica descritiva utilizada em ontologias. A validação deste método foi testada com a transcrição de definições formais da Blood Ontology e o auxílio de motores de inferências para ontologias. Além disso, propôs-se a criação de um sistema online (nomeado como OntoLogica) voltado à didática do ensino da lógica para sistemas de ontologia aplicada. Este sistema, que foi prototipado, auxiliará aos usuários na criação e transcrição de definições para a lógica utilizando-se do método proposto neste trabalho. Os resultados das transcrições geraram inferências corretas e outras que foram consideradas equivocadas. Conclui-se que as transcrições são importantes por deduzirem conhecimento novo, mas que o campo precisa de cuidados para definição de definições biomédicos e também para sua transcrição.

Palavras-chave: ontologia aplicada, lógica descritiva, inferências automáticas, editores de ontologia.

Abstract

The field of applied ontology is increasing due to modernization of information systems and need of its semantization, that includes from specialist domains to the web. So is done a great effort to formalize the contents in a way that allow it be processed by computers and transform explicit information into implicit one. Nevertheless, the building process of ontologies needs an effort that requires interdisciplinarity of the responsible staff, because it involves specialist knowledge. In addition to the knowledge of applied ontology, the science information professional needs to count on knowledge from information systems and technology, knowledge from descriptive logic and support from a specialist in the realm to be built. Only with the support of these three bases will be possible to build an ontology well-grounded which attends all requisites of information representation. However, is realized that science information professionals not always do courses that teach the logical content that involves ontologies construction. The logic belongs to the core of ontology theory, but it's not used in a suitable way to improve its potential. Therefore this work is an initiative to elucidate the use of logic in ontology based systems. To reach that, it was created a method of transcription of formal definitions based in description logic used by ontologies. The validation of the method was tested with the transcriptions of some definitions of Blood Ontology and checked with help of reasoners. Furthermore, it was proposed an online system (called OntoLogica) to help the teaching of logic in applied ontology systems. OntoLogica, which was prototyped, will help its users to create and transcript formal definitions to logic using the method proposed in this work. The results of the transcriptions made generated correct and also incorrect inferences. The conclusion of this work was the transcriptions are important because they acquire new knowledge, but the field still needs cautions to define and transcript formal definitions.

Keywords: applied ontology, description logic, automated inferences, ontology editors.

Lista de ilustrações

Figura 1 – Arquitetura da representação do conhecimento baseado em lógica descritiva.	41
Figura 2 – Representação gráfica de classes, indivíduos e propriedades	51
Figura 3 – Representação gráfica dos recursos abrangidos por cada sub-linguagem da OWL	52
Figura 4 – Estrutura da OWL 2	53
Figura 5 – Arquitetura do Pellet	62
Figura 6 – Protégé 5.0	68
Figura 7 – NeOn Toolkit	69
Figura 8 – WebODE	70
Figura 9 – Jena API no Eclipse	71
Figura 10 – Ontologias BLO	73
Figura 11 – Fluxograma para a transcrição de definições em uma ontologia	78
Figura 12 – Definição de <i>disorder of hemostasis</i> no Protégé	87
Figura 13 – Atribuição de condições para <i>blood disorder</i> no Protégé	87
Figura 14 – indivíduos para <i>disorder of hemostasis</i> no Protégé	88
Figura 15 – Inconsistências no Protégé	89
Figura 16 – Dedução de conhecimento novo no Protégé	90
Figura 17 – Definição de <i>bleeding disorder associated with acquired disorders of coagulation</i> no Protégé	95
Figura 18 – Definição de <i>bleeding disorder after extracorporeal circulation</i> no Protégé	100
Figura 19 – Delimitadores de alcance para a propriedade <i>transfusing</i>	103
Figura 20 – Definição de <i>bleeding disorder after massive transfusion</i> no Protégé . .	105
Figura 21 – Definição de <i>bleeding disorder associated with accelerated destruction of coagulation factors</i> no Protégé	109
Figura 22 – Definição de <i>Bleeding disorder caused by fibrinolysis</i> no Protégé	112
Figura 23 – Indivíduos para a classe <i>bleeding disorder associated with acquired disorders of coagulation</i>	113
Figura 24 – Indivíduos para a classe <i>bleeding disorder after extracorporeal circulation</i>	114
Figura 25 – Indivíduos para a classe <i>bleeding disorder after massive transfusion</i> . .	115
Figura 26 – Indivíduos para a classe <i>bleeding disorder associated with accelerated destruction of coagulation factors</i>	116
Figura 27 – Indivíduos para a classe <i>bleeding disorder caused by fibrinolysis</i>	116
Figura 28 – Validação de consistência das classes no Protégé	117

Figura 29 –Inferências sobre a classe <i>bleeding disorder after extracorporeal circulation</i> no Protégé	117
Figura 30 –Inferências sobre a classe <i>bleeding disorder after massive transfusion</i> no Protégé	118
Figura 31 –Inferências realizadas pelo Protégé baseados nas definições de BFO	118
Figura 32 –Inferência realizada pelo Protégé sobre o indivíduo <i>bdcf2</i>	119
Figura 33 –Inferência realizada pelo Protégé sobre os indivíduos <i>bc1</i> , <i>bc3</i> , <i>wb1</i> e <i>wb2</i>	119
Figura 34 –Inferência realizada pelo Protégé sobre os indivíduos <i>bdamtind1</i> e <i>bdamtind2</i>	120
Figura 35 –Modelo da página com a licença GNU GPLv3 do OntoLogica	122
Figura 36 –Modelo da página inicial do OntoLogica	123
Figura 37 –Modelo do perfil do usuário do OntoLogica	123
Figura 38 –Modelo do processo de recuperação de senha do OntoLogica	124
Figura 39 –Modelo da página de termos de serviço do OntoLogica	124
Figura 40 –Modelo da página de políticas de privacidade do OntoLogica	125
Figura 41 –Modelo da aba de edição de classes do OntoLogica	127
Figura 42 –Modelo de exclusão de classes do OntoLogica	127
Figura 43 –Modelo de busca de classes do OntoLogica	128
Figura 44 –Modelo de funcionamento do botão <i>Add New Annotation</i> do OntoLogica	128
Figura 45 –Modelo da aba de edição de propriedades do OntoLogica	129
Figura 46 –Modelo da aba de edição de indivíduos do OntoLogica	130
Figura 47 –Modelo da aba de inferências do OntoLogica	131
Figura 48 –Modelo da aba de visualização de classes OntoLogica	131
Figura 49 –Modelo da aba de transcrição de termos do OntoLogica	132

Lista de tabelas

Tabela 1	– Lógicas formais, representações e conclusões.	28
Tabela 2	– Operadores da lógica proposicional	29
Tabela 3	– Operadores da lógica descritiva	36
Tabela 4	– Sintaxe da Lógica Descritiva \mathcal{EL}^{++}	38
Tabela 5	– Regras de inferência das lógicas proposicional e de primeira ordem . . .	40
Tabela 6	– Tradução das sintaxes da OWL para lógica descritiva	55
Tabela 7	– Características dos motores de inferências	60
Tabela 8	– Características dos editores de ontologias	66
Tabela 9	– Etapas da metodologia de pesquisa	75
Tabela 10	– Operadores da lógica descritiva e as relações ontológicas	82
Tabela 11	– Tabela-verdade para os termos $A \leftrightarrow B$ e $(A \wedge \neg B) \vee (B \wedge \neg A)$	84
Tabela 12	– Condições necessárias e suficientes e suas relações ontológicas	84
Tabela 13	– Características da GNU GPLv3	121

Lista de abreviaturas e siglas

ABox	Assertion Box
AIFB	Institute of Applied Informatics and Formal Description Methods
API	Application program interface
BACK	Berlin Advanced Computational Knowledge Representation System
BLO	Blood Ontology
CLIPS	C Language Integrated Production System
CLO	Cell Line Ontology
CoE	Center of Excellence in Bioinformatics
CORBA	Common Object Request Broker Architecture
DAML	DARPA Agent Markup Language
DBMS	Database Management System
ECI	Escola de Ciência da Informação
FaCT	Fast Classification of Terminologies
FZI	Forschungszentrum Informatik
FLogic	Frame Logic
FMA	Foundational Model of Anatomy Ontology
GPU	General Public License
HEMONTO	Hemocomponents and Hemoderivatives Ontology
HL7	Health Level Seven
HP	Hewlett-Packard
HPO	Human Phenotype Ontology
HTML	Hyper Text Markup Language
ICPC2P	International Classification of Primary Care - 2 PLUS

IDE	Integrated development environment
IRI	Internationalized Resource Identifier
IRS	Internet Reasoning Service
KRIS	Knowledge Representation and Inference System
MEDDRA	Medical Dictionary for Regulatory Activities
MESH	Medical Subject Headings
MMI	Marine Metadata Interoperability
NCI	National Cancer Institute
NCIT	National Cancer Institute Thesaurus
NRQL	New Relic Query Language
OIL	Ontology Inference Layer
OKBC	Open Knowledge Base Connectivity
OBI	Ontology for Biomedical Informations
OWL	Web Ontology Language
PHP	Hypertext Preprocessor
Prolog	Programming Logic
RACER	Renamed ABoxes and Concept Expression Reasoner
RCD	Read Codes
RDF	Resource Description Framework
RDFa	Resource Description Framework in Attributes
RDFS	Resource Description Framework Schema
RO	Relations Ontology
SNOMED CT	Systematized Nomenclature of Medicine - Clinical Terms
SWRL	Semantic Web Rule Language
SWWS	Semantic Web and Web Service technologies
TBox	Terminological Box

Turtle	Terse RDF Triple Language
URI	Uniform Resource Identifier
W3C	World Wide Web Consortium
WSMO	Web Service Modelling Ontology
WSMX	Web Service Modelling eXecution environment
XML	Extensible Markup Language

Lista de símbolos

\sim	Negação
\neg	Negação
\cup	União de conjunto
\cap	Interseção de conjunto
\vee	"Ou" lógico
\wedge	"E" lógico
\rightarrow	Implicação ou condicional lógica
\Rightarrow	Implicação forte ou condicional lógica
\leftrightarrow	Bicondicional ou equivalência lógica
\Leftrightarrow	Bicondicional forte ou equivalência lógica
\forall	Para todo
\exists	Existe pelo menos um
$\exists!$	Existe apenas um
\top	Verdade/Tautologia
\perp	Falso/Contradição
\subseteq	É subconjunto de
\leq	É menor ou igual que
\geq	É maior ou igual que
\sqsubseteq	Inclusão
\doteq	Equivalência lógica
$:$	Tal que
Δ	Não infinitesimal
\emptyset	Conjunto vazio

\in	Pertente à
\times	Produto vetorial
\circ	Composição
\equiv	Equivalência

Sumário

1	Introdução	23
2	Lógica para ontologias: conceitos, inferências e simbolização	27
2.1	Lógica proposicional e o conhecimento básico para ler sentenças lógicas	28
2.2	A linguagem natural e sua simbolização	29
2.2.1	Termos	30
2.2.2	Predicados	30
2.2.3	Predicado de n variáveis	31
2.2.4	Quantificadores	32
2.2.5	Variáveis livres e variáveis ligadas	33
2.2.6	Exemplo Final	33
2.3	Lógicas Descritivas	34
2.4	Inferências Lógicas	39
2.4.1	<i>TBox</i> e <i>ABox</i>	41
2.4.1.1	<i>TBox</i>	42
2.4.1.2	<i>ABox</i>	42
2.4.2	Inferências na lógica descritiva	43
2.4.3	Método dos <i>tableaux</i> analíticos	44
3	Ontologias no meio computacional: conceitos, linguagens e ferramentas	47
3.1	Ontologia e Tecnologia	47
3.2	<i>Web Ontology Language – OWL</i>	50
3.3	A sintática e a semântica da OWL	54
3.3.1	Sintaxe RDF/XML	56
3.3.2	Sintaxe <i>Turtle</i>	57
3.3.3	Sintaxe OWL/XML	57
3.3.4	Sintaxe Manchester OWL	58
3.4	Motores de inferência	59
3.4.1	Pellet	61
3.4.2	RACER	61
3.4.3	KAON2	62
3.4.4	CRACK	63
3.4.5	KRIS	63
3.4.6	HermiT	64
3.4.7	Fact++	64
3.5	Editores de ontologias para a linguagem OWL	65

3.5.1	Protégé	67
3.5.2	NeOn Toolkit	68
3.5.3	WebODE	69
3.5.4	Jena	70
4	Metodologia	72
4.1	A ontologia objeto da pesquisa	73
4.2	Descrição da metodologia de pesquisa	74
4.3	Prototipação da Plataforma Online	75
4.3.1	Análise dos principais editores de ontologias	76
4.3.2	Levantamento de Requisitos	76
4.3.3	Prototipação	77
4.4	Coleta de dados	77
4.5	Transcrição de definições	77
4.5.1	Anotação da definição textual	77
4.5.1.1	Obter a versão final da definição em linguagem natural	78
4.5.1.2	Associar os substantivos com termos já existentes em ontologias	80
4.5.1.3	Criar um novo termo	81
4.5.1.4	Associar as relações com termos já existentes em ontologias e/ou lógica descritiva	81
4.5.1.5	Criar uma nova propriedade	83
4.5.1.6	Identificar a existência de condições necessárias e suficientes para os termos	83
4.5.2	Definição formal em OWL	85
4.5.2.1	Reescrevendo a definição no editor de ontologia	85
4.6	Avaliação	87
4.6.1	Inserção de indivíduos	87
4.6.2	Verificação de inconsistência	88
4.6.3	Avaliação de inferências	89
5	Análise e Resultados	91
5.1	Transcrição e avaliação dos termos	91
5.1.1	<i>Bleeding disorder associated with acquired disorders of coagulation</i>	91
5.1.1.1	Identificando os substantivos para <i>bleeding disorder associated with acquired disorders of coagulation</i>	92
5.1.1.2	Criando novos termos para <i>bleeding disorder associated with acquired disorders of coagulation</i>	93
5.1.1.3	Identificando as relações do termo <i>bleeding disorder associated with acquired disorders of coagulation</i>	93

5.1.1.4	Criando novas propriedades para o termo <i>bleeding disorder associated with acquired disorders of coagulation</i>	94
5.1.1.5	Identificando as condições necessárias e suficientes para o termo <i>bleeding disorder associated with acquired disorders of coagulation</i>	94
5.1.1.6	Reescrevendo o termo <i>bleeding disorder associated with acquired disorders of coagulation</i> no editor de ontologias	94
5.1.2	<i>Bleeding disorder after extracorporeal circulation</i>	95
5.1.2.1	Identificando os substantivos para <i>bleeding disorder after extracorporeal circulation</i>	96
5.1.2.2	Criando novos termos para <i>Bleeding disorder after extracorporeal circulation</i>	97
5.1.2.3	Identificando relações para o termo <i>Bleeding disorder after extracorporeal circulation</i>	97
5.1.2.4	Criando propriedades para o termo <i>bleeding disorder after extracorporeal circulation</i>	98
5.1.2.5	Identificando relações necessárias e suficientes para o termo <i>bleeding disorder after extracorporeal circulation</i>	98
5.1.2.6	Reescrevendo o termo <i>bleeding disorder after extracorporeal circulation</i> no editor de ontologias	99
5.1.3	<i>Bleeding disorder after massive transfusion</i>	100
5.1.3.1	Identificando os substantivos para <i>Bleeding disorder after massive transfusion</i>	100
5.1.3.2	Criando novos termos para <i>bleeding disorder after massive transfusion</i>	101
5.1.3.3	Identificando relações para o termo <i>Bleeding disorder after massive transfusion</i>	102
5.1.3.4	Criando propriedades para o termo <i>bleeding disorder after massive transfusion</i>	102
5.1.3.5	Identificando relações necessárias e suficientes para o termo <i>bleeding disorder after massive transfusion</i>	103
5.1.3.6	Reescrevendo o termo <i>bleeding disorder after massive transfusion</i> no editor de ontologias	104
5.1.4	<i>Bleeding Disorder Associated With Accelerated Destruction of Coagulation Factors</i>	105
5.1.4.1	Identificando os substantivos para <i>Bleeding Disorder Associated With Accelerated Destruction of Coagulation Factors</i>	105
5.1.4.2	Criando novos termos para <i>bleeding disorder associated with accelerated destruction of coagulation factors</i>	106

5.1.4.3	Identificando relações para o termo <i>bleeding disorder associated with accelerated destruction of coagulation factors</i>	106
5.1.4.4	Criando propriedades para o termo <i>bleeding disorder associated with accelerated destruction of coagulation factors</i>	107
5.1.4.5	Identificando relações necessárias e suficientes para o termo <i>bleeding disorder associated with accelerated destruction of coagulation factors</i>	107
5.1.4.6	Reescrevendo o termo <i>bleeding disorder associated with accelerated destruction of coagulation factors</i> no editor de ontologias	108
5.1.5	<i>Bleeding disorder caused by fibrinolysis</i>	109
5.1.5.1	Identificando os substantivos para <i>bleeding disorder caused by fibrinolysis</i>	109
5.1.5.2	Criando novos termos para <i>bleeding disorder caused by fibrinolysis</i>	110
5.1.5.3	Identificando relações para o termo <i>bleeding disorder caused by fibrinolysis</i>	110
5.1.5.4	Criando propriedades para o termo <i>bleeding disorder caused by fibrinolysis</i>	110
5.1.5.5	Identificando relações necessárias e suficientes para o termo <i>bleeding disorder caused by fibrinolysis</i>	110
5.1.5.6	Reescrevendo o termo <i>bleeding disorder caused by fibrinolysis</i> no editor de ontologias	111
5.1.6	Inserindo indivíduos para as classes transcritas	112
5.1.6.1	Indivíduos para <i>bleeding disorder associated with acquired disorders of coagulation</i>	112
5.1.6.2	Indivíduos para <i>bleeding disorder after extracorporeal circulation</i>	113
5.1.6.3	Indivíduos para <i>bleeding disorder after massive transfusion</i>	114
5.1.6.4	Indivíduos para <i>bleeding disorder associated with accelerated destruction of coagulation factors</i>	115
5.1.6.5	Indivíduos para <i>bleeding disorder caused by fibrinolysis</i>	116
5.1.7	Validação e raciocínio lógico dos termos segundo os motores de inferências	116
5.1.7.1	Validação de consistência sobre as classes modificadas	117
5.1.7.2	Inferências realizadas sobre as classes modificadas	117
5.1.7.3	Inferências realizadas sobre os indivíduos adicionados	118
5.2	OntoLogica: a plataforma online	120
5.2.1	Disponibilidade	121

5.2.2	Usuário	122
5.2.3	Privacidade e Segurança	123
5.2.4	Navegação	125
5.2.5	Funcionalidades	125
5.2.5.1	Carregar e excluir de ontologias	126
5.2.5.2	Criando, editando e excluindo classes	126
5.2.5.3	Criando, editando e excluindo propriedades	128
5.2.5.4	Criando, editando e excluindo indivíduos	129
5.2.5.5	Realizando inferências	130
5.2.5.6	Recursos de visualização	131
5.2.5.7	Instruções para a transcrição de termos	132
6	Conclusões e trabalhos futuros	133
6.1	Motivações iniciais	133
6.2	Conclusões sobre os resultados	134
6.3	Dificuldades encontradas	135
6.4	Trabalhos futuros	137
	Referências	138

1 Introdução

A teoria da representação do conhecimento dispõe de várias técnicas para a captura do conhecimento existente e a transformação do mesmo em conhecimento disponível e acessível. O conhecimento definido primeiramente por Platão¹ como sendo uma crença verdadeira e justificada, é uma indagação contínua e inerente do ser humano e traduz-se no interesse do mesmo acerca das atividades fenomênicas da realidade que o cerca. A teoria do conhecimento formalizada por Locke (1841)² em 1690, traz uma abordagem de um modelo sistemático quanto à origem, essência e a certeza do conhecimento humano. A relação de conhecimento do ser humano se traduz entre a relação do sujeito e do objeto. Hessen (1999) explica que a apreensão do objeto ocorre quando se defrontam a consciência do sujeito e o objeto. Estes objetos, que são classificados como reais³ e ideais⁴, exigem que o conhecimento de verdade também seja atrelado à teoria do conhecimento. O conhecimento passa então a depender do conceito da verdade⁵. O conhecimento lida com a correlação que o sujeito faz do objeto intencionado⁶. Já a verdade está relacionada com as possibilidades do conhecimento, ou seja, às crenças adotadas pelo sujeito que leva a acreditar ou não no objeto intencionado (TARSKI, 1956).

O conceito de verdade e conhecimento são abordagens constantes na teoria da representação do conhecimento, que procura sistemicamente a materialização deste conhecimento, de forma que se transformem em ações, produtos e serviços estratégicos (SOWA et al., 2000). A busca da verdade e do conhecimento traduz-se em diversificadas maneiras e técnicas de apreender o conhecimento gerado pela espécie humana. Entre estas técnicas utilizadas, as ontologias se destacaram como uma alternativa significativa para a representação deste conhecimento (ALMEIDA; BARBOSA, 2009).

As ontologias são modelos de representação do conhecimento dedicado a compreender um domínio e classificar termos com vocabulários controlados. Esses vocabulários são usados para gerar inferências em sistemas de informação. São sistemas de representação de domínios específicos para gerar inferências e produzir conhecimento novo. As inferências geradas dessas ontologias são produzidas através da semântica de estruturas definidas por Kripke (1963). Essas semânticas são úteis para que sejam definidos sistemas de derivação ou inferência do conhecimento obtido e é feita por especialistas que dominam e estudam a

¹ Platão — filósofo e matemático grego (428/427 a.C. - 348/347 a.C.)

² John Locke — filósofo inglês (1632 - 1704)

³ Os objetos reais são aqueles em que se conhecem através de experiências empíricas

⁴ Os objetos ideais são aqueles que são irrealis, meramente imaginados através da consciência humana

⁵ Verdade-como-correspondência: teoria para definição de verdade para uma linguagem que estabelece que, "a verdade é uma correspondência com a realidade" (PEREIRA, 2011)

⁶ A intencionalidade, é uma característica de todos os atos de consciência, fenômenos "psíquicos" ou "mental", pelo qual poderia ser separada dos fenômenos "físico" ou "natural" (BRENTANO, 2014).

área da lógica. Para simular este profissional que toma as decisões através de extenuantes estudos, a literatura oferece como solução a construção de sistemas especialistas.

Os sistemas especialistas têm como objetivo simular um profissional de uma área específica do qual obtém expertise (JACKSON, 1999). A construção destes sistemas é abordada por Hayes-Roth, Waterman e Lenat (1983) como uma necessidade advinda da evolução da tecnologia e das demandas geradas pelas empresas para resolver assuntos como o diagnóstico de doenças. São sistemas desenvolvidos para resolver problemas complexos baseados em conhecimento. Pertence a área de Inteligência Artificial e pode ser dividido em dois subsistemas: motor de inferência e base de conhecimento (BARR; FEIGENBAUM, 1981). A base de conhecimento é onde são armazenados os fatos sobre um determinado assunto. O motor de inferência são regras que serão aplicadas a estas bases de conhecimento e deduzir conhecimento novo.

O uso de ontologias em sistemas especialistas é citado por Middleton, Shadbolt e Roure (2004) como uma grande vantagem para esses sistemas. A possibilidade de inferências serem empregadas como uma alternativa de descoberta de novas ações/interesses dão as ontologias credibilidade para serem usadas como alternativa a outros métodos de representação da informação, tais como tesouros⁷ e taxonomias⁸. A construção de uma ontologia como um sistema especialista exige esforços de diversos lados. É necessário um cientista da informação com conhecimentos sólidos em construção de ontologias, um especialista no domínio do qual será representado por uma ontologia e um cientista da computação, capaz de transformar o conhecimento formalizado em lógica de computador. Tamanho esforço exige um trabalho interdisciplinar em sintonia com esses campos do saber que nem sempre é possível devido à especialização dos profissionais envolvidos.

A exemplo do trabalho de Aganette (2015) é demonstrado como a formalização de definições em uma ontologia é um processo custoso que exige uma metodologia criteriosa e bastante delicada. Quando o assunto envolve ontologias médicas, a abordagem requer ainda mais cuidado. O conhecimento necessário para a formalização de uma ontologia em sistemas de informação, nem sempre é contemplado por um profissional da ciência da informação ou da biblioteconomia. São exigidos conhecimentos de técnicas de programação e de lógica computacional para que a formalização obedeça a critérios lógicos a fim de evitar problemas como ambiguidade e inconsistência entre as definições formalizadas (SMITH, 2003).

Levando em consideração o uso de lógica para a derivação de conhecimento novo e o uso de ontologias como ferramental da lógica e estruturação formal deste conhecimento, o presente trabalho limitou-se a responder as seguintes perguntas: **"Como é feita a derivação de conhecimento novo em ontologias?"** e **"Como é feita a tradução de**

⁷ Tesouro: compilação do léxico de uma língua ou de uma área do saber.

⁸ Taxonomia: teoria ou nomenclatura das descrições e classificações científicas.

definições formais em ontologias biomédicas para lógica formal?". Tais perguntas são levantadas com o intuito de ajudar a vida do profissional de ciência da informação a lidar com o cotidiano de desenvolver ontologias formais. As dificuldades encontradas durante este desenvolvimento devem ser amenizadas com métodos factíveis e utilizáveis por qualquer profissional da ciência informação que necessite de criar e/ou manusear uma ontologia já existente.

O autor faz valer seu conhecimento em Ciência da Computação para contribuir de forma interdisciplinar com a Ciência da Informação. Por isto, o presente trabalho tem como objetivo geral:

- Elucidar o funcionamento de sistemas de inferência no campo da ontologia aplicada na ECI.

Espera-se que esse método ajude os profissionais da ciência informação que não possuem conhecimento necessário em lógica de computadores a construir e formalizar as ontologias com as quais trabalham. Como objetivos específicos o trabalho se propôs:

- A criação de um método, dirigido aos profissionais da ciência da informação para a transcrição de definições da linguagem natural para a lógica utilizada em ontologia aplicada, ou seja, passível de leitura por computadores;
- Transcrever as definições da *Blood Ontology* que foram definidos por outros pesquisadores utilizando-se o método proposto;
- Validar a transcrição utilizando-se motores de inferências estudados, e;
- A prototipação de uma plataforma online para transcrições de definições e testes lógicos.

Por motivos diversos, e talvez por falta de uma formação em lógica, esses pesquisadores não puderam realizar a transcrição destas definições para a lógica processada por computadores. A transcrição destas definições ontológicas possibilitou a verificação de sua consistência, pois com o uso de derivadores de conhecimento foram adicionadas novas instâncias para que sejam executadas verificações automáticas, demonstrando assim a sua utilidade. As definições formalizados foram validadas em uma lista contendo as plataformas mais utilizadas por desenvolvedores de ontologias. Um protótipo baseado nesta lista foi desenvolvido e descreve uma plataforma online onde usuários poderão carregar e editar suas ontologias para fazerem verificações e testes formais de validações lógicas. Estes testes contribuirão com a inclusão do desenvolvimento de ontologias computacionais em campos da Ciência da Informação para usuários menos experientes onde esta disciplina

se mostra ainda pouco acessível. Para que isto fosse possível, o presente trabalho adotou uma metodologia de pesquisa aplicada, qualitativa, exploratória e experimental.

O presente trabalho foi dividido da seguinte maneira: o Capítulo 2 introduz a lógica como campo da filosofia e aborda seu uso em sistemas computacionais. São apresentados desde os conceitos primários da lógica proposicional, até a lógica descritiva utilizada por sistemas semânticos e ontologia aplicada. Ainda neste capítulo é abordado sobre inferências na lógica, como elas são feitas, seus métodos e papel nas lógicas proposicional e descritiva. É explicado e relacionado como as inferências são importantes na ontologia aplicada, oferecendo inferências e dedução de fatos novos a um domínio do conhecimento. Por fim, discute-se a linguagem natural e como simbolizá-la, parte do processo teórico para a transcrição de definições ontológicas.

O Capítulo 3 trata do papel da ontologia no meio computacional. É apresentada a ontologia como artefato produzido e trabalhado em computadores e as tecnologias que foram desenvolvidas para que estes artefatos pudessem existir. É apresentada a OWL e suas variações, assim como as sintaxes disponíveis para seu desenvolvimento. São apresentados e comparados os principais motores de inferência para ontologias. São explicados a importância destes motores no papel do desenvolvimento de ontologia aplicada. Ainda neste capítulo é feita uma análise dos principais editores de ontologias do mercado com o intuito de levantar requisitos para o protótipo de editor de ontologias.

O Capítulo 4 descreve a metodologia usada na pesquisa. O Capítulo 5 transcreveu as definições e testou-os segundo motores de inferências analisados. Ainda neste capítulo é feito um protótipo da plataforma online que foi pensada para comportar o trabalho feito anteriormente. O protótipo é composto por esquetes de telas explicando a funcionalidade de cada uma. O Capítulo 6 discute as conclusões tiradas do capítulo anterior, lista as dificuldades encontradas durante a realização do trabalho e lista também os possíveis trabalhos futuros, ideias que surgiram diante das dificuldades encontradas.

2 Lógica para ontologias: conceitos, inferências e simbolização

A lógica teve origem na filosofia de Aristóteles (*Organon*¹), usada para abstrair as sentenças de uma linguagem (OWEN et al., 1901). A lógica aristotélica foi a base para a lógica moderna, a qual sofreu influência de outros pensadores até os dias de hoje. No século XI, o filósofo Avicena² fez uma contribuição significativa ao introduzir as noções de silogismo hipotético, temporalidade, modalidade e indução como alternativas à lógica aristotélica (AHMED, 2010). No século XVIII, Kant³ definiu lógica como julgamento científico, de forma que inferências válidas deveriam seguir julgamentos estruturados (KANT; YOUNG, 2004). No século XX, as descobertas de Kant foram rompidas por Frege (1991), Whitehead (1901) e Russell (2007) e reestruturada com a pesquisa em lógica simbólica e a matematização.

As sentenças declarativas usadas na lógica são estruturadas através de um conjunto de declarações avaliadas através de sistemas formais de dedução. Estes sistemas lidam com proposições (conjuntos de sentenças declarativas) capazes de expressar um valor de verdade com relação aos fatos que esta lógica representa (MORTARI, 2001). Whitehead e Russell (1912) descrevem o estudo da lógica como o estudo de métodos e princípios utilizados para distinguir o raciocínio correto. Além de sua aplicação na Filosofia, a lógica tem grande uso na Matemática e na Ciência da Computação.

Existem vários tipos de lógica, em geral organizadas em duas categorias principais: clássica e não-clássica (BALLARD; COPI, 1978). Cada um desses tipos de lógica envolve variações, gerando dezenas de diferentes tipos. Um tipo de lógica pode ser diferenciado pela sua capacidade de representação, ou seja, sua expressividade em relação às sentenças que descrevem os fatos. No mundo digital, quanto maior a expressividade da lógica, menor seu poder computacional.

A lógica em nível mental, citada na filosofia como linguagem natural, não é passível de ser computada, uma vez que os algoritmos para computadores são compostos por sequências lógicas formais. Dipert et al. (1994) fazem alusão sobre como programas de computador são representados em uma linguagem lógica: dados (vetores, variáveis, etc.) representam fatos do mundo e os processos (funções, chamadas, etc.) são sentenças lógicas.

A Tabela 1 mostra alguns tipos de lógicas, o que são capazes de representar e a conclusão epistemológica que podem abstrair a partir das representações.

¹ *Organon* é o nome atribuído a um conjunto de obras sobre lógica de Aristóteles

² Abū 'Alī al-usayn ibn 'Abd Allāh ibn Sīnā ou Avicena – polímata persa (980 – 1037)

³ Immanuel Kant – filósofo prussiano (1724 – 1804)

Tabela 1 – Lógicas formais, representações e conclusões.

Linguagem	Representação do Mundo	Conclusão Epistemológica
Lógica proposicional	Fatos	Verdadeiro e falso
Lógica de primeira ordem	Fatos, objetos e relações	Verdadeiro, falso e desconhecido
Lógica temporal	Fatos, objetos, relações e tempo	Verdadeiro, falso e desconhecido
Lógica difusa	Fatos com valor de confiança entre 0 e 1	Valor de confiança entre 0 e 1

Fonte: (RUSSELL; NORVIG, 2002)

2.1 Lógica proposicional e o conhecimento básico para ler sentenças lógicas

A lógica proposicional é também chamada de lógica sentencial, pois ela se concentra em apenas sentenças conectadas por operadores lógicos. Uma sentença em lógica é o conjunto de pelo menos um sujeito conectado por pelo menos um predicado. Essas conexões entre sujeito e predicado são feitas pelos operadores lógicos. O resultado desse conjunto possui um valor, podendo ser verdadeiro, falso ou desconhecido, segundo a Tabela 1. A lógica proposicional estuda meios de junção e/ou disjunção de sentenças, proposições e constatações de forma a derivar novas constatações, sentenças e proposições tal que novos relacionamentos são extraídos, dado um conjunto de sentenças de um domínio.

A lógica proposicional tem origem clássica com os filósofos estoicos e contribuições de Crisipo de Solis⁴ (MUELLER et al., 1979), mas somente no século XX ganhou força com a criação da lógica simbólica e contribuições de Boole (1854)⁵. Russell e Norvig (2002) definem a lógica proposicional como uma linguagem declarativa, pois ela lida com as semânticas baseadas numa relação de verdade⁶ entre as sentenças e os mundos possíveis⁷ de Kripke (1972)⁸. Além disso, a lógica proposicional possui a propriedade de composicionalidade, ou princípio de Frege⁹. Isto significa que se os termos de uma sentença forem removidos, o que restará são as regras da composição desta mesma sentença. Isto garante a computabilidade da lógica proposicional, uma vez que é garantida a estrutura semântica das sentenças.

⁴ Crisipo de Solis – filósofo grego (280 a.C. – 208 a.C.)

⁵ George Boole – matemático, educador, filósofo e lógico inglês (1815 – 1864)

⁶ A verdade lógica é uma declaração verdadeira que permanece verdadeira independentemente das diferentes interpretações sobre seus componentes.

⁷ Os mundos possíveis se referem às situações que não aconteceram, mas que poderiam ter acontecido (STALNAKER, 1976).

⁸ Saul Kripke — filósofo e lógico estadunidense (1940-)

⁹ Friedrich Ludwig Gottlob Frege – matemático, lógico e filósofo alemão (1848 – 1925)

De forma a entender como funciona a lógica, é preciso considerar certo nível de formalização. Apresentam-se aqui alguns conceitos básicos necessários para entender e ler sentenças em lógica. Primeiro, é preciso explicar certas definições da lógica, por exemplo, lógica de primeira ordem, predicados, quantificadores e operadores lógicos. As definições apresentadas aqui são baseadas principalmente em Gensler (2010).

Os operadores lógicos são essenciais para o entendimento da conexão de elementos em sentenças lógicas. Existe mais que um símbolo para representar estes operadores. A notação tradicional é representada na Tabela 2. Como mostrado na Tabela 1, a lógica proposicional representa fatos, sendo atribuído apenas a sentenças, não garantindo a relação entre estes fatos como na lógica de primeira ordem.

Tabela 2 – Operadores da lógica proposicional

Nome do Operador	Representação em Lógica Proposicional	Representação em Linguagem Natural	Exemplo
Negação	\sim ou \neg	Não	$\sim a$, $\neg a$
Conjunção	\cap , \wedge ou $\&$	E	$a \cap b$, $a \wedge b$, $a \& b$
Disjunção	\cup , \vee ou $ $	Ou	$a \cup b$, $a \vee b$, $a b$
Implicação ou Condicional	\rightarrow ou \Rightarrow	Se, então	$a \rightarrow b$, $a \Rightarrow b$
Bicondicional ou Equivalência	\leftrightarrow ou \Leftrightarrow	Se, e somente se	$a \leftrightarrow b$, $a \Leftrightarrow b$

De forma básica, qualquer lógica consiste de uma linguagem mais um sistema dedutivo. Em geral, a linguagem é um subconjunto da linguagem natural, como o português ou o inglês. O sistema dedutivo objetiva capturar as inferências corretas para certa linguagem. No entanto, a lógica proposicional constitui-se da forma mais simples de lógica a ser processada por computadores. Lógicas mais expressivas como a lógica descritiva se faz necessário quando o assunto é modelos mais complexos de representação do conhecimento.

2.2 A linguagem natural e sua simbolização

O uso de conectivos lógicos mostrados nas Tabela 2 e Tabela 3 somado ao uso de símbolos e parênteses da lógica simbólica, não são suficientes para expressar distinções lógicas mais simples e elementares. A lógica simbólica não é capaz de simbolizar substantivos próprios e comuns, pronomes, verbos ou adjetivos. Em geral, as variações gramaticais mais comuns não podem ser distinguidas pela lógica (SUPPES, 2012).

Para isto deve-se apresentar três conceitos correspondentes aos tipos de expressões utilizados em uma linguagem comum: termos, predicados e quantificadores. Deve-se entender como é possível traduzir sentenças em linguagem natural do dia a dia em uma linguagem consistindo apenas conectivos sentenciais, termos, predicados, quantificadores e parênteses. Mostra-se nas seções posteriores a comparação da lógica simbólica com a gramática da linguagem natural do dia a dia.

2.2.1 Termos

Um termo é o um objeto associado ao universo do discurso¹⁰. É um componente básico da proposição. Para Aristóteles, o termo é simplesmente algo que representa uma parte da proposição. Não pode ser verdadeiro ou falso, tem um significado neutro sendo apenas algo na realidade, como por exemplo, "homem" ou "mortal". Um termo pode ser um nome ou a descrição de um objeto (Ex: $a = b$), ou um nome ou descrição de um objeto que pode ser substituído em uma expressão (Ex: $x = b$ onde x pode ser uma variável qualquer). Existem três tipos de termos diferentes:

1. Variável: São símbolos como x , y , z ou letras com subscritos como x_1 , x_2 , x_3 , etc. A função gramatical das variáveis é similar aos dos pronomes e substantivos comuns da linguagem natural. Por exemplo: "Uma coisa é vermelha ou não vermelha" equivale a sentença "Para cada x , x é vermelho ou x não é vermelho", onde a variável x corresponde ao substantivo "coisa";
2. Nome próprio: usa as primeiras letras do alfabeto em caixa baixa para designar nomes próprios. Por exemplo, $a = \text{Isaac Newton}$ na sentença: " a é o maior matemático dos últimos três séculos", e;
3. Descrição precisa: "o maior matemático dos últimos três séculos" é uma descrição precisa que funciona como um nome próprio formando assim uma nova classe de termos. Aqui também usa-se as letras iniciais do alfabeto em caixa baixa para designar descrições. Por exemplo, $b = \text{"o maior matemático dos últimos três séculos"}$. Desta forma, podemos definir uma sentença formado por a é b , onde o verbo "é" é o que garante identidade à sentença e pode ser traduzido pelo símbolo $=$. Nomes de descrições de objetos são às vezes nomeados constantes, por exemplo: a , b e "Isaac Newton" são constantes.

2.2.2 Predicados

Um predicado segundo a lógica aristotélica, é uma sentença declarativa em que o valor pode ser verdadeiro ou falso, dependendo do valor de suas variáveis. Predicados são

¹⁰ O universo do discurso é uma ferramenta analítica usada na lógica dedutiva, especialmente na lógica de predicados. Indica o conjunto relevante de entidades as quais os quantificadores se referem.

utilizados para falar sobre propriedades de objetos, definindo os conjuntos que possuem certa propriedade em comum (Ex: O carro é azul é uma sentença declarativa que afirma que a cor do carro é azul).

Usa-se letras em caixa alta para designar os predicados. Por exemplo, a letra V substitui o predicado "é vermelho" na sentença: "Para cada x , $Vx \vee \neg Vx$ ". Na gramática tradicional o predicado é uma ou mais palavras em uma sentença ou cláusula que expressa o que é dito pelo sujeito (INFANTE; NETO, 2008). No português, o predicado deve ser formado com a presença de um verbo, ou de uma locução verbal. Já no inglês, o predicado pode ser formado de diferentes formas: apenas um verbo, um verbo junto a um advérbio ou uma cláusula adverbial que o modifique, a junção de verbo com um substantivo ou um adjetivo.

Na lógica, o predicado não é diferenciado de acordo com as partes gramaticais em que eles são construídos. Não existe lugar para advérbios, adjetivos e substantivos comuns. Apenas letras sozinhas para predicados são representados, e não há símbolos que represente advérbios e os demais componentes de um predicado. Predicados têm um amplo papel na lógica em seu uso comum. Por exemplo:

"Todo homem é um animal" será traduzido como: (2.1)

Para cada x , se x é um homem, então x é um animal (2.2)

Na gramática "é um animal" é um predicado da Equação 2.1. Sua tradução na Equação 2.2 tem o predicado adicional "é um homem" que substitui o substantivo comum "homem" na Equação 2.1. Utilizando a letra H para o predicado "é um homem" e a letra A para o predicado "é um animal", a Equação 2.1 pode ser transformada em:

Para cada x , $Hx \rightarrow Ax$ (2.3)

A Equação 2.3 exemplifica o padrão de sentenças lógicas.

2.2.3 Predicado de n variáveis

Um predicado de n variáveis são predicados que envolvem um ou mais sujeitos de uma sentença. Em geral, um predicado de n variáveis é um combinado de um ou mais termos que compõe uma sentença. Por exemplo:

Grúchenka estava muito feliz (2.4)

O senhor Dmitri era consideravelmente mais velho do que Grúchenka (2.5)

Transformando os sujeitos Grúchenka e Dmitri nas variáveis g e d respectivamente, e, usando as letras F e V para os predicados "estava muito feliz" e "era consideravelmente mais velho do que", a Equação 2.4 e Equação 2.5 são simbolizadas por Fg e Vdg . O predicado F é chamado de predicado de uma variável, enquanto o predicado V é chamado de predicado de duas variáveis. Outro exemplo:

Grúchenka estava muito feliz e bonita (2.6)

Na Equação 2.6, o predicado "muito feliz e bonita" é chamado na gramática de sentença composta. Utilizando a letra B para o predicado "estava bonita", a Equação 2.6 pode ser representada como:

$$Fg \wedge Bg \quad (2.7)$$

De outra maneira, a Equação 2.7 pode ser simbolizada transformando o predicado "estava feliz e bonita" na letra D e a sentença ficaria conforme demonstrado na Equação 2.8. A lógica não faz distinção entre sentenças simples ou compostas. A escolha de representação cabe de acordo com a conveniência do contexto.

$$Dg \quad (2.8)$$

2.2.4 Quantificadores

Os quantificadores mostrados na Tabela 3 são construções que especifica a quantidade de indivíduos dentro do universo do discurso e que se aplicam a (ou satisfazem) uma fórmula aberta. Uma fórmula aberta são fórmulas contendo variáveis que não são nem verdadeiras nem falsas. Por exemplo:

$$x \text{ ama } y \quad (2.9)$$

$$x + y = z + 2 \quad (2.10)$$

Se x e y são substituídos por "Grúchenka" e "Dmitri", o resultado da Equação 2.9 será verdadeiro dentro do universo do discurso de Dostoiévski¹¹. Se na Equação 2.10 as

¹¹ Fiódor Mikhailovitch Dostoiévski – filósofo e escritor russo (1821 – 1881)

letras x , y e z são substituídas pelos números 2, 3 e 4 respectivamente, a sentença obtida será falsa dentro do universo do discurso da aritmética.

Na lógica, não é preciso substituir as variáveis para atestar se uma sentença é verdadeira ou falsa. Outro método utilizado para essa verificação é criar prefixos para as sentenças, como por exemplo: "para cada x ", "Existe um x tal que", "para todo x ", etc. Estas frases são chamadas de quantificadores universais. Por exemplo: a sentença "para todo x , x é azul" pode ser escrito na Equação 2.11:

$$\forall(x)(x \rightarrow a) \quad (2.11)$$

Na gramática existem frases do tipo: "algum x ", "existe um x tal que", "há pelo menos um x ". Estas frases são chamadas de quantificadores existenciais. Desta forma, a fórmula " $x \geq 0$ " é uma sentença verdadeira quando o quantificador existencial "há um x tal que" é prefixado a ele. A Equação 2.12 demonstra o uso deste quantificador

$$\exists(x)(x \geq 0) \quad (2.12)$$

Desta forma pode-se dizer que uma fórmula sem os quantificadores não possuem necessariamente um valor de verdade. Mas ao adicionar estes quantificadores, é dito que existem valores tais quais uma sentença é satisfeita e dita como verdade, independentemente se os valores são revelados ou não.

2.2.5 Variáveis livres e variáveis ligadas

Uma ocorrência de uma variável em uma sentença pode ser feita de duas maneiras: variáveis que têm ocorrência controlada por um quantificador ou sem este controle. Por exemplo, a Equação 2.11 possui a variável controlada x que está conectada pelo quantificador universal \forall . A variável a no entanto, é uma variável livre, pois não depende de nenhum quantificador na sentença.

2.2.6 Exemplo Final

Considere a seguinte sentença escrita em linguagem natural (SUPPES, 2012, p.55):

*If one instant of time is after a second,
then there is an instant of time after the second and before the first* (2.13)

O melhor caminho para começar a tradução para lógica é introduzir as variáveis antes de traduzir a sentença. Na Equação 2.13 foram identificados três variáveis diferentes que são mostrados na Equação 2.14:

*If one instant of time x is after an instant of time y ,
then there is an instant of time z after that z is after y and z is before x* (2.14)

O próximo passo é reescrever a sentença utilizando o adjetivo descritivo *instant of time* pelo predicado correspondente. O resultado é uma sentença que se mostra estranha para a linguagem natural, mas que é mais fácil de ser entendido no ponto de vista lógico.

*If x is an instant of time, y is an instant of time, and x is after y ,
then there is a z such that z is an instant of time, z is after y , and z is before x .*
(2.15)

A Equação 2.15 é então simbolizada substituindo-se os predicados *time*, *is after*, *is before* por T , A e B respectivamente. O resultado é mostrado na Equação 2.16

$$Tx \wedge Ty \wedge Axy \rightarrow (\exists z)(Tz \wedge Azy \wedge Bzx) \quad (2.16)$$

O último passo é identificar as variáveis livres e as variáveis conectadas, adicionando os quantificadores universais. Conforme mostrado na Equação 2.16, as variáveis x e y estão livres, conseqüentemente sendo impossível saber seus valores de verdade. Adicionando-se os quantificadores para transformá-los em variáveis conectadas obtém-se a Equação 2.17.

$$\forall(x)(y)[Tx \wedge Ty \wedge Axy \rightarrow (\exists z)(Tz \wedge Azy \wedge Bzx)] \quad (2.17)$$

2.3 Lógicas Descritivas

A lógica de primeira ordem é um sistema de lógica simbólica que usa predicados e quantificadores. Para entender predicados, considere-se uma sentença contendo um sujeito mais um predicado: “o cão está latindo”. O sujeito da sentença é “o cão” e o predicado é “está latindo”. Então, um símbolo pode representar uma constante que é atribuída a um sujeito, tal como c para “cão”; e outro símbolo pode ser atribuído ao predicado, por exemplo, Lx para “ x está latindo”. Assim pode-se simbolizar a sentença “o cão está latindo” por Lc . Mais sobre predicados é falado na subseção 2.2.2.

Com relação aos quantificadores, consideram-se dois: o quantificador universal e o quantificador existencial. O primeiro é simbolizado pelo símbolo \forall , pode ser lido “para cada x ...” e significa “tudo”; o segundo é simbolizado pelo símbolo \exists , pode ser lido “existe pelo menos um y ...” e significa “algum(ns)”. Um exemplo de uso do quantificador universal

é a sentença formal $\forall x : (Px \rightarrow Vx)$, a qual pode ser lida, por exemplo, como “para cada x , se x é uma batata, então x é um vegetal” ou simplesmente “todas as batatas são vegetais”. Um exemplo de uso do quantificador existencial é a sentença formal $\exists x : (Px \wedge Cx)$, a qual pode ser lida, por exemplo, como “existe pelo menos um x tal que x é uma batata, e x é batata frita” ou simplesmente “algumas batatas são fritas”. Mais sobre quantificadores é falado na subseção 2.2.4.

Lógicas descritivas são subconjuntos da lógica de primeira ordem, sendo assim menos expressivas do que ela, mas mais expressivas que a lógica proposicional. A lógica de primeira ordem, também chamada de lógica de predicados, expressa um modelo genérico capaz de expressar a matemática, mas tal expressividade a torna ineficiente do ponto de vista computacional. Em contrapartida a lógica descritiva é mais restrita em termos de expressividade, mas garante o processamento computacional (BRACHMAN; LEVESQUE, 1984).

As lógicas descritivas são famílias da representação do conhecimento que podem ser usadas para representar conhecimento de um domínio de maneira bem estruturada (BAADER; HORROCKS; SATTLER, 2008). Elas permitem a realização de inferências automáticas eficientes. São representações do conhecimento que fornecem suporte para linguagens modernas de construção de ontologias como a OWL. Com elas é possível construir ferramentas para raciocínio automático (comumente conhecido como motores de inferência) capazes de realizar inferências, ou seja, gerar conhecimento novo a partir de conhecimento existente por via computacional.

Segundo Russell e Norvig (2002), a lógica descritiva consiste de notações adequadas para descrever definições e propriedades de categorias, ou seja, são projetadas para que seja mais fácil a expressão sobre coisas e objetos. Como um fragmento da lógica de primeira ordem que é, a lógica descritiva representa descrições formais de classes, indivíduos e as relações entre eles (ALLEMANG; HENDLER, 2011).

Como descrito na Tabela 1, a lógica descritiva é derivada da lógica de primeira ordem e trata também de fatos, objetos e relações.

Os sistemas que utilizam a lógica descritiva evoluíram a partir da *web* semântica como uma forma de formalizar o que redes semânticas representavam, além da necessidade de padronização e organização. As principais verificações realizadas pela lógica descritiva são a subordinação (verificar se uma categoria é um subconjunto de outra) e classificação (verificar se um objeto pertence a uma categoria). Alguns sistemas que se utilizam de lógica descritiva fazem verificações adicionais como, por exemplo, a consistência de uma categoria, satisfatibilidade¹², entre outros.

A lógica descritiva é importante para a pesquisa em ontologias computacionais

¹² Ver mais na subseção 2.4.2.

porque permite a construção e o uso de motores de inferência, as ferramentas que vão verificar a sua consistência. Além disso, a lógica descritiva possui uma sintaxe que facilita a construção de axiomas lógicos para representar um domínio de conhecimento. Além dos operadores da lógica proposicional descritos na Tabela 2, a lógica de primeira ordem ganha novos operadores que são adicionados para ganhar mais expressividade. Estes operadores são mostrados na Tabela 3.

Tabela 3 – Operadores da lógica descritiva

Nome do Operador	Representação em Lógica Descritiva	Representação em Linguagem Natural	Exemplo
Quantificação Universal	\forall	Para todo	$\forall x: P(x) \rightarrow Q(x)$
Quantificação Existencial	\exists	Existe pelo menos um	$\exists x: P(x) \rightarrow Q(x)$
Quantificação Única	$\exists!$	Existe exatamente um	$\exists! x: P(x) \rightarrow Q(x)$
Equivalência	\equiv	É equivalente a	$P \equiv Q$
Indivíduo ou Instância	\top	Indivíduo ou instância	$A \vee \neg A \equiv \top$
Conceito Vazio	\perp	Vazio	$A \wedge \neg A \equiv \perp$
Inclusão	\sqsubseteq	Está incluído em	$P \sqsubseteq Q$
Definição	\doteq	Está definido para ser igual a	$P \doteq Q$
Atribuição conceitual	:	É	$a : P$
Atribuição de papel	:	está relacionado por X à x	$(a,b) : P$

Fonte: (BAADER, 2003)

Através dos operadores listados na Tabela 2 e na Tabela 3 é possível derivar diferentes linguagens usadas pelos motores de inferência para a realização de inferências automáticas. Para definir a sintaxe que uma linguagem suporta, criou-se uma convenção que descreve os operadores permitidos (HARMELEN; LIFSCHITZ; PORTER, 2008). Abaixo estão listadas algumas dessas convenções (BAADER, 2003):

- \mathcal{AL} : Linguagem atributiva, que permite:
 - Negação atômica¹³;
 - Conjunção;
 - Quantificação Universal, e;

¹³ Uma fórmula atômica é uma fórmula básica sem o uso de conectivos lógicos. A negação atômica é a negação de uma fórmula atômica.

- Quantificação Existencial.
- \mathcal{C} : Negação Complexa¹⁴;
- \mathcal{S} : Uma abreviação de \mathcal{ALC} adicionando-se o silogismo hipotético;
- \mathcal{H} : Papel hierárquico;
- \mathcal{O} : Propriedades nominais (PITTS, 2003);
- \mathcal{I} : Propriedades inversas;
- \mathcal{N} : Restrições de cardinalidade;
- \mathcal{R} : Inclusão de axiomas; propriedades reflexivas e papéis disjuntos;
- \mathcal{Q} : Restrições de cardinalidade qualificadas (restrições que não se enquadram no conceitos de \top e \perp);
- (\mathcal{D}): Uso de tipos de dados, valores ou suas propriedades;
- \mathcal{EL} : Linguagem existencial, que permite:
 - Conjunção, e;
 - Restrições existenciais ou quantificação existencial completa.
- \mathcal{EL}^{++} : junção das linguagens \mathcal{ELRO} .

A combinação destas convenções formam linguagens descritivas que são utilizadas para a construção de motores de inferências e outras aplicações que exigem a expressividade para modelagem conceitual. Utilizando essas convenções, pode-se descrever a OWL como uma lógica descritiva do tipo $\mathcal{SHOIN}^{(D)}$. As linguagens mais utilizadas são \mathcal{SHIQ} , $\mathcal{SHOIN}^{(D)}$ (Utilizada pela OWL-DL), $\mathcal{SHIF}^{(D)}$ (Utilizada pela OWL-Lite e OWL 2 QL), $\mathcal{SROIQ}^{(D)}$ (Utilizada pela OWL 2) e \mathcal{EL}^{++} (Utilizada pela sub-linguagem OWL 2 EL) (BAADER; HORROCKS; SATTNER, 2009). A sintaxe das operações permitidas pela lógica descritiva \mathcal{EL}^{++} , usada por ontologias biomédicas em OWL 2 EL, é apresentada na Tabela 4.

¹⁴ Uma fórmula complexa é uma fórmula composta de fórmulas atômicas conectadas por conectivos lógicos. A negação complexa é a negação de uma fórmula complexa.

Tabela 4 – Sintaxe da Lógica Descritiva \mathcal{EL}^{++} .

Constructo	Sintaxe	Semântica
Instância	\top	Δ^x
Vazio	\perp	\emptyset
Propriedades nominais	$\{a\}$	$\{a^X\}$
Conjunção	$C \wedge D$	$C \cap D$
Restrição existencial	$\exists r.C$	$\{x \in \Delta^x \mid \exists y \in \Delta^x : (x, y) \in r^x \wedge y \in C^x\}$
Domínios concretos ¹⁵	$p(f_1, \dots, f_k)$ para $p \in P^{\mathcal{D}_j}$	$\{x \in \Delta^x \mid \exists y_1, \dots, y_k \in \Delta^{\mathcal{D}_j} : f_i^X(x) = y_i \text{ para } 1 \leq i \leq k \wedge (y_1, \dots, y_k) \in p^{\mathcal{D}_j}\}$
Inclusão	$C \sqsubseteq D$	$C \subseteq D$
Restrição	$r_1, \circ \dots \circ, r_k \sqsubseteq r$	$r_1^x, \circ \dots \circ, r_k^x \subseteq r^x$
Restrição de domínio	$\text{dom}(r) \sqsubseteq C$	$r^x \subseteq C^x \times \Delta^x$
Restrição de alcance	$\text{ran}(r) \sqsubseteq C$	$r^x \subseteq \Delta^x \times C^x$
Atribuição conceitual	$C(a)$	$a^x \in C^x$
Atribuição de papel	$r(a, b)$	$(a^x, b^x) \in r^x$

Fonte: (BAADER; BRANDT; LUTZ, 2008)

Baader, Brandt e Lutz (2008) ressalta que o constructo da inclusão mostrada na Tabela 4 é uma generalização de vários meios expressivos importantes para a ontologia aplicada. Eles são:

- Hierarquias como $r \sqsubseteq s$ são expressadas como $r \sqsubseteq s$
- Equivalências como $r \equiv s$ são expressadas como $r \sqsubseteq s, s \sqsubseteq r$
- Transições são expressadas como $r \circ r \sqsubseteq r$
- Reflexões são expressadas como $\mathcal{E}^{16} \sqsubseteq r$
- Elementos neutros¹⁷ de esquerda e direita são expressados como $r \circ s \sqsubseteq s$ e $r \circ s \sqsubseteq r$ respectivamente;

Além disso, a combinação da sintaxe \perp com as sintaxes de inclusão podem ser usadas pra expressar diferentes propriedades de uma linguagem. Elas são:

- Disjunções são expressadas como $C \wedge D \sqsubseteq \perp$
- Identidades de dois indivíduos são expressadas como $\{a\} \sqsubseteq \{b\}$
- Distinção de dois indivíduos são expressados como $\{a\} \wedge \{b\} \sqsubseteq \perp$

¹⁵ Conjuntos concretos como números reais, *strings*, etc. Contém também predicados para lidar com estes conjuntos tais como \geq e comparação de *strings*

¹⁶ Representação para inclusão.

¹⁷ Um elemento é dito neutro quando a sua operação com qualquer outra unidade resulta na unidade com a qual foi operado.

2.4 Inferências Lógicas

O conhecimento humano é produzido através de processos cognitivos que chegam à nossa mente por meios intuitivos ou através da razão ou raciocínio lógico. O raciocínio lógico exige que demonstrações e provas sejam apresentadas para que possa ser estabelecida determinada conclusão sobre um conjunto de fatos. As inferências são etapas do raciocínio lógico produzidas por premissas até se chegar a uma conclusão. Segundo Peirce (1974) existem na ciência três tipos de raciocínio lógico: dedução, indução e abdução.

A dedução é o processo que examina uma ou mais premissas e através das regras de dedução chega-se a um resultado lógico possível (verdadeiro, falso ou desconhecido) (CHENG; HOLYOAK, 1985). As premissas funcionam como uma ligação para as conclusões lógicas. Ou seja, se todas as premissas e regras de um dado conjunto forem verdadeiras, a sua conclusão necessariamente precisa ser verdadeira. Seu método é examinar o estado do sujeito atribuído nas premissas. Então forma-se um diagrama com o estado do sujeito e então percebe-se nas premissas constatações implícitas sobre o sujeito que não foram mencionadas. Testa-se as constatações notadas sempre se satisfazem em um dado cenário e então conclui-se a sua veracidade ou não. O método dedutivo tem uma abordagem de solução *top-down*, ou seja, aplica-se regras de dedução gerais às premissas dadas para ir estreitando as possibilidades, chegando-se a um único e possível resultado (LAKATOS; MARCONI, 2010).

Na lógica proposicional o uso da dedução traduziu-se em regras de inferências que passaram a fazer parte dos princípios da lógica. Verificadas através da tabela verdade¹⁸, conjuntos de regras de inferências foram criadas para derivar expressões lógicas complexas em expressões lógicas simples. A Tabela 5 mostra as principais regras relacionadas a lógica proposicional e também para a lógica descritiva que foram derivadas utilizando-se da tabela verdade.

¹⁸ Tabela verdade é um conjunto de linhas e colunas que relacionam uma ou mais variáveis e seus valores de verdade. É utilizada pelos lógicos para determinar se uma expressão é válida.

Tabela 5 – Regras de inferência das lógicas proposicional e de primeira ordem

Regra de Inferência	Operadores	Demonstração em linguagem natural
Modus ponens	$a \rightarrow b$	Se tenho a , logo tenho b
Modus tolens	$a \rightarrow b$	Se não tenho b , logo não tenho a
Silogismo disjuntivo	$a \vee b$	Se não tenho a , logo tenho b
Adição	a	Logo tenho $a \vee b$ ou tenho $b \vee a$
Simplificação	$a \wedge b$	Se tenho a e b , logo tenho a e tenho b
Conjunção	a e b	Tenho a e tenho b ; Logo tenho $a \wedge b$
Silogismo hipotético	$a \rightarrow b$; $b \rightarrow c$	Se tenho $a \rightarrow b$ e $b \rightarrow c$, logo tenho $a \rightarrow c$
Eliminação universal	$\forall x : P(x) \rightarrow Q(x)$	Logo $P(a) \rightarrow Q(a)$, onde a é uma constante
Eliminação existencial	$\exists x : P(x) \rightarrow Q(x)$	Logo $P(a) \rightarrow Q(a)$, onde a é uma constante

A indução é o oposto da dedução. É considerada como uma abordagem de solução *bottom-up*, ou seja, baseada em evidências. O método indutivo verifica se casos exploratórios dados obedecem uma regra natural e formula a partir daí soluções que são subordinadas à todos os casos dados. A indução também necessita que o raciocínio lógico seja empregado para determinar a validade do método indutivo. Caso contrário a indução feita será considerada falsa. Segundo Peirce (1974), a conclusão realizada na indução é aproximada da verdade, mas não necessariamente totalmente verdadeira. Isso ocorre porque a conclusão da indução deve ser verdade para todos os futuros casos possíveis, ainda que nem todos sejam examinados. O método indutivo é utilizado também para determinar o comportamento futuro de um evento dado as ocorrências passadas. Teorias da probabilidade como inferências bayesianas foram baseadas na indução lógica (HARTIGAN, 2012).

A abdução, ou hipotético, foge da linha do raciocínio e parte para a linha da intuição. É a busca de uma conclusão através da interpretação de fatos e sinais fornecidos dentro de um universo. Apesar de ser intuitivo, a conclusão não se dá imediatamente após o fornecimento dos fatos. É feita uma análise passo a passo até chegar a uma conclusão possível intuitiva. A abdução procura por explicações simples e direta a um acontecimento, ou seja, a mais plausível e coerente. Peirce (1974) diz que o método abdução é a forma inicial do raciocínio quando este se encontra com uma nova área de estudo ainda não estudada. Uma pessoa pode entender a abdução como "inferência através da melhor explicação", pois dado um conjunto de fatos F ; T como sendo uma teoria que explicaria F caso seja verdade; então não há outra explicação possível para F que não seja T ; logo F é verdade (JOSEPHSON; JOSEPHSON, 1996).

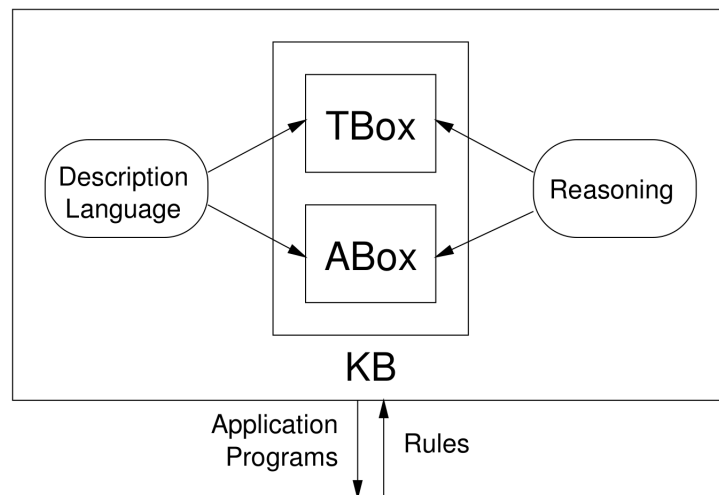
Inferências baseadas nos métodos de raciocínio apresentados permeiam a lógica e se traduzem em diversas aplicações utilizadas para a produção de conhecimento novo. Cada método de raciocínio possui seu tipos de derivações de conhecimento e regras de infe-

rências como silogismo hipotético, inferência causal, predição, etc. possuindo propriedades específicas. O entendimento e processamento das inferências dita o avanço da lógica em sistemas computacionais. A capacidade do processamento dessas regras em tempo hábil permite que sistemas baseados em lógica sejam construídos e sua aplicação seja feita.

2.4.1 *Tbox* e *Abox*

As propriedades da sintaxe de uma lógica descritiva permite a construção de modelos para inferência. A lógica descritiva é baseada em quatro aspectos fundamentais: o conjunto de constructos; as atribuições de conceitos *TBox* (componente terminológico); as atribuições de instâncias *ABox* (componentes atributivos) e; mecanismos de inferências que atuam tanto nas *TBoxes* quanto nas *Aboxes* (GIACOMO; LENZERINI, 1996). Segundo Gruber (2009), o *TBox* é definido por uma equação atômica de sujeito e predicado e *ABox* é uma especialização de inclusão de sujeito ou conjunto e predicado ou subconjunto. Assim como na representação do conhecimento há a definição de conhecimento intensional¹⁹ e extensional²⁰, uma linguagem descritiva pode de maneira análoga ser comprimida em um *Tbox* (intensional) e um *Abox* (extensional) (BAADER, 2003). A Figura 1 mostra o papel do *Abox* e do *Tbox* como o intermediário da lógica descritiva e os sistemas de inferência. A representação do conhecimento comprimida em *TBox* e *ABox* permite que aplicações sejam executadas utilizando a linguagem descritiva.

Figura 1 – Arquitetura da representação do conhecimento baseado em lógica descritiva.



Fonte: (BAADER, 2003)

¹⁹ "Intensional: "...compreende a menção ao conceito genérico mais próximo (o conceito superordenado) já definido ou supostamente conhecido e às características distintivas que delimitam o conceito a ser definido;"(LARA, 2004).

²⁰ Extensional: "...descreve o conceito pela enumeração exaustiva dos conceitos aos quais se aplica (conceitos subordinados), que correspondem a um critério de divisão..."(LARA, 2004).

2.4.1.1 *TBox*

O *TBox* é a representação intensional de um conhecimento em forma de terminologia. É construído através de declarações que descrevem as propriedades gerais dos termos. As terminologias que formam as *TBoxes* são formadas pelas junções de termos que já foram previamente definidos. Por exemplo:

$$\text{Mulher} \equiv \text{Pessoa} \wedge \text{Sexo Feminino} \quad (2.18)$$

A Equação 2.18 demonstra um *TBox* que define o conceito de *Mulher* utilizando-se dos conceitos prévios de *Pessoa* e *SexoFeminino*. Estas terminologias são bastante fortes quando comparadas com outros modelos de representação do conhecimento e por isso são recursos característicos de modelos baseados em lógica descritiva (BAADER, 2003). Dois pressupostos comumente são atribuídos às terminologias em lógica descritiva:

- Apenas uma definição por termo é permitida e;
- Definições são acíclicas, ou seja, não podem ser definidas por elas mesmas, nem por termos que façam referências indiretas ao termo definido.

As restrições apontadas acima garantem que uma linguagem na lógica descritiva tenha seus termos expansíveis de maneira que expressões complexas se tornem um conjunto de expressões atômicas, se seus termos forem substituídos pelos correspondentes da direita. Nebel (1990) mostrou que a expansão dos termos permitidas pelas *TBoxes* não afetam a complexidade computacional, o que garante esta abordagem seja computacionalmente factível.

Em linhas gerais a tarefa primordial da construção de uma terminologia é a sua classificação, ou seja, definir uma nova expressão terminológica adequadamente respeitando a hierarquia de terminologias. Classificações podem ser realizadas através das verificações das relações obtidas entre os conceitos definidos na hierarquia.

Inferências em *Tboxes* baseiam-se principalmente em implicação lógica. São verificadas se relações genéricas, como subsunção entre dois termos, são resultados de consequência lógica e partir daí deduz-se conhecimento novo.

2.4.1.2 *ABox*

O *ABox* é formado por atribuições de instâncias aos termos definidos pelas *TBoxes*. A Equação 2.19 atesta que Anna é uma pessoa do sexo feminino. Este caso é chamado de atribuição terminológica.

$$\text{Pessoa}(\text{Anna}) \wedge \text{Sexo Feminino} \quad (2.19)$$

Na Equação 2.20 é atestado que Anna possui um filho chamado Douglas. Neste caso se tem uma atribuição de papel, ou seja, estamos designando uma função para Anna e para Douglas. Na linguagem natural dizemos que Anna é mãe de Douglas e Douglas é filho de Anna.

$$\text{temFilho}(\text{Anna}, \text{Douglas}) \quad (2.20)$$

No primeiro caso, expressões terminológicas são permitidas pela linguagem, pois está se fazendo uma atribuição em forma de conceitos primários definidos. No segundo caso essas expressões não são permitidas. O fato da expressão *temFilho* ser uma atribuição de papel, definida por um conjunto de expressões, torna a linguagem da qual ela representa mais expressiva que as linguagens apenas com atribuições terminológicas.

Inferências em *ABoxes* baseiam-se principalmente na verificação de indivíduos. É verificado se um indivíduo *pertence_a* a um termo especificado. A verificação de indivíduos é derivada em várias outras. Elas são:

- Consistência: verifica se cada terminologia tem pelo menos um indivíduo;
- Realização: verifica se há um indivíduo na terminologia mais específica e;
- Recuperação: verifica se indivíduos são instâncias de um dado termo.

Donini et al. (1994) mostraram que a presença de indivíduos aumenta a complexidade de inferências do ponto de vista computacional, o que exige que técnicas de inferências sejam feitas em nível de *TBox* para que haja simplificação. Embora a verificação do *ABox* seja derivada em outras formas, quando uma *TBox* não consegue lidar com todas as questões de simplificação, toda a base de conhecimento (*TBoxes* e *ABoxes*) são processadas por mecanismo de inferências, tornando a tarefa mais complexa.

2.4.2 Inferências na lógica descritiva

Inferências na lógica descritiva são divididos em três partes diferentes: inferências dos operadores lógicos; inferências para *TBoxes* e; inferências para *ABoxes*. Os termos de uma expressão ligada por operadores lógicos estão sujeitos à quatro tipos diferentes de inferência: satisfatibilidade, subsunção, equivalência e disjunção (BAADER, 2003).

O conceito de subsunção representado por $C \sqsubseteq D$ é uma inferência básica realizada na lógica descritiva. Subsunção é a validação se o termo subsumidor (D) é considerado mais geral que o termo subsumido (C). É a verificação se o primeiro termo pertence ao conjunto denotado pelo segundo termo. A satisfatibilidade é determinar se um termo é vazio ou não. A equivalência verifica se um termo possui as mesmas propriedades de um

outro termo, ou seja, são iguais se tem suas expressões expandidas. Por último, a disjunção. Esta verifica se a interseção de dois termos em um dado conjunto é vazio.

Em regras gerais toda inferência realizada na lógica descritiva pode ser reduzida a um caso de subsunção. Baader (2003) mostra que dado os termos C e D se tem:

- C é insatisfazível $\Leftrightarrow C$ é subsumido por \perp ;
- C e D são equivalentes $\Leftrightarrow C$ é subsumido por D e D é subsumido por C ;
- C e D são disjuntos $\Leftrightarrow C \wedge D$ é subsumido por \perp .

As reduções para insatisfatibilidade permite que métodos de decisão sejam construídos em cima destes conceitos. Estes métodos podem fornecer além da satisfatibilidade das expressões, também conjunção e negação destas expressões. Esta possibilidade motivou os pesquisadores que estudavam a lógica descritiva, pois um conceito pode ser representado por um outro conceito que é a negação do primeiro. A redução dos casos para subsunção levaram cientistas a criarem abordagens de novos métodos para verificação de satisfabilidade de uma expressão, conhecido como métodos especializados de cálculo de *tableau*. Motores de inferências para lógica descritiva são baseados principalmente na verificação de satisfatibilidade dos termos.

2.4.3 Método dos *tableaux* analíticos

A partir da lógica proposicional a tabela verdade não tem uso para a dedução de conhecimento novo. O aumento no número de termos de uma expressão aumenta exponencialmente o tamanho de uma tabela, tornando-a inviável como ferramenta de dedução e constatação. A partir daí utiliza-se de outros métodos dedutivos mais eficazes para o raciocínio lógico. Na lógica descritiva utiliza-se principalmente dos métodos dos *tableaux* analíticos. Este método apresentado primeiramente por Schmidt-Schauß e Smolka (1991) para linguagem \mathcal{ALC} demonstrou sua utilidade para verificação de satisfabilidade da linguagem. Desde então, este método foi utilizado para as verificações de correção²¹ e completude²² (logo também de subsunção) para várias extensões diferentes da lógica descritiva. Hollunder (1990), Hollunder e Baader (1991), Donini et al. (1994) e Haarslev e Möller (2000) estenderam o método de *tableaux* analíticos para as *ABoxes* fazendo verificações de consistência. As *TBoxes* também obtiveram avanços em métodos de inferências utilizando-se os *tableaux* analíticos com Buchheit, Donini e Schaerf (1993) e Baader, Buchheit e Hollander (1996) que verificaram inclusão de conjuntos, entre outros.

²¹ Um sistema lógico é correto se todas as suas fórmulas são válidas do ponto de vista semântico de seu sistema.

²² Um sistema lógico é completo se todas as suas fórmulas podem ser demonstradas como sendo verdadeiras e/ou falsas.

D'Agostino et al. (2013) dizem que os *tableaux* foram descobertos por Jaakko Hintikka²³ e Evert Willem Beth²⁴ de forma independente. Aperfeiçoado por Raymond Smullyan²⁵ e por Melvin Fitting²⁶ é tido como o método de provas mais popular, uma vez que apresenta um sistema de provas com abordagens semânticas se tornando bastante intuitivo aos iniciantes de lógica. O *tableau* é um método formal de dedução lógica que está presente em diferentes tipos de linguagens lógicas, mas sempre obedecendo os seguintes procedimentos (D'AGOSTINO et al., 2013):

1. Refuta-se a sua fórmula, ou seja, através de uma expressão declarada intencionalmente falsa, deve-se provar sua falsidade;
2. A fórmula refutada é particionada sintaticamente em várias expressões menores;
3. Em seguida estas fórmulas são expandidas em conectivos lógicos primários (geralmente usa-se a generalização disjuntiva²⁷ para isso);
4. São aplicadas regras de fechamento²⁸ nas expressões expandidas;
5. Se a expressão menor é um caso fechado, então toda a expressão maior será um caso fechado, concluindo-se assim a prova através de dedução.

A Equação 2.21 para o conjunto $\{(a \vee \neg b) \wedge b, \neg a\}$ pode ter sua insatisfatibilidade provada através de um *tableau*. Para que isto aconteça, todas as expressões do conjunto devem ser insatisfazíveis. As regras de inferência de silogismo disjuntivo e simplificação

²³ Kaarlo Jaakko Juhani Hintikka — filósofo e lógico finlandês (1929 - 2015).

²⁴ Evert Willem Beth — filósofo e lógico holandês (1908 - 1964).

²⁵ Raymond Merrill Smullyan — matemático, pianista, lógico, filósofo taoísta e mágico estadunidense (1919 - 2017).

²⁶ Melvin Fitting — lógico estadunidense (1942-).

²⁷ Generalização disjuntiva é a conversão de termos lógicos até que todos estejam conectados apenas pelo conectivo lógico \vee .

²⁸ As regras de fechamento verificam se os termos da expressão são insatisfazíveis, ou seja, contém um termo e a sua negação na mesma expressão.

mostradas na Tabela 5 são usadas para inferências do conjunto.

$$\begin{array}{c}
 (a \vee \neg b) \wedge b \\
 | \\
 \neg a \\
 | \\
 a \vee \neg b \\
 | \\
 b \\
 \swarrow \quad \searrow \\
 a \qquad \neg b
 \end{array}
 \tag{2.21}$$

As linhas um e dois da equação representam as expressões contidas no conjunto onde se quer provar a insatisfatibilidade. As linhas três e quatro da equação são resultados da expansão da expressão $(a \vee \neg b) \wedge b$ usando a regra de inferência da simplificação. A linha cinco é a expansão da expressão $a \vee \neg b$ usando a regra de inferência do silogismo disjuntivo. Assim que todas as expressões forem expandidas em termos atômicos, verifica-se então a contradição destes termos. Na Equação 2.21 a contradição a é demonstrada na linha dois e no nodo à esquerda da linha cinco. O termo b é demonstrado contraditório nas linhas quatro e no nodo da direita da linha cinco. Logo conclui-se que o conjunto $\{(a \vee \neg b) \wedge b, \neg a\}$ é insatisfazível. O método analítico de *tableau* é usado por muitos motores de inferências que atuam sobre as ontologias aplicadas.

3 Ontologias no meio computacional: conceitos, linguagens e ferramentas

A necessidade de definir o que existe em nosso mundo e classificar esta realidade de acordo com aspectos filosóficos remonta a antiguidade com Aristóteles¹ e a sua busca pela categorização das coisas existentes. Surgindo como um ramo da filosofia dentro da metafísica, a Ontologia é uma disciplina que se ocupa no estudo da nossa realidade como a conhecemos. É a ciência do que "é", de tipos e estruturas de objetos, propriedades, eventos, processos e relações em todas as áreas da realidade (SMITH, 2003). Filósofos se utilizam desta disciplina para se orientarem a respeito do que existe; suas categorias, composições e as relações que os termos têm entre si (LOWE, 2006). Smith (2003) ainda considera que na filosofia a ontologia é uma "coisa", uma "substância" que trata de representar o mundo como é visto.

Na Ciência da Informação as ontologias aparecem a partir do século XX como um estudo para a representação do conhecimento. Na década de 1990 o estudo de ontologias ganha força em um momento em que os métodos e a classificação de informação começaram a também fazer parte do escopo da Ciência da Computação (VICKERY, 1997). As ontologias são utilizadas em diversas áreas para organizar o conteúdo das fontes de dados. Elas são criadas por especialistas e definem regras que regulam a combinação entre termos e relações em um domínio do conhecimento (ALMEIDA; BAX, 2003). As ontologias constituem-se como uma alternativa para a gestão do conhecimento e tem seu uso sido aumentado com bastante notoriedade ao longo dos anos (HOLSAPPLE; JOSHI, 2004).

3.1 Ontologia e Tecnologia

Tendo como base uma estrutura que apreende o conhecimento as ontologias começam a ser notadas entre cientistas da computação. Isto porque a evolução dos ambientes digitais torna inevitável o uso de tecnologias da informação como um auxílio para solucionar questões de pesquisa da Ciência da Informação. As relações entre Ciência da Computação e Ciência da Informação tornam-se cada vez mais estreitas e complementares ao decorrer dos anos (OLIVEIRA, 2005).

Existem muitas citações sobre ontologias no campo da Ciência da Computação. Segundo (GUARINO; OBERLE; STAAB, 2009), uma ontologia é para a computação um objeto ou artefato informacional, constituído de um vocabulário específico utilizado para descrever um domínio real do conhecimento. Dias e Costa (2012) definem ontologias como

¹ Aristóteles — filósofo grego (384 a.C. - 322 a.C.)

instrumentos de controle terminológico utilizados na área da recuperação da informação para domínios específicos. Para Gruber (1993), Gruber (1995), a representatividade de uma ontologia em Inteligência Artificial é dizer o que existe e o que pode ser representado. Gruber diverge bastante do conceito filosófico de ontologias. Para ele, ontologias são apenas um conjunto de conceitos e definições. O principal propósito da construção de ontologias na computação é permitir que o conhecimento fosse compartilhado e reutilizado de acordo com princípios de reaproveitamento de código² e informação que são disseminados nesta área. Para isto, Guarino (1995) propôs uma criação de definições formais para conceitos que transformou ontologias em mecanismos de representação a um determinado domínio do conhecimento e não como um todo.

As ontologias na computação surgem em meio ao contexto da Inteligência Artificial e do seu subdomínio a Representação do Conhecimento. Em 1960 o termo começou a ser usado como um significado para estruturas de conceitos representados por um vocabulário lógico. Na década de 1970 cientistas do campo da inteligência artificial reconheceram que capturar e organizar o conhecimento são a peça chave para construção de sistemas inteligentes robustos ao qual poderia ser aplicada lógica matemática (MCCARTHY, 1987). O termo ontologia ganhou força entre pesquisadores da área de Inteligência Artificial que utilizaram ontologias para se referir à teoria de modelos de representação do conhecimento. As ontologias começaram a ser usadas pela computação como uma forma de aplicação dos conceitos definidos pela filosofia (GRUBER, 2009).

Para Almeida (2013) as ontologias apresentadas no campo da Ciência da Computação propõem-se como um sistema para criar vocabulários para representações em sistemas e gerar inferências. O objetivo dos cientistas e engenheiros de *softwares* é desenvolver sistemas que contêm afirmativas representando fatos governados por regras, ou seja, relações entre indivíduos que compõem o sistema. Estes sistemas também denominados como bases de conhecimento são construídas e mantidas por profissionais chamados de engenheiros do conhecimento que tem como função formalizar o conhecimento passado por especialistas e representar isto através de uma ontologia que permita que inferências automáticas sejam feitas. As ontologias são um conjunto de termos que é utilizado para formular consultas por especialistas. Estes especialistas utilizam de termos específicos para encontrar a informação adequada dentro de uma ontologia.

Com a evolução da tecnologia e o acesso da computação em todos os meios, facilitou a entrada de recursos computacionais em várias áreas do conhecimento, incluindo a Ciência da Informação. Isso levou o aumento da popularidade dos computadores para resolver tarefas humanas, e assim a área do estudo de ontologias ganhou ferramentas que ajudasse na sua construção, manutenção e divulgação. A tecnologia e a computação

² Na Ciência da Computação a reutilização de código é uma prática que acompanha a área desde o seu surgimento. Trata-se da construção de novos *softwares* utilizando-se de códigos fontes já existentes.

contribuíram para a área do estudo de ontologias fornecendo linguagens semânticas para a sua construção, editores para a sua visualização e manutenção, assim como também motores de inferência ou derivadores lógicos que contribuíram para a descoberta de conhecimento novo (BERNERS-LEE et al., 2001).

A tecnologia sem dúvida alguma se mostra uma aliada no processo de construção de ontologias. Porém o campo da tecnologia está mais ligado ao campo da Ciência da Computação do que o campo da Ciência da Informação onde ela tem seu estudo mais aprofundado. Para ilustrar este exemplo podem-se comparar diferentes ontologias encontradas na *web*. Percebe-se que ontologias feitas por cientistas da computação muitas vezes carecem de metodologia apropriada para a construção da mesma. Isso reflete em sua formação precária no assunto que contempla apenas a parte lógica e funcional, que engloba práticas de desenvolvimento de código.

A nova geração do estudo de ontologias trouxe com ela pontos positivos e negativos. A parte negativa se encontra na visão da ontologia como um artefato informacional e não como o modelo filosófico de representação do conhecimento. Outro ponto negativo se encontra na limitação que computadores têm para entender e interpretar os dados disponibilizados nestas bases do conhecimento. Em contrapartida a digitalização das ontologias trouxe consigo um aliado muito forte. A disseminação destas ontologias, seu reaproveitamento e reuso fazem da tecnologia um caminho sem volta. O uso de editores visuais que permitem visualizar a construção de uma ontologia ajuda muito os profissionais que ficam responsáveis por sua criação e manutenção. A utilização de ferramentas lógicas que são capazes de tomar decisões e fazer inferência automática se torna indispensável para grandes ontologias bem construídas.

Além de tudo as ontologias vêm para ajudar a tecnologia também. Problemas como excesso de informação na *web* podem ser contornados se utilizados modelos de dados bem elaborados que ajudem na recuperação de informação como o caso de ontologias. Modelos de representação de dados e organização do conhecimento estão sendo cada vez mais necessitados em meio ao *boom* informacional que sofre a internet.

As ontologias em sistemas de informação e tecnologia contam com apoio de diversas ferramentas que contribuem para seu uso e disseminação. Entre elas encontram-se as linguagens de programação que fornecem meios de construir uma ontologia em um computador; editores de ontologias que é a plataforma onde as ontologias são construídas e dão suportes diversos, como visualização de classes, etc. e; motores de inferência que analisa a ontologia criada verificando sua validade, consistência e inferindo novos fatos que se encontravam implícitos aos olhares humanos. Esse ferramental contribui para o avanço da área de ontologia aplicada, possibilitando cada vez mais precisão na formalização de domínios.

3.2 *Web Ontology Language – OWL*

A OWL é uma variação de lógica descritiva que surgiu da necessidade de uma linguagem de modelagem de ontologias mais expressiva. Isso levou a uma iniciativa para criação da DAML+OIL³ em 2001, a qual foi posteriormente nomeada como OWL. A OWL é o ponto de partida do *World Wide Web Consortium (W3C) Web Ontology Working Group* para definir uma linguagem padronizada e amplamente aceita como a linguagem para ontologias na *web* semântica (ANTONIOU; HARMELEN, 2004). A OWL é tida por Patel-Schneider et al. (2004), como um componente da atividade da *web* Semântica e que tem como objetivo tornar os recursos da *web* mais acessíveis ao processamento por máquina, ao adicionar informação sobre recursos que descrevem ou fornecem conteúdo.

A OWL define e instancia ontologias na *web*. Uma ontologia em OWL pode incluir descrições de classes, propriedades e as instâncias. Possui um conjunto de operadores lógicos, e é baseada em um modelo que possibilita definir conceitos da forma em que são descritos (HORRIDGE, 2011). Dada uma ontologia, a semântica formal em OWL especifica como derivar consequências lógicas, como por exemplo, fatos que não estão explícitos na ontologia, mas implicados na semântica da linguagem de representação (WELTY; MCGUINNESS; SMITH, 2004).

A OWL foi desenvolvida para uso quando a informação contida em documentos precisa ser processada por aplicações, o que é o oposto das situações em que o conteúdo precisa ser apenas apresentado aos usuários (HARMELEN; MCGUINNESS, 2004). Surgindo como necessidade de aplicações para processamento do conteúdo de ontologias, a OWL provê um tipo de interpretador para processamento por máquinas, se mostrando mais efetiva do que outras linguagens como XML, RDF e RDFS. A principal diferença entre a OWL e o XML é que a segunda é uma linguagem para formatos de mensagens, enquanto a primeira fornece um vocabulário com uma semântica bem definida.

As ontologias OWL possuem três elementos que são usados para representar um domínio do conhecimento:

- **Indivíduos:** são instâncias de classes e representam objetos em um domínio. Podem ser relacionados entre si através das propriedades. Exemplo: Arthur, cachorro, lápis, etc;
- **Propriedades:** são relações binárias entre indivíduos ou de um indivíduo para um valor. Exemplo: a propriedade *pertence_a* pode ligar o indivíduo cachorro ao indivíduo Arthur. As propriedades podem ser inversas também. Por exemplo: Arthur *possui* cachorro é inverso da propriedade cachorro *pertence_a* Arthur. As propriedades ainda são divididas em funcionais (quando um indivíduo não pode se

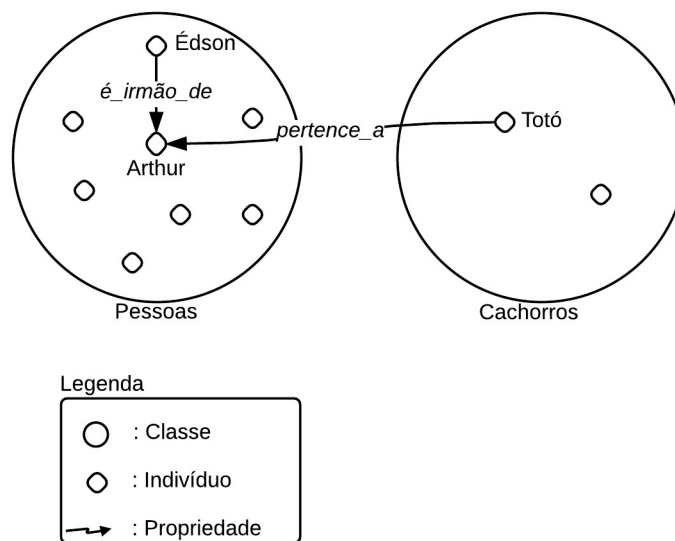
³ <http://www.daml.org/language/>

relacionar com mais de um indivíduo ao mesmo tempo), transitivas (Um indivíduo A se relaciona com B que se relaciona com C . Logo o indivíduo A se relaciona com C) ou simétricas (O indivíduo A se relaciona com B e B se relaciona com A), e;

- **Classes:** consiste de um grupo de indivíduos que pertencem a um mesmo conjunto porque compartilham propriedades. As classes podem ser organizadas em hierarquias e divididas em subclasses. Exemplo: animal é uma classe e cachorro é sua subclasse.

A Figura 2 mostra a representação de indivíduos, propriedades e classes.

Figura 2 – Representação gráfica de classes, indivíduos e propriedades

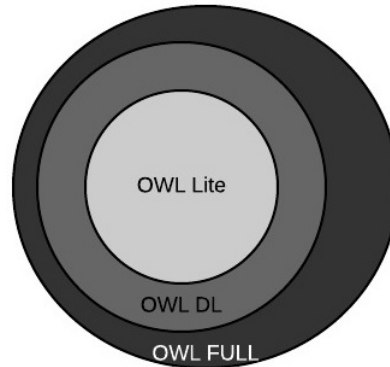


A OWL possui três sub-linguagens desenvolvidas e projetadas para uso por comunidades específicas e por usuários finais:

- **OWL Lite:** é a mais simples das três derivações da OWL por ser sintaticamente mais restrita. Atende as situações em que se necessita de uma classificação hierárquica e recursos simples para restrição (DEAN et al., 2004). Um exemplo de aplicação da OWL Lite é uma forma de migração dos tesauros;
- **OWL DL:** é a mais utilizada com fins computacionais em sistemas de apoio à decisão. Atende as situações em que se deseja expressividade sem perder a capacidade de processamento. A OWL DL contém a OWL Lite, mas traz restrições, tais como separação por tipos (uma classe não pode ser uma propriedade ou um indivíduo e nem vice-versa).
- **OWL Full:** é a mais expressiva das sub-linguagens da OWL. Em OWL Full, as classes podem ser propriedades e indivíduos, além de permitir o acréscimo de um vocabulário pré-definido. Porém, isto implica em que dificilmente sistemas de apoio à decisão suportam os recursos da OWL Full.

Cada destas sub-linguagens são uma extensão da outra nos quesitos que elas são expressas e válidas como uma ontologia. A Figura 3 ilustra com mais clareza este exemplo.

Figura 3 – Representação gráfica dos recursos abrangidos por cada sub-linguagem da OWL



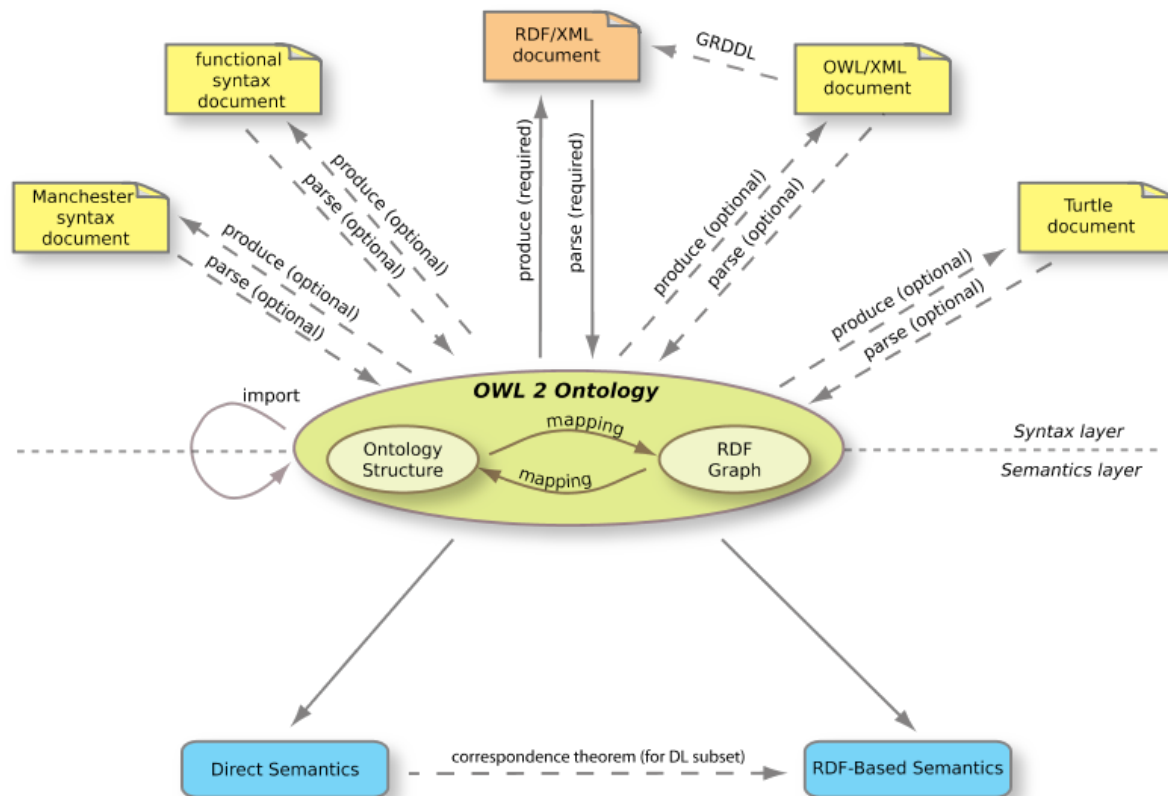
A versão mais atual é a OWL 2 homologada em 2012 por Consortium et al. (2012). Foi apresentada como uma continuação da OWL (que é chamada desde então de OWL 1) e foi desenvolvida com os mesmos preceitos de facilitar o desenvolvimento de ontologias e tornar o conteúdo da *web* mais acessível por máquinas. Apesar de ter obtido bastante sucesso, a OWL 1 trouxe problemas que precisavam ser corrigidos. Segundo Krötzsch (2012), a OWL 1 originalmente apresentava dificuldades para computar todas as conclusões lógicas possíveis de uma ontologia, tornando a derivação lógica multiexponencial e até mesmo incomputável. Para lidar com estes problemas, três novas sub-linguagens foram criadas para compor a OWL 2. A OWL EL, OWL RL e OWL QL são sub-linguagens leves e computacionalmente processáveis. A OWL 2 foi lançada com a proposta de realizar grandes avanços em sua performance e escalabilidade.

A Figura 4 mostra como é a estrutura da OWL 2. No centro encontra-se a representação abstrata do mapeamento de uma ontologia, ou seja, como ela é construída independente da sintaxe utilizada pelo desenvolvedor. A parte de cima mostra as diferentes sintaxes⁴ que são utilizadas para a construção concreta destas ontologias. Estas sintaxes fornecem interoperabilidade entre as diversas ontologias disponíveis na *web*. Na parte de baixo encontram-se as semânticas que definem o significado sintático construído pela sintaxe utilizada. Essas semânticas são divididas em duas partes: diretas e baseadas em RDF. As semânticas baseadas em RDF transformam os axiomas escritos em OWL em triplas⁵ em RDF. As semânticas diretas relacionam os axiomas escritos em OWL para a lógica descritiva, permitindo assim derivações lógicas computáveis. Na prática a maioria dos usuários fazem uso de apenas uma sintaxe e uma semântica. O centro da figura é compreendido pelos computadores que processam a ontologia construída pelos usuários.

⁴ A sintaxe em uma linguagem de programação são conjuntos de regras estabelecidas (palavras reservadas) que formam uma linguagem. Ex: "var", "if", "else", etc...

⁵ Uma tripla é uma declaração contendo sujeito, predicado e objeto, essenciais para relações semânticas.

Figura 4 – Estrutura da OWL 2



Fonte: (CONSORTIUM et al., 2012)

Diferentemente das variações da OWL 1 apresentadas previamente, as sub-linguagens da OWL 2 são independentes entre si. Cada uma delas foi desenvolvida para um cenário e um público específico que usarão e construirão suas ontologias de acordo com a estrutura solicitada (MOTIK et al., 2016). Elas são apresentadas a seguir:

- **OWL 2 EL:** foi desenvolvida para a construção de grandes ontologias contendo massiva quantidade de classes e propriedades. É usada amplamente por ontologias biomédicas pelo fato de ser capaz de processar inferências em tempo polinomial⁶ referente ao tamanho da ontologia;
- **OWL 2 QL:** foi desenvolvida para ontologias contendo quantidades massivas de indivíduos. Estas ontologias priorizam mais consultas de indivíduos, onde seu desempenho é melhor do que a realização de inferências sobre as classes e propriedades. Esta sub-linguagem permite que um sistema de buscas seja implementado utilizando-se técnicas de banco de dados tradicionais. Fornece uma aplicação em base de dados com uma camada ontológica para consulta. Assim como a OWL 2 EL, esta sub-linguagem

⁶ Na Ciência da Computação um problema polinomial é aquele capaz de ser processado por um computador em tempo hábil embora não exato.

é capaz de processar inferências em tempo polinomial embora não seja tão expressiva quanto a primeira, e;

- **OWL 2 RL:** Esta sub-linguagem é voltada para aplicações que requerem processamento de inferências escalável, ou seja, possível em ontologias grandes, sem perder expressividade. Desenvolvida para que as aplicações que usem estas ontologias possam intercambiar a expressividade da linguagem OWL 2 sem perder eficiência. Adequada para ontologias que dão valor semântico ao conteúdo da *web*, pois tem mais poder de consultas.

Conhecer a OWL e suas variações é importante para que o profissional da ciência informação saiba em qual cenário ele se encaixa melhor. Cada linguagem e sub-linguagem apresentada possui um propósito que deve ser respeitado se os objetivos tem que ser cumpridos. Ontologias biomédicas como o caso da BLO requerem que sua linguagem seja capaz de realizar inferências em tempo hábil, pois precisa ser verificado que esta ontologia seja consistente. Caso contrário problemas como derivação de conhecimento errôneos podem ser produzidos e constatações falsas serão feitas, contrapondo o rigor metodológico e científico exigido nas áreas médicas.

3.3 A sintática e a semântica da OWL

Por trás de toda ontologia desenvolvida por profissionais da ciência informação utilizando a OWL existe um código fonte que torna os computadores capazes de interpretar parte das informações contidas naquela ontologia. As variações da OWL fazem parte das linguagens de representação do conhecimento. Segundo Bobrow e Winograd (1977), estas linguagens são tentativas de dar entendimento ao conhecimento para os sistemas computacionais através de declarações. Estas declarações permitem que estas linguagens expressem conhecimentos através de estruturas lógicas para um domínio do conhecimento. A estruturação lógica de uma base de conhecimento gera inferências e constatações lógicas, além de controlar a acessibilidade destes fatos e descrições, aumentando seu significado.

Como uma linguagem, a OWL é composta de sintaxe, semântica e pragmática. Por sintaxe entende-se do conjunto de palavras, frases e sentenças que compõem a linguagem. No entanto apenas o conjunto de palavras e sentenças desordenadamente não constitui significado. O significado de uma linguagem é composto por sua semântica, ou seja, a ordenação das palavras, frases e sentenças de tal forma que elas ganham significados para um sistema, no caso das linguagens de computadores, o próprio computador. A pragmática se refere aos aspectos práticos que a linguagem fornece para que sejam construídos objetivos (PIERCE, 2002). É o conjunto de semântica e sintática reunidos para expressar funções de um determinado objetivo, no caso da OWL, definir e relacionar termos, propriedades e classes.

Sintaticamente a OWL pode ser escrita de várias maneiras, dividido em duas categorias principais: baseados em RDF/XML e; baseados em OWL. A variedade de sintaxes existentes difere nas diferentes propostas e aplicações, no entanto todas as linguagens não são definidas por uma sintaxe concreta, mas sim por uma especificação de alto nível que mapeia em várias sintaxes concretas. Embora a recomendação do W3C especifica a sintaxe baseada em RDF/XML como padrão, as demais sintaxes são suportadas pelas principais ferramentas de ontologias (HORRIDGE, 2010).

A OWL é parcialmente baseado na lógica descritiva, que também faz parte das famílias de linguagens de representação do conhecimento. Os blocos de construção da OWL são bem parecidos com da lógica descritiva, como os axiomas que são construídos a partir de elementos do vocabulário e os construtores são baseados em operadores lógicos. Ontologias em lógica descritiva são usualmente chamadas de bases de conhecimento; classes e propriedades em lógica descritiva são chamadas de conceitos e papeis (KRÖTZSCH, 2012). A Tabela 6 faz a tradução das sintaxes da OWL para a lógica descritiva, como demonstração do relacionamento destas duas linguagens.

Tabela 6 – Tradução das sintaxes da OWL para lógica descritiva

Expressões	Sintaxe OWL	Lógica Descritiva
Axiomas	<i>SubClassOf(C D)</i>	$C \sqsubseteq D$
	<i>ClassAssertion(C a)</i>	$C(a)$
	<i>ObjectPropertyAssertion(P a b)</i>	$P(a,b)$
Classes	<i>ObjectIntersectionOf(C D)</i>	$C \wedge D$
	<i>ObjectUnionOf(C D)</i>	$C \vee D$
	<i>ObjectComplementOf(C)</i>	$\neg C$
	<i>owl:Thing</i>	\top
	<i>owl:Nothing</i>	\perp
	<i>ObjectSomeValuesFrom(P C)</i>	$\exists P.C$
	<i>ObjectAllValuesFrom(P C)</i>	$\forall P.C$
Expressões	<i>ObjectInverseOf(P)</i>	P^-

Fonte: (KRÖTZSCH, 2012)

Consortium et al. (2012) descreve a ontologia sob o ponto de vista estrutural, ou seja, uma ontologia pode ser nomeada com um IRI⁷ e um conjunto de axiomas que descrevem as suas relações. Isto é feito em alto nível, o que possibilita a descrição de funcionalidades essenciais da OWL sem se prender a detalhes técnicos que variam entre estas sintaxes (HORRIDGE, 2010). O conhecimento das sintaxes e de como a informação é

⁷ IRI é uma extensão do URI, sequência de caracteres usadas para identificar um objeto na internet, que é composto do conjunto universal de caracteres.

construída e disponibilizada por trás de uma ontologia, revelará a forma como os motores de inferência trabalharão para analisar os dados, extrair informação, aplicar métodos de dedução e derivar conhecimento novo.

3.3.1 Sintaxe RDF/XML

RDF/XML é uma linguagem para edições em códigos extensos com muitas classes. Tem como principais vantagens: é amplamente suportada por ferramentas de ontologias por ser a linguagem padrão recomendada pela W3C e; sintaxes baseadas em RDF são boas para representações de *ABox* em ontologias, ajudando motores de inferências que trabalham diretamente neste tipo de componente.

As desvantagens dessa linguagem são: bastante prolixa, o que a torna difícil de ser lida; a representação de classes mais complexas de axiomas em OWL podem tornar o código ainda mais poluído e difícil de ser lido e; o *design* desta linguagem foi baseado na OWL 1 o que torna a análise das ontologias representadas em RDF um processo não trivial. A análise das ontologias quase sempre requer que sua verificação seja feita em dois passos, gastando assim muita memória computacional, deixando os processos lentos e inviabilizando o uso de ontologias extremamente grandes (HORRIDGE, 2010). O Código 3.1 mostra a representação de um adolescente entre 12 e 20 anos em XML/RDF.

Código 3.1 – Descrição de um adolescente entre 12 e 20 anos em RDF/XML

```
<owl:Class rdf:about="http://www.semanticweb.org/ontologies/ontogenesis/
example#Teenager">
  <owl:equivalentClass>
    <owl:Class>
      <owl:intersectionOf rdf:parseType="Collection">
        <rdf:Description rdf:about="http://www.semanticweb.org/
ontologies/ontogenesis/example#Person"/>
        <owl:Restriction>
          <owl:onProperty rdf:resource="http://www.semanticweb.
org/ontologies/ontogenesis/example#hasAge"/>
          <owl:someValuesFrom>
            <rdfs:Datatype>
              <owl:onDatatype rdf:resource="http://www.w3.org
/2001/XMLSchema#integer"/>
              <owl:withRestrictions rdf:parseType="Collection
">
                <rdf:Description>
                  <xsd:maxExclusive rdf:datatype="http://
www.w3.org/2001/XMLSchema#integer
">20</xsd:maxExclusive>
                </rdf:Description>
              <rdf:Description>
```

```

        <xsd:minExclusive rdf:datatype="http://
            www.w3.org/2001/XMLSchema#integer
            ">12</xsd:minExclusive>
    </rdf:Description>
</owl:withRestrictions>
</rdfs:Datatype>
</owl:someValuesFrom>
</owl:Restriction>
</owl:intersectionOf>
</owl:Class>
</owl:equivalentClass>
</owl:Class>

```

3.3.2 Sintaxe *Turtle*

A sintaxe *Turtle* é levemente menos prolixa que a baseada em RDF/XML, porém não oferece muitos recursos para que seja lida por pessoas. É amplamente suportada. Ferramentas como Jena são capazes de produzir e consumir ontologias escritas em *Turtle*. As desvantagens são as mesmas citadas na subsecção 3.3.1 para a linguagem RDF/XML. O Código 3.2 mostra a representação de um adolescente entre 12 e 20 anos em *Turtle*.

Código 3.2 – Descrição de um adolescente entre 12 e 20 anos em *Turtle*

```

:Teenager rdf:type owl:Class ;
  owl:equivalentClass [ rdf:type owl:Class ;
    owl:intersectionOf ( : Person
      [ rdf:type owl:Restriction ;
        owl:onProperty :hasAge ;
        owl:someValuesFrom [ rdf:type rdfs:Datatype ;
          owl:onDatatype xsd:integer ;
          owl:withRestrictions (
            [ xsd:maxExclusive "20"^^xsd:integer ]
            [ xsd:minExclusive "12"^^xsd:integer ]
          )
        ]
      ]
    )
  ] .

```

3.3.3 Sintaxe OWL/XML

O RDF/XML não obedece aos padrões da sintaxe XML, tornando-o assim difícil de ser conjugado em aplicações exclusivas para o XML. Então a OWL/XML surgiu devido à necessidade de padronizar as ontologias dentro de um formato já conhecido, o XML, ao qual já existiam diversas ferramentas que poderiam ser utilizadas para lidar com a informação disponibilizada.

As vantagens dessa sintaxe são: seu modelo é sujeito aos padrões do XML Schema o que permite que ferramentas que não são para ontologias, mas que fazem análise de XML funcionem nestes documentos e; é facilmente escrito e analisado através de ferramentas externas próprias para XML. As desvantagens é que ainda é uma sintaxe prolixa. Ontologias podem ficar muito grandes, fazendo com que análise e inferência de outros aplicativos sejam lentos e demorados. O Código 3.3 mostra a representação de um adolescente entre 12 e 20 anos em OWL/XML.

Código 3.3 – Descrição de um adolescente entre 12 e 20 anos em OWL/XML

```
<EquivalentClasses>
  <Class IRI="#Teenager"/>
  <ObjectIntersectionOf>
    <Class IRI="#Person"/>
    <DataSomeValuesFrom>
      <DataProperty IRI="#hasAge"/>
      <DatatypeRestriction>
        <Datatype abbreviatedIRI="xsd:integer"/>
        <FacetRestriction facet="http://www.w3.org/2001/XMLSchema#
          maxExclusive">
          <Literal datatypeIRI="http://www.w3.org/2001/XMLSchema#
            integer">20</Literal>
        </FacetRestriction>
        <FacetRestriction facet="http://www.w3.org/2001/XMLSchema#
          minExclusive">
          <Literal datatypeIRI="http://www.w3.org/2001/XMLSchema#
            integer">12</Literal>
        </FacetRestriction>
      </DatatypeRestriction>
    </DataSomeValuesFrom>
  </ObjectIntersectionOf>
</EquivalentClasses>
```

3.3.4 Sintaxe Manchester OWL

A principal motivação desta sintaxe é produzir algo que fosse utilizado para editar expressões em ferramentas como o Protégé. Desde então foi estendido de forma que é possível construir ontologias inteiras utilizando esta sintaxe.

Sua principal vantagem é ser compacto, fácil de ser lido e escrito. Como desvantagem, o Manchester OWL se mostra complicado para representar axiomas em OWL que têm sua representação do lado esquerdo da equação, como representação de subclasses. O Código 3.4 mostra a representação de um adolescente entre 12 e 20 anos em Manchester OWL.

Código 3.4 – Descrição de um adolescente entre 12 e 20 anos em Manchester OWL

```
Class: Teenager
```

EquivalentTo: Person and (hasAge some integer[> 12 , < 20])

3.4 Motores de inferência

Os motores de inferência são *softwares* computacionais que tem a capacidade de inferirem consequências lógicas baseados em um conjunto de axiomas definidos por meio de lógica descritiva. Segundo Antoniou e Harmelen (2004), os motores de inferência transformam conhecimento implícito em conhecimento explícito através de deduções de um conhecimento que lhe foi dado. As linguagens formais de ontologias possuem diversas ferramentas que podem auxiliar os engenheiros do conhecimento, porém a mais importante delas, os motores de inferência, são as que possibilitam o maior ganho do uso de tecnologias na construção de sistemas ontológicos.

Uma ontologia ganha relevância para estes motores de inferências quando seus axiomas estão explicitamente vinculados entre si. Ou seja, o significado de uma ontologia é dado por todas as conclusões que podem ser extraídas delas, não importando se elas estão explicitamente estabelecidas (KRÖTZSCH, 2012). Portanto a tarefa dos motores de inferências se tornam importantes para o desenvolvimento de ontologias. A proposta de um motor de inferência é computar vinculações particulares feitas em uma ontologia para verificações lógicas de consistência e constatações e derivações de fatos implícitos.

Mishra e Kumar (2011) dividiu os motores de inferência em quatro gerações distintas de acordo com as aplicações para as quais eles foram desenvolvidas. Estas gerações vão desde 1975 com KL-ONE(BRACHMAN; SCHMOLZE, 1985) até os dias atuais com HermiT(SHEARER; MOTIK; HORROCKS, 2008). O RACER(HAARSLEV; MÖLLER, 2001) que apareceu em 2001 foi o primeiro motor de inferência pra OWL e pertence à terceira geração dos motores de inferência. Khamparia e Pandey (2017) analisou 23 motores de inferências diferentes e avaliou-os segundo aspectos de inferência e versatilidade como suporte a completude, tempo de resposta, expressividade, etc., como mostrado na Tabela 7. Concluiu-se que apenas 21,7% dos motores analisados possuem 80% ou mais dos aspectos analisados. Os motores de inferências considerados robustos por Khamparia e Pandey (2017) são: Pellet, RACER, KAON2, CRACK e KRIS.

Além destes foram considerados o HermiT e o FaCT++. Apesar de serem considerados como motores de inferências medianos por Khamparia e Pandey (2017) (entre 40% e 80%), estes motores tem um forte apelo popular por conta de serem nativos na ferramenta Protégé.

⁸ Justificação é assumir que uma hipótese é verdadeira sem prová-la para usá-la como prova de um teorema maior.

Tabela 7 – Características dos motores de inferências

Recurso/ Motor de Inferência	Pellet	RACER	KAON2	CRACK	KRIS	HermiT	FaCT++
Metodologia	<i>tableau</i>	<i>tableau</i>	<i>SHIQ</i>	<i>tableau</i>	<i>tableau</i>	<i>hyper-tableau</i>	<i>tableau</i>
Correção	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Completude	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Expressividade	Alta	Alta	Alta	Média	Baixa	Alta	Alta
Tempo de resposta	Baixo	Alto	Alto	Baixo	Alto	Alto	Baixo
Regras de inferência	Sim	Sim	Sim	Sim	Sim	Sim	Não
Plataforma	Todas	Todas	Todas	Todas	Todas	Todas	Todas
Justificação ⁸	Sim	Sim	Sim	Sim	Sim	Não	Não
Inferência em <i>ABox</i>	Sim	Sim	Sim	Sim	Sim	Sim	Sim
OWL API	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Suporte ao Protégé	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Suporte ao Jena	Sim	Não	Sim	Não	Sim	Não	Não
Implementação	Java	LISP	Java	Common LISP	Common LISP	Java	C++
Compreensão	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Atualizável	Sim	Sim	Sim	Não	Não	Sim	Sim

Fonte: adaptado de (KHAMPARIA; PANDEY, 2017)

3.4.1 Pellet

O Pellet é um projeto *open source* desenvolvido pelo grupo *Mind Swap*⁹ baseado em Java¹⁰ para OWL-DL. Foi o primeiro motor de inferência a incorporar métodos de inferências para todas as linguagens oferecidas pela OWL-DL e a OWL 2 (SIRIN et al., 2007). É um motor de inferência baseado nos métodos de *tableau* analítico que foi desenvolvido para lógicas descritivas de expressividade (PARSIA; SIRIN, 2004). É suportado pela OWL DL por tratar todos os axiomas desta linguagem. Para isso, é feita uma combinação de algoritmos de raciocínio lógico existentes baseados em lógica descritiva. Estes algoritmos são utilizados com a meta de alcançar uma ampla cobertura de todos os casos de inferências em ontologias. O Pellet garante a completude e a correção lógica incorporando procedimentos de decisão da linguagem *SHOIQ* (HORROCKS et al., 2005). Seu principal foco de inferência são baseados nas *TBoxes* e *ABoxes* como é mostrado na Figura 5. O Pellet armazena os axiomas sobre classes na *TBox* e também as asserções sobre os indivíduos nos componentes da *ABox*. As inferências são feitas a partir das regras padrões do *tableau* junto com otimizações estratégicas que garantem a consistência das inferências. Elas verificam a subsunção, satisfatibilidade, classificação, recuperação e consultas conjuntivas. O Pellet foi implementado em Java sob a licença do MIT e tem seu código aberto ao público.

3.4.2 RACER

O motor de inferência RACER, foi desenvolvido para a lógica descritiva *SHIQ*^(D) baseado em *tableaux* para realização de derivações (HAARSLEV; MÖLLER; WESSEL, 2004). Mishra e Kumar (2011) cita o RACER como sendo o primeiro motor de inferência que surgiu para a linguagem OWL. O RACER, mais tarde chamado de RACERPro, fornece serviços para o conhecimento de termos bem como o conhecimento sobre os indivíduos (HAARSLEV et al., 2012). Para domínios concretos, este motor de inferência suporta equações polinomiais sobre números inteiros e reais, e também a comparação de palavras e validação de igualdade (BOCK et al., 2008). Assim como o Pellet, o RACERPro também foca em fazer inferências sobre as *ABoxes* e *TBoxes*. Para isto, ele lida com ambos sob uma única distinção. Além de tarefas básicas de inferências como satisfatibilidade e subsunção, o RACERPro também fornece consultas otimizadas em NRQL¹¹. Foi implementado na linguagem de programação Lisp¹² que é utilizada principalmente para representações matemáticas e de lógica (BOCK et al., 2008). O W3C detalhou através de *feedbacks* que este motor de inferência é utilizado por centenas de projetos e também dentro do contexto

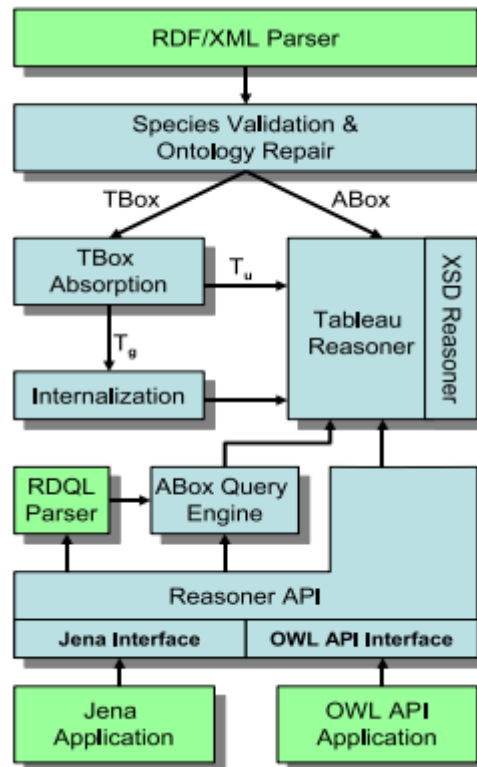
⁹ <http://www.mindswap.org/>

¹⁰ <https://www.java.com>

¹¹ <https://docs.newrelic.com/docs/insights/new-relic-insights/using-new-relic-query-language/nrql-reference>

¹² <https://www.common-lisp.net/>

Figura 5 – Arquitetura do Pellet



Fonte: (PARSIA; SIRIN, 2004)

industrial.

3.4.3 KAON2

KAON2 é um motor de inferência grátis (para uso não comercial) baseado em Java apresentado por Motik e Studer (2005). Foi desenvolvido para a linguagem $SHIQ^{(D)}$, SWRL e FLogic. Sua implementação é baseada em procedimentos voltados para *TBoxes* (subsunção, satisfatibilidade e classificação) e *ABoxes* (recuperação e consultas). Seu processo é baseado na redução da base de conhecimento da linguagem $SHIQ^{(D)}$ em um conjunto de dados disjuntos, aumentando a eficácia das consultas de dados (MOTIK; SATTLER, 2006). Essa eficiência se dá porque a maioria dos motores de inferências lidam com uma quantidade grande de dados, e o KAON2 faz inferências em cada parte disjunta dos dados (MOTIK, 2008). O KAON2 é o sucessor do KAON, porém não são compatíveis entre si. Foi desenvolvido como um esforço mútuo da FZI (*Forschungszentrum Informatik*) e da AIFB (*Institute of Applied Informatics and Formal Description Methods*). Tem suporte para RDF-S com características como extensão de cardinalidade, modularização e metamodelagem. Além disso, o KAON também suporta as propriedades simétricas, transitivas, inversas e representação léxica da informação (MISHRA; KUMAR, 2011).

3.4.4 CRACK

Apresentado por Bresciani, Franconi e Tessaris (1995), teve como principal meta a implementação da verificação de correção e completude para tratar inferências sobre indivíduos e termos. Esta implementação utiliza-se de métodos analíticos de *tableau* primeiramente para a linguagem $\mathcal{ALCRIFO}$ e posteriormente realiza inferências para as demais lógicas descritivas mais expressivas (BAADER, 2003). Sua abordagem semântica é diferente dos métodos tradicionais, onde indivíduos são mapeados dentro de conjuntos de objetos de um domínio. Ao invés disto, o CRACK mapeia os indivíduos como elementos do domínio. Esta abordagem permite que correferências sejam consideradas como conceitos (BRESCIANI; FRANCONI; TESSARIS, 1995). O CRACK fornece para uma interface *web* que foi implementada em CommonLISP. Assim como o KRIS, as informações disponibilizadas pelo motor de inferência são filtradas e as mais explícitas são exploradas pela arquitetura do motor através de análise dos *tableaux*. Seu tempo de resposta baixo, se deve ao tratamento da informação de maneira direta, interferindo assim na decidibilidade do algoritmo que foi desenvolvido para a derivação manual dos resultados. Nestes casos os *tableaux* não eram os mais indicados para este tipo de abordagem (BAADER, 2003).

3.4.5 KRIS

Apresentado por Baader e Hollunder (1991a), KRIS (acrônimo para *Knowledge Representation and Inference System*) foi o primeiro motor de inferência à suportar verificação de completude e correção em múltiplas *ABoxes* para lógicas descritivas. Sua proposta foi voltada para lógicas descritivas expressivas com o objetivo de desenvolver técnicas de otimização para inferências em *TBoxes*. A lógica descritiva do KRIS é \mathcal{ALCNF} além de fornecer tipos enumerados e inferências em domínios concretos (BAADER; HOLLUNDER, 1991b) (BAADER; HANSCHKE, 1991) (BAADER; HANSCHKE, 1992). Sua implementação foi feita na linguagem CommonLISP e assim como o CRACK também fornece uma interface *web*. KRIS consegue lidar com múltiplas instâncias de *ABox* e *TBox* ao mesmo tempo. Apesar de ter sido lançado na década de noventa, KRIS é ainda usado por sistemas modernos de lógica descritiva (BAADER, 2003).

A ideia por trás deste motor de inferência é otimizar as classificações da *TBox* afim de explorar as informações explícitas disponibilizada pela linguagem. Muitas destas informações se mostram uma conjunção de indivíduos e termos que podem ser subsumidos se expandidos. Outro ponto importante do KRIS é evitar o cálculo de relações de subsunção detectadas nos passo anterior. Desta maneira, processos são economizados, aumentando a eficiência computacional. Embora testes realizados por Baader et al. (1994) mostrem que o desempenho do KRIS para inferências em *TBox* fosse comparável aos demais motores da sua época, a sua resolução de *tableaux* por métodos não-determinísticos provou ter desempenho superior para situações de decidibilidade computacional (BAADER, 2003).

3.4.6 HermiT

O HermiT foi o primeiro motor de inferência para ontologias que foi disponibilizado para OWL (GLIMM et al., 2009). Foi escrito utilizando da própria linguagem OWL. Ele é baseado em uma nova arquitetura que endereça as fontes de complexidade de uma ontologia (SHEARER; MOTIK; HORROCKS, 2008). Dado um arquivo escrito em OWL, o HermiT é capaz de determinar consistência, subordinação, dentre outros. Foi o primeiro motor para ontologias a ser baseado em *hyper tableau*¹³ e por isso fornece inferências mais eficientes que os demais algoritmos que utilizam-se apenas de *tableau* (HORROCKS; MOTIK; WANG, 2012).

Este motor de inferência se destacou pelo seu alto desempenho e pela capacidade de fazer inferências em ontologias complexas que até a data de seu lançamento ainda não existia nenhum concorrente que conseguisse fazer o mesmo. Em sua versão mais atual já é capaz de realizar todos os testes de conformidade na OWL 2. É atualmente o motor de inferência padrão do Protégé.

3.4.7 Fact++

Apresentado por Tsarkov e Horrocks (2006), FaCT (acrônimo para *Fast Classification of Terminologies*) implementa um *tableau* para decisões em lógica descritiva com suporte adicional para tipos de dados como strings e números inteiros. Pode ser usado para testes de satisfabilidade tanto para a lógica descritiva quanto para a lógica modal. O FaCT realiza inferências em duas principais linguagens: *SHF* e *SHIQ*. O *tableau* implementado no sistema do FaCT faz verificações de completude e correção para lógicas descritivas expressivas, o que faz com que seja eficiente em grandes bases de conhecimento (BAADER, 2003). Sua arquitetura de desenvolvimento é baseado no cliente-servidor CORBA que suporta inferências em TBoxes e parcialmente em OWL-Lite (GUO; PAN; HEFLIN, 2005).

Evoluído do FaCT, o FaCT++ foi desenvolvido em virtude do grande aumento do número de ontologias e de sua usabilidade. É uma evolução do sistema FaCT e por isto incorpora técnicas de otimizações advindas do mesmo além de outras novas (HORROCKS, 1998). O seu foco é em bases de conhecimento que são pré-processadas (onde várias otimizações já são feitas), normalizadas e transformadas em uma representação interna. Então o motor de inferência faz uma classificação para avaliação de subordinação que reduz o número de testes realizados. O classificador faz uma checagem nas bases de conhecimento utilizando o método de satisfabilidade para resolver possíveis problemas na hora de inferir as subordinações hierárquicas. O FaCT++ é um motor de inferência para OWL DL que usa a arquitetura de *ToDo List* para incorporar um amplo alcance de otimizações a serem feitas. O FaCT++ suporta as linguagens *SHOIQ*, OWL-DL e *SROIQ*. O FaCT++

¹³ *Hyper tableau* é um método de cálculo com as propriedades de completude e correção para as famílias da lógica de primeira ordem (BAUMGARTNER, 1998).

implementa um método de *tableau* para TBoxes (subsunção, satisfabilidade e classificação) e parcialmente para ABoxes (busca) (SATTLER, 2007).

3.5 Editores de ontologias para a linguagem OWL

Os editores são aplicações desenvolvidas para ajudar na criação e manipulação de ontologias. Permitem habilitar, inspecionar, visualizar, codificar e modificar ontologias, além de suportar o desenvolvimento e manutenção de tarefas inerentes (SURE, 2002).

Existem muitos editores, os quais variam em sua proposta, complexidade de modelos do conhecimento, usabilidade, escalabilidade, dentre outros (DAS; WU; MCGUINNESS, 2001). Sua grande variedade se encontra por causa da demanda em se suprir vários aspectos durante a construção de uma ontologia sendo a manutenção e a portabilidade os mais críticos. Atualmente as ontologias são desenvolvidas através destes editores exigindo-lhes cada vez mais funcionalidades.

Um editor de ontologias precisa fornecer uma interface adequada ao profissional da ciência da informação para construção e modificação de ontologias (STOJANOVIC; MOTIK, 2002). Além disso, existem funcionalidades que podem aumentar significativamente a usabilidade de um editor de ontologias e melhorar a experiência dos usuários.

Segundo um estudo apresentado por Cardoso (2007), mostra que os editores de ontologias mais usados por estudantes, professores, profissionais da ciência informação dentre outros são: Protégé¹⁴, SWOOP¹⁵, OntoEdit (Nomeado posteriormente para OntoStudio¹⁶), Editores de texto¹⁷, Altova SemanticWorks¹⁸, OilEd¹⁹, IsaViz²⁰, WebODE²¹, OntoBuilder²², WSMO Studio²³, Top Braid Composer²⁴, pOWL²⁵, dentre outros. Além disso, Khamparia e Pandey (2017) analisou trinta ferramentas para ontologias diferentes e avaliou-os segundo aspectos de funcionalidades disponíveis. Concluiu-se que apenas 20% das ferramentas analisados possuem 80% ou mais das funcionalidades disponíveis, como mostrado na Tabela 8. Destes 20% apenas 67% são editores de ontologias. Estes editores são: Protégé, Neon Toolkit²⁶, WebODE e Jena²⁷.

¹⁴ <http://protege.stanford.edu/>

¹⁵ <https://github.com/ronwalf/swoop>

¹⁶ <http://www.semafora-systems.com/en/products/ontostudio/>

¹⁷ Alguns usuários não usam nenhum tipo de editor de ontologia. Ao invés disso, eles preferem utilizar editores de texto simples.

¹⁸ <http://altova-semanticworks.soft112.com/>

¹⁹ <http://oiled.semanticweb.org/building/>

²⁰ <https://www.w3.org/2001/11/IsaViz/>

²¹ <http://mayor2.dia.fi.upm.es/oeg-upm/index.php/en/old-technologies/60-webode>

²² <http://ontobuilder.bitbucket.org/>

²³ http://www.wsmo.org/wsmo_tools.html

²⁴ <http://www.topquadrant.com/tools/modeling-topbraid-composer-standard-edition/>

²⁵ <http://powl.sourceforge.net/index.php>

²⁶ http://neon-toolkit.org/wiki/Main_Page.html

²⁷ <https://jena.apache.org/>

Tabela 8 – Características dos editores de ontologias

Recurso/Editor de Ontologias	Protégé	Neon Toolkit	WebODE	AeroDAML	Cerebra	Jena
Formato de importação	XML, RDF e XML Schema	RDF e OWL	RDF + RDFS	RDF e DAML	RDF + OWL	UML, RDF e OWL
Formato de exportação	XML, CLIPS, FLogic e RDF	RDF e OWL	RDF + RDFS	RDF e DAML	RDF + OWL	UML, RDF e OWL
Verificação de consistência	Sim	Sim	Sim	Sim	Sim	Sim
Extensibilidade	Sim	Sim	Não	Sim	Sim	Sim
Disponibilidade	Grátis	Grátis	Grátis	Grátis	Grátis	Grátis
Inferência embutida	Sim	Sim	Sim	Sim	Sim	Sim
Armazenamento de ontologias	Arquivos e DBMS	Arquivos	DBMS	Arquivos	Arquivos	Arquivos
Multiusuários	Sim	Não	Sim	Sim	Sim	Sim
Suporte <i>web</i>	Sim	Sim	Sim	Sim	Sim	Sim
Mesclagem	Sim	Sim	Sim	Sim	Sim	Sim
Versionamento	NI ²⁸	Sim	NI	Não	Sim	Sim
Consultas	Sim	Sim	Sim	NI	Sim	Sim
Motor de inferência	Pellet	Pellet2 e HermiT	NI	FaCT	Pellet	Pellet HermiT
Implementação	Java	Eclipse	Java	Java	Java	Java
Qualidade pragmática	Sim	Sim	Sim	Sim	Sim	Sim
Qualidade semântica	Sim	Sim	Sim	Sim	Sim	Sim
Qualidade sintática	Sim	Sim	Sim	Sim	Sim	Não
Colaborativo	Sim	Sim	Sim	Sim	Sim	Sim
Taxonomia gráfica	Sim	Sim	Sim	Sim	Sim	Sim
Arquitetura	<i>Standalone</i> ²⁹	<i>Standalone</i>	Cliente/Servidor	<i>Standalone</i>	<i>Standalone</i>	<i>Standalone</i>
Suporte à ontologias grandes	Sim	Sim	Sim	Sim	Sim	Sim
Bibliotecas de ontologias	NI	Sim	Não	Sim	Sim	Sim

Fonte: adaptado de (KHAMPARIA; PANDEY, 2017)

Este trabalho se propôs a analisar apenas as ferramentas consideradas como robustas, por conterem todas ou maioria das funcionalidades requisitadas por um profissional da ciência da informação. Tal análise tem como propósito o levantamento de requisitos para o protótipo de editor de ontologias.

3.5.1 Protégé

O Protege surgiu como uma ferramenta para eliciação de conhecimento médico (GENNARI et al., 2003) em 1987. Seu propósito inicial era a construção ferramentas para aquisição de conhecimento em planejamento médico. Em seguida, o Protégé evoluiu e se tornou em uma plataforma extensível para desenvolvimento e pesquisa de bases de conhecimento. As versões mais recentes podem ser executadas em uma variedade de plataformas, suportando ainda uma alta gama de personalizações. Incorpora o modelo de conhecimento OKBC³⁰, interage com modelos tradicionais de armazenamento como banco de dados relacionais, XML, e RDF e tem sido usados por muitas pessoas e grupos de pesquisa.

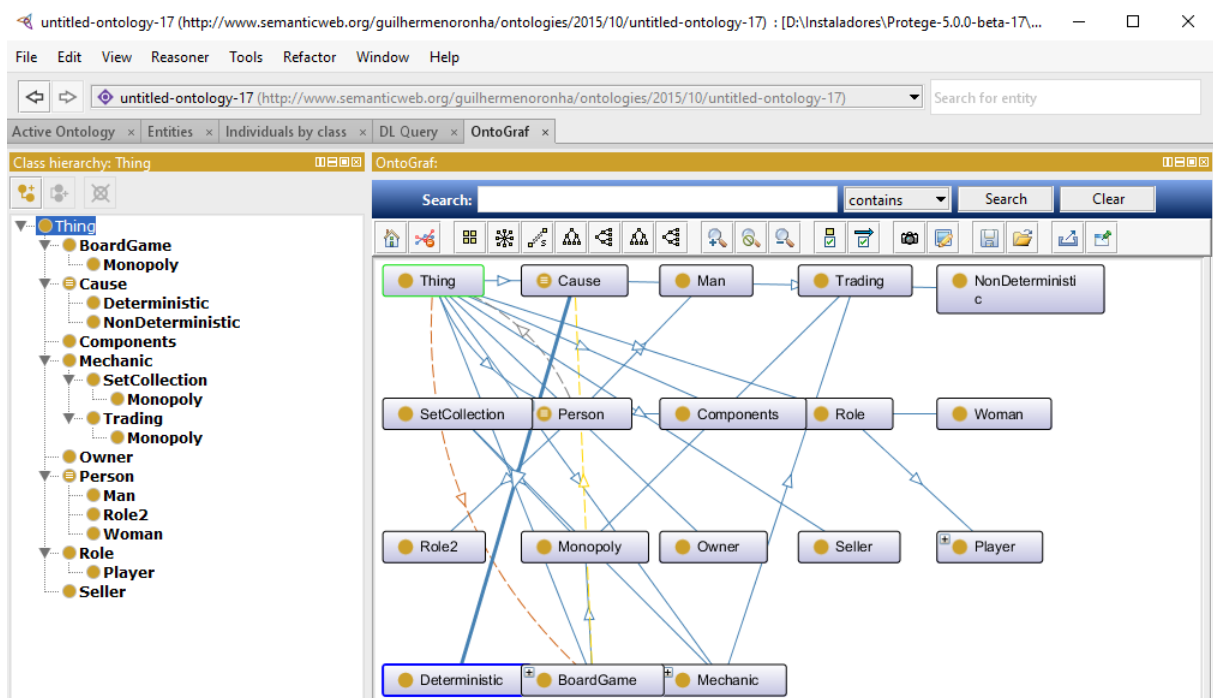
É uma ferramenta *open source* bastante suportada pela comunidade que fornece uma variedade grande de *plugins* que ajuda no entendimento e desenvolvimento de ontologias, assim como também na utilização de *reasoners* para fazer inferência nos sistemas de conhecimento. É escrita em Java e por isso oferece uso em alta variedade de sistemas. É mantido e desenvolvido pela universidade de Stanford junto com a universidade de Manchester. Contém mais de 200.000 usuários registrados e é até então o mais usado editor de ontologias (GASEVIĆ; DJURIĆ; DEVEDZIĆ, 2009). A Figura 6 mostra a versão 5 do Protégé.

²⁹ NI: Não identificado

²⁹ Programas são ditos *standalones* quando não necessitam de nenhum suporte para ser executado, dependendo apenas de si mesmo.

³⁰ <http://www.ai.sri.com/~okbc/>

Figura 6 – Protégé 5.0



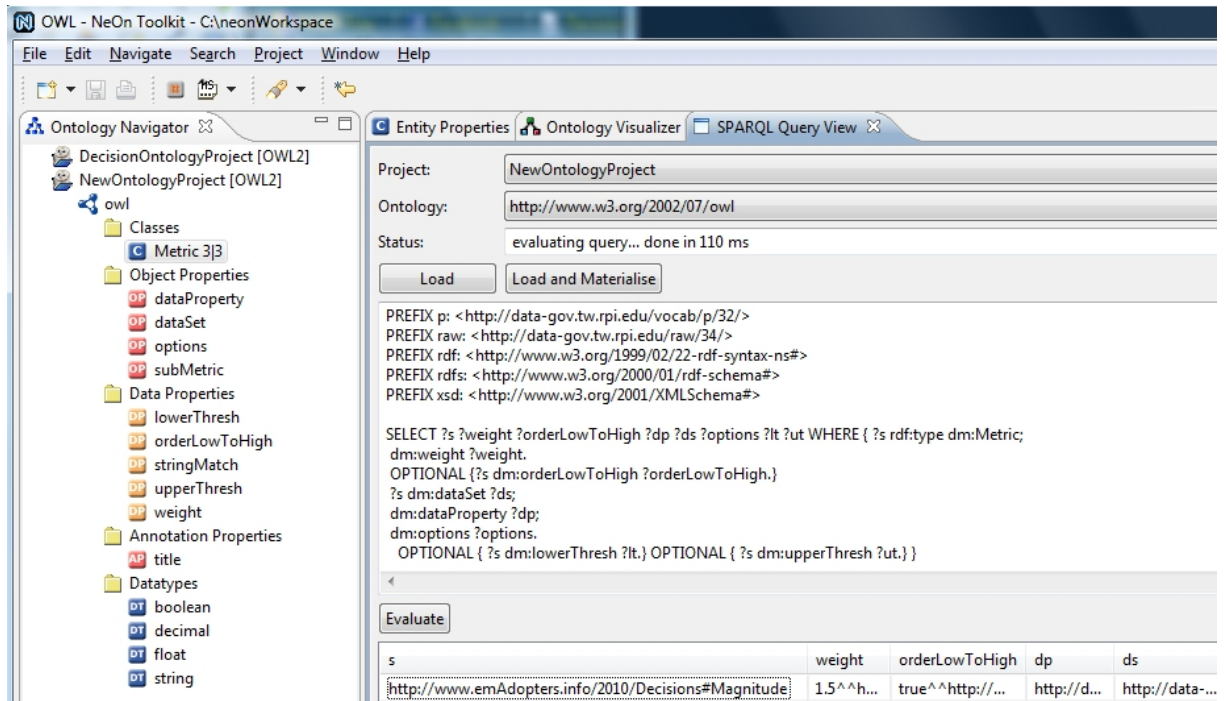
3.5.2 NeOn Toolkit

O NeOn Toolkit é uma nova geração de editor de ontologias que combina força industrial e robustez com um suporte para o ciclo de vida da engenharia de ontologias (HAASE et al., 2008). A comunidade alvo para este editor de ontologias inclui pesquisadores da *web* semântica bem como usuários profissionais com um respaldo comercial.

Construído em cima do Eclipse³¹, o NeOn Toolkit possui o núcleo de seu código-fonte disponibilizado *open source*. Foi desenvolvido em cima de uma arquitetura modular que inclui serviços de infraestrutura como repositórios e suporta componentes distribuídos para gerenciamento de ontologias como motores de inferência e colaboração mútua em redes de desenvolvimento. O NeOn Toolkit tem forte ênfase em gerenciamento de ontologias em rede. Por exemplo, uma ontologia é incorporada em uma rede de ontologias via relacionamentos semânticos, incluindo modelos para ontologias modulares e para mapeamento de ontologias. A Figura 7 mostra a tela do NeOn Toolkit.

³¹ <https://eclipse.org/>

Figura 7 – NeOn Toolkit



3.5.3 WebODE

O WebODE foi um editor de ontologias criado em 1999 e descontinuado em 2006. Ele era um editor de ontologias baseados em uma aplicação contida em um servidor. O seu núcleo era localizado em um servidor que era utilizado para ter acesso às ontologias que estavam conectadas em seu servidor (ARPÍREZ et al., 2001). O WebODE permitia editar e navegar por ontologias em um browser, e era baseado em HTML e Java *applets*³².

Sua construção foi feita para ser uma plataforma escalável, extensível e integrada que suporta a maior parte das atividades envolvidas no processo de desenvolvimento de ontologias. Entre esses processos se encontram a conceptualização, *reasoning*, exportação, etc. O editor conta ainda com uma série de serviços relacionados que permitem a interoperabilidade com outros sistemas de informação (ARPÍREZ et al., 2003).

As linguagens suportadas pelo WebODE são: XML, RDFS, OIL, DAML+OIL, OWL, CARIN (LEVY; ROUSSET, 1996), FLogic, Jess (FRIEDMAN-HILL, 2003) e Prolog. Com essas linguagens é possível realizar diversas tarefas como avaliação, documentação, mesclagem de ontologias e inferências. A Figura 8 mostra a tela deste editor na versão 2.0.

³² Applet é um pequeno *software* com função específica que executa dentro de uma aplicação maior (geralmente em browsers *web*)

Figura 8 – WebODE

The screenshot displays the WebODE 2.0 interface. At the top, there is a navigation bar with buttons for 'Show Term Properties', 'Graphical Taxonomy Edition', 'Intermediate Representations', 'Inference Engine', 'Instances', 'ODEClient', and 'Back'. A 'Clipboard' icon is circled in red. Below the navigation bar, the main content area is titled 'Instance Attributes for Term Travel'. On the left, a 'Browsing area' shows a tree view of the ontology with 'Travel' selected. The main area is an 'Edition area' with a table of instance attributes and a form for editing them.

Instance Attribute Name	Description	Value Type	Cardinality	Measurement Unit	Precision	Value Interval
arrival Date	Date of arrival of the trip	Date	(0, 1)			
company Name	Transportation company or companies in charge of a trip	String	(0, N)			
departure Date	Date of departure of the trip	Date	(0, 1)			
single Fare	Fare of a single ticket	Float	(0, 1)	US Dollar	0.01	0 -

The 'Edition area' form includes the following fields:

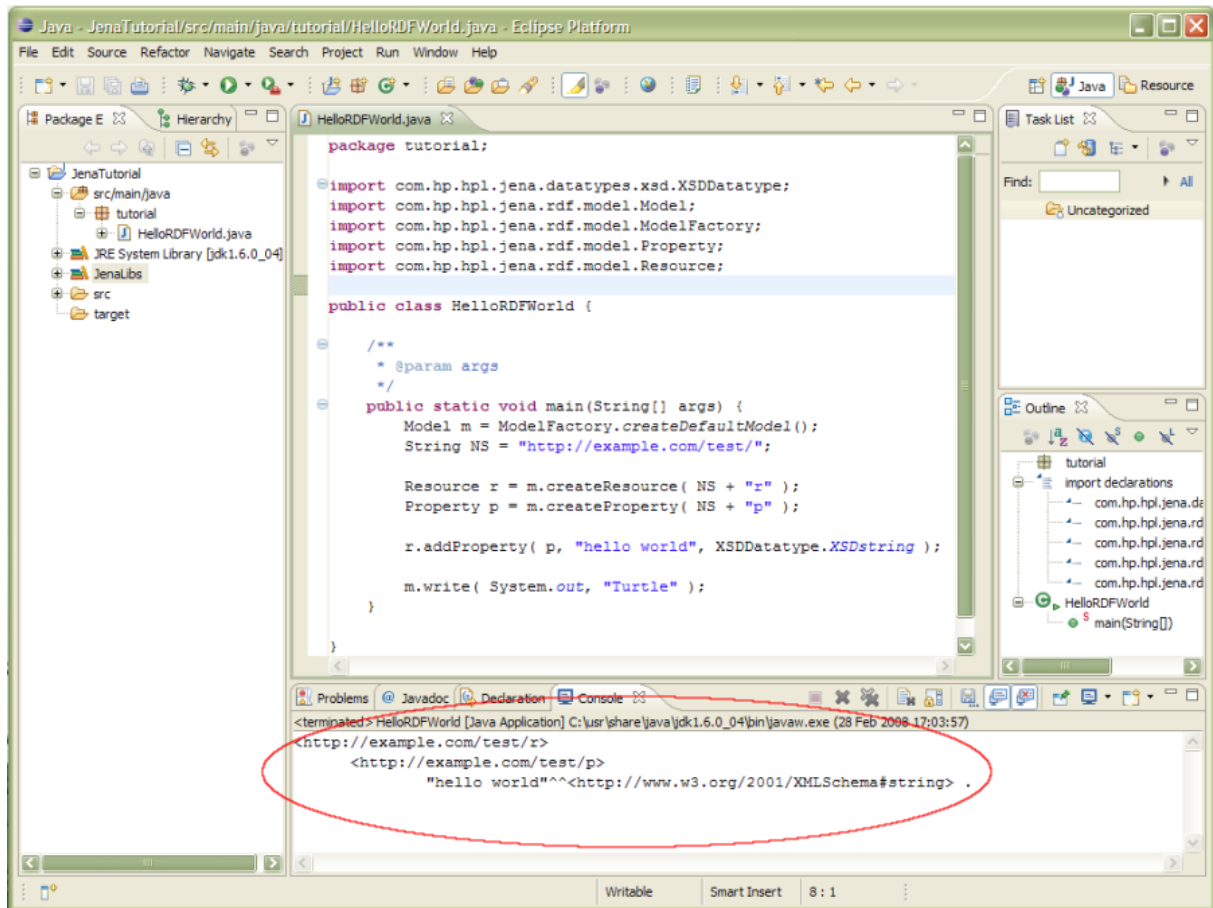
- Term Name: Travel
- Instance Attribute Name: return fare
- Description: Fare of a return ticket
- Value Type: Float
- Minimum/Maximum Cardinality: 0 | 1
- Measurement Unit: US Dollar
- Precision: 0.01
- Minimum Value: 0
- Maximum Value: (empty)

3.5.4 Jena

Jena é um *framework* baseado em Java com a proposta de desenvolvimento de aplicações de *web* semântica. Ele fornece uma extensiva quantidade de bibliotecas em Java para desenvolver código em RDF, RDFS, RDFa, OWL e SPARQL, com normas alinhadas com as recomendações da W3C. Desenvolvido em 2000 por McBride (2002) nos laboratórios da HP, o Jena surgiu como uma ferramenta *open source* com o intuito de facilitar o desenvolvimento de aplicações semânticas desenvolvidas para a *web*. Por ser bibliotecas escritas em Java, o Jena permite que seu uso seja portátil para IDEs que suportem o desenvolvimento em Java, como Eclipse e NetBeans³³. A Figura 9 mostra o funcionamento do Jena na IDE do Eclipse.

³³ <https://netbeans.org/>

Figura 9 – Jena API no Eclipse



Este editor de ontologias inclui um motor de inferência baseado em ontologias OWL e RDF, assim como uma variedade de estratégias para armazenagem de triplas RDF. Em 2009 o Jena recebeu novo foco pela HP para desenvolvimento da API e em 2010 o seu projeto de desenvolvimento foi adotado pela fundação Apache³⁴ onde tem seu desenvolvimento continuado.

³⁴ <https://www.apache.org/>

4 Metodologia

A revisão de literatura feita nos capítulos anteriores abordou os aspectos necessários para que a metodologia deste capítulo fosse construída. O Capítulo 2 tratou de lógica, suas origens, as bases para que se possa entender as sentenças lógicas e demonstrações de como é feita a inferência e derivação de conhecimento através de métodos dedutivos. Na seção 2.4 foi falado de inferências automáticas e como elas acontecem em linguagens descritivas. Também foi falado de regras de dedução lógica contidas em sistemas lógicos das quais serão a base de estudo para a criação de regras de transcrição de definições ontológicas. O Capítulo 3 abordou o conceito de ontologias e sua importância em sistemas de informação tecnológicos. Na seção 3.5 listou-se e estudou-se os principais editores de OWL. Estes editores serão a base para a prototipação da plataforma online onde será feita a transcrição de definições ontológicas.

Levando em consideração Lakatos e Marconi (2010), a metodologia da presente pesquisa foi classificada da seguinte maneira:

- Quanto a sua natureza: é uma pesquisa aplicada, pois tem como objetivo gerar aplicação prática aos usuários finais da informação.
- Quanto à abordagem do problema: ela é uma pesquisa qualitativa, pois os benefícios de um método de transcrição de definições ontológicas não pode ser mensurado.
- Quanto aos objetivos: ela é exploratória, pois ela faz uso da literatura para criar novas maneiras de lidar com o objeto de pesquisa.
- Quanto aos procedimentos técnicos ela é bibliográfica por fazer uso de material publicado e experimental por definir formas de tratar definições ainda não transcritos para linguagem lógica.

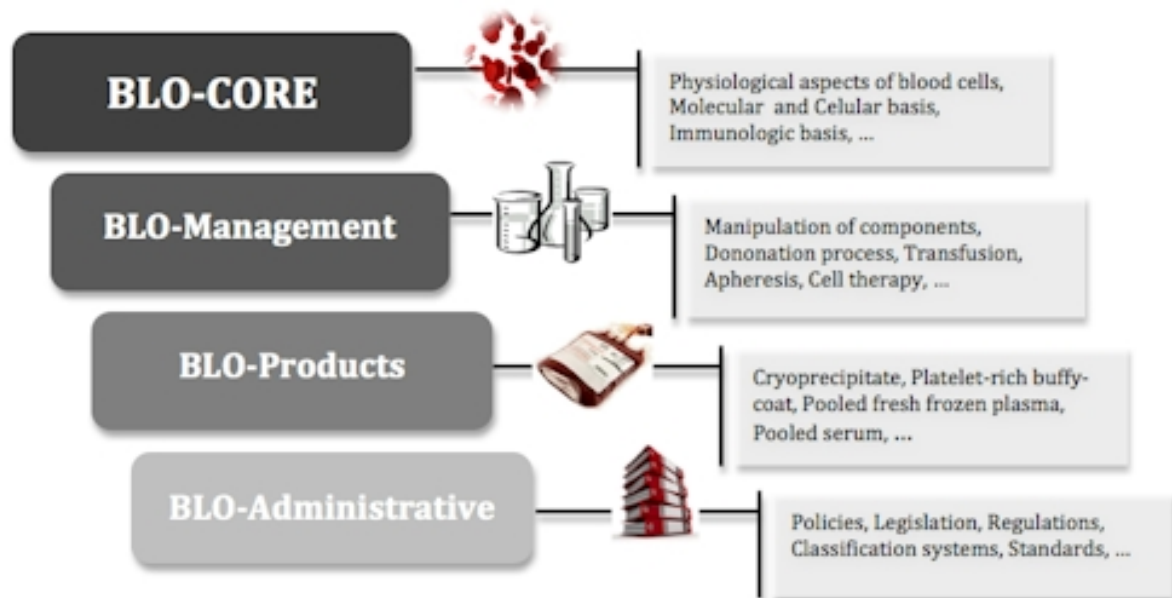
O restante deste capítulo está organizado da seguinte forma: a seção 4.1 apresenta a BLO, a ontologia que será o objeto dessa pesquisa; a seção 4.2 mostra a descrição da metodologia de pesquisa com o passo a passo de cada etapa realizada; a seção 4.3 entra no mérito tecnológico, descrevendo o passo a passo para o desenvolvimento de um software que será a plataforma online em questão; A seção 4.4 fala da coleta dos definições que vão ser formalizados e quais os critérios usados para a coleta dos mesmos; a seção 3.5 explica como é feita a transcrição de definições em uma ontologia biomédica, detalhando os procedimentos usados, e por fim; a seção 4.6 explica como serão feitos os testes de usuários e das definições transcritas definidas nesta metodologia.

4.1 A ontologia objeto da pesquisa

A ontologia que será usada no contexto empírico para exemplificar a metodologia proposta no presente trabalho será a *Blood Ontology*. A BLO é uma ontologia formal designada a descrever e formalizar os conceitos relacionados ao sangue. O projeto nasceu devido a constatação da falta de uma fonte sistematizada que descrevesse e organizasse dados relacionados ao sangue humano como dados sobre hematologia e hemoterapia. A falta destes dados causa problemas aos médicos, biólogos, pesquisadores e técnicos ao realizarem suas tarefas diárias envolvendo manipulação de sangue bem como outros assuntos relacionados ao sangue (ALMEIDA et al., 2011).

A BLO é um conjunto de várias ontologias relacionadas ao sangue, como mostrado na Figura 10. Cada uma destas ontologias tem como foco uma área específica relacionada ao sangue. A BLO-Core é onde estão formalizados os conceitos sobre as células do sangue, base imunológica, etc. A BLO-Management descreve procedimentos relacionados a manipulação de sangue como transfusão, terapia celular, etc. A BLO-Products está relacionada a formalização de "produtos" criados em consequência da manipulação de sangue. A BLO-Administrative está responsável pela formalização de políticas, legislação e regulamentos que dizem respeito ao sangue.

Figura 10 – Ontologias BLO



A Blood Ontology foi escolhida por ser uma ontologia que nasceu dentro da UFMG em parcerias com a Hemominas¹, *Ontology Research Group* (da universidade estadual de *New York*) e o CoE. Existem trabalhos recentes realizados na formalização de definições

¹ <http://www.hemominas.mg.gov.br/>

da BLO como por exemplo, o trabalho de Aganette (2015), que ainda necessita de uma formalização em lógica descritiva. Isto reforça a escolha da BLO como a ontologia objeto da presente pesquisa.

4.2 Descrição da metodologia de pesquisa

A revisão de literatura somada com os trabalhos correlatos feitos nos capítulos anteriores mostraram diversos problemas relacionados com a construção e manutenção de ontologias aplicadas. Existe a necessidade de se realizar um trabalho interdisciplinar que permeia até aos problemas em lidar com a lógica.

O extenso referencial teórico mostrou-se apenas uma parte do assunto referente ao desenvolvimento de ontologia aplicada. É necessário dominar linguagens OWL, escolher o editor de ontologia mais adequado, saber como funciona os motores de inferência, conhecer os fundamentos da representação da informação, conhecimento do domínio a ser representado e ainda por cima, conhecimento em lógica.

Os capítulos anteriores mostraram a relevância de desvendar a metodologia que envolve a transcrição de definições ontológicas. O estudo da lógica se mostrou bastante complexo, ainda que na sua visão mais rasa. É composto de muitos termos e simbologias que causam dificuldade e confusão aos usuários mais desacostumados do assunto. Para aliviar a gama de conhecimentos aprofundados que o ontologista carrega, se fez necessário a criação de um método de transcrição para definições formais escritos em linguagem natural para a linguagem lógica.

Para que os profissionais da ciência da informação mais desacostumados à tarefa de construir uma ontologia, se fez necessário também a construção de uma plataforma online cujo acesso esteja disponível para a didática da transcrição de definições ontológicas. Se fez necessário que esta plataforma seja simples e usual, feita para usuários leigos que desejam transcrever suas ontologias para lógica formal e verificar a sua consistência.

Para isto foi elaborado um plano metodológico para o presente trabalho. A Tabela 9 mostra os procedimentos metodológicos que serão aplicados nesta pesquisa. Estes procedimentos que estão detalhados brevemente nesta tabela serão explicados com mais detalhes nas seções posteriores.

Tabela 9 – Etapas da metodologia de pesquisa

Etapa	Sub-etapa	Breve descrição da etapa
Prototipação da plataforma online	Análise dos principais editores de ontologias	Verificar nos principais editores de ontologias similaridades e ideias relacionadas com os requisitos levantados
	Levantamento de requisitos	Descrição de todos requisitos a serem implementados no sistema
	Prototipação	Desenho do protótipo das telas da plataforma de acordo com os requisitos levantados
Coleta de dados	Coleta de definições a serem formalizados	Ler e analisar estudos de formalização feitos na BLO nos últimos cinco anos e selecionar definições que foram formalizadas em linguagem natural, mas não em lógica descritiva.
Transcrição das definições	Anotação da definição textual	Destacar expressões nominais e verbais da definição bem como a identificação de condições necessárias e suficientes
	Definição formal em OWL	Adicionar a definição anotada na ontologia de acordo com o editor de ontologias especificado
Avaliação	Inserção de indivíduos	Inserir indivíduos e relacioná-los de acordo com as definições transcritas
	Verificação de consistência	Avaliar se as transcrições geraram classes inconsistentes
	Avaliação de inferência	Verificar junto aos motores de inferência se conhecimento novo foi deduzido das transcrições

4.3 Prototipação da Plataforma Online

A plataforma online deverá funcionar nos *browsers* mais comuns como Microsoft Edge², Mozilla Firefox³, Google Chrome⁴ e Opera⁵. Seu design deve ser responsivo e capaz de rodar em tipos diferentes de dispositivos, com foco principal para *desktop*.

A prototipação da tela foi feita utilizando o *software* Balsamiq⁶, uma ferramenta para esboços de telas que especificarão o design de um software. O código desenvolvimento deverá ser depositado em uma plataforma de controle de versão *online* e gratuita. O código

² <https://www.microsoft.com/en-us/windows/microsoft-edge>

³ <https://www.mozilla.org/en-US/firefox/products/>

⁴ <https://www.google.com/chrome/>

⁵ <http://www.opera.com/pt-br>

⁶ <https://balsamiq.com/>

fonte de desenvolvimento deverá ser *open source* visando a contribuição para a comunidade científica e deverá estar sob a licença GPLv3⁷.

4.3.1 Análise dos principais editores de ontologias

A análise dos editores de ontologias foi feita com base no levantamento teórico levantado na seção 3.5. Para isso foram analisadas três características consideradas mais relevantes nestes editores:

- **Plataforma:** qual o tipo de plataforma que o editor disponibiliza para seus usuários (instalador, *standalone*, cliente/servidor, etc.);
- **Funcionalidades:** quais as funcionalidades que o editor suporta para o manuseio de ontologias, e;
- **Usabilidade:** o quão fácil e intuitivo é o uso do editor para usuários novatos.

4.3.2 Levantamento de Requisitos

O levantamento de requisitos para o sistema foram baseados em três aspectos considerados mais relevantes para o aprendizado de lógica. Eles são:

- **Privacidade e segurança:** política de privacidade e controle de dados fornecidos pelos usuários;
- **Navegação:** meios de navegação entre uma página e outra para facilidade do usuário;
- **Funcionalidades:** funções primárias de editores de ontologias:
 - **Manipulação de arquivos:** criar, editar e excluir ontologias e seus componentes como classes, propriedades e indivíduos;
 - **Inferências:** realização de inferências e resultados obtidos de inferências realizadas por motores de terceiros;
 - **Visualização:** ferramenta para visualização de classes, propriedades e indivíduos em forma de grafos, e;
 - **Tutorial:** explicação didática para transcrição de definições em linguagem natural.

⁷ <http://www.gnu.org/licenses/gpl-3.0.en.html>

4.3.3 Prototipação

O produto resultante das subseções 4.3.1 e 4.3.2 será externalizado em formato de modelos gerados pelo Balsamiq. Estes modelos são baseados em *websites* e representam uma ideia de como ficará o produto final do desenvolvimento da plataforma online.

4.4 Coleta de dados

O processo para a coleta de definições a serem formalizadas foi definido em conjunto com o orientador. A escolha dos dados se deu pela verificação de trabalhos feitos na escola que envolveram a formalização de definições da BLO em linguagem natural e sem a sua transcrição para lógica. O recorte utilizado se deu pelo trabalho mais atual neste âmbito com transcrições suficientes para a realização de testes.

4.5 Transcrição de definições

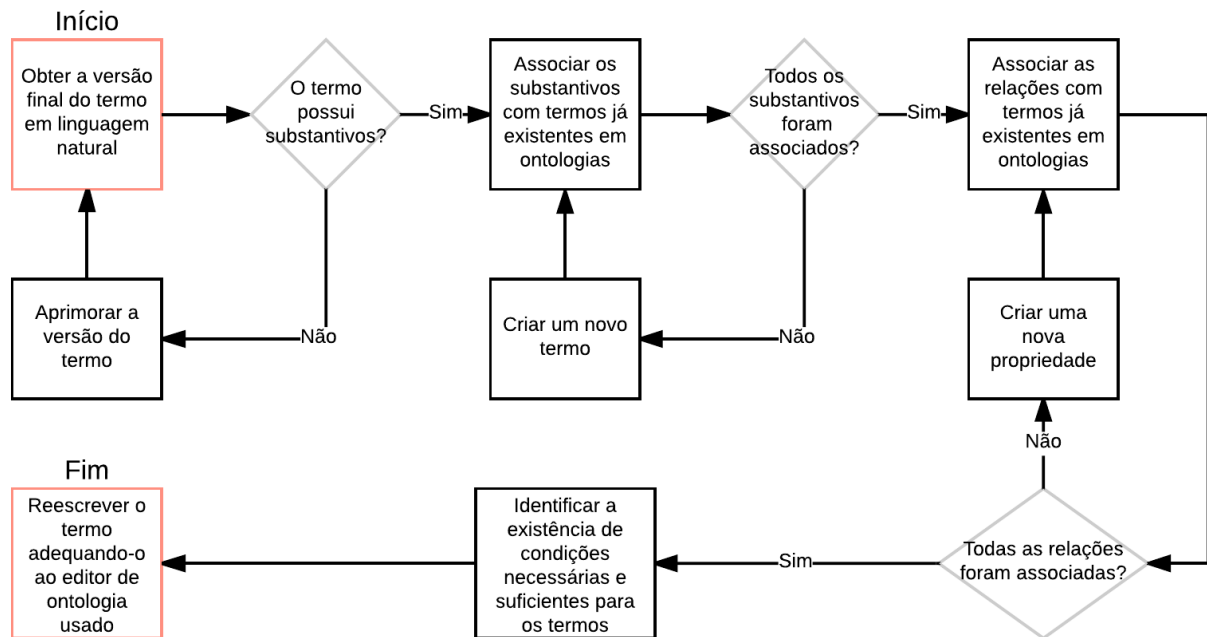
Conforme citado anteriormente neste trabalho, a metodologia pretendeu sistematizar a transcrição de definições da linguagem natural e depois verificar sua consistência e as inferências possíveis. Ou seja, dado uma proposição em linguagem natural, o mesmo deverá ser transformado em lógica descritiva que seja capaz de ser processada por um computador.

4.5.1 Anotação da definição textual

Esta metodologia envolve a identificação e marcação de termos e propriedades, seguindo o princípio de reuso de código (ontologias), criação de novos termos e propriedades, identificação de símbolos lógicos e condições necessárias e suficientes, e por fim, a reescrita da definição utilizando-se dos passos anteriores e adequando-os ao editor de ontologias utilizado.

A Figura 11 mostra um fluxograma resultante desta metodologia para a execução da transcrição destas definições.

Figura 11 – Fluxograma para a transcrição de definições em uma ontologia



4.5.1.1 Obter a versão final da definição em linguagem natural

O primeiro passo para a transcrição da definição é obter uma versão final descrevendo o significado deste em linguagem natural. Para isto recorre-se as metodologias de definição que reúnem condições necessárias e suficientes que atestem o significado de uma definição (SMITH et al., 2014). A versão final de uma descrição para uma ontologia deve ser uma definição que contém requisitos necessários e suficientes para sua unicidade e distinção perante as demais definições. Uma definição é constituída de sua definição e relações entre outras definições. Segundo Smith et al. (2014 apud AGANETTE; ALMEIDA, 2015), baseando-se no método aristotélico os seguintes passos são adotados para a construção de uma definição:

1. Separar o termo;
2. Obter a elucidação sobre o significado do termo de alguma fonte;
3. Estabelecer o *genus* superior no contexto de uso do termo;
4. Estabelecer características essenciais das coisas (sentido aristotélico), distinguindo o *genus* das espécies;
5. Formular e escrever a primeira versão da definição;
6. Verificar se a definição é uma declaração de condições necessárias e suficientes:
 - 6.1. Ser um *A* é condição **suficiente** para ser um *B*, então cada *A* é um *B* e;

- 6.2. Ser um A é condição **necessária** para ser um B , então cada B é um A .
7. Verificar deficiências na primeira versão da definição;
 - 7.1. Verificar princípio da não circularidade;
 - 7.2. Verificar e eliminar heranças múltiplas e;
 - 7.3. Verificar princípio da substituição.
8. Refinar e propor uma versão final.

Para a definição de uma relação entre os termos utiliza-se da mesma metodologia proposta por Smith et al. (2014 apud AGANETTE; ALMEIDA, 2015). Os passos adotados são:

1. Definir os tipos de relações;
2. Especificar quais expressões serão usadas para designar quais entidades são os *relata* nas relações correspondentes;
3. Analisar as limitações sobre os tipos de relações que podem ocorrer entre os relata;
4. Definir as primitivas no nível das instâncias;
5. Definir relação de subsunção em universais;
6. Definir a relação parte-todo em universais;
7. Definir a relação de participação em universais;
8. Definir a relação de localização em universais e;
9. Estabelecer a lógica das relações.

O resultado final, ou o produto resultante, desta etapa é uma anotação em linguagem natural, oriunda de um idioma, tal como português ou inglês. Exemplo da definição de *Disorder of hemostasis* retirada da BLO:

A disorder of hemostasis is a blood disorder related with the stoppage of blood flow through a blood vessel or body part. (4.1)

4.5.1.2 Associar os substantivos com termos já existentes em ontologias

O próximo passo da etapa é a identificar e associar os substantivos que compõe a definição criada na subseção 4.5.1.1. Os substantivos segundo Infante e Neto (2008) são nomeações ou classificações de objetos, fenômenos, lugares, qualidades, dentre outros e podem variar quanto ao seu número, grau e gênero. Os substantivos são classificados como comum, próprio, simples, composto, concreto, abstrato, primitivo, derivado e coletivo e descrevem as características que moldam o mundo e o pensamento.

Para a definição exemplificada na subseção 4.5.1.1 os substantivos identificados são: ***disorder of hemostasis***, ***blood disorder***, ***stoppage of blood flow***, ***blood vessel*** e ***body part***. Cada substantivo deverá ser identificado com uma definição pré-existente de outras ontologias já criadas ou identificados por definições oriundas da própria ontologia.

Para a verificação da existência de uma definição, é utilizado mecanismos de recuperação da informação voltados para ontologias. Os principais segundo Alexander (2011) são: MMI Ontology Registry and Repository⁸, BioPortal⁹, Protégé Ontology Library¹⁰ e Swoogle¹¹. Para este trabalho, utilizou-se o BioPortal por conter um vasto repositório de ontologias biomédicas aprovados por uma equipe especializada. Para os substantivos identificados foram encontradas as definições:

- ***Disorder of hemostasis***: é o termo a ser definido, seu resultado será a conjunção das definições dos outros substantivos;
- ***Blood disorder***: retirado de BLO, *disposition of hematopoietic human system to have abnormal conditions under certain bodiliy circumstances*;
- ***Stoppage of blood flow***: definição não encontrada;
- ***Blood vessel***: retirado da FMA¹², *Segment of blood vessel tree organ which is a tubular structure the lumen of which contains blood; together with other spatially associated blood vessels constitute blood vessel tree organ, e*;
- ***Body part***: retirado da NCIT¹³, *Any part of an organism*.

O termo inicial produto da subseção 4.5.1.1 deve ser reescrito, adicionando-se as ontologias e o nome do termo dado naquela ontologia entre colchetes. Os termos definidos pela própria ontologia e/ou não definidos são marcados, porém não se adiciona os colchetes:

⁸ <http://mmisw.org/orr/>

⁹ <http://bioportal.bioontology.org/>

¹⁰ http://protegewiki.stanford.edu/wiki/Protege_Ontology_Library

¹¹ <http://swoogle.umbc.edu/>

¹² <http://si.washington.edu/projects/fma>

¹³ <https://ncit.nci.nih.gov/ncitbrowser/>

*A disorder of hemostasis is a blood disorder related with the stoppage of blood flow through a blood vessel[FMA: *blood vessel*] or body part[NCIT: *body part*].*
(4.2)

4.5.1.3 Criar um novo termo

Os substantivos com definições não identificadas na subseção 4.5.1.2 deverão ser transformados em novos termos da ontologia. Estes termos serão criados utilizando-se da metodologia aristotélica mostrada na subseção 4.5.1.1 e a sua transcrição para lógica deverá ser expandida futuramente.

A criação de novos termos, ainda que sem sua transcrição lógica, permite que a transcrição do termo atual proceda para as etapas subsequentes. De acordo com o exemplo, os termos ***stoppage of blood flow***, ***body part*** e ***blood vessel*** foram criados em BLO, sendo os dois últimos importados de outras ontologias.

4.5.1.4 Associar as relações com termos já existentes em ontologias e/ou lógica descritiva

O produto resultante da subseção 4.5.1.1 foi a versão do termo final com os substantivos marcados por definições importadas de outras ontologias. Os demais termos que não tiveram suas definições encontradas em base de dados biomédicas, foram criados na ontologia corrente.

O próximo passo da etapa é semelhante ao apontado na subseção 4.5.1.1. Serão identificados as relações que os termos têm entre si, ou seja, os verbos e conectivos que ligam um substantivo ao outro. Estas definições são encontradas em ontologias como a RO¹⁴ e a BFO¹⁵. A lógica descritiva suporta naturalmente algumas relações em sua linguagem. Essas relações devem ser identificadas e substituídas por operadores lógicos, ao invés de definições formais importadas de outras ontologias. A Tabela 10 mostra estas relações¹⁶:

¹⁴ <http://obofoundry.org/ontology/ro/>

¹⁵ <http://ifomis.uni-saarland.de/bfo/>

¹⁶ A Tabela 10 contém apenas algumas das relações ontológicas possíveis para os operadores lógicos.

Tabela 10 – Operadores da lógica descritiva e as relações ontológicas

Operadores	Relações ontológicas	Definição
\equiv	<i>Same, equivalent, equal etc.</i>	Equivalência entre duas sentenças complexas
\sqsubseteq	<i>Subset, part of, including, etc.</i>	Um termo é subconjunto e está contido dentro de outro termo
:	<i>Is a</i>	Definição de um termo através de uma sentença complexa
\wedge	<i>And</i>	Conjunção entre duas sentenças complexas
\vee	<i>Or</i>	Disjunção entre duas sentenças complexas
\neg	<i>Not</i>	Negação de uma definição
P^-	<i>Inverse of, contrary, antonym, etc.</i>	Definição de um termo como sendo exatamente oposto a outro

No exemplo supracitado as relações encontradas foram: ***is a, related with, through a*** e ***or***.

- ***Is a:*** relação de equivalência lógica que é substituído pelo símbolo \equiv ;
- ***Related with:*** retirado da *HL7 Reference Information Model*¹⁷, *An action taken with respect to a subject Entity by a regulatory or authoritative body with supervisory capacity over that entity. The action is taken in response to behavior by the subject Entity that body finds to be undesirable;*
- ***Through a:*** retirado da NCIT (*over*), definição não encontrada, e;
- ***Or:*** relação de disjunção lógica que é substituído pelo símbolo \vee .

O termo resultante da subseção 4.5.1.4 é então reescrito adicionando-se os novos termos identificados e os operadores lógicos. Os termos são adicionados em negrito para diferenciá-los dos adicionados na subseção 4.5.1.2.

A disorder of hemostasis : blood disorder

related with [HL7: *related*] the stoppage of blood flow ***through a***[NCIT: *over*]

(blood vessel[FMA: *blood vessel*] \vee body part[NCIT: *body part*]). (4.3)

¹⁷ <http://www.hl7.org/implement/standards/rim.cfm>

4.5.1.5 Criar uma nova propriedade

As relações identificadas na subseção 4.5.1.4 deverão ser transformadas em novas propriedades da ontologia. Estas propriedades serão criadas utilizando-se da metodologia aristotélica mostrada na subseção 4.5.1.1 e a sua configuração deverá ser expandida futuramente.

As propriedades criadas, ainda que sem suas configurações definidas, permite que a transcrição do termo atual proceda para as etapas subsequentes. De acordo com o exemplo, as relações *related* e *over* foram criados, importados de suas ontologias respectivamente.

4.5.1.6 Identificar a existência de condições necessárias e suficientes para os termos

As condições necessárias e suficientes para os termos é a identificação das variáveis livres e ligadas que compõem a sentença, conforme mostrado na subseção 2.2.5. No entanto tal identificação que é eficaz para a lógica de primeira ordem, mostra-se insuficiente para a lógica descritiva. Neste caso utilizou-se dos conceitos de necessidade e suficiência aliados à prova lógica para a definição dos termos. Devlin (2003) faz três suposições sobre necessidade e suficiência dos termos:

- A é necessário para B , se $B \rightarrow A$;
- A é suficiente para B , se $A \rightarrow B$, e;
- A é suficiente e necessário para B se $A \leftrightarrow B$.

A relação de necessidade em ontologia aplicada é dada pela relação classe-subclasse. O fato de uma classe ser subclasse de outra já implica em uma relação de necessidade.

A relação de suficiência é dado pela relação do quantificador existencial \exists . Uma classe é suficiente de outra classe se ela possui um ou mais atributos (alguns) correspondentes da classe.

A relação de suficiência e necessidade é dada pela relação do quantificador universal \forall . Uma classe é necessária e suficiente de outra classe se ela possui todos os atributos correspondentes da classe.

Para determinar se um termo é necessário, suficiente ou ambos utilizou se da negação da suficiência e necessidade. Seja $A \leftrightarrow B$ a representação da condição necessária e suficiente, sua negação pode ser dada equação $(A \wedge \neg B) \vee (B \wedge \neg A)$. A Tabela 11 mostra a prova da bicondicional e sua negação.

A Tabela 11 mostra que se um termo não é necessário e suficiente de outro, ou ele é necessário, ou suficiente. Com base nisto, serão verificadas nas descrições formalizadas a

Tabela 11 – Tabela-verdade para os termos $A \leftrightarrow B$ e $(A \wedge \neg B) \vee (B \wedge \neg A)$.

A	B	$A \leftrightarrow B$	$(A \wedge \neg B) \vee (B \wedge \neg A)$
V	V	V	F
V	F	F	V
F	V	F	V
F	F	V	F

presença de necessidade e suficiência dos termos. Em caso de ausência deste, será verificada ou a necessidade, ou a suficiência.

Além das relações de necessidade e suficiência, existem também as limitações de cardinalidade que são associados com $\exists!$ para exatamente um e $\geq, >, \leq, <$ e $=$ para intervalos numéricos. A Tabela 12 mostra os tipos de restrições necessárias e suficientes.

Tabela 12 – Condições necessárias e suficientes e suas relações ontológicas

Condições	Relações ontológicas	Definição
\forall	<i>All, For all, every, etc.</i>	Todos os termos correspondem às definições estabelecidas
\exists	<i>Exists, there is, etc.</i>	Existe pelo menos um termo que corresponde às definições estabelecidas
$\exists!$	<i>Only one, exactly one, etc.</i>	Existe exatamente um termo que corresponde às definições estabelecidas
$=$	<i>Equal, even, like, etc.</i>	A cardinalidade de um termo é igual a um número real x
$>$	<i>Greater, more than, bigger, etc.</i>	A cardinalidade de um termo é maior que um número real x
\geq	<i>Greater or equal, etc.</i>	A cardinalidade de um termo é igual ou maior do que um número real x
$<$	<i>Less, smaller, etc.</i>	A cardinalidade de um termo é menor que um número real x
\leq	<i>Less or equal, etc.</i>	A cardinalidade de um termo é menor ou igual a um número real x

O exemplo supracitado não dá indícios em sua definição das condições necessárias e suficientes. Neste caso recorre-se a literatura onde foi feita a definição em busca destes indícios. Na ausência de definições concludentes recorre-se ao especialista do domínio para tal identificação. Na ausência deste, cabe ao autor a interpretação da definição.

A relação *is a* explicita a condição de necessidade entre *disorder of hemostasis* e *blood disorder*. As relações *related with* e *through a* com suas respectivas classes esclarece apenas a suficiência entre *disorder of hemostasis* e algumas das características de *blood disorder* que identificam esta causa. Neste caso, a desordem do sangue têm como algumas

causas: estar relacionada com a interrupção do fluxo sanguíneo através dos vasos sanguíneos ou parte do corpo.

Portanto foram associadas às classes *stoppage of blood flow*, *blood vessel* e *body part* o quantificador existencial \exists como mostra a Equação 4.4:

A disorder of hemostasis : *blood disorder*

$$\textit{related with} [HL7: \textit{related}] \textit{ the } \exists \textit{ stoppage of blood flow through a} [NCIT: \textit{over}] \\ \exists (\textit{blood vessel} [FMA: \textit{blood vessel}] \vee \exists \textit{ body part} [NCIT: \textit{body part}]). \quad (4.4)$$

4.5.2 Definição formal em OWL

Uma vez definidas as anotações dos termos que serão importados e/ou criados na própria BLO, a próxima e última etapa do fluxograma mostrado na Figura 11 contempla a inserção do termo transcrito na ontologia. Para isto, usou-se um editor de ontologias dentre os analisados na seção 3.5 para que fosse feita a inserção em OWL.

4.5.2.1 Reescrevendo a definição no editor de ontologia

Para reescrever a definição utilizou-se um processo de três etapas:

1. **Identificar a definição a ser reescrito:** localizar em BLO o termo que será reescrito. Caso não seja encontrado ele deverá ser criado utilizando-se a metodologia da subseção 4.5.1.3;
2. **Reescrever a definição em linguagem natural:**¹⁸ etapa dividida em cinco sub-etapas:
 - 2.1. Substituir os símbolos da lógica descritiva por palavras correspondentes;
 - 2.2. Substituir os termos originais pelos termos que serão importados de outras ontologias;
 - 2.3. Substituir as propriedades originais pelas propriedades que serão importadas de outras ontologias;
 - 2.4. Conectar as propriedades com cada termo à qual ela se relaciona, e;
 - 2.5. Conectar as diferentes propriedades entre si utilizando-se o conectivo lógico *and* (\wedge).
3. **Editar o termo:** o termo deverá ser editado utilizando-se relações de equivalência e hierarquia de classes presentes nos editores de ontologias. Termos como *is_a* devem

¹⁸ Esta etapa visa a facilitação do entendimento para a transcrição do termo em um editor de ontologias.

ser relacionados como sendo subclasse de outra. As propriedades devem ser associadas com suas classes e suas respectivas restrições de cardinalidade, como quantificadores universais e existenciais.

Para a transcrição do termo da Equação 4.4, utilizou-se o Protégé na versão 5.2.0. A primeira etapa consistiu localizar o termo *disorder of hemostasis*. Ela é subclasse de *blood disorder* e é relacionada com as demais.

A segunda etapa consistiu em reescrever o termo:

1. \exists e \wedge foram substituídos por *some* e *and* respectivamente;
2. Nenhuma classe foi substituída;
3. As propriedades *related with* e *through a* foram substituídas por *related* e *over* respectivamente;
4. A propriedade *related* está relacionada à classe *stoppage of blood flow* e *over* está relacionada às classes *blood vessel* e *body part*, e;
5. As propriedades *related* e *over* foram conectadas pela palavra *and*.

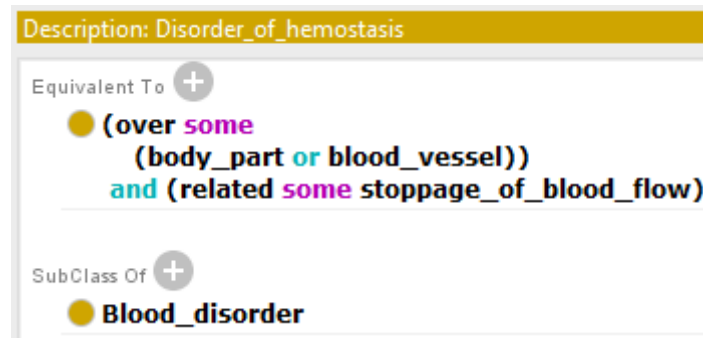
O produto resultante da segunda etapa foi a seguinte sentença:

$$\begin{aligned}
 & \textit{Disorder of hemostasis is a subclass of blood disorder} \\
 & \quad \textit{and is related some stoppage of blood flow and over} \\
 & \quad \quad \textit{some blood vessel or over some body part.} \quad (4.5)
 \end{aligned}$$

O quantificador existencial utilizado nas classes associadas tem seus termos explicitados. A relação entre *disorder of hemostasis* como sendo subclasse de *blood disorder* também é explicitada. Nota-se que a relação de subclasse com as demais relações são conectadas pelo conectivo lógico *and* (\wedge). As relações entre os termos *related* e *over* são conectadas entre si pelo conectivo lógico *and* (\wedge). Desta maneira temos uma sentença logicamente conectada que pode ser transportada para um editor que é realizado na terceira etapa.

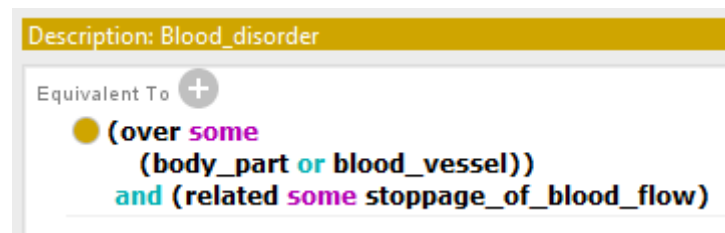
No Protégé foi utilizado o *Class expression editor* para alterar a relação de equivalência da classe que é sustentada na Equação 4.4 pelo símbolo ":". O resultado é mostrado na Figura 12.

Figura 12 – Definição de *disorder of hemostasis* no Protégé



A subseção 4.5.1.6 também definiu parcialmente o termo *blood disorder* listando algumas de suas características como: interrupção do fluxo sanguíneo e através dos vasos sanguíneos ou parte do corpo. Fez-se necessário atribuir esta definição na classe em questão como é mostrado na Figura 13.

Figura 13 – Atribuição de condições para *blood disorder* no Protégé



4.6 Avaliação

De acordo com a metodologia descrita na seção 4.5, os testes deverão ser feitos sobre os termos transcritos pela seção 4.5 avaliando sua consistência e inferência. Os termos transcritos seguindo a metodologia apresentada na seção 4.5, serão incorporados à BLO. Após esta incorporação, será feita uma avaliação junto aos motores de inferência analisados na seção 3.4. Esses motores verificarão a estrutura da BLO após a adição dos novos indivíduos, identificando se há problemas de inconsistência. Além disso, será observado se os motores serão capazes de inferir novos relacionamentos entre os indivíduos e termos adicionados.

4.6.1 Inserção de indivíduos

A inserção de indivíduos contemplará todas as classes ao qual o termo transcrito se relaciona. É preciso criar uma teia de relacionamentos que contemple dois casos considerados importantes para verificação de equivalência:

- **Indivíduo da definição principal:** o indivíduo será definido como pertencente à classe ao qual foi renomeada, e;

- **Indivíduos como conjunto de definições:** o indivíduo será definido como pertencente ao resultado da transcrição da classe original.

Para o exemplo supracitado foram criados oito indivíduos. Sua nomenclatura é uma abreviação de suas classes principais. Eles são:

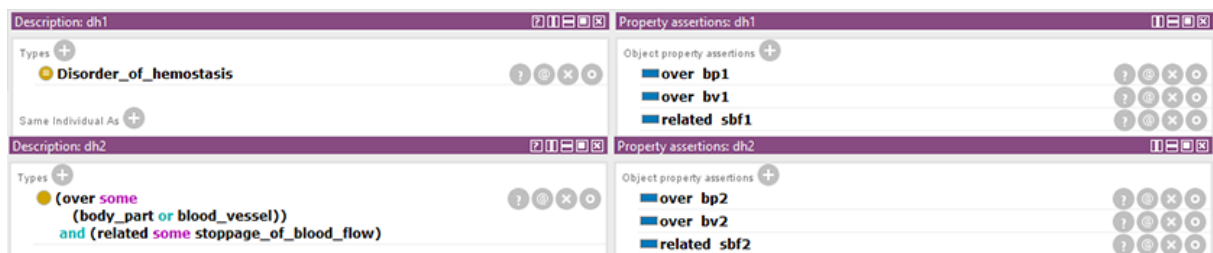
- **dh1 e dh2:** o primeiro como indivíduo de *disorder of hemostasis* e o segundo como indivíduo da equivalência de *disorder of hemostasis* mostrada na Figura 12;
- **sbf1 e sbf2:** indivíduos da classe *stoppage of blood flow*;
- **bv1 e bv2:** indivíduos da classe *blood vessel*, e;
- **bp1 e bp2:** indivíduos da classe *body part*.

Após a sua criação estes indivíduos precisaram ser relacionados um com os outros de forma que obedeçam a regra lógica criada para o termo *disorder of hemostasis*. Neste caso ficou:

- **dh1:** *related* sbf1, *over* bv1 e *over* bp1, e;
- **dh2:** *related* sbf2, *over* bv2 e *over* bp2.

O resultado da inserção é mostrado na Figura 14:

Figura 14 – indivíduos para *disorder of hemostasis* no Protégé

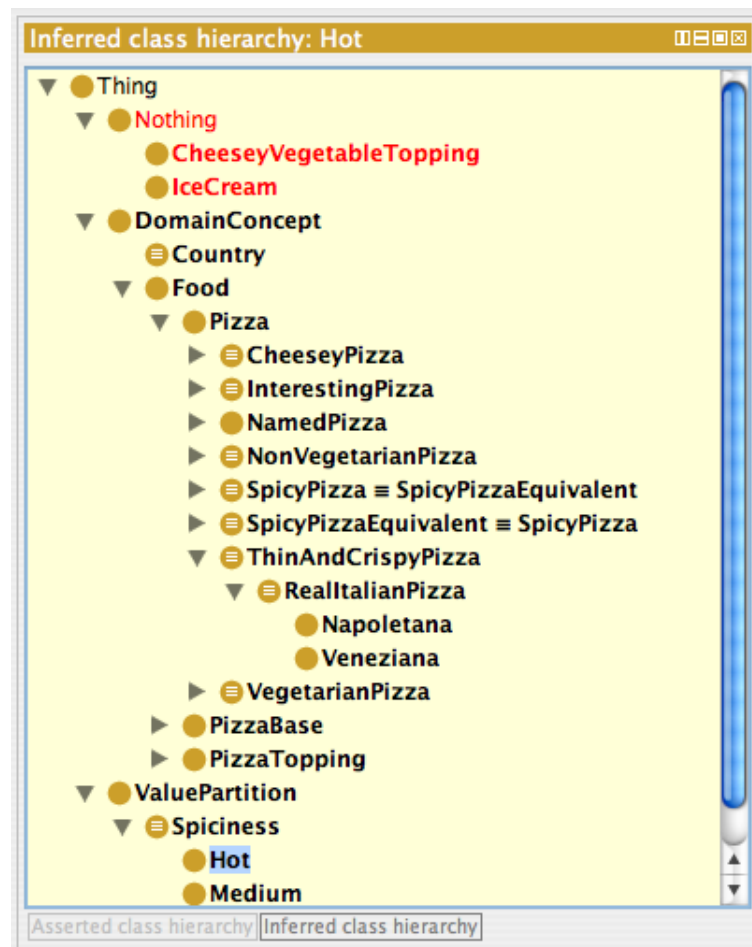


4.6.2 Verificação de inconsistência

Uma ontologia é consistente se e somente se todos os seus termos não desobedecem às regras de satisfabilidade, equivalência e subsunção mostrados na subseção 2.4.2. Se uma classe não respeita uma dessas regras, diz-se que ela é inconsistente. Uma ontologia inconsistente é dita se ela possui uma ou mais classes inconsistentes.

A inconsistência é mostrada no Protégé após a sua inicialização do motor de inferência. As classes inconsistentes são marcadas de vermelhas pelo Protégé como mostra a Figura 15.

Figura 15 – Inconsistências no Protégé

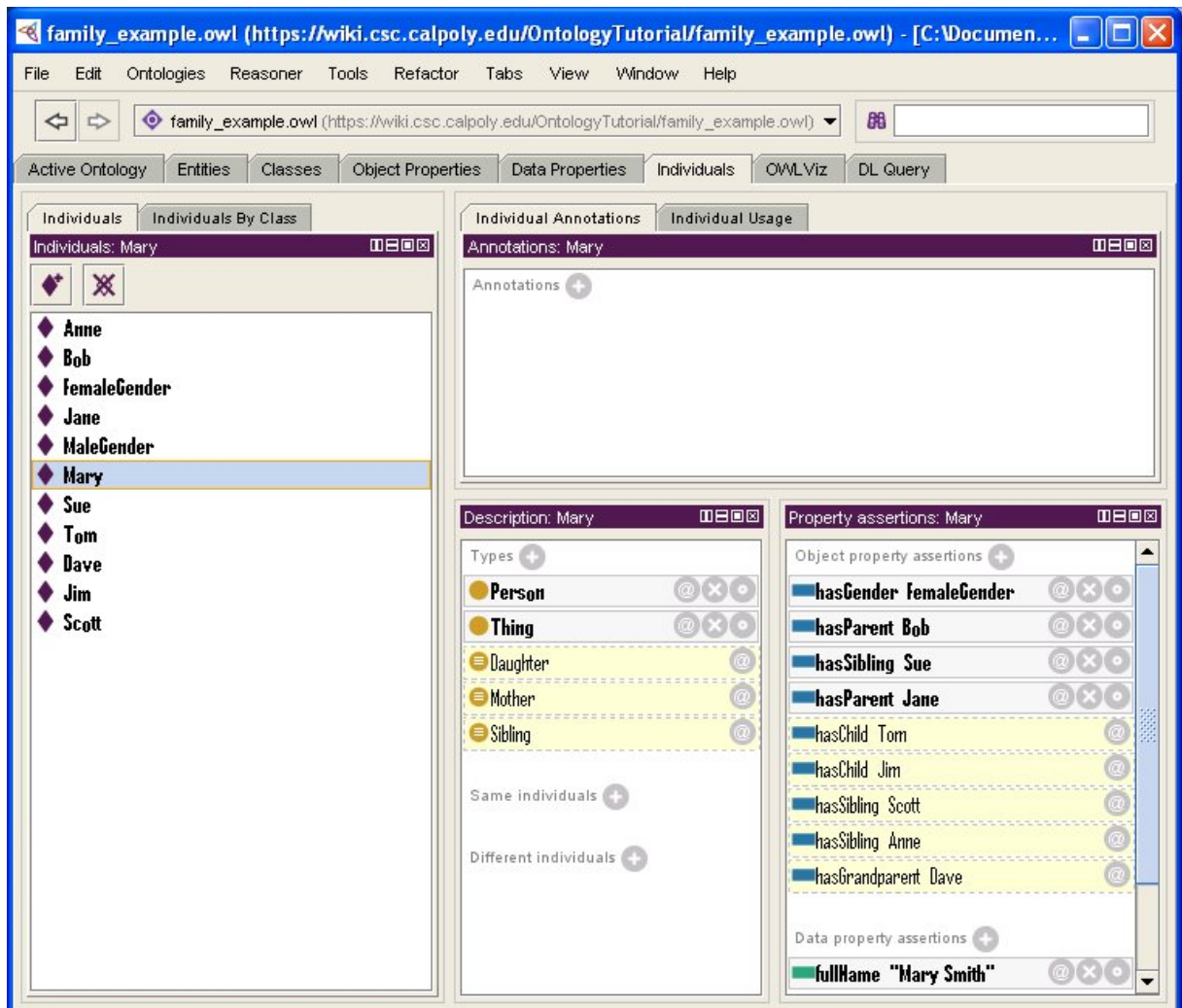


4.6.3 Avaliação de inferências

Esta etapa de avaliação consistiu na utilização de motores de inferências analisados na seção 3.4 que são compatíveis com o editor de ontologia escolhido para os testes, no caso, o Protégé. Os motores disponíveis que são compatíveis com a versão 5.2.0 do Protégé foram o Pellet, Hermit e FaCT++.

Os motores de inferência foram executados e as classes modificadas, assim como os indivíduos adicionados, foram analisados para a constatação de dedução de conhecimento novo. A Figura 16 mostra a disposição do Protégé para sinalizar conhecimento novo deduzido dos motores de inferência. O conhecimento novo deduzido está grifado de amarelo.

Figura 16 – Dedução de conhecimento novo no Protégé



5 Análise e Resultados

A metodologia proposta no Capítulo 4 desenvolveu-se neste capítulo em duas partes distintas. A seção 5.1 trata da formalização da transcrição de termos biomédicos, que foram pré-definidos em linguagem natural, para a linguagem lógica utilizada em editores de ontologias. Os termos recolhidos do trabalho de Aganette (2015) foram transcritos utilizando-se o editor de ontologia Protégé avaliado na subseção 3.5.1. Sua avaliação foi feita da análise de motores de inferências analisados na seção 3.4 que tem suporte para o Protégé.

A seção 5.2 trata da prototipação de uma ferramenta online que deverá ser capaz de carregar, editar e fazer raciocínio de ontologias. Tendo como proposta principal a didática do ensino de lógica e transcrição de termos biomédicos, o OntoLogica (assim nomeado), foi desenhado para os profissionais da ciência da informação iniciantes na área de ontologia aplicada. Seu protótipo foi desenhado baseado nos editores analisados na seção 3.5 e suas funcionalidades explicadas para um desenvolvimento futuro.

5.1 Transcrição e avaliação dos termos

Baseado nas reflexões feitas na seção 2.2, a presente seção dedicou-se em formular um guia para a transcrição de definições feitas para termos em uma ontologia. Para isto seguiu uma metodologia padronizada para a transcrição destes termos, de forma que todos os casos sejam contemplados.

Os termos recolhidos para transcrição estavam contidos no trabalho de Aganette (2015) e são:

- *Bleeding disorder associated with acquired disorders of coagulation*
- *Bleeding disorder after extracorporeal circulation*
- *Bleeding disorder after massive transfusion*
- *Bleeding disorder associated with accelerated destruction of coagulation factors*
- *Bleeding disorder caused by fibrinolysis*

5.1.1 *Bleeding disorder associated with acquired disorders of coagulation*

Segundo Aganette (2015), o termo foi definido da seguinte maneira:

Uma hemorragia associada com desordens de coagulação adquiridas (HADCA) é um transtorno de coagulação do sangue o qual é causada por um defeito, o qual prende o fluxo do sangue dentro dos vasos, de natureza composta. (5.1)

Neste caso fez-se necessário traduzir o termo para o inglês, língua padrão da BLO:

Bleeding disorder associated with acquired disorders of coagulation is a disorder of blood coagulation which is caused by a defect, which retains blood flow inside the vessels, of compound nature. (5.2)

A partir daí segue-se o fluxograma mostrado na Figura 11.

5.1.1.1 Identificando os substantivos para *bleeding disorder associated with acquired disorders of coagulation*

Os substantivos identificados foram: ***bleeding disorder associated with acquired disorders of coagulation***, ***blood coagulation disorder***, ***defect***, ***blood flow*** e ***vessel***. Para os substantivos encontrados foram dadas as seguintes definições:

- ***Bleeding disorder associated with acquired disorders of coagulation***: termo a ser definido. Seu resultado será a conjunção das definições de outros substantivos;
- ***Blood coagulation disorder***: retirado de BLO (*disorder of blood coagulation*), *disorder of homeostasis characterized by the disruption of the body's ability to control blood clotting*;
- ***Defect***: retirado de NCIT, *a failing or deficiency*;
- ***Blood flow***: retirado de NCIT (*blood circulation*), *the movement of blood to and from the heart and to the rest of the body's organs, tissues, and cells through the network of arteries, veins, and capillaries*, e;
- ***Vessel***: retirado de BLO (*blood vessel*), *segment of blood vessel tree organ which is a tubular structure the lumen of which contains blood; together with other spatially associated blood vessels constitute blood vessel tree organ*.

Associando os termos identificados com suas respectivas ontologias obteve-se o termo 5.3:

Bleeding disorder associated with acquired disorders of coagulation is a blood coagulation disorder[BLO: disorder of blood coagulation] which is caused by a defect[NCIT: defect], which retains blood flow[NCIT: blood circulation] inside the vessels[BLO: blood vessel], of compound nature. (5.3)

5.1.1.2 Criando novos termos para *bleeding disorder associated with acquired disorders of coagulation*

Dos substantivos identificados, os termos **defect** e **blood flow** foram importados de suas respectivas ontologias. Os termos foram criados como sendo subclasses de *Uncategorized*.

5.1.1.3 Identificando as relações do termo *bleeding disorder associated with acquired disorders of coagulation*

As relações encontradas foram: **is a**, **caused by**, **retains**, **inside** e **of compound nature**.

- **Is a:** relação de equivalência lógica que é substituído pelo símbolo ::
- **Caused by:** retirado de NCIT (*due to*), *caused by or resulting from*;
- **Retain:** retirado de NCIT, *to maintain possession of something*;
- **Inside:** retirado de NCIT (*inner*), *inside or closer to the inside of the body or object*, e;
- **Of compound nature:** definição não encontrada;

Associando as relações identificadas com suas respectivas ontologias obteve-se o termo 5.4:

*Bleeding disorder associated with acquired disorders of coagulation : blood coagulation disorder[BLO: disorder of blood coagulation] **caused by**[SNMI:due to] a defect[NCIT: defect], which **retains**[NCIT:retain] blood flow[NCIT: blood circulation] **inside**[NCIT:inner] the vessels[BLO: blood vessel], **of compound nature**.*

(5.4)

5.1.1.4 Criando novas propriedades para o termo *bleeding disorder associated with acquired disorders of coagulation*

Das propriedades identificadas na subseção 5.1.1.3 os termos *due to*, *retain* e *inner* foram importados de suas respectivas ontologias. O termo *compound nature* não foi encontrado em nenhuma ontologia e portanto foi criado em BLO.

5.1.1.5 Identificando as condições necessárias e suficientes para o termo *bleeding disorder associated with acquired disorders of coagulation*

Não há indícios em sua definição das condições necessárias e suficientes. Neste caso recorreu-se a literatura. Aganette (2015, p. 170) faz a seguinte observação sobre o termo:

"Nessa versão da definição, a parte da sentença após o termo 'o qual' foi considerado a característica essencial."

A característica essencial do termo é o que define a necessidade e suficiência do mesmo. Neste caso a definição "o qual" que liga o termo transtorno de coagulação do sangue aos demais é necessária e suficiente.

A relação *retain blood flow inside the vessel* é uma definição separada para a classe *defect*. Esta relação não deixa explícito a necessidade e suficiência. Entende-se que apenas a relação de suficiência é atendida. Reter fluxo sanguíneo dentro dos vasos é condição suficiente para ser um defeito. Desta forma, é feita a seguinte definição:

\forall (*bleeding disorder associated with acquired disorders of coagulation*)

(*disorder of blood coagulation*)

Bleeding disorder associated with acquired disorders of coagulation :

blood coagulation disorder[BLO: *disorder of blood coagulation*] **caused by**[SNMI:*due to*]
 \exists *defect*[NCIT: *defect*], which **retains**[NCIT:*retain*] \exists *blood flow*[NCIT: *blood circulation*]
inside[NCIT:*inner*] \exists *vessels*[BLO: *blood vessel*], **of compound nature**.

(5.5)

5.1.1.6 Reescrevendo o termo *bleeding disorder associated with acquired disorders of coagulation* no editor de ontologias

Bleeding disorder associated with acquired disorders of coagulation é uma subclasse de *disorder of blood coagulation*. O termo foi reescrito seguindo as etapas:

1. \forall foi substituído por *for all* e \exists foi substituído por *some*;

2. As classes *blood coagulation disorder*, *blood flow* e *vessel* foram substituídas por *disorder of blood coagulation*, *blood circulation* e *blood vessel* respectivamente;
3. As propriedades *caused by* e *inside* foram substituídas por *due to* e *inner* respectivamente;
4. As propriedades *due to*, *retain*, *inner* e *of compound nature* estão relacionadas com as classe *defect*, *blood flow* e *bleeding disorder associated with acquired disorders of coagulation*, e;
5. As propriedades *due to*, *retain*, *inner* e *of compound nature* foram conectadas pela palavra *and*.

O resultado obtido em linguagem natural é mostrado na Equação 5.6 e sua inserção no Protégé é mostrado na Figura 17.

*For all bleeding disorder associated with acquired disorders of coagulation
of compound nature is a subclass of disorder of blood coagulation and is due to
some defect which retain some blood circulation inner some blood vessel. (5.6)*

Figura 17 – Definição de *bleeding disorder associated with acquired disorders of coagulation* no Protégé



5.1.2 *Bleeding disorder after extracorporeal circulation*

Segundo Aganette (2015), o termo foi definido da seguinte maneira:

Uma hemorragia após circulação extracorpórea (BDEC) é uma desordem do sangue o qual está associada com múltiplas causas, incluindo dificuldade do processo de coagulação pela diluição do sangue em água, neutralização inadequada de heparina, disfunção plaquetária adquirida, e trombocitopenia. (5.7)

Neste caso fez-se necessário traduzir o termo para inglês, língua padrão da BLO:

Bleeding disorder after extracorporeal circulation is a blood disorder with is related with multiple causes, including difficulting process of coagulation by diluting blood in water, inadequate neutralization of heparin, acquired platelet dysfunction and thrombocytopenia. (5.8)

A partir daí segue-se o fluxograma mostrado na Figura 11.

5.1.2.1 Identificando os substantivos para *bleeding disorder after extracorporeal circulation*

Os substantivos identificados foram: ***bleeding disorder after extracorporeal circulation, blood disorder, cause, difficulting process of coagulation by diluting blood in water, inadequate neutralization of heparin, acquired platelet dysfunction e thrombocytopenia.***

- ***Bleeding disorder after extracorporeal circulation:*** termo a ser definido. Seu resultado será a conjunção das definições de outros substantivos;
- ***Blood disorder:*** retirado de BLO, *disposition of hematopoietic human system to have abnormal conditions under certain bodily circumstances;*
- ***Cause:*** retirado de NCIT (Etiology), *the cause of a disease or abnormal condition;*
- ***Difficulting process of coagulation by diluting blood in water:*** definição não encontrada;
- ***Inadequate neutralization of heparin:*** definição não encontrada;
- ***Acquired platelet dysfunction:*** definição não encontrada, e;
- ***Thrombocytopenia:*** retirado de NCIT, *A laboratory test result indicating that there is an abnormally small number of platelets in the circulating blood.*

Associando os termos identificados com suas respectivas ontologias obteve-se o termo 5.9:

*Bleeding disorder after extracorporeal circulation is a blood disorder
with is related with multiple causes[NCIT: etiology], including difficulting process
of coagulation by diluting blood in water, inadequate neutralization of
heparin, acquired platelet dysfunction
and thrombocytopenia. [NCIT: thrombocytopenia]* (5.9)

5.1.2.2 Criando novos termos para *Bleeding disorder after extracorporeal circulation*

Dos substantivos indicados, os termos *difficulting process of coagulation by diluting blood in water, inadequate neutralization of heparin* e *acquired platelet dysfunction* não foram encontrados em nenhuma ontologia e serão criados para BLO. Os termos *etiology* e *thrombocytopenia* foram importados das ontologias NCIT e RCD respectivamente. O termo *etiology* foi adicionado como subclasse de *Uncategorized* e os demais termos como subclasse de *etiology*, pois a descrição do termo indica que fazem parte de um conjunto de causas.

5.1.2.3 Identificando relações para o termo *Bleeding disorder after extracorporeal circulation*

As relações encontradas foram: *is a, related, including* e *and*.

- **Is a:** relação de equivalência lógica que é substituído pelo símbolo $;$;
- **Related:** retirado de BLO, *an action taken with respect to a subject Entity by a regulatory or authoritative body with supervisory capacity over that entity. The action is taken in response to behavior by the subject Entity that body finds to be undesirable. Suspension, license restrictions, monetary fine, letter of reprimand, mandated training, mandated supervision, etc.*;
- **Including:** relação de subconjunto lógico que é substituído pelo símbolo \sqsubseteq , e;
- **And:** relação de conjunção lógica que é substituído pelo símbolo \wedge .

Associando as relações identificadas com suas respectivas ontologias obteve-se o termo 5.10:

$$\begin{aligned}
& \underline{\textit{Bleeding disorder after extracorporeal circulation}} : \underline{\textit{blood disorder}} \\
& \textit{with is related with multiple causes[NCIT: etiology]} / (\underline{\textit{difficulting process}} \\
& \quad \underline{\textit{of coagulation by diluting blood in water}} \wedge \underline{\textit{inadequate neutralization of}} \\
& \quad \quad \underline{\textit{heparin}} \wedge \underline{\textit{acquired platelet dysfunction}} \\
& \wedge \underline{\textit{thrombocytopenia.[NCIT: thrombocytopenia]}} \sqsubseteq \underline{\textit{causes[NCIT:etiology]}} \quad (5.10)
\end{aligned}$$

5.1.2.4 Criando propriedades para o termo *bleeding disorder after extracorporeal circulation*

Das propriedades identificadas, todas já estavam contidas em BLO e nenhuma propriedade precisou ser criada.

5.1.2.5 Identificando relações necessárias e suficientes para o termo *bleeding disorder after extracorporeal circulation*

Não há indícios em sua definição das condições necessárias e suficientes. Neste caso recorreu-se a literatura. Aganette (2015, p. 173) faz a seguinte observação sobre o termo:

"Nessa versão da definição, a parte da sentença após o termo “o qual” foi considerado a característica essencial."

Foi identificado também que a definição de *bleeding disorder after extracorporeal circulation* é uma desordem do sangue relacionada com múltiplas causas. As causas, ou etiologias, são um subconjunto de outros termos que compõem este termo. Ou seja, uma causa pode ser dificuldade no processo de coagulação por diluição de sangue em água, neutralização inadequada de heparina, disfunção adquirida de plaquetas, trombocitopenia ou todas estas causas juntas.

No caso específico para o termo a ser definido a condição é que as causas são algumas das causas possíveis, sendo pelo menos as quatro listadas, devendo-se considerar possíveis causas não identificadas fora do universo do discurso. Feita as considerações obteve-se o seguinte termo:

$$\begin{aligned}
& \forall(\textit{bleeding disorder after extracorporeal circulation})(\textit{blood disorder}) \\
& \quad \underline{\textit{Bleeding disorder after extracorporeal circulation}} : \underline{\textit{blood disorder}} \\
& \quad \textit{with is related with} \exists [\underline{\textit{causes[NCIT: etiology]}} / (\underline{\textit{difficulting process}} \\
& \quad \underline{\textit{of coagulation by diluting blood in water}} \wedge \underline{\textit{inadequate neutralization of}} \\
& \quad \quad \underline{\textit{heparin}} \wedge \underline{\textit{acquired platelet dysfunction}} \\
& \quad \wedge \underline{\textit{thrombocytopenia.[NCIT: thrombocytopenia]}} \sqsubseteq \underline{\textit{causes[NCIT: etiology]}}) \quad (5.11)
\end{aligned}$$

5.1.2.6 Reescrevendo o termo *bleeding disorder after extracorporeal circulation* no editor de ontologias

Bleeding disorder after extracorporeal circulation é uma subclasse de *bleeding disorder associated with acquired disorders of coagulation*. O termo foi reescrito seguindo as etapas:

1. \forall foi substituído por *for all* e \exists foi substituído por *some*;
2. A classe *causes* foi substituída por *etiology*;
3. Nenhuma propriedade foi alterada;
4. A propriedades *related* foi associada à classe *etiology*. À classe *etiology* foi incluída como um conjunto das classes *difficulting process of coagulation by diluting blood in water*, *inadequate neutralization of heparin*, *acquired platelet disorder* e *thrombocytopenia*, e;
5. Nenhuma propriedade precisou ser conectada pelo conectivo *and*.

O resultado obtido é mostrado na Equação 5.12 e sua inserção no Protégé é mostrado na Figura 18.

$$\begin{aligned}
 & \textit{For all bleeding disorder after extracorporeal circulation is a subclass of} \\
 & \quad \textit{blood disorder with some etiologies: difficulting process of coagulation} \\
 & \quad \textit{by diluting blood in water and inadequate neutralization of heparin and} \\
 & \quad \textit{acquired platelet dysfunction and thrombocytopenia.} \quad (5.12)
 \end{aligned}$$

Figura 18 – Definição de *bleeding disorder after extracorporeal circulation* no Protégé

Description: *Bleeding disorder after extracorporeal circulation*

Equivalent To +

- **related only etiology**

SubClass Of +

- **Bleeding_disorder_associated_with_acquired_disorders_of_coagulation**
- **Blood_disorder**

General class axioms +

SubClass Of (Anonymous Ancestor)

- **(over some (body_part or blood_vessel)) and (related some stoppage_of_blood_flow)**
- **(over some (body_part or blood_vessel)) and (related some stoppage_of_blood_flow)**
- **of_compound_nature only Bleeding_Disorder_Associated_with_Acquired_Coagulation_Disorders_of_Bleeding and due_to some defect and retain some blood_circulation and inner some blood_vessel**

Nota-se que apenas o termo *etiology* foi considerado na definição do Protégé. Suas subclasses, ou seja, as etiologias correspondentes tiveram as suas relações consideradas a nível de indivíduo, como mostradas na subseção 5.1.6.2.

5.1.3 *Bleeding disorder after massive transfusion*

Segundo Aganette (2015), o termo foi definido da seguinte maneira:

A bleeding disorder after massive transfusion is a bleeding disorder consists in the process of transferring whole blood or blood components from one person (donor) to another (recipient) indicated in patients with bleeding with platelet dysfunction irrespective of platelet count. (5.13)

5.1.3.1 Identificando os substantivos para *Bleeding disorder after massive transfusion*

Os substantivos identificados foram:

- ***Bleeding disorder after massive transfusion***: termo a ser definido. Seu resultado será a conjunção das definições de outros substantivos;
- ***Bleeding Disorder***: retirado de ICPC2P¹, definição não encontrada;

¹ <http://www.who.int/classifications/icd/adaptations/icpc2/en/>

- **Process**: retirado da BLO, *a processual entity that is a maximally connected spatiotemporal whole and has bona fide beginnings and endings corresponding to real discontinuities*;
- **Whole blood**: retirado de NCIT, *blood that has not been separated into its various components; blood that has not been modified except for the addition of an anticoagulant*;
- **Blood component**: retirado de SNOMED CT², definição não encontrada;
- **Donor**: retirado de OBI³, *a role which inheres in an organism or part thereof from which any part including cell, organ or tissue is removed with the intention that the donated part will be placed into another organism and/or cultured in vitro*;
- **Recipient**: retirado de NCIT, *a person who gets something*;
- **Patient**: retirado de NCIT, *a person who receives medical attention, care, or treatment, or who is registered with medical professional or institution with the purpose to receive medical care when necessary, e*;
- **Bleeding with platelet dysfunction irrespective of platelet count**: definição não encontrada.

Associando os termos identificados com suas respectivas ontologias obteve-se o termo Equação 5.14:

A bleeding disorder after massive transfusion is a bleeding disorder [ICPC2P: bleeding disorder] consists in the process of transferring whole blood[NCIT: whole blood] or blood components[SNOMED CT: blood component] from one person (donor)[OBI: donor] to another (recipient)[NCIT: recipient] indicated in patients[NCIT: patient] with bleeding with platelet dysfunction irrespective of platelet count. (5.14)

5.1.3.2 Criando novos termos para *bleeding disorder after massive transfusion*

Dos substantivos indicados, o termo ***bleeding with platelet dysfunction irrespective of platelet count*** não foi encontrado em nenhuma ontologia e será criado para BLO. Os termos ***bleeding disorder***, ***blood component*** e ***donor*** foram importados das ontologias ICPC2P, SNOMED CT E OBI respectivamente. Os termos ***whole blood***, ***recipient*** e ***patient*** foram importados de NCIT. Os termos criados foram adicionados como subclasses de *Uncategorized*.

² <http://searchhealthit.techtarget.com/definition/SNOMED-CT>

³ <http://obi-ontology.org/>

5.1.3.3 Identificando relações para o termo *Bleeding disorder after massive transfusion*

As relações encontradas foram: *is a*, *consists*, *transferring*, *or*, *from*, *to*, *indicated* e *with*. As relações *from* e *to* estão vinculadas com a relação *transfusing* ou seja, transferindo algo de para de. Neste caso é transferido sangue ou componentes do sangue de um doador para um receptor. As propriedades *from* e *to* são delimitadores de alcance da propriedade *transferring*.

- ***Is a***: relação de equivalência lógica que é substituído pelo símbolo \equiv ;
- ***Consists in***: definição não encontrada;
- ***Transferring***: retirado de GALEN⁴(*transfusing*), definição não encontrada;
- ***Or***: relação de disjunção lógica que é substituído pelo símbolo \vee ;
- ***Indicated***: retirado de SNOMED CT, definição não encontrada, e;
- ***With***: retirado de NCIT. *used to indicate the presence of something or someone.*

Associando as relações identificadas com suas respectivas ontologias obteve-se o termo 5.16:

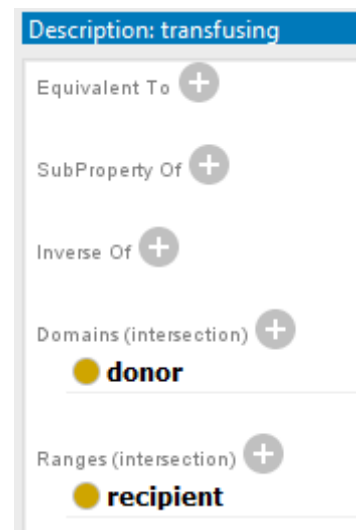
$$\begin{aligned}
 & \text{A } \underline{\textit{bleeding disorder after massive transfusion}} : \underline{\textit{bleeding disorder}} \\
 & [\textit{ICPC2P: bleeding disorder}] \textbf{consists in} \textit{ the } \underline{\textit{process}} \textit{ of } \textbf{transferring}[\textit{GALEN: transfusing}] \\
 & \underline{\textit{whole blood}}[\textit{NCIT: whole blood}] \vee \underline{\textit{blood components}}[\textit{SNOMED CT: blood component}] \\
 & \textit{ from one } \underline{\textit{person (donor)}}[\textit{OBI: donor}] \textit{ to another (recipient)}[\textit{NCIT: recipient}] \\
 & \textbf{indicated}[\textit{SNOMED CT: indicated}] \textit{ in } \underline{\textit{patients}}[\textit{NCIT: patient}] \textbf{with}[\textit{NCIT: with}] \\
 & \underline{\textit{bleeding with platelet dysfunction irrespective of platelet count.}} \quad (5.15)
 \end{aligned}$$

5.1.3.4 Criando propriedades para o termo *bleeding disorder after massive transfusion*

Das propriedades identificadas na Equação 5.16 os termos *transfusing*, *indicated* e *with* foram importados de suas respectivas ontologias. O termo *consists in* não foi encontrado em nenhuma ontologia e portanto foi criado em BLO.

A propriedade *transfusing* foi criada com os delimitadores de alcance de *donor* para *recipient* como mostra a Figura 19.

⁴ <http://www.opengalen.org/>

Figura 19 – Delimitadores de alcance para a propriedade *transfusing*

5.1.3.5 Identificando relações necessárias e suficientes para o termo *bleeding disorder after massive transfusion*

O termo resultante da Equação 5.16 não dá indícios em sua definição das condições necessárias e suficientes. Neste caso recorreu-se a literatura. Aganette (2015, p. 179) faz a seguinte observação sobre o termo:

"Parece que a tanto a condição necessária quanto a condição suficiente são atendidas."

Foi identificado também que os termos *donor* e *recipient* estão ligados à propriedade *transfusing* e limitados à quantidade de um. O termo foi então reescrito da seguinte maneira:

$$\begin{aligned} &\forall (bleeding\ disorder\ after\ massive\ transfusion)(bleeding\ disorder)(process)(transfusing) \\ &(whole\ blood)(blood\ component)(patient)(bleeding\ with\ platelet\ dysfunction\ irrespective \\ &\quad of\ platelet\ count) \underline{bleeding\ disorder\ after\ massive\ transfusion} : \underline{bleeding\ disorder} \\ &[ICPC2P: bleeding\ disorder] \mathbf{consists\ in\ the\ process\ of\ transferring}[GALEN: transfusing] \\ &\underline{whole\ blood}[NCIT: whole\ blood] \vee \underline{blood\ components}[SNOMED\ CT: blood\ component] \\ &\quad \exists! (person\ (donor)[OBI: donor])\ to\ \exists! (another\ (recipient)[NCIT: recipient]) \\ &\quad \mathbf{indicated}[SNOMED\ CT: indicated] \mathbf{in\ patients}[NCIT: patient] \mathbf{with}[NCIT: with] \\ &\quad \underline{bleeding\ with\ platelet\ dysfunction\ irrespective\ of\ platelet\ count}. \quad (5.16) \end{aligned}$$

5.1.3.6 Reescrevendo o termo *bleeding disorder after massive transfusion* no editor de ontologias

Bleeding disorder after massive transfusion é uma subclasse de *bleeding disorder associated with acquired disorders of coagulation*. O termo foi reescrito seguindo as etapas:

- \forall foi substituído por *for all* e $\exists!$ substituído por *one*;
- As classes *person(donor)* e *another(recipient)* foram substituídas pelas classes *donor* e *recipient*;
- A propriedade *transferring* foi substituída por *transfusing*;
- As propriedades *consists in*, *transfusing*, *indicated* e *with* foram associadas às classes *process*, *whole blood* e *blood components*, *patient* e *bleeding with platelet dysfunction irrespective of platelet count* respectivamente. As classes *donor* e *recipient* não tiveram nenhuma propriedade associada por serem delimitadores de alcance da propriedade *transfusing*, e;
- As propriedades *consists in*, *transfusing*, *indicated* e *with* foram conectadas pelo conectivo *and*.

O resultado obtido é mostrado na Equação 5.17 e sua inserção no Protégé é mostrado na Figura 20.

For all bleeding disorder after massive transfusion is a subclass of bleeding disorder consists in all the process and transfusing whole blood or blood components (from one donor to one recipient) and indicated in all patients and with bleeding with platelet dysfunction irrespective of platelet count. (5.17)

Figura 20 – Definição de *bleeding disorder after massive transfusion* no Protégé

Description: *bleeding_disorder_after_massive_transfusion*

Equivalent To +

- (consists_in only process)
- and (indicated only patient)
- and (transfusing only
- (blood_component or whole_blood))
- and (with only bleeding_with_platelet_dysfunction_irrespective_of_platelet_count)

SubClass Of +

- Bleeding_disorder_associated_with_acquired_disorders_of_coagulation
- Blood_disorder

General class axioms +

SubClass Of (Anonymous Ancestor)

- (over some
- (body_part or blood_vessel))
- and (related some stoppage_of_blood_flow)

- (over some
- (body_part or blood_vessel))
- and (related some stoppage_of_blood_flow)

- of_compound_nature only Bleeding_Disorder_Associated_with_Acquired_Coagulation_Disorders_of_Bleeding and
- due_to some defect and
- retain some blood_circulation and
- inner some blood_vessel

5.1.4 *Bleeding Disorder Associated With Accelerated Destruction of Coagulation Factors*

Segundo Aganette (2015), o termo foi definido da seguinte maneira:

A bleeding disorder associated with accelerated destruction of coagulation factors is a Disorder and function as: mediators of one of the body's natural means of stopping bleeding (homeostasis), activated by damage to the vessel lining and anticoagulation. (5.18)

5.1.4.1 Identificando os substantivos para *Bleeding Disorder Associated With Accelerated Destruction of Coagulation Factors*

Os substantivos identificados foram:

- ***Bleeding disorder associated with accelerated destruction of coagulation factors***: termo a ser definido. Seu resultado será a conjunção das definições de outros substantivos;
- ***Disorder***: retirado de CLO⁵, *a disease that involves a disturbance of function*,

⁵ <http://www.clo-ontology.org/>

structure, or both of any part, organ, or system of the body resulting from a genetic or embryonic failure in development;

- **Homeostasis**: retirado de MESH⁶, *the processes whereby the internal environment of an organism tends to remain balanced and stable;*
- **Damage**: retirado de NCIT, *physical harm reducing the value, operation, or usefulness of something;*
- **Vessel lining**: definição não encontrada, e;
- **Anticoagulation**: retirado de GALEN, definição não encontrada.

Nota-se que o texto *mediators of one of the body's natural means of stopping bleeding* foi retirado da definição original, pois o mesmo se trata de uma definição de *homeostasis*.

Associando os termos identificados com suas respectivas ontologias obteve-se o termo Equação 5.19:

A bleeding disorder associated with accelerated destruction of coagulation factors is a disorder[CLO: disorder] and function as homeostasis[MESH: homeostasis], activated by damage[NCIT: damage] to the vessel lining and anticoagulation[GALEN: anticoagulation]
(5.19)

5.1.4.2 Criando novos termos para *bleeding disorder associated with accelerated destruction of coagulation factors*

Dos substantivos indicados, o termo inicial ***bleeding disorder associated with accelerated destruction of coagulation factors*** e o termo ***vessel lining*** não foi encontrado em nenhuma ontologia e será criado para BLO. Os termos ***disorder***, ***homeostasis***, ***damage*** e ***anticoagulation*** foram importados das ontologias CLO, MESH, NCIT e GALEN respectivamente. Os termos criados foram adicionados como subclasses de *Uncategorized*.

5.1.4.3 Identificando relações para o termo *bleeding disorder associated with accelerated destruction of coagulation factors*

As relações encontradas foram: ***is a***, ***function as***, ***activated by***, ***to***, e ***and***.

- ***Is a***: relação de equivalência lógica que é substituído pelo símbolo ***;***

⁶ <https://www.ncbi.nlm.nih.gov/mesh>

- **Function as**: definição não encontrada;
- **Activated by**: retirado de HPO⁷ (*triggered by*), *an external factor that leads to the manifestation of a sign or symptom in a person with a susceptibility to developing that manifestation*;
- **To**: substituído por *cause harm*, pois a definição de *damage* é causar prejuízo a alguma coisa. Definição não encontrada, e;
- **And**: relação de conjunção lógica que é substituída pelo símbolo \wedge .

Associando as relações identificadas com suas respectivas ontologias obteve-se o termo 5.20:

*A bleeding disorder associated with accelerated destruction of coagulation factors :
disorder[CLO: disorder] and **function as** homeostasis[MESH: homeostasis],
activated by[HPO: triggered by] damage[NCIT: damage]
to[cause harm] the vessel lining \wedge anticoagulation[GALEN: anticoagulation]* (5.20)

5.1.4.4 Criando propriedades para o termo *bleeding disorder associated with accelerated destruction of coagulation factors*

Das propriedades identificadas na subseção 5.1.4.3 o termo **triggered by** foi importado de sua respectiva ontologia. O termo **function as** não foi encontrado em nenhuma ontologia e portanto foi criado em BLO.

5.1.4.5 Identificando relações necessárias e suficientes para o termo *bleeding disorder associated with accelerated destruction of coagulation factors*

O termo resultante da subseção 5.1.4.3 não dá indícios em sua definição das condições necessárias e suficientes. Neste caso recorreu-se a literatura. Aganette (2015, p. 183) faz a seguinte observação sobre o termo:

"Parece que a condição necessária não é atendida. Não basta uma entidade ser mediadora de coagulação e anticoagulação do sangue para ser uma BDAADCF. Deve haver outras entidades que são mediadoras de coagulação e anticoagulação do sangue. A definição atende apenas às condições suficientes."

Sob essas condições, o termo foi reescrito da seguinte forma:

⁷ <http://human-phenotype-ontology.github.io/>

*A bleeding disorder associated with accelerated destruction of coagulation factors:
disorder[CLO: disorder] and **function as** \exists **homeostasis**[MESH: homeostasis]
activated by[HPO: triggered by] \exists damage[NCIT: damage]
to[cause harm] **the** \exists (vessel lining \wedge anticoagulation[GALEN: anticoagulation])* (5.21)

5.1.4.6 Reescrevendo o termo *bleeding disorder associated with accelerated destruction of coagulation factors* no editor de ontologias

Bleeding disorder associated with accelerated destruction of coagulation factors é uma subclasse de *disorder*. O termo foi reescrito seguindo as etapas:


- \exists foi substituído por *some*;
- Nenhuma classe precisou ser substituída;
- As propriedades *activated by* e *to* foram substituídas por *triggered by* e *cause harm* respectivamente;
- As propriedades *function as*, *triggered by* e *cause harm* foram associadas às classes *homeostasis*, *damage* e *vessel lining* e *anticoagulation* respectivamente, e;
- As propriedades *activated by*, *triggered by* e *cause harm* foram conectadas pelo conectivo *and*.

O resultado obtido é mostrado na Equação 5.22 e sua inserção no Protégé é mostrado na Figura 21.


Bleeding disorder associated with accelerated destruction of coagulation factors
is a subclass of *disorder* that **function as** **some** **homeostasis** and **triggered by** **some** **damage**
(that cause harm to some vessel lining and some anticoagulation) (5.22)

Figura 21 – Definição de *bleeding disorder associated with accelerated destruction of coagulation factors* no Protégé


Description: **Bleeding_disorder_associated_with_accelerated_destruction_of_coagulation_factors**

Equivalent To 

- (cause_harm some (anticoagulation and vessel_lining)) and (function_as some homeostasis) and (triggered_by some damage)

SubClass Of 

- Bleeding_disorder_associated_with_acquired_disorders_of_coagulation
- disorder

General class axioms 

SubClass Of (Anonymous Ancestor)

- (over some (body_part or blood_vessel)) and (related some stoppage_of_blood_flow)
- (over some (body_part or blood_vessel)) and (related some stoppage_of_blood_flow)

5.1.5 *Bleeding disorder caused by fibrinolysis*

Segundo Aganette (2015), o termo foi definido da seguinte maneira:

*A Bleeding Disorder Caused by Fibrinolysis is a bleeding disorder,
which is considered a blood vessels defense system. (5.23)*

5.1.5.1 Identificando os substantivos para *bleeding disorder caused by fibrinolysis*

Os substantivos identificados foram *bleeding disorder caused by fibrinolysis*, *bleeding disorder* e *blood vessels defense system*:

- *Bleeding disorder caused by fibrinolysis*: termo a ser definido. Seu resultado será a conjunção das definições de outros substantivos;
- *Bleeding disorder*: retirado de BLO, definição não encontrada, e;
- *Blood vessels defense system*: definição não encontrada.

Associando os termos identificados com suas respectivas ontologias obteve-se o termo Equação 5.24:

*A bleeding disorder caused by fibrinolysis is a bleeding disorder,
which is considered a blood vessels defense system. (5.24)*

5.1.5.2 Criando novos termos para *bleeding disorder caused by fibrinolysis*

Dos substantivos indicados, o termo ***blood vessels defense system*** não foi encontrado em nenhuma ontologia e será criado para BLO. O termo criado foi adicionado como subclasse de *Uncategorized*.

5.1.5.3 Identificando relações para o termo *bleeding disorder caused by fibrinolysis*

As relações encontradas foram: ***is a*** e ***considered***.

- ***Is a***: relação de equivalência lógica que é substituído pelo símbolo :, e;
- ***Considered***: substituído por *due to* de BLO, *caused by or resulting from*. A substituição se deu por causa da interpretação do termo. É entendido que a desordem hemorrágica se dá devido ao sistema de defesa dos vasos.

Associando as relações identificadas com suas respectivas ontologias obteve-se o termo 5.25:

*A bleeding disorder caused by fibrinolysis : bleeding disorder,
which is **considered**[BLO: due to] a blood vessels defense system. (5.25)*

5.1.5.4 Criando propriedades para o termo *bleeding disorder caused by fibrinolysis*

Nenhuma das propriedades identificadas na subseção 5.1.5.3 precisou ser criada.

5.1.5.5 Identificando relações necessárias e suficientes para o termo *bleeding disorder caused by fibrinolysis*

O termo resultante da subseção 5.1.5.3 não dá indícios em sua definição das condições necessárias e suficientes. Neste caso recorreu-se a literatura. Aganette (2015, p. 187) faz a seguinte observação sobre o termo:

"Parece que a condição necessária não é atendida. Não basta uma entidade ser considerada um sistema de defesa vascular para ser uma BDCF. Deve haver outras entidades consideradas um sistema de defesa vascular que não são BDCF."

Dada estas considerações o termo foi reescrito da seguinte maneira:

$$A \text{ bleeding disorder caused by fibrinolysis : bleeding disorder,} \\ \text{which is **considered**[BLO: due to] } \exists \text{ blood vessels defense system.} \quad (5.26)$$

5.1.5.6 Reescrevendo o termo *bleeding disorder caused by fibrinolysis* no editor de ontologias

Bleeding disorder caused by fibrinolysis é uma subclasse de *bleeding disorder*. O termo foi reescrito seguindo as etapas:

- \exists foi substituído por *some*;
- Nenhuma classe precisou ser substituída;
- As propriedade *considered* foi substituída por *due to*;
- A propriedade *due to* foi associada à classe *blood vessels defense system*, e;
- Nenhuma propriedade precisou ser conectada pelo conectivo *and*.

O resultado obtido é mostrado na Equação 5.27 e sua inserção no Protégé é mostrado na Figura 22.

$$\text{Bleeding disorder caused by fibrinolysis is a bleeding disorder due to some} \\ \text{blood vessels defense system.} \quad (5.27)$$

Figura 22 – Definição de *Bleeding disorder caused by fibrinolysis* no Protégé

Description: *Bleeding_disorder_caused_by_fibrinolysis*

Equivalent To +

- **due_to_some blood_vessels_defense_system**

SubClass Of +

- **bleeding_disorder**
- **Bleeding_disorder_associated_with_accelerated_destruction_of_coagulation_factors**

General class axioms +

SubClass Of (Anonymous Ancestor)

- **(over some (body_part or blood_vessel)) and (related some stoppage_of_blood_flow)**
- **(over some (body_part or blood_vessel)) and (related some stoppage_of_blood_flow)**
- **of_compound_nature only Bleeding_Disorder_Associated_with_Acquired_Coagulation_Disorders_of_Bleeding and due_to_some defect and retain some blood_circulation and inner some blood_vessel**

5.1.6 Inserindo indivíduos para as classes transcritas

A inserção de indivíduos como parte da avaliação de inferências de BLO segue a metodologia proposta em subseção 4.6.1. Para cada classe criada durante a transcrição, foi criado pelo menos dois indivíduos, sendo um deles como indivíduo da definição principal e o outro como indivíduo da transcrição realizada pelos indivíduos.

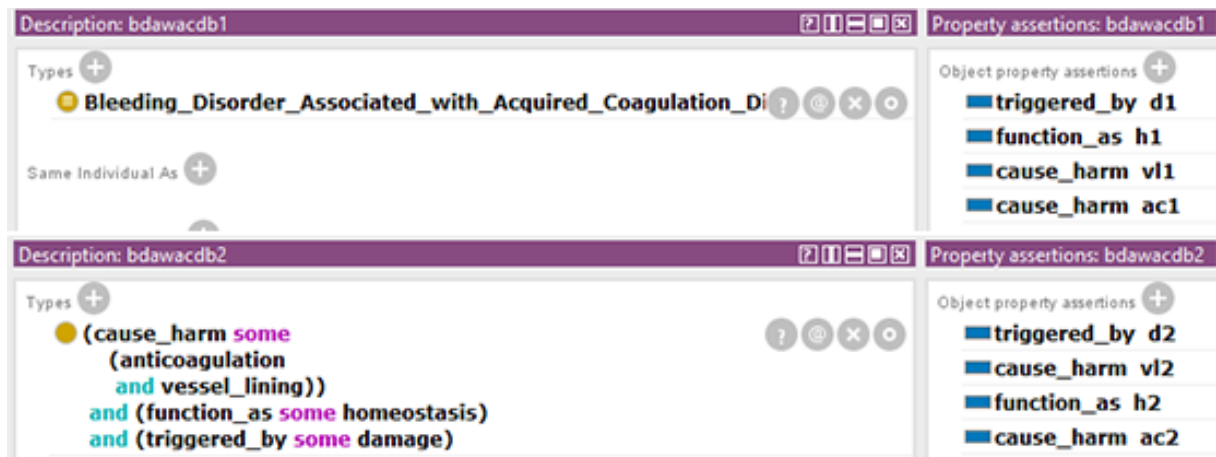
5.1.6.1 Indivíduos para *bleeding disorder associated with acquired disorders of coagulation*

De acordo com a metodologia citada na subseção 4.6.1, foram criados seis indivíduos para a classe *bleeding disorder associated with acquired disorders of coagulation*. Os indivíduos criados foram:

- **bdawadc1 e bdawadc2:** indivíduos da classe *bleeding disorder associated with acquired disorders of coagulation* e relacionados com a propriedade *of compound nature*;
- **d1 e d2:** indivíduos da classe *defect* e relacionados com a propriedade *due to*;
- **bc1 e bc2:** indivíduos da classe *blood circulation* e relacionados com a propriedade *retain*, e;
- não foi preciso criar indivíduos da classe *blood vessel*, os indivíduos *bv1* e *bv2* já criados anteriormente foram relacionados com a propriedade *inner*.

A inserção destes indivíduos são mostrados na Figura 23.

Figura 23 – Indivíduos para a classe *bleeding disorder associated with acquired disorders of coagulation*

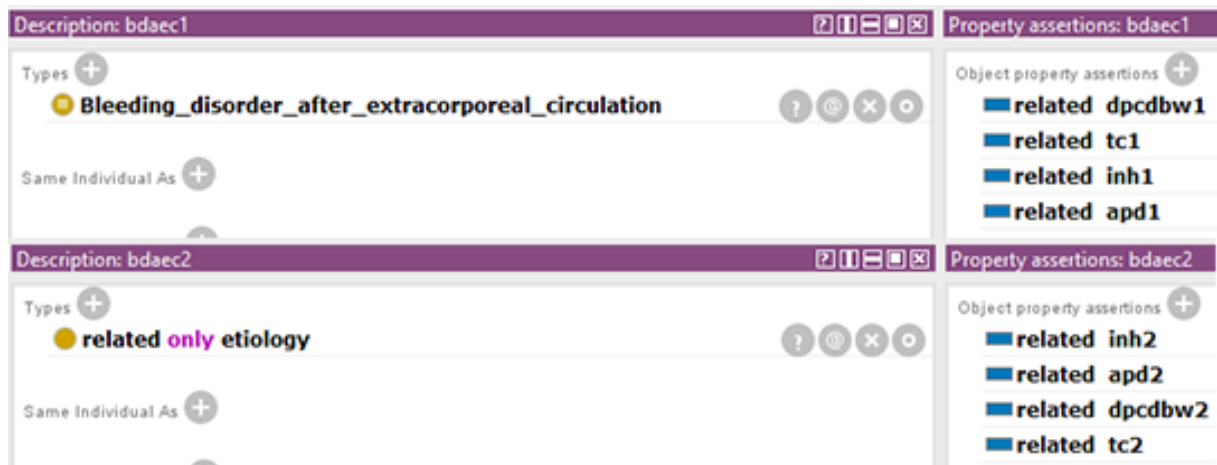


5.1.6.2 Indivíduos para *bleeding disorder after extracorporeal circulation*

De acordo com a metodologia citada na subseção 4.6.1, foram criados dez indivíduos para a classe *bleeding disorder associated with acquired disorders of coagulation*. Os indivíduos criados foram:

- **bdaec1 e bdaec2:** indivíduos da classe *bleeding disorder after extracorporeal circulation*;
- **apd1 e apd2:** indivíduos da classe *acquired platelet dysfunction* e relacionados com a propriedade *related*;
- **dpcdbw1 e dpcdbw2:** indivíduos da classe *difficulting process of coagulation by diluting blood in water* e relacionados com a propriedade *related*;
- **inh1 e inh2:** indivíduos da classe *inadequate neutralization of heparin* e relacionados com a propriedade *related*, e;
- **tc1 e tc2:** indivíduos da classe *thrombocytopenia* e relacionados com a propriedade *related*.

A inserção destes indivíduos são mostrados na Figura 24.

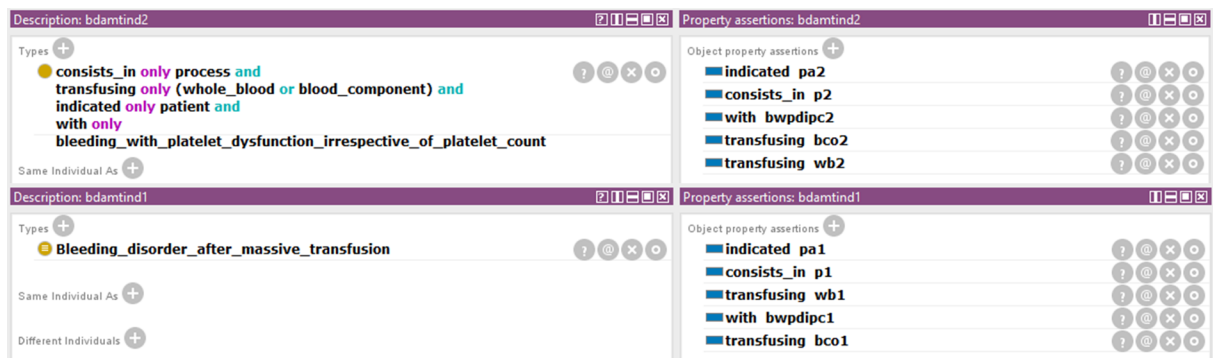
Figura 24 – Indivíduos para a classe *bleeding disorder after extracorporeal circulation*

5.1.6.3 Indivíduos para *bleeding disorder after massive transfusion*

De acordo com a metodologia citada na subseção 4.6.1, foram criados doze indivíduos para a classe *bleeding disorder after massive transfusion*. Os indivíduos criados foram:

- **bdamtind1 e bdamtind2:** indivíduos da classe *bleeding disorder after massive transfusion*;
- **p1 e p2:** indivíduos da classe *process* e relacionados com a propriedade *consists in*;
- **wb1 e wb2:** indivíduos da classe *whole blood* e relacionados com a propriedade *transfusing*;
- **bco1 e bco2:** indivíduos da classe *blood component* e relacionados com a propriedade *transfusing*;
- **pa1 e pa2:** indivíduos da classe *patient* e relacionados com a propriedade *indicated*, e;
- **bwpdipc1 e bwpdipc2:** indivíduos da classe *bleeding with platelet dysfunction irrespective of platelet count* e relacionados com a propriedade *with*.

A inserção destes indivíduos são mostrados na Figura 25.

Figura 25 – Indivíduos para a classe *bleeding disorder after massive transfusion*

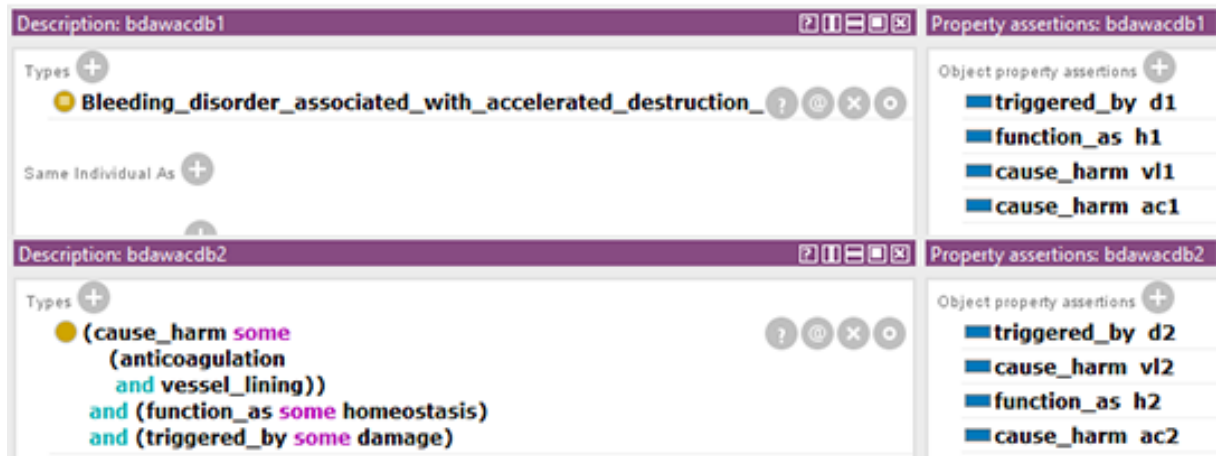
5.1.6.4 Indivíduos para *bleeding disorder associated with accelerated destruction of coagulation factors*

De acordo com a metodologia citada na subsecção 4.6.1, foram criados dez indivíduos para a classe *bleeding disorder associated with accelerated destruction of coagulation factors*. Os indivíduos criados foram:

- **bdawacdb1 e bdawacdb2:** indivíduos da classe *bleeding disorder associated with accelerated destruction of coagulation factors*;
- **h1 e h2:** indivíduos da classe *homeostasis* e relacionados com a propriedade *function as*;
- **dam1 e dam2:** indivíduos da classe *damage* e relacionados com a propriedade *triggered by*;
- **vl1 e vl2:** indivíduos da classe *vessel lining* e relacionados com a propriedade *cause harm*, e;
- **ac1 e ac2:** indivíduos da classe *anticoagulation* e relacionados com a propriedade *cause harm*.

A inserção destes indivíduos são mostrados na Figura 26.

Figura 26 – Indivíduos para a classe *bleeding disorder associated with accelerated destruction of coagulation factors*



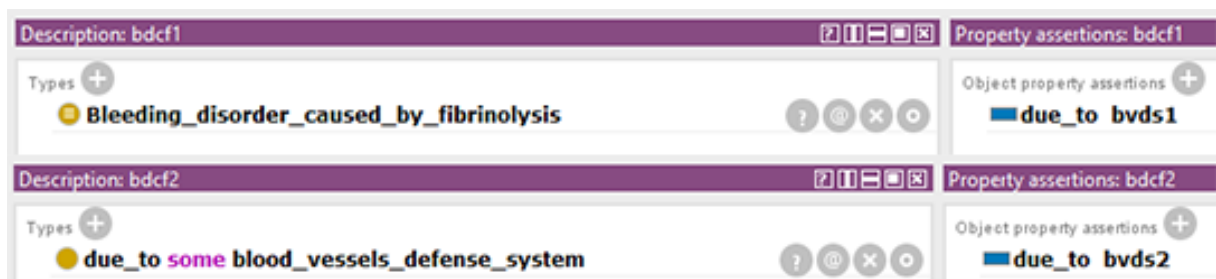
5.1.6.5 Indivíduos para *bleeding disorder caused by fibrinolysis*

De acordo com a metodologia citada na subseção 4.6.1, foram criados quatro indivíduos para a classe . Os indivíduos criados foram:

- **bdcf1 e bdcf2:** indivíduos da classe *bleeding disorder caused by fibrinolysis*, e;
- **bvds1 e bvds2:** indivíduos da classe *blood vessels defense system* e relacionados com a propriedade *due to*.

A inserção destes indivíduos são mostrados na Figura 27.

Figura 27 – Indivíduos para a classe *bleeding disorder caused by fibrinolysis*



5.1.7 Validação e raciocínio lógico dos termos segundo os motores de inferências

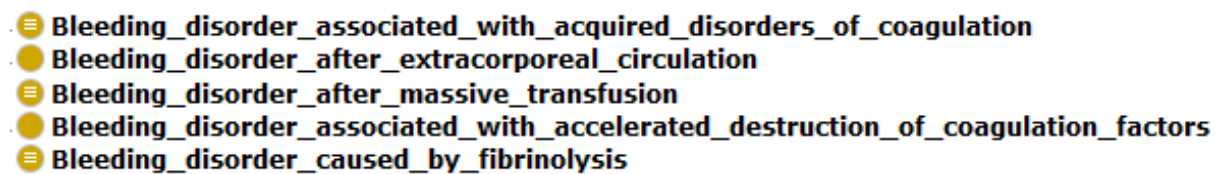
Dos motores de inferências listados na seção 3.4 apenas o Pellet, FaCT++ e HermiT são compatíveis com a versão do Protégé utilizada neste trabalho. Para cada motor de

inferência foi feito três testes distintos: consistência, inferências em TBoxes e inferências em ABoxes.

5.1.7.1 Validação de consistência sobre as classes modificadas

A validação de consistência realizada pelos motores de inferências citados verificou se as definições criadas na seção 5.1 produziram alguma irregularidade lógica em BLO. Os motores de inferências Pellet, FaCT++ e HermiT não encontraram nenhuma inconsistência nas classes criadas e modificadas na seção 5.1 como mostra a Figura 28.

Figura 28 – Validação de consistência das classes no Protégé

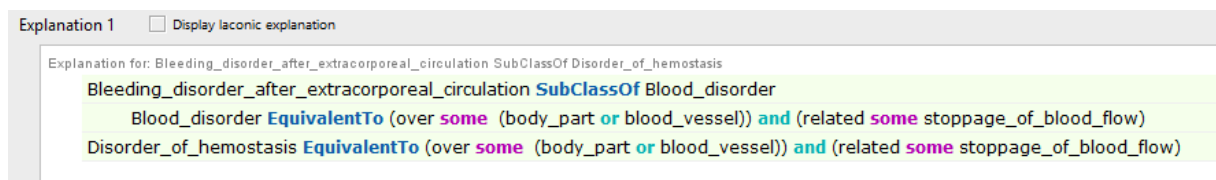


5.1.7.2 Inferências realizadas sobre as classes modificadas

Para cada uma das cinco classes modificadas na seção 5.1 foram executados os três motores de inferências listados. Depois foi feita a verificação se conhecimento novo foi deduzido por estes motores de inferência.

O FaCT++ detectou que *bleeding disorder after extracorporeal circulation* é equivalente a *blood disorder*. A explicação lógica da inferência mostrada na Figura 29 é resultante da definição feita na Figura 13. Uma desordem do sangue possui características de ocorrer sobre parte do corpo ou sobre os vasos sanguíneos e são relacionadas com a interrupção do fluxo sanguíneo. A mesma definição é encontrada na desordem do sangue após circulação extracorpórea.

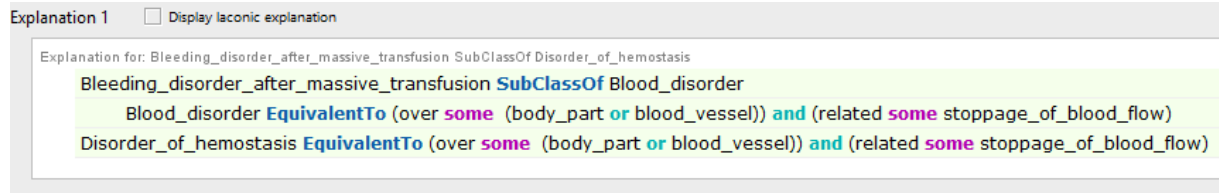
Figura 29 – Inferências sobre a classe *bleeding disorder after extracorporeal circulation* no Protégé



O FaCT++ também detectou que *bleeding disorder after massive transfusion* é uma subclasse de *disorder of hemostasis* como mostra a Figura 30. Esta dedução é uma consequência da dedução anterior. Uma vez que *bleeding disorder after massive transfusion* é subclasse de *blood disorder* e *blood disorder* é equivalente a *bleeding disorder after extracorporeal circulation*, logo *bleeding disorder after massive transfusion* é subclasse de

bleeding disorder after extracorporeal circulation. Uma demonstração do funcionamento do silogismo hipotético em OWL.

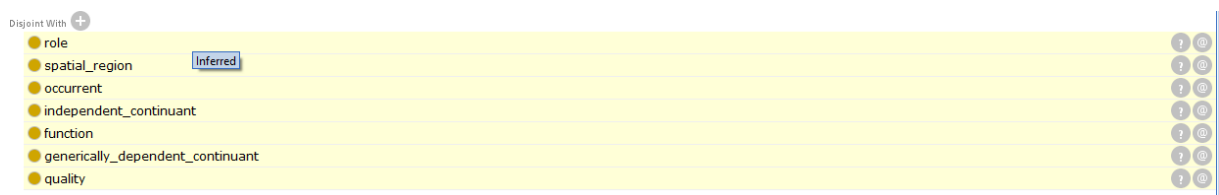
Figura 30 – Inferências sobre a classe *bleeding disorder after massive transfusion* no Protégé



As inferências mostradas na Figura 29 e na Figura 30 são consideradas incorretas pelo autor do ponto de vista ontológico. O fato da definição de *blood disorder* estar incompleta (com apenas as alterações sugeridas na subseção 4.5.2.1), faz com que a equivalência entre os dois termos seja um fator decisivo para dedução de conhecimento novo. Com a adição de novas características para *blood disorder* somados com as já existentes fará com que o próprio motor de inferência, no caso o FaCT++ corrija estas deduções equivocadas.

Além disso, o conhecimento novo gerado pelos motores de inferências também foram relacionados às classes de BFO que foi incorporada pela BLO, como mostra a Figura 31. As inferências feitas foram consequências por heranças das classes superiores às classes modificadas, todas oriundas de BFO. Todas as inferências deduzidas foram feitas por disjunção, ou seja, se uma classe pai é disjunta de outra classe, logo as classes filhas também serão. Como exemplo a classe *bleeding disorder caused by fibrinolysis* é subclasse de *disposition* que é disjunta da classe *role*. Todos os motores de inferência chegaram às mesmas conclusões de dedução de conhecimento, não tendo variações de um para o outro.

Figura 31 – Inferências realizadas pelo Protégé baseados nas definições de BFO



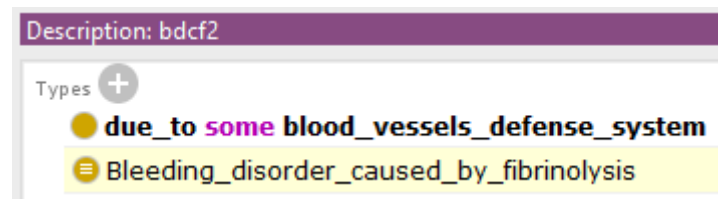
5.1.7.3 Inferências realizadas sobre os indivíduos adicionados

Para cada um dos indivíduos adicionados na subseção 5.1.6 foram executados os três motores de inferências listados. Depois foi feita a verificação se conhecimento novo foi deduzido destes motores de inferência.

Os três motores inferiram que os indivíduos que foram adicionados como conjuntos de definições (e não como um tipo específico de uma classe) são pertencentes às classes

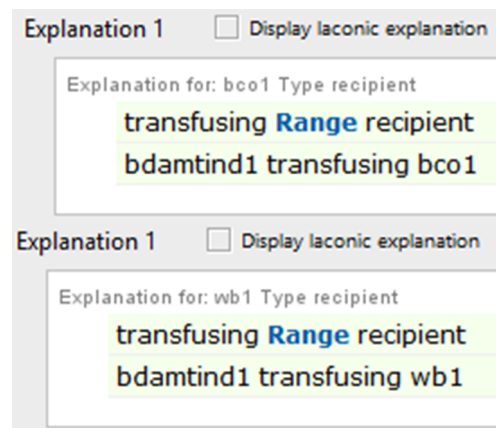
as quais possuem estes conjuntos de definições em suas relações de equivalência. Por exemplo, a Figura 32 mostra que o indivíduo *bdcf2* é também um indivíduo da classe *blood disorder caused by fibrinolysis*. Inferências do tipo reforçam a possibilidade de dedução de conhecimento novo baseado em definições. Por exemplo, um indivíduo não é necessariamente uma desordem causada por fibrinolíticos, porém seu conjunto de definições faz com que seja considerado como tal. Uma vez que a definição de desordem causada por fibrinolíticos esteja correta, as demais inferências também estarão.

Figura 32 – Inferência realizada pelo Protégé sobre o indivíduo *bdcf2*



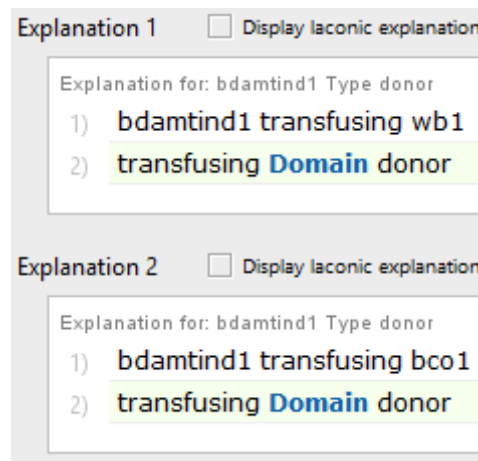
Os três motores de inferências inferiram que *bc1*, *bc2*, *wb1* e *wb2*, indivíduos correspondentes à classe *blood component* e *whole blood* respectivamente, também pertencem à classe *recipient*. Sua explicação mostrada na Figura 34 se dá pelo alcance de transfusão contemplar o recipiente (receptor da doação). Interpretando em linguagem natural, pode-se dizer que componente sanguíneo faz parte do recipiente.

Figura 33 – Inferência realizada pelo Protégé sobre os indivíduos *bc1*, *bc3*, *wb1* e *wb2*



Igualmente foi inferido pelos três motores que os indivíduos *bdamtind1* e *bdamtind2*, pertencentes à classe *blood disorder after massive transfusion*, também são indivíduos da *donor*. Sua explicação mostrada na Figura 34 se dá pelo domínio de transfusão contemplar o doador. Os indivíduos *bdamtind1* e *bdamtind2* fazem transfusão tanto de *blood component* quanto de *whole blood*. Interpretando em linguagem natural, pode-se dizer que: desordem do sangue após transfusão massiva faz parte do doador.

Figura 34 – Inferência realizada pelo Protégé sobre os indivíduos *bdamtind1* e *bdamtind2*



O autor entende que esta inferência é equivocada. A definição de *blood disorder after massive transfusion* dá a entender que esta desordem é constituída da transfusão total de sangue ou de componentes oriundas do doador. Percebe-se que a definição, ainda que considerada necessária e suficiente por Aganette (2015), carece de uma maior especificação quanto aos domínios e alcances e também quanto às regras para necessidade e suficiência do termo. O correto não seria que *blood disorder after massive transfusion* seria consequência do processo de transfusão realizado? Tal pergunta só pode ser respondida por um especialista do domínio.

Os três motores de inferências inferiram que *dh1* e *dh2*, pertencentes à classe *disorder of hemostasis*, são também indivíduos do tipo *blood disorder*. Sua explicação é oriunda do mesmo fator citado na Figura 29. Como citado no exemplo da equivalência entre *blood disorder* e *disorder of hemostasis*, as definições incompletas geram inferências equivocadas naquele instante em que são equivalentes.

5.2 OntoLogica: a plataforma online

Batizada como OntoLogica, um sistema para didática e transcrição de termos em linguagem natural para lógica descritiva em ontologias, esta plataforma online foi inspirada principalmente nos editores de ontologias listados na seção 3.5. Os principais editores analisados para a composição dos requisitos do OntoLogica foram: Protégé e webODE. O Protégé por ser uma ferramenta completa e abordar vários aspectos da engenharia de ontologias e; o webODE por ter sido uma plataforma online baseada em cliente e servidor, mesma proposta deste sistema.

Os **requisitos principais** do sistema são: ser uma plataforma simples onde os usuários serão capazes de carregar e editar as principais características de suas ontologias; oferecer suporte para a transcrição de termos em linguagem natural para a linguagem OWL e; fazer verificações e inferências das ontologias carregadas. O detalhamento destes

requisitos está acompanhado por modelos de telas desenhados no Balsamiq que ilustram a organização e composição do OntoLogica.

5.2.1 Disponibilidade

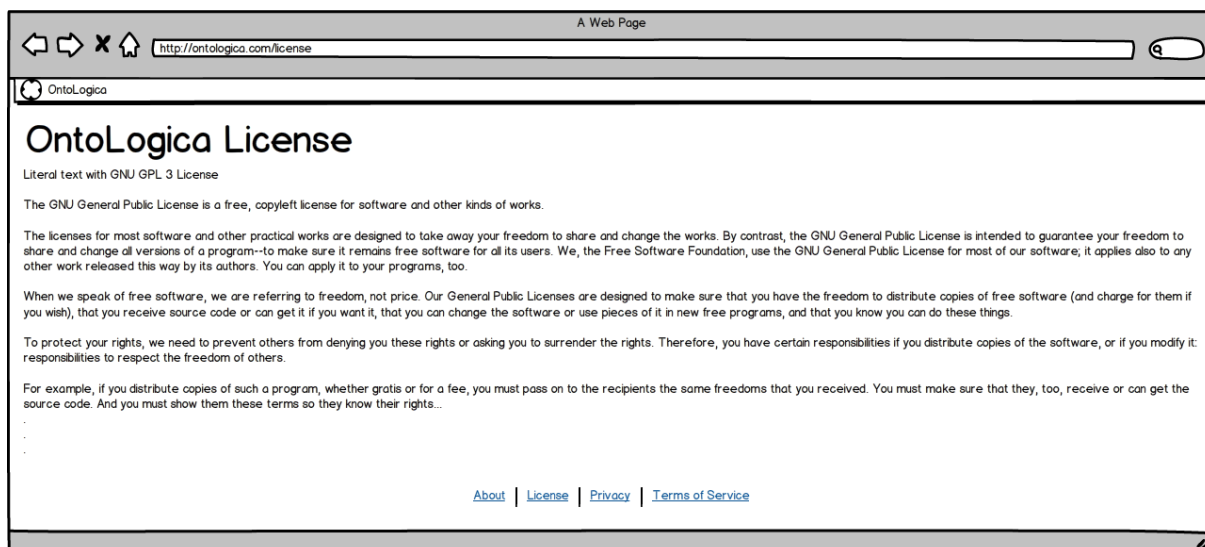
O código fonte da construção do OntoLogica deve ser *open source* sob a proteção da licença GNU GPLv3, pois o autor acredita que a ciência deve ser livre podendo ser usada e aprimorada por todos. Esta licença requer que a distribuição original do código fonte seja feita sob os mesmos termos e condições, além de fornecer uma garantia expressa dos direitos dos autores (STALLMAN et al., 1998). A Tabela 13 mostra as principais características desta licença:

Tabela 13 – Características da GNU GPLv3

Tipo	Característica	Descrição
Permissões	Uso comercial	O código e derivados podem ser usados para finalidades comerciais.
	Distribuição	A distribuição pode ser feita por qualquer pessoa.
	Modificação	A modificação pode ser feita por qualquer pessoa.
	Patente	O contribuinte com o sistema garante ao contribuidor direitos sobre a patente do sistema.
	Uso privado	O código fonte pode ser modificado para uso privado sem precisar ser distribuído.
Condições	Divulgação do código fonte	O código fonte deve estar disponível quando o sistema estiver disponível.
	Licença e direitos	O sistema deve ser acompanhado de uma cópia da licença.
	Licença idêntica	As modificações feitas por outros usuários devem ser distribuídas sob a mesma licença do original.
	Notificação de modificação	As modificações feitas devem notificadas.
Limitações	Responsabilidade	A responsabilidade do uso indevido por terceiros é limitada.
	Garantia	Não é fornecida nenhuma garantia ao usuário final.

Para que se assegure parte dos direitos fornecidos pela licença supracitada, o código original de desenvolvimento do OntoLogica deve estar disponível em repositório online, gratuito de acesso à todos. Uma página dedicada à licença do código deve estar presente no sistema, como na Figura 35.

Figura 35 – Modelo da página com a licença GNU GPLv3 do OntoLogica



5.2.2 Usuário

Cada usuário deverá ter a sua sessão única. Para isto é necessário que o mesmo faça cadastro no OntoLogica. As informações requeridas pelo OntoLogica ao usuário para cadastro serão e-mail e senha. Posteriormente será requerido que o mesmo preencha informações como nome completo e instituição de trabalho/estudo. Visando a praticidade, será permitido também ao usuário que faça cadastro utilizando redes sociais como Facebook⁸ e o Google⁹. As Figura 36 e Figura 37 mostra a tela inicial do OntoLogica com informações para autenticação e cadastro e também a tela pós-autenticação onde novos dados deverão ser incluídos.

⁸ <https://www.facebook.com/>

⁹ <https://www.google.com>

Figura 36 – Modelo da página inicial do OntoLogica

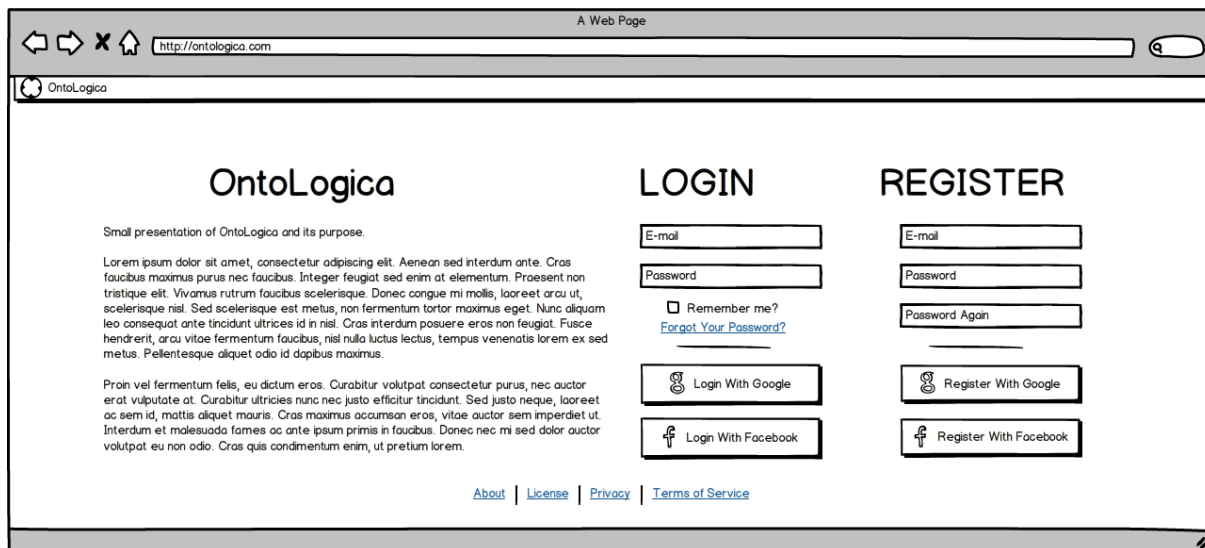
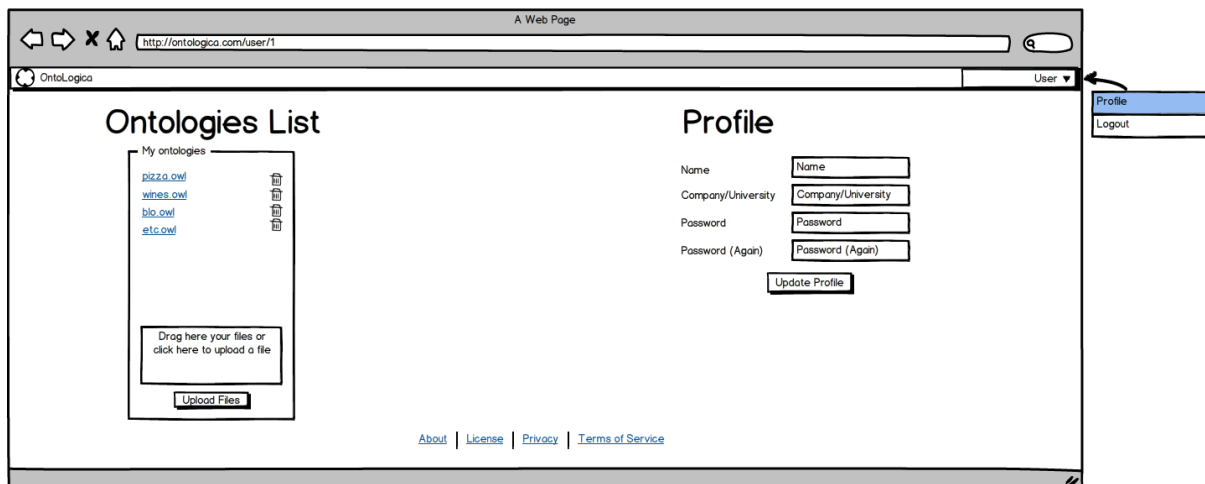


Figura 37 – Modelo do perfil do usuário do OntoLogica



O cadastro e autenticação do usuário permitirá que o mesmo faça uso de trabalhos personalizados, mantenha e use suas ontologias privadas no sistema.

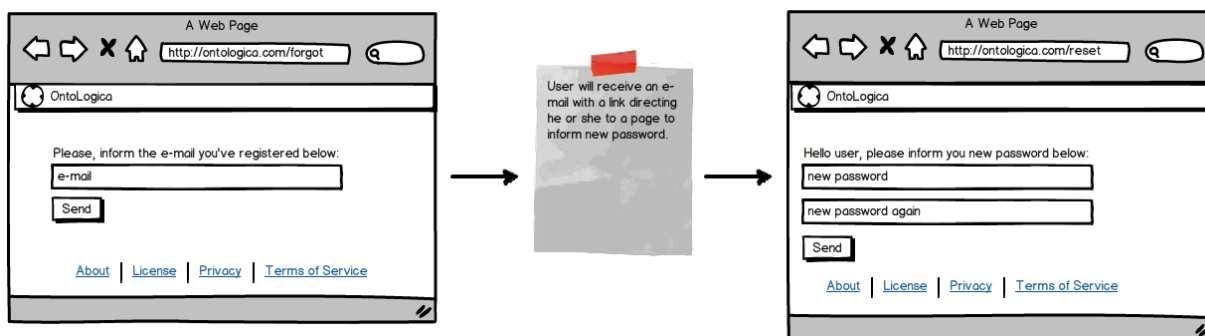
5.2.3 Privacidade e Segurança

Está vedada a coleta de dados dos usuários para fins comerciais, como oferta de produtos, propagandas em e-mail, etc. A única coleta permitida ao OntoLogica é para fins de pesquisa, como pesquisa de satisfação, perfil dos usuários, etc. Os dados dos usuários deverão ser sigilosos à terceiros e as senhas deverão ser criptografadas antes de serem armazenadas em banco de dados.

Dados pessoais dos usuários como as ontologias que eles carregarão no OntoLogica deverão ser de uso privado e exclusivo destes usuários. A recuperação de uma senha perdida

deverá ser feita via solicitação por e-mail e redefinição de senha através de um *link* único fornecido pelo OntoLogica. A Figura 38 mostra as telas e o fluxo para recuperação de senhas perdidas no sistema.

Figura 38 – Modelo do processo de recuperação de senha do OntoLogica



Duas páginas devem estar disponíveis ao usuário para que ele esteja ciente das políticas de privacidade e termos de serviço do OntoLogica. Os modelos estão ilustrados nas figuras 39 e 40.

Figura 39 – Modelo da página de termos de serviço do OntoLogica

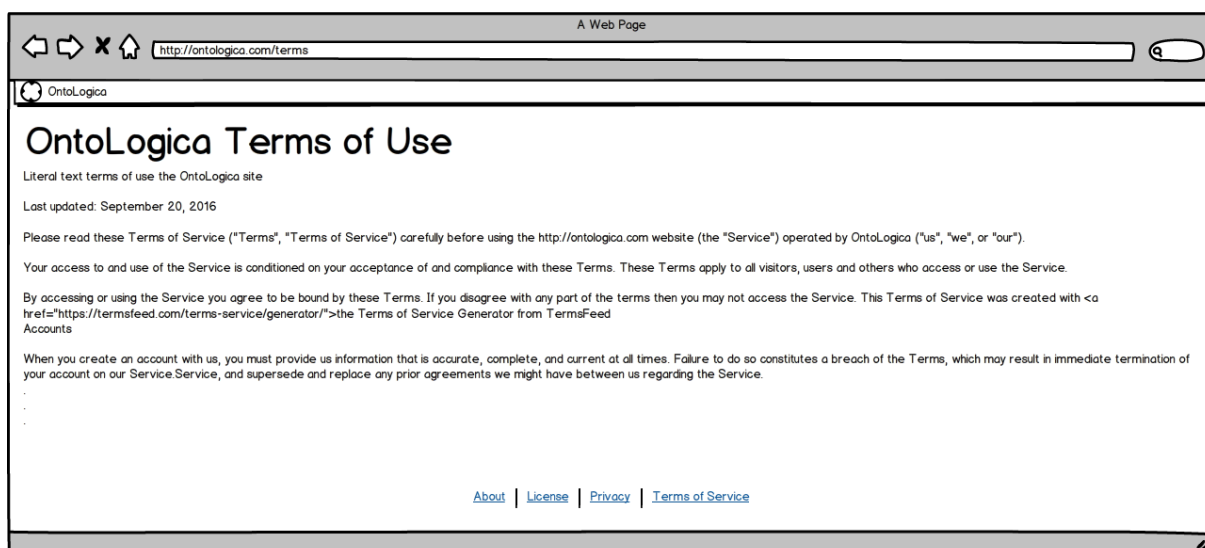


Figura 40 – Modelo da página de políticas de privacidade do OntoLogica



As ontologias carregadas pelos usuários, assim como qualquer informação editada posteriormente deve ser automaticamente salva pelo OntoLogica. O sistema deverá fazer a identificação de alterações e o salvamento automático destas alterações para que os mesmos não se percam.

5.2.4 Navegação

A navegação dentro do OntoLogica obedecerá procedimentos simples e intuitivos aos usuários. Na página inicial do sistema estará disponível ao usuários meios de autenticação como *login* via rede social ou digitando e-mail e senha. Uma vez autenticado, o sistema redirecionará o usuário para a página de seu perfil. Como mostrado na Figura 37, também irá aparecer um botão na parte superior da direita com um menu *dropdown* onde o usuário poderá ir para a página de seu perfil, caso esteja em outra página, ou encerrar a sua sessão. Para acessar uma ontologia que ele carregou, basta que clique no *link* com o nome da ontologia e o sistema irá redirecionar o usuário até as telas de edição referente àquela ontologia. Uma vez editando uma ontologia, o usuário poderá alternar entre as suas funcionalidades clicando nas abas de serviço, como demonstrado na Figura 41.

Para voltar à página inicial, bastará que o usuário clique no botão OntoLogica localizado na parte superior esquerda demonstrada na Figura 37.

5.2.5 Funcionalidades

Como descrito anteriormente, as funcionalidades primordiais para o OntoLogica são: carregar e editar ontologias; fazer inferências; fornecer didática para transcrição de

termos da linguagem natural para a linguagem OWL. Esta subseção irá detalhar cada processo e seu fluxo.

5.2.5.1 Carregar e excluir de ontologias

Na tela do perfil do usuário, mostrado na Figura 37 é fornecido a opção para que o mesmo carregue suas ontologias no sistema. O OntoLogica deverá ser capaz de fazer o *upload* de arquivos somente em OWL e com limite máximo de dois *megabytes* para evitar sobre carregamento do sistema. Os arquivos carregados pelos usuários devem ser salvos em um banco de dados e estar disponíveis sempre que o usuário autenticar no sistema. Estes arquivos devem ser privados e de uso exclusivo do usuário que fez seu *upload*. Não há limites de ontologias para o usuário carregar.

Para excluir uma ontologia carregada, basta que o usuário clique no ícone em formato de lixeira logo a direita do nome da ontologia a ser excluída. Um *popup* deverá abrir pedindo a confirmação do usuário para fazer a exclusão, alertando-o que o procedimento não pode ser revertido. Uma vez confirmado, os registros desta ontologias devem ser excluídas do banco de dados do OntoLogica.

5.2.5.2 Criando, editando e excluindo classes

O OntoLogica possui uma aba para a manipulação de classes em OWL. Esta aba que é mostrada na Figura 41 permite a criação, edição e exclusão de classes. No canto esquerdo ficará um menu que mostrará a árvore das classes, começando com a classe mais básica *Thing*. Além disto conterà os botões *Create*, *Delete* e *Search*. O primeiro criará uma classe nova como sendo subclasse da classe selecionada. Caso nenhuma classe tenha sido selecionada, a nova classe será subclasse de *Thing*. O segundo botão deletará a classe selecionada com a confirmação do usuário, como mostrado na Figura 42. O terceiro botão fará uma busca entre as classes criadas, para a localização rápida em uma ontologia grande, como mostrado na Figura 43. Ao localizar uma classe que foi buscada através da ferramenta, o OntoLogica deverá selecioná-la e expandir todos os ramos necessários até chegar na classe selecionada.

Na parte central da tela conterà quatro campos correspondente às características das classes como mostrada na Figura 41. Eles são:

1. ***Display Name:*** Nome da classe;
2. ***Property:*** propriedades criadas pela aba *Properties* (ver subseção 5.2.5.3) e/ou propriedades pré-definidas pela linguagem RDF e OWL;
3. ***Value:*** Valor da propriedade de acordo com o tipo e;
4. ***Language:*** Linguagem que está sendo documentada a classe.

Figura 41 – Modelo da aba de edição de classes do OntoLogica

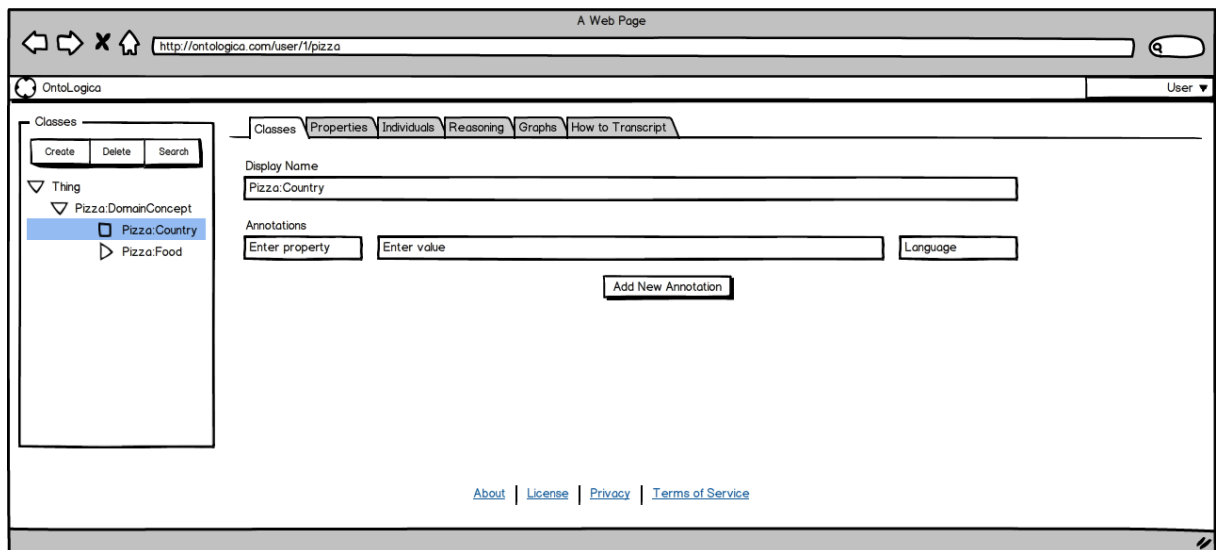


Figura 42 – Modelo de exclusão de classes do OntoLogica

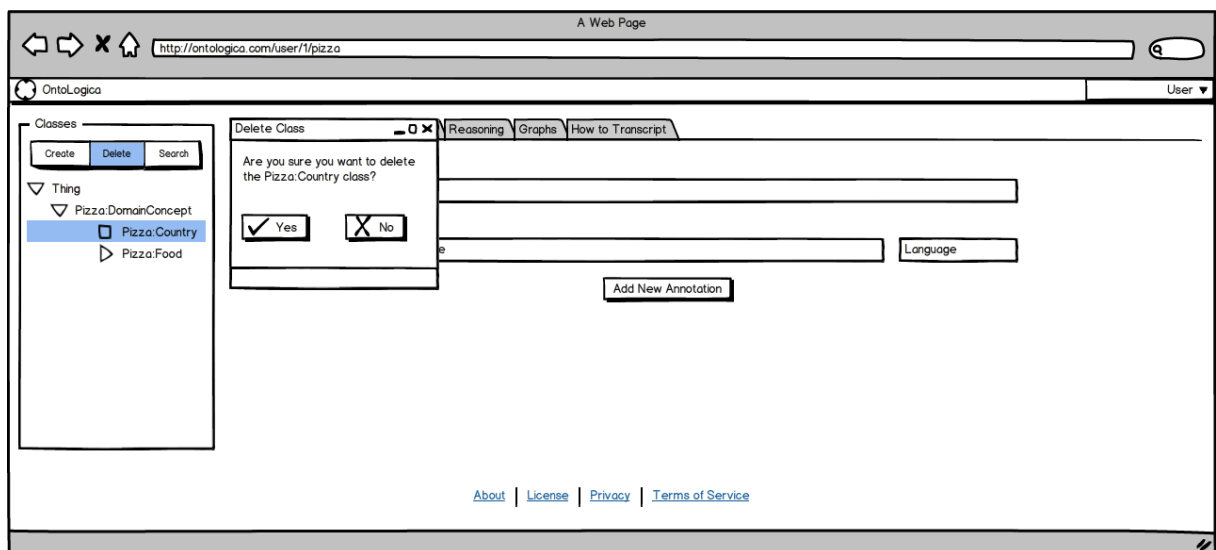
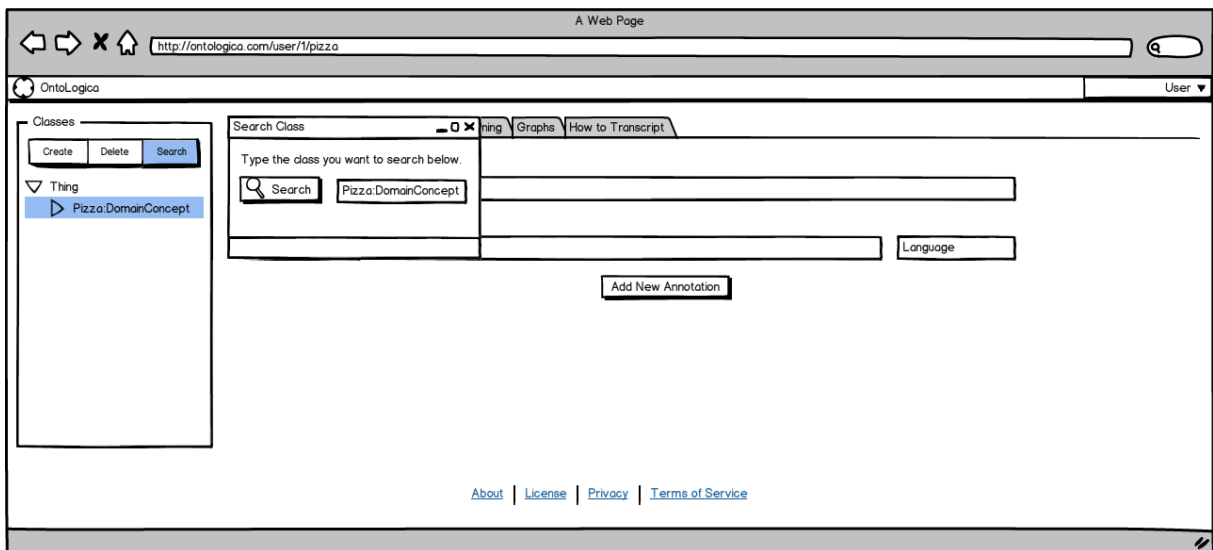
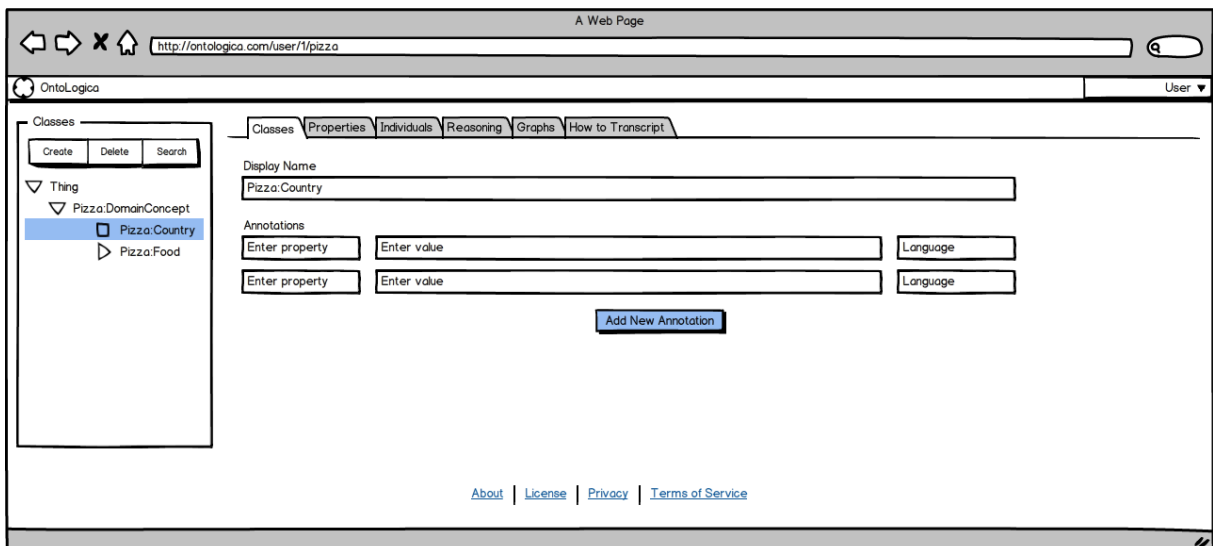


Figura 43 – Modelo de busca de classes do OntoLogica



Abaixo dessas informações há um botão chamado *Add New Annotation* que adicionará novos campos mostrados acima, para adicionar novas informações a uma classe. Este fluxo está representado na Figura 44.

Figura 44 – Modelo de funcionamento do botão *Add New Annotation* do OntoLogica

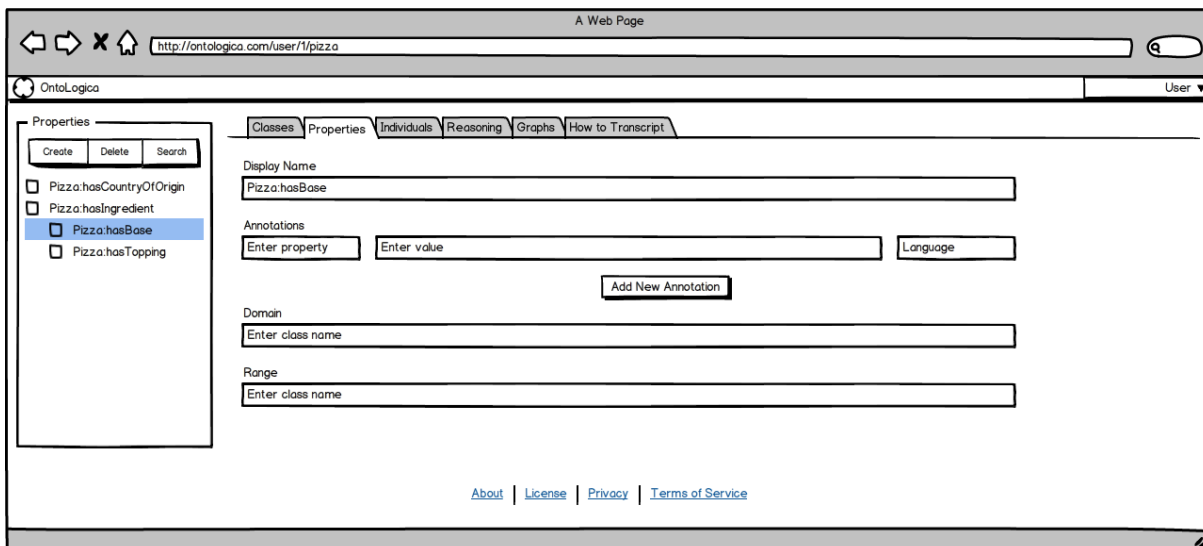
5.2.5.3 Criando, editando e excluindo propriedades

A aba de propriedades mostrada na Figura 45 contempla a criação, edição, remoção e busca de propriedades. As funcionalidades de criar, deletar e buscar são idênticas das mostradas na subseção 5.2.5.2. A função de edição ganha dois campos novos:

1. **Domain:** classe primária origem da propriedade e;

2. **Range:** classe secundária destino da propriedade.

Figura 45 – Modelo da aba de edição de propriedades do OntoLogica



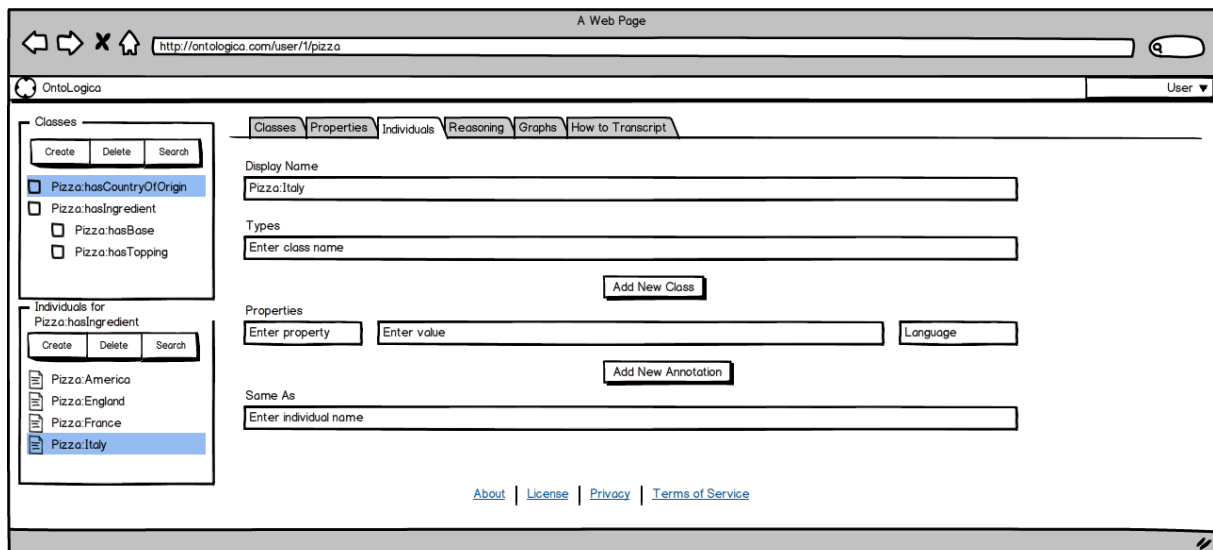
Os novos campos serão usados para construir os relacionamentos das classes criadas anteriormente. O menu à esquerda mostra as propriedades criadas em formato de árvore, ou seja, cada propriedade pode abrigar uma ou mais propriedades que herdarão as mesmas propriedades e anotações do pai.

5.2.5.4 Criando, editando e excluindo indivíduos

A aba de indivíduos mostrada na Figura 46 contempla a criação, edição, remoção e busca de indivíduos. As funcionalidades de criação, remoção e buscas são idênticas das mostradas na subseção 5.2.5.2. As diferenças principais se encontram na disposição do *layout* da página. No canto esquerdo serão apresentados dois menus diferentes: as classes da ontologia que serão utilizadas para definir à qual classe um indivíduo pertencerá e; o menu de indivíduos, onde estarão listados todos os indivíduos pertencentes a uma determinada classe selecionada no menu acima. A função de edição de indivíduos contém dois campos diferentes dos mostrados nas subseções anteriores:

1. **Types:** adiciona uma nova classe ao qual este indivíduo pertencerá. Um indivíduo poderá pertencer a mais de uma classe. Para isto há um botão chamado *Add New Class* que criará um campo novo de classe para ser editado e;
2. **Same As:** relação de equivalência entre um indivíduo e outro.

Figura 46 – Modelo da aba de edição de indivíduos do OntoLogica

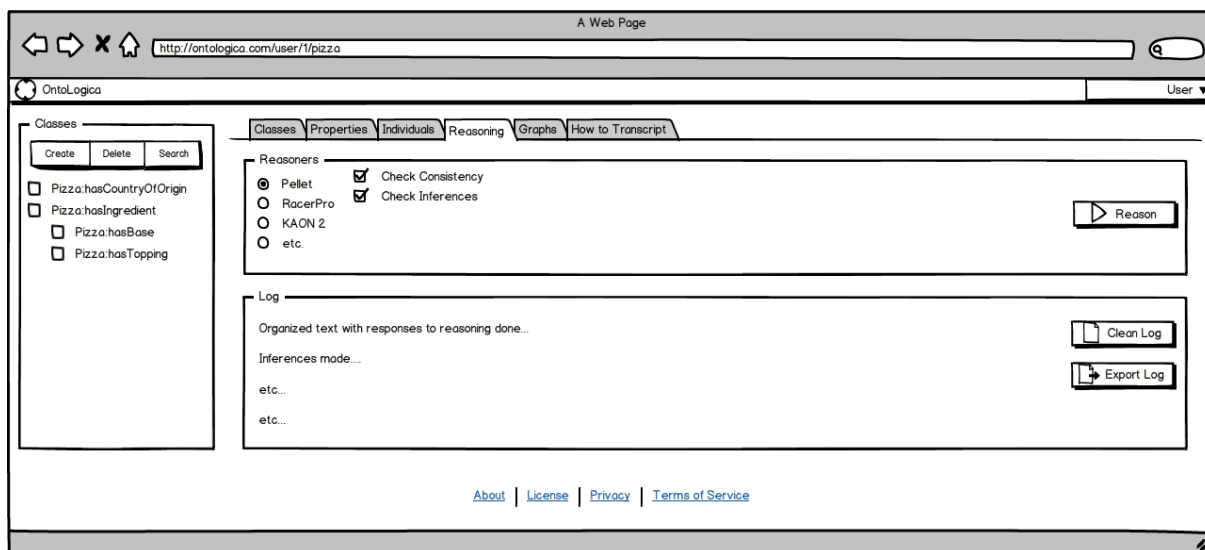


5.2.5.5 Realizando inferências

A aba de inferências denominada como *Reasoning* na Figura 47 é a parte do OntoLogica onde verificações serão feitas pela ontologia construída. Nesta aba há um menu de motores de inferências como Pellet, KAON2 e outros. Ainda nesta aba o usuário poderá escolher se ele deseja que o motor verifique apenas consistência ou realize também inferências lógicas. O botão *Reason* executará o motor de inferência e um *log* será produzido no menu logo abaixo.

O *log* conterà todas as verificações que foram feitas pelo motor de inferência e os resultados, como a ontologia sendo consistente ou não, etc. O usuário terá a opção de apagar o *log* ou exporta-lo para um documento do tipo texto. As informações do *log* devem ser traduzidas para uma linguagem capaz de compreensão e análise humana.

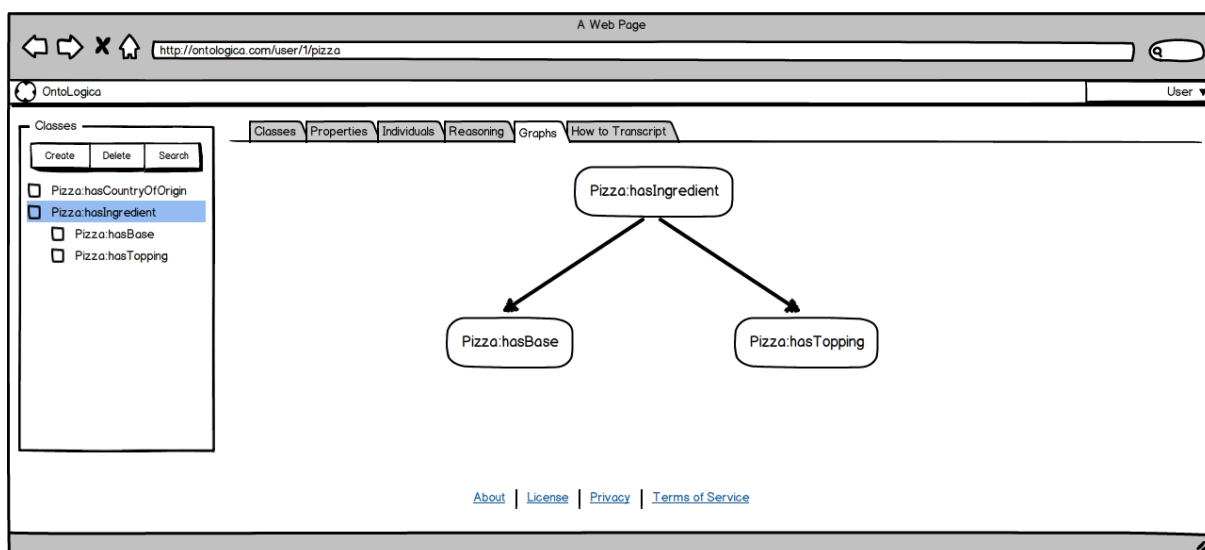
Figura 47 – Modelo da aba de inferências do OntoLogica



5.2.5.6 Recursos de visualização

O OntoLogica contará também com recursos de visualização gráfica como mostrado na Figura 48. Esta aba mostrará as classes criadas pela ontologia e a ligação q elas possuem através das propriedades criadas na subseção 5.2.5.3. A visualização da ontologia será através de grafos onde as classes serão os nodos e as propriedades as arestas que conectam as classes. Esta aba funcionará apenas para motivos didáticos, não sendo possível a manipulação de classes ou propriedades.

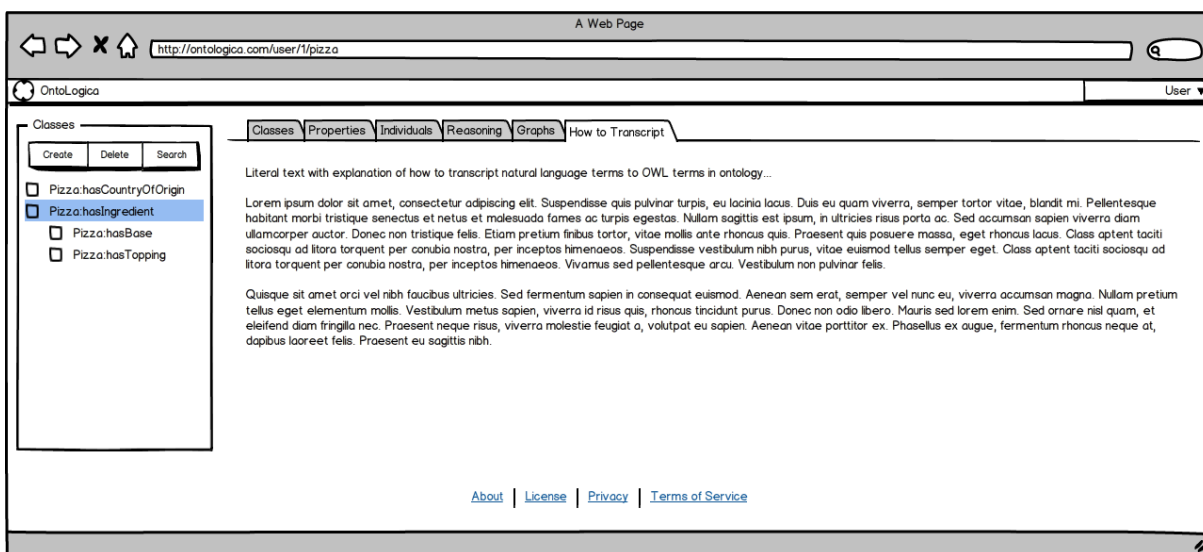
Figura 48 – Modelo da aba de visualização de classes OntoLogica



5.2.5.7 Instruções para a transcrição de termos

A aba de transcrição de termos mostrada na Figura 49 seguirão os modelos criados na seção 5.1. Nela deverá estar disponível todas as informações para que o usuário compreenda e seja capaz de realizar a transcrição de termos de uma ontologia e portá-lo para o seu sistema. As informações devem estar disponíveis de forma didática e categórica, pontuando as regras e o passo-a-passo do usuário para fazer uma transcrição de termos sem equívocos.

Figura 49 – Modelo da aba de transcrição de termos do OntoLogica



6 Conclusões e trabalhos futuros

O Capítulo 5 mostrou o desenvolvimento resultante dos trabalhos levantados nos capítulos 2 e 3 e da metodologia proposta no Capítulo 4. O presente capítulo propõe-se a discutir as motivações iniciais que levaram ao desenvolvimento deste trabalho e debater sobre os resultados apresentados pelo Capítulo 5. Também foram apresentados neste capítulo as dificuldades encontradas durante o desenvolvimento deste trabalho bem como os trabalhos futuros a serem realizados. Trabalhos estes que foram originados das dificuldades encontradas.

6.1 Motivações iniciais

As principais motivações que originou o seu desenvolvimento foram: elucidar ao público da Ciência da Informação o uso e as vantagens de modelos como ontologias e; ajudar na didática da construção de ontologias oferecendo um método de transcrição de termos e um protótipo de ferramenta para estas transcrições. Estas motivações surgiram em virtude do pouco conhecimento que profissionais da ciência da informação tem sobre o campo da ontologia aplicada.

É sabido que o campo da ontologia aplicada oferece vantagens quanto a outros sistemas de modelagem informacional. Uma ontologia possui os mesmos recursos de uma taxonomia e de um tesouro e em adição, oferece que os termos definidos tenham relações entre si. No entanto não fica claro as possibilidades além, que uma ontologia oferece com o auxílio de artefatos computacionais. Este trabalho propôs a mostrar ao leitor o potencial que uma ontologia agrega a um modelo de dados bem definido.

Neste sentido pensou-se e trabalhou-se para elucidar o funcionamento da lógica dentro de um sistema ontológico. A motivação encontrou-se no potencial oferecido pelo uso da lógica na ontologia aplicada e seus inúmeros benefícios. No entanto, assim como os motivos de uma ontologia aplicada eram pouco conhecidos por usuários da informação, a lógica também é um recurso pouco explorado por estes usuários.

Procurou então desvendar o funcionamento da lógica, além de explicar ao leitor o necessário e suficiente para o entendimento da lógica e como ela atua nos sistemas ontológicos. Sua potencialidade foi demonstrada através da explicação sobre inferências e constatações lógicas que são feitas com auxílio de motores próprios para isso.

Como o ensino da lógica é uma área pertencente à outro campo de ensino, notou-se também a necessidade do ensino da mesma para aplicação em ontologias. Por isto este trabalho propôs-se a criar um método para uso da lógica em transcrições de termos em

sistemas ontológicos. Este método permite ao usuário fazer uso da lógica enriquecendo estes sistemas, fazendo uso de seu potencial que é dar valor semântico aos termos e relações.

Tendo em vista que as ferramentas computacionais para auxílio de pessoas com pouco conhecimento em ontologias e em lógica são pouco amigáveis, surgiu a motivação de criar uma ferramenta com tal propósito. Tal ferramenta auxiliará na didática de novos usuários que ingressarem no campo da ontologia aplicada. A motivação maior para criação desta ferramenta é a criação de termos e propriedades definidas com lógica e testes de inferências auxiliadas por programas de terceiros. Esta plataforma foi prototipada para trabalho futuro.

6.2 Conclusões sobre os resultados

Do ponto de vista das motivações citadas na seção 6.1 considera-se que este trabalho atingiu seu objetivo. O Capítulo 2 explicou o funcionamento da lógica, desde os conceitos básicos até conceitos avançados de dedução como o método dos *tableaux* analíticos. O Capítulo 3 explicou a relação da ontologia aplicada à computação. Seus artefatos, usos e benefícios trazidos. Com base nisso o Capítulo 4 criou um método para uso da lógica em ontologias assim como sugeriu uma nova ferramenta que ajudasse na didática desse método.

O ponto positivo observado nos resultados foi a ausência de inconsistências apresentadas nos termos transcritos. Isto deu uma garantia de que, para os casos selecionados, o método funcionou corretamente. No entanto, os resultados encontrados pelos motores de inferência para avaliar os termos criados e modificados mostraram-se insatisfatórios.

Os resultados para inferências obteve resultados positivos e também negativos. Positivamente temos as inferências baseadas em definições, que identificaram que determinados indivíduos são pertencentes à uma determinada classe. Isto dá ao profissional da ciência da informação liberdade para caracterizar os indivíduos de acordo com suas propriedades mais evidentes. Um indivíduo x é caracterizado por ser uma reação do sistema de defesa dos vasos sanguíneos. Um motor de inferência será capaz de inferir que este indivíduo é uma desordem do sangue causada por fibrinolíticos. Uma vez que as definições sejam feitas corretamente e indicados as condições necessárias e suficientes, os sistemas de apoio ao usuário será capaz de deduzir conhecimento novo.

No entanto, as transcrições geraram também resultados que foram considerados equivocados. Tal equívoco foi gerado pela incompletude das transcrições feitas e também por definições formais ambíguas e muito extensas. Estes equívocos acendem um alerta para a definição formal do conhecimento. Ainda é preciso ter muita cautela para que estas definições sejam feitas sem gerar múltiplas interpretações. A lógica requer que as sentenças sejam mais objetivas, como por exemplo: "um triângulo tem exatamente três lados". No

entanto, as definições biomédicas são sentenças compostas de múltiplos substantivos conectados por múltiplos verbos. Há uma necessidade de quebrar estas sentenças em sentenças menores para elucidar melhor a definição.

De modo geral os resultados foram positivos por permitirem que a lógica e a ontologia aplicada fossem demonstradas como uma ferramenta única que auxilia os domínios do conhecimento, em especial a medicina, a construir modelos fortes e bem estruturados sob regras rígidas, que é a lógica descritiva. O modelo de transcrição criado junto com a prototipação de uma ferramenta didática, lança novos olhares para esse casamento da lógica e ontologia aplicada. Dado isso, espera-se que novos pesquisadores e usuários desses sistemas de modelagem façam mais uso desse ferramental que a ontologia oferece, criando assim modelos mais consistentes e confiáveis.

6.3 Dificuldades encontradas

O estudo da ontologia aplicada perpassa por diversos campos do conhecimento além da Ciência da Informação. Áreas como a ciência da computação, lógica e medicina também precisam ser contemplados em virtude da qualidade do trabalho proposto. Isto abre um leque grande de conhecimento que precisa ser abrangido pelo pesquisador durante a execução do trabalho. Temas muito complexos precisam ser aprofundados e isto constantemente coloca o pesquisador em um dilema: aprofundar-se em apenas alguns temas deixando os demais de lado ou, tratar cada tema sob um ponto de vista raso e superficial, mas contemplar todos eles? O presente trabalho precisou ser aprofundado em quatro áreas distintas: ontologia aplicada, lógica, computação e medicina do sangue.

Do ponto de vista lógico precisou debruçar-se sobre os conceitos da lógica, suas terminologias e funcionamento. Quais as famílias lógicas existentes? Quais dessas famílias se enquadrariam no uso de ontologia aplicada? Qual o papel da lógica nas ontologias? Como a lógica atua nesses sistemas? Quais seus benefícios e o que o profissional da ciência informação precisa saber para utilizar desses benefícios em seu favor? Essas foram as principais questões que precisaram ser respondidas durante o estudo. O campo da lógica demonstrou ser vasto e profundo, com muito a que estudar e muito a se falar.

A computação se mostra como um campo bastante volátil. As tecnologias são mutáveis assim como a natureza. Em um curto espaço de tempo uma tecnologia nasce e morre. Algumas delas continuam a ser citadas por anos, mas não existem mais vestígios "físicos" que permita ao pesquisador utilizar e testar determinada tecnologia. Não há rastros. Nenhum arquivo pode ser encontrado. Há páginas que não são atualizadas há muito tempo, mas continuam sendo o único recurso disponível sobre aquele assunto. Este trabalho se propôs a responder algumas perguntas durante sua execução. Qual o relacionamento da ontologia aplicada com a tecnologia? Qual o ferramental tecnológico disponível para o

profissional da ciência da informação trabalhar com ontologia aplicada em um ambiente informatizado? Como as ferramentas auxiliam estes profissionais? Qual o estado da arte e as possíveis melhorias que podem ser apresentadas para o desenvolvimento da ontologia aplicada?

Durante o desenvolvimento deste trabalho para responder a tais perguntas notou-se que muitas das ferramentas ficaram estagnadas no tempo. Ferramentas como WebODE, CRACK, KRIS, etc. que são consideradas robustas por outros pesquisadores não têm atualizações desde a década passada. Há uma certa exaltação a um ferramental que se tornou obsoleto perante o tempo. Novos materiais científicos são publicados e a lista de componentes desatualizados persiste entre estes materiais, criando-se assim um ciclo sem fim. Notou-se que as pesquisas ferramentais para ontologia aplicada carece de um filtro atualizado das ferramentas disponíveis.

As ferramentas disponíveis exigem do profissional um *know-how* para manipulá-las. Algumas ferramentas como Protégé, Jena, etc. exigem do usuário um conhecimento prévio para manipulá-las uma vez que não são familiares à primeira vista. O mesmo vale para o desenvolvimento da plataforma online sugerida na seção 5.2. O desenvolvimento de software requer do desenvolvedor conhecimentos amplos sobre diversas ferramentas como APIs, *frameworks*, banco de dados, linguagens de programação, redes de computadores, plataforma cliente-servidor, etc. O campo do ferramental disponível e o conhecimento necessário para administrar cada uma destas ferramentas demonstrou ser vasto e profundo, com muito a ser estudado.

A teoria da ontologia apresentada por Aganette (2015) somada à rigidez imposta pelo campo da medicina consegue produzir definições formais para termos biomédicos. No entanto tais definições esbarram no sentido ambíguo que as palavras conseguem produzir na linguagem natural. Durante a transcrição dos termos houve dificuldade para a tomada de certas decisões. Por exemplo, desordem hemorrágica causada por fibrinolíticos é considerada um sistema de defesa dos vasos sanguíneos. O que quer dizer exatamente a palavra "considerada"? Segundo o Priberam¹ o significado de considerar pode ser: examinar atentamente; ter em consideração e; ponderar, calcular, reputar. Neste caso, o que deveria ser considerado? Que a desordem hemorrágica é devido ao uso de fibrinolíticos? O sistema de defesa dos vasos sanguíneos age de acordo com o uso destes fármacos? Se sim, qual deveria ser o verdadeiro uso para a palavra considerar? Tais decisões se tornam complexas quando não se tem o conhecimento ou o auxílio do especialista na área biomédica. Além de sua complexidade, tais decisões podem comprometer o raciocínio lógico dos motores de inferências, uma vez que, associando um termo erroneamente a outro, demais associações serão inferidas erroneamente.

¹ <https://www.priberam.pt/>

6.4 Trabalhos futuros

As dificuldades listadas na seção 6.3 motivaram uma série de ideias para contribuir com o campo da ontologia aplicada. Estas ideias contemplam duas linhas de raciocínio diferentes: a didática do ensino de ontologia aplicada vinculada ao uso da lógica, e; meios de facilitação para a construção de ontologias.

A primeira ideia contempla o protótipo feito na seção 5.2 e outros meios de ensino de lógica em ontologias. Ficou claro que a lógica é uma área pouco contemplada na Ciência da Informação. Essa área deveria estar mais presente, principalmente vinculada ao ensino de ontologia aplicada, uma vez que seu uso garante maior confiabilidade às ontologias construídas.

A construção do OntoLogica junto à comunidade interessada, também colaborará com a disseminação e uso da lógica por usuários modeladores de ontologias. Além da plataforma poder contar com meios didáticos de ensino de transcrição de termos, o que auxiliará ainda mais na disseminação do uso da lógica em sistemas ontológicos.

A segunda ideia visa a facilitação do profissional da ciência da informação ao construir ontologias. Como foi citado durante este trabalho, o processo de construção de ontologias envolvem diversas áreas do conhecimento distintas e que nem sempre compete ao profissional envolvido. Tais empecilhos elevam o nível de dificuldade para a construção destes artefatos, uma vez que o usuário precisará recorrer a outros profissionais para a execução de determinadas etapas da construção de uma ontologia.

A facilitação deste profissional envolve a automatização de determinados processos ao construir uma ontologia. Por exemplo, o processo de definição, de criação de novas classes e propriedades, de transcrição de termos da linguagem natural para lógica, de definir relações, etc. podem e devem ter seus meios facilitados. Ainda que uma tarefa difícil, o processo de automatização permitirá que o profissional da ciência da informação tome menos decisões difíceis durante o desenvolvimento de uma ontologia. Essa economia cerebral permitirá que o usuário concentre-se em aspectos importantes do desenvolvimento de uma ontologia.

Espera-se contribuições ao campo da ontologia aplicada, tanto do ponto de vista didático, quanto do ponto de vista processual. Trabalhando nestes aspectos espera-se que o campo evolua de forma progressiva e que produza resultados mais satisfatórios. Que as ontologias possam cumprir o seu papel de modelar o conhecimento disponível e dar semântica aos dados.

Referências

- AGANETTE, E. C. *Representação do conhecimento biomédico: uma investigação sobre a teoria da terminologia e a teoria da ontologia aplicada no domínio do sangue humano*. Tese (Doutorado) — Universidade Federal de Minas Gerais, 2015. Citado 14 vezes nas páginas 24, 74, 91, 94, 95, 98, 100, 103, 105, 107, 109, 110, 120 e 136.
- AGANETTE, E. C.; ALMEIDA, M. B. Avanços teórico-empíricos na comparação entre teoria da ontologia e teoria da terminologia. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, v. 20, n. 44, p. 3–24, 2015. Citado 2 vezes nas páginas 78 e 79.
- AHMED, A. Q. *The Deliverance: Logic*. [S.l.]: OUP Oxford, 2010. Citado na página 27.
- ALEXANDER, P. *Finding Ontologies*. 2011. Disponível em: <<http://marinemetadata.org/guides/vocabs/ont/existing/finding>>. Citado na página 80.
- ALLEMANG, D.; HENDLER, J. *Semantic web for the working ontologist: effective modeling in RDFS and OWL*. [S.l.]: Elsevier, 2011. Citado na página 35.
- ALMEIDA, M. B. Revisiting ontologies: a necessary clarification. *Journal of the American Society for Information Science and Technology*, Wiley Online Library, v. 64, n. 8, p. 1682–1693, 2013. Citado na página 48.
- ALMEIDA, M. B.; BARBOSA, R. R. Ontologies in knowledge management support: A case study. *Journal of the American Society for Information Science and Technology*, Wiley Subscription Services, Inc., A Wiley Company, v. 60, n. 10, p. 2032–2047, 2009. ISSN 1532-2890. Disponível em: <<http://dx.doi.org/10.1002/asi.21120>>. Citado na página 23.
- ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. *Ciência da Informação, Brasília*, SciELO Brasil, v. 32, n. 3, p. 7–20, 2003. Citado na página 47.
- ALMEIDA, M. B. et al. The blood ontology: An ontology in the domain of hematology. In: CITESEER. *ICBO*. [S.l.], 2011. Citado na página 73.
- ANTONIOU, G.; HARMELEN, F. V. *A semantic web primer*. [S.l.]: MIT press, 2004. Citado 2 vezes nas páginas 50 e 59.
- ARPÍREZ, J. C. et al. Webode: a scalable workbench for ontological engineering. In: ACM. *Proceedings of the 1st international conference on Knowledge capture*. [S.l.], 2001. p. 6–13. Citado na página 69.
- ARPÍREZ, J. C. et al. Webode in a nutshell. *AI magazine*, v. 24, n. 3, p. 37, 2003. Citado na página 69.
- BAADER, F. *The description logic handbook: Theory, implementation and applications*. [S.l.]: Cambridge university press, 2003. Citado 7 vezes nas páginas 36, 41, 42, 43, 44, 63 e 64.

BAADER, F.; BRANDT, S.; LUTZ, C. Pushing the envelope further. Citeseer, 2008. Citado na página 38.

BAADER, F.; BUCHHEIT, M.; HOLLANDER, B. Cardinality restrictions on concepts. *Artificial Intelligence*, Elsevier, v. 88, n. 1, p. 195–213, 1996. Citado na página 44.

BAADER, F.; HANSCHKE, P. A scheme for integrating concrete domains into concept languages. In: MORGAN KAUFMANN. *12th International Joint Conference on Artificial Intelligence*. [S.l.], 1991. p. 452–457. Citado na página 63.

BAADER, F.; HANSCHKE, P. Extensions of concept languages for a mechanical engineering. Saarländische Universitäts-und Landesbibliothek, 1992. Citado na página 63.

BAADER, F.; HOLLUNDER, B. Kris: Knowledge representation and inference system. *Acm Sigart Bulletin*, ACM, v. 2, n. 3, p. 8–14, 1991. Citado na página 63.

BAADER, F.; HOLLUNDER, B. A terminological knowledge representation system with complete inference algorithms. In: *Processing Declarative Knowledge*. [S.l.]: Springer, 1991. p. 67–86. Citado na página 63.

BAADER, F. et al. An empirical analysis of optimization techniques for terminological representation systems. *Applied Intelligence*, Springer, v. 4, n. 2, p. 109–132, 1994. Citado na página 63.

BAADER, F.; HORROCKS, I.; SATTTLER, U. Description logics. *Foundations of Artificial Intelligence*, Elsevier, v. 3, p. 135–179, 2008. Citado na página 35.

BAADER, F.; HORROCKS, I.; SATTTLER, U. Description logics. In: *Handbook on ontologies*. [S.l.]: Springer, 2009. p. 21–43. Citado na página 37.

BALLARD, K. E.; COPI, I. M. *Study Guide for Copi Introduction to Logic*. [S.l.]: Monograph Collection (Matt - Pseudo), 1978. Citado na página 27.

BARR, A.; FEIGENBAUM, E. *The Handbook of Artificial Intelligence*. [S.l.]: Heuristech Press, 1981. (The Handbook of Artificial Intelligence, v. 1). ISBN 9780865760882. Citado na página 24.

BAUMGARTNER, P. Hyper tableau—the next generation. In: SPRINGER. *International Conference on Automated Reasoning with Analytic Tableaux and Related Methods*. [S.l.], 1998. p. 60–76. Citado na página 64.

BERNERS-LEE, T. et al. The semantic web. *Scientific american*, New York, NY, USA., v. 284, n. 5, p. 28–37, 2001. Citado na página 49.

BOBROW, D. G.; WINOGRAD, T. An overview of krl, a knowledge representation language. *Cognitive science*, Wiley Online Library, v. 1, n. 1, p. 3–46, 1977. Citado na página 54.

BOCK, J. et al. Benchmarking owl reasoners. In: TENERIFE. *ARea2008-Workshop on Advancing Reasoning on the Web: Scalability and Commonsense*. [S.l.], 2008. Citado na página 61.

BOOLE, G. *An investigation of the laws of thought: on which are founded the mathematical theories of logic and probabilities*. [S.l.]: Dover Publications, 1854. Citado na página 28.

BRACHMAN, R. J.; LEVESQUE, H. J. The tractability of subsumption in frame-based description languages. In: *AAAI*. [S.l.: s.n.], 1984. v. 84, p. 34–37. Citado na página 35.

BRACHMAN, R. J.; SCHMOLZE, J. G. An overview of the kl-one knowledge representation system. *Cognitive science*, Elsevier, v. 9, n. 2, p. 171–216, 1985. Citado na página 59.

BRENTANO, F. *Psychology from an empirical standpoint*. [S.l.]: Routledge, 2014. Citado na página 23.

BRESCIANI, P.; FRANCONI, E.; TESSARIS, S. Implementing and testing expressive description logics: Preliminary report. In: *Proc. of the 1995 Description Logic Workshop (DL'95)*. [S.l.: s.n.], 1995. p. 131–139. Citado na página 63.

BUCHHEIT, M.; DONINI, F. M.; SCHAERF, A. Decidable reasoning in terminological knowledge representation systems. *JAIR*, v. 1, p. 109–138, 1993. Citado na página 44.

CARDOSO, J. The semantic web vision: Where are we? *Intelligent Systems, IEEE, IEEE*, v. 22, n. 5, p. 84–88, 2007. Citado na página 65.

CHENG, P. W.; HOLYOAK, K. J. Pragmatic reasoning schemas. *Cognitive psychology*, Elsevier, v. 17, n. 4, p. 391–416, 1985. Citado na página 39.

CONSORTIUM, W. W. W. et al. Owl 2 web ontology language document overview. Word Wide Web Consortium, 2012. Disponível em: <<https://www.w3.org/TR/owl2-overview/>>. Citado 3 vezes nas páginas 52, 53 e 55.

D'AGOSTINO, M. et al. *Handbook of tableau methods*. [S.l.]: Springer Science & Business Media, 2013. Citado na página 45.

DAS, A.; WU, W.; MCGUINNESS, D. L. Industrial strength ontology management. In: *The Emerging Semantic Web*. [S.l.: s.n.], 2001. Citado na página 65.

DEAN, M. et al. Owl web ontology language reference. *W3C Recommendation February*, v. 10, 2004. Citado na página 51.

DEVLIN, K. *Sets, functions, and logic: an introduction to abstract mathematics*. [S.l.]: CRC Press, 2003. Citado na página 83.

DIAS, E. A. V.; COSTA, H. G. Mapeamento da produção científica no escopo da ontologia. *Sistemas & Gestão*, v. 6, n. 4, p. 481–507, 2012. Citado na página 47.

DIPERT, R. R. et al. *An Introduction to Logic*. [S.l.]: JSTOR, 1994. Citado na página 27.

DONINI, F. M. et al. Queries, rules and definitions as epistemic sentences in concept languages. In: *Foundations of Knowledge Representation and Reasoning*. [S.l.]: Springer, 1994. p. 113–132. Citado 2 vezes nas páginas 43 e 44.

FREGE, G. *Collected Papers on Mathematics, Logic, and Philosophy*. [S.l.]: Wiley-Blackwell, 1991. Citado na página 27.

FRIEDMAN-HILL, E. *JESS in Action*. [S.l.]: Manning Greenwich, CT, 2003. Citado na página 69.

- GASEVIĆ, D.; DJURIĆ, D.; DEVEDZIĆ, V. *Model driven engineering and ontology development*. [S.l.]: Springer Science & Business Media, 2009. Citado na página 67.
- GENNARI, J. H. et al. The evolution of protégé: an environment for knowledge-based systems development. *International Journal of Human-computer studies*, Elsevier, v. 58, n. 1, p. 89–123, 2003. Citado na página 67.
- GENSLER, H. J. *Introduction to logic*. [S.l.]: Routledge, 2010. Citado na página 29.
- GIACOMO, G. D.; LENZERINI, M. Tbox and abox reasoning in expressive description logics. *KR*, v. 96, n. 316-327, p. 10, 1996. Citado na página 41.
- GLIMM, B. et al. Hermit: reasoning with large ontologies. *Computing Laboratory, Oxford University*, 2009. Citado na página 64.
- GRUBER, T. Ontology. encyclopedia of database systems, ling liu and m. tamer özsü. *Springer-Verlag. Cloud Computing Economics: Democratization and Monetization of Services. Journal of Business & Economics Research*, v. 7, n. 6, p. 1–11, 2009. Citado 2 vezes nas páginas 41 e 48.
- GRUBER, T. R. A translational approach to portable ontologies. *Knowledge Acquisition*, v. 5, n. 2, p. 199–220, 1993. Citado na página 48.
- GRUBER, T. R. Toward principles for the design of ontologies used for knowledge sharing? *International journal of human-computer studies*, Elsevier, v. 43, n. 5, p. 907–928, 1995. Citado na página 48.
- GUARINO, N. Formal ontology, conceptual analysis and knowledge representation. *International journal of human-computer studies*, Elsevier, v. 43, n. 5, p. 625–640, 1995. Citado na página 48.
- GUARINO, N.; OBERLE, D.; STAAB, S. What is an ontology? In: *Handbook on ontologies*. [S.l.]: Springer, 2009. p. 1–17. Citado na página 47.
- GUO, Y.; PAN, Z.; HEFLIN, J. Lubm: A benchmark for owl knowledge base systems. *Web Semantics: Science, Services and Agents on the World Wide Web*, Elsevier, v. 3, n. 2, p. 158–182, 2005. Citado na página 64.
- HAARSLEV, V. et al. The racerpro knowledge representation and reasoning system. *Semantic Web*, IOS Press, v. 3, n. 3, p. 267–277, 2012. Citado na página 61.
- HAARSLEV, V.; MÖLLER, R. Expressive abox reasoning with number restrictions, role hierarchies, and transitively closed roles. 2000. Citado na página 44.
- HAARSLEV, V.; MÖLLER, R. Racer system description. In: SPRINGER. *International Joint Conference on Automated Reasoning*. [S.l.], 2001. p. 701–705. Citado na página 59.
- HAARSLEV, V.; MÖLLER, R.; WESSEL, M. Querying the semantic web with racer+nrql. In: *Proceedings of the KI-2004 International Workshop on Applications of Description Logics (ADL'04), Ulm, Germany*. [S.l.: s.n.], 2004. v. 24. Citado na página 61.
- HAASE, P. et al. The neon ontology engineering toolkit. *WWW*, 2008. Citado na página 68.

- HARMELEN, F. V.; LIFSCHITZ, V.; PORTER, B. *Handbook of knowledge representation*. [S.l.]: Elsevier, 2008. Citado na página 36.
- HARMELEN, F. V.; MCGUINNESS, D. L. Owl web ontology language overview. *World Wide Web Consortium (W3C) Recommendation*, 2004. Citado na página 50.
- HARTIGAN, J. A. *Bayes theory*. [S.l.]: Springer Science & Business Media, 2012. Citado na página 40.
- HAYES-ROTH, F.; WATERMAN, D. A.; LENAT, D. B. *Building Expert Systems*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1983. ISBN 0-201-10686-8. Citado na página 24.
- HESSEN, J. *Teoria do conhecimento*. [S.l.]: Martins fontes, 1999. Citado na página 23.
- HOLLUNDER, B. Hybrid inferences in kl-one-based knowledge representation systems. In: SPRINGER. *GWAI-90 14th German Workshop on Artificial Intelligence*. [S.l.], 1990. p. 38–47. Citado na página 44.
- HOLLUNDER, B.; BAADER, F. Qualifying number restrictions in concept languages. *KR*, v. 91, p. 335–346, 1991. Citado na página 44.
- HOLSAPPLE, C. W.; JOSHI, K. D. A formal knowledge management ontology: Conduct, activities, resources, and influences. *Journal of the American Society for Information Science and Technology*, Wiley Online Library, v. 55, n. 7, p. 593–612, 2004. Citado na página 47.
- HORRIDGE, M. *OWL Syntaxes*. Ontogenesis, 2010. Disponível em: <<http://ontogenesis-knowledgeblog.org/88>>. Citado 2 vezes nas páginas 55 e 56.
- HORRIDGE, M. A practical guide to building owl ontologies using protégé 4 and co-ode tools edition 1.3. University of Manchester, 2011. Citado na página 50.
- HORROCKS, I. Using an expressive description logic: Fact or fiction? *KR*, v. 98, p. 636–645, 1998. Citado na página 64.
- HORROCKS, I.; MOTIK, B.; WANG, Z. The hermit owl reasoner. *ORE*, v. 858, 2012. Citado na página 64.
- HORROCKS, I. et al. Owl rules: A proposal and prototype implementation. *Web Semantics: Science, Services and Agents on the World Wide Web*, Elsevier, v. 3, n. 1, p. 23–40, 2005. Citado na página 61.
- INFANTE, U.; NETO, P. C. *Gramática da língua portuguesa*. [S.l.]: Scipione, 2008. 584 p. ISBN 8526270761. Citado 2 vezes nas páginas 31 e 80.
- JACKSON, P. *Introduction to expert systems*. [S.l.]: Addison-Wesley, 1999. (International computer science series). ISBN 9780201876864. Citado na página 24.
- JOSEPHSON, J. R.; JOSEPHSON, S. G. *Abductive inference: Computation, philosophy, technology*. [S.l.]: Cambridge University Press, 1996. Citado na página 40.
- KANT, I.; YOUNG, J. M. *Lectures on logic*. [S.l.]: Cambridge University Press, 2004. Citado na página 27.

- KHAMPARIA, A.; PANDEY, B. Comprehensive analysis of semantic web reasoners and tools: a survey. *Education and Information Technologies*, Springer, p. 1–25, 2017. Citado 4 vezes nas páginas 59, 60, 65 e 66.
- KRIPKE, S. A. Semantical analysis of modal logic i normal modal propositional calculi. *Mathematical Logic Quarterly*, Wiley Online Library, v. 9, n. 5-6, p. 67–96, 1963. Citado na página 23.
- KRIPKE, S. A. Naming and necessity. In: *Semantics of natural language*. [S.l.]: Springer, 1972. p. 253–355. Citado na página 28.
- KRÖTZSCH, M. Owl 2 profiles: An introduction to lightweight ontology languages. In: SPRINGER. *Reasoning Web International Summer School*. [S.l.], 2012. p. 112–183. Citado 3 vezes nas páginas 52, 55 e 59.
- LAKATOS, E. M.; MARCONI, M. d. A. Fundamentos da metodologia científica. In: *Fundamentos da metodologia científica*. [S.l.]: Atlas, 2010. Citado 2 vezes nas páginas 39 e 72.
- LARA, M. L. Diferenças conceituais sobre termos e definições e implicações na organização da linguagem documentária. *Ciência da Informação*, SciELO Brasil, v. 33, n. 2, p. 91–96, 2004. Citado na página 41.
- LEVY, A. Y.; ROUSSET, M.-C. Carin: A representation language combining horn rules and description logics?. In: CITESEER. *ECAI*. [S.l.], 1996. p. 323–327. Citado na página 69.
- LOCKE, J. *An essay concerning human understanding*. [S.l.: s.n.], 1841. Citado na página 23.
- LOWE, E. J. The four-category ontology: A metaphysical foundation for natural science. 2006. Citado na página 47.
- MCBRIDE, B. Jena: A semantic web toolkit. *IEEE Internet computing*, IEEE, v. 6, n. 6, p. 55–59, 2002. Citado na página 70.
- MCCARTHY, J. Circumscription—a form of nonmonotonic reasoning. *ℳ & ℒ*, 1987. Citado na página 48.
- MIDDLETON, S. E.; SHADBOLT, N. R.; ROURE, D. C. D. Ontological user profiling in recommender systems. *ACM Transactions on Information Systems (TOIS)*, ACM, v. 22, n. 1, p. 54–88, 2004. Citado na página 24.
- MISHRA, R. B.; KUMAR, S. Semantic web reasoners and languages. *Artificial Intelligence Review*, Springer, v. 35, n. 4, p. 339–368, 2011. Citado 3 vezes nas páginas 59, 61 e 62.
- MORTARI, C. A. *Introdução à lógica*. [S.l.]: SciELO-Editora UNESP, 2001. Citado na página 27.
- MOTIK, B. Kaon2-scalable reasoning over ontologies with large data sets. *ERCIM News*, v. 2008, n. 72, 2008. Citado na página 62.
- MOTIK, B. et al. Owl 2 web ontology language profiles (2012). *W3C Recommendation*, 2016. Citado na página 53.

- MOTIK, B.; SATTTLER, U. Practical dl reasoning over large aboxes with kaon2. *Submitted for publication*, 2006. Citado na página 62.
- MOTIK, B.; STUDER, R. Kaon2—a scalable reasoning tool for the semantic web. In: *Proceedings of the 2nd European Semantic Web Conference (ESWC'05), Heraklion, Greece*. [S.l.: s.n.], 2005. v. 17. Citado na página 62.
- MUELLER, I. et al. The completeness of stoic propositional logic. *Notre Dame Journal of Formal Logic*, University of Notre Dame, v. 20, n. 1, p. 201–215, 1979. Citado na página 28.
- NEBEL, B. Terminological reasoning is inherently intractable. *Artificial Intelligence*, Elsevier, v. 43, n. 2, p. 235–249, 1990. Citado na página 42.
- OLIVEIRA, M. de. *Ciência da informação e biblioteconomia: novos conteúdos e espaços de atuação*. [S.l.]: Editora UFMG, 2005. Citado na página 47.
- OWEN, O. F. et al. *The Organon, Or Logical Treatises of Aristotle: With the Introduction of Porphyry Literally Translated, with Notes, Syllogistic Examples, Analysis, and Introduction*. [S.l.]: G. Bell and Sons, 1901. Citado na página 27.
- PARSIA, B.; SIRIN, E. Pellet: An owl dl reasoner. In: *Third International Semantic Web Conference-Poster*. [S.l.: s.n.], 2004. v. 18, p. 13. Citado 2 vezes nas páginas 61 e 62.
- PATEL-SCHNEIDER, P. F. et al. Owl web ontology language semantics and abstract syntax. *W3C recommendation*, v. 10, 2004. Citado na página 50.
- PEIRCE, C. S. *Collected papers of charles sanders peirce*. [S.l.]: Harvard University Press, 1974. Citado 2 vezes nas páginas 39 e 40.
- PEREIRA, R. M. A concepção da verdade-como-correspondência. *Anais do VII Seminário de Pós-Graduação em Filosofia da UFSCar*, p. 374–184, 2011. Citado na página 23.
- PIERCE, B. C. *Types and programming languages*. [S.l.]: MIT press, 2002. Citado na página 54.
- PITTS, A. M. Nominal logic, a first order theory of names and binding. *Information and computation*, Elsevier, v. 186, n. 2, p. 165–193, 2003. Citado na página 37.
- RUSSELL, B. *Logic and knowledge*. [S.l.]: Spokesman Books, 2007. Citado na página 27.
- RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence: A Modern Approach (International Edition)*. [S.l.]: Pearson US Imports & PHIPes, 2002. Paperback. ISBN 0130803022. Citado 2 vezes nas páginas 28 e 35.
- SATTTLER, U. *Description Logic Reasoners*. 2007. Citado na página 65.
- SCHMIDT-SCHAUSS, M.; SMOLKA, G. Attributive concept descriptions with complements. *Artificial intelligence*, Elsevier, v. 48, n. 1, p. 1–26, 1991. Citado na página 44.
- SHEARER, R.; MOTIK, B.; HORROCKS, I. Hermit: A highly-efficient owl reasoner. In: *OWLED*. [S.l.: s.n.], 2008. v. 432, p. 91. Citado 2 vezes nas páginas 59 e 64.

- SIRIN, E. et al. Pellet: A practical owl-dl reasoner. *Web Semantics: science, services and agents on the World Wide Web*, Elsevier, v. 5, n. 2, p. 51–53, 2007. Citado na página 61.
- SMITH, B. *Ontology*. 2003. Citado 2 vezes nas páginas 24 e 47.
- SMITH, B. et al. *Basic Formal Ontology 2.0 Draft Specification and User's Guide*. 2014. Citado 2 vezes nas páginas 78 e 79.
- SOWA, J. F. et al. *Knowledge representation: logical, philosophical, and computational foundations*. [S.l.]: MIT Press, 2000. Citado na página 23.
- STALLMAN, R. et al. *The GNU project*. 1998. Citado na página 121.
- STALNAKER, R. C. Possible worlds. *Noûs*, JSTOR, p. 65–75, 1976. Citado na página 28.
- STOJANOVIC, L.; MOTIK, B. Ontology evolution within ontology editors. In: *Proceedings of the OntoWeb-SIG3 Workshop*. [S.l.: s.n.], 2002. p. 53–62. Citado na página 65.
- SUPPES, P. *Introduction to logic*. [S.l.]: Courier Corporation, 2012. Citado 2 vezes nas páginas 29 e 33.
- SURE, Y. On-to-knowledge ontology based knowledge management tools and their application. *KI*, v. 16, n. 1, p. 35–37, 2002. Citado na página 65.
- TARSKI, A. The concept of truth in formalized languages. *Logic, semantics, metamathematics*, Oxford, v. 2, p. 152–278, 1956. Citado na página 23.
- TSARKOV, D.; HORROCKS, I. Fact++ description logic reasoner: System description. In: SPRINGER. *International Joint Conference on Automated Reasoning*. [S.l.], 2006. p. 292–297. Citado na página 64.
- VICKERY, B. C. Ontologies. *Journal of information science*, Sage Publications, v. 23, n. 4, p. 277–286, 1997. Citado na página 47.
- WELTY, C.; MCGUINNESS, D. L.; SMITH, M. K. Owl web ontology language guide. *W3C recommendation, W3C*, 2004. Citado na página 50.
- WHITEHEAD, A. N. Memoir on the algebra of symbolic logic. *American Journal of Mathematics*, JSTOR, v. 23, n. 2, p. 139–165, 1901. Citado na página 27.
- WHITEHEAD, A. N.; RUSSELL, B. *Principia mathematica*. [S.l.]: University Press, 1912. Citado na página 27.