



PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

"Uma Ontologia para as Doenças Tropicais  
Negligenciáveis - NTDO"

POR

**FILIPPE SANTANA DA SILVA**

*DISSERTAÇÃO DE MESTRADO*



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
POSGRADUACAO@CIN.UFPE.BR  
WWW.CIN.UFPE.BR/~POSGRADUACAO

RECIFE, MAR/2012



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE INFORMÁTICA  
PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

FILIPPE SANTANA DA SILVA

“UMA ONTOLOGIA PARA AS DOENÇAS TROPICAIS  
NEGLIGENCIÁVEIS - NTDO ”

*ESTE TRABALHO FOI APRESENTADO À PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO DO CENTRO DE INFORMÁTICA DA UNIVERSIDADE FEDERAL DE PERNAMBUCO COMO REQUISITO PARCIAL PARA OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIA DA COMPUTAÇÃO.*

ORIENTADOR: FREDERICO L. GONÇALVES DE FREITAS  
CO-ORIENTADOR: STEFAN SCHULZ

RECIFE, MARÇO/2012

**Catálogo na fonte**  
**Bibliotecária Jane Souto Maior, CRB4-571**

**Silva, Filipe Santana da**  
**Uma ontologia para as doenças tropicais negligenciáveis / Filipe Santana da Silva. - Recife: O Autor, 2012.**

**xi, 131 p. : il., fig., tab., quadro**

**Orientador: Frederico Luís Gonçalves de Freitas.**  
**Dissertação (mestrado) - Universidade Federal de Pernambuco. CIn, Ciência da Computação, 2012.**  
**Inclui bibliografia.**

**1. Inteligência artificial. 2. Representação de conhecimento. 3. Ontologia. I. Freitas, Frederico Luís Gonçalves de (orientador). II. Título.**

**006.3**

**CDD (23. ed.)**

**MEI2012 – 054**

Dissertação de Mestrado apresentada por **Filipe Santana da Silva** à Pós-Graduação em Ciências da Computação do Centro de Informática da Universidade Federal de Pernambuco, sob o título "**Uma Ontologia para as Doenças Tropicais Negligenciáveis**", orientada pelo **Prof. Frederico Luís Gonçalves de Freitas** e aprovada pela Banca Examinadora formada pelos professores:

---

Profa. Anjolina Grisi de Oliveira  
Centro de Informática / UFPE

---

Profa. Renata Wassermann  
Instituto de Informática e Estatística / USP

---

Prof. Frederico Luís Gonçalves de Freitas  
Centro de Informática / UFPE

Visto e permitida a impressão.  
Recife, 5 de maro de 2012.

---

**Prof. Nelson Souto Rosa**  
Coordenador da Pós-Graduação em Ciência da Computação do  
Centro de Informática da Universidade Federal de Pernambuco.

*“O significado das coisas não está nas coisas em si,  
mas em nossa atitude em relação à elas.”*

Antoine de Saint-Exupéry

## AGRADECIMENTOS

Com o desenrolar de dois anos de trabalho, gostaria de agradecer, inicialmente a professora Zulma Medeiros, por ter me concedido a oportunidade de conhecer o meu orientador Fred Freitas, ao qual sou igualmente grato, pela confiança, ensinamentos, empenho e disponibilidade.

Nesse mesmo meio, não posso esquecer as contribuições e a amizade do Prof. Stefan Schulz, da Roberta Fernandes e do Daniel Schober. Trabalhando com todas essas pessoas, e por causa deles, tive a oportunidade de expandir meus conhecimentos e conhecer lugares e outras pessoas fantásticas.

Gostaria, também, de agradecer ao meu pai, Silvio Santana, e a minha irmã, Mariana Santana, pelo suporte, troca de idéias, conselhos e palavras de conforto. E também, a Eliane Leal e Família, por ter sempre auxiliado nos momentos de dificuldade com muito zelo e paciência. Antes de tudo, tenho que pedir desculpa aos meus demais familiares pela ausência freqüente.

Ainda, devo também agradecimentos as Profas. Maria Tereza Cartaxo e Maria do Socorro Cavalcanti, do Instituto de Ciências Biológicas da Universidade de Pernambuco, por terem me introduzido na vida acadêmica e terem contribuído substancialmente para a minha formação, tanto como pesquisador quanto como pessoa.

Por fim, gostaria de lembrar aqueles que já se foram, que marcam, e irão marcar a minha vida eternamente, em caráter e pensamentos principalmente. Sou grato a elas: minha falecida mãe Lucienne Santana; e a minha falecida tia-avó Edite Nunes, por não saírem da minha mente em momento algum e por ter na lembrança delas o meu alicerce, minha admiração e grande parte da minha motivação.

Sem mais, sou eternamente grato a todos.

## RESUMO

Muitas aplicações não conseguem tratar a ambiguidade presente em fontes de dados e informação. Tal fato ganhou maior notoriedade a partir do desenvolvimento de tecnologias relacionadas à web semântica, principalmente com as ontologias. O estudo de modelos com certo grau de complexidade representacional relacionados às doenças infecciosas, especificamente as Doenças Tropicais Negligenciáveis (DTNs), vem gradualmente ganhando interesse por parte dos pesquisadores. O presente estudo visa representar um conjunto de conhecimento complexo sobre a transmissão de Doenças Tropicais Negligenciáveis e os possíveis processos que ocorrem a partir do desenvolvimento destas, como o falecimento de indivíduos, em uma ontologia: a *NTDO (Neglected Tropical Disease Ontology)*. A partir do modelo básico de transmissão de doenças, incluindo vetores artrópodes, e do conteúdo tabular com a representação de vetores, patógenos, hospedeiros, locais de ocorrência e doenças causadas, foi possível descrever um Padrão de Projeto Ontológico (PPO) para a representação de tais processos, refinados e testados segundo consultas em Lógica de Descrições. Outros resultados foram encontrados a partir da representação de processos complexos relacionados ao falecimento de indivíduos por causas específicas. No presente estudo, conhecimento acerca das DTNs foi descrito a partir de informações legadas presentes em tabelas, e puderam ser expressas em uma ontologia formal. A NTDO evidencia eventos complexos com marcações temporais e sequência de processos, desde a transmissão de um patógeno ao falecimento de um indivíduo por uma doença. Assim, a NTDO pode permitir a construção de consultas inteligentes em bancos de dados de Morbidade e Mortalidade. Ainda, pode permitir uma inovação no que concerne a vigilância de casos de doenças relacionados a infecções por doenças, principalmente negligenciadas, por possibilitar o estudo de um amplo conjunto de variáveis, inerentes aos registros de morbidade e mortalidade, e a conseqüente construção de novo conhecimento sobre os dados de saúde.

**Palavras-chave:** Ontologia, Doenças Tropicais Negligenciáveis, PPO, Representação Tabular, Mortalidade.

## ABSTRACT

Many applications are not able to handle the ambiguity present in data and information sources. These facts gained notoriety due to the development of semantic web related technologies, mainly due to ontologies, that appear to be a solution for those problems. The study of complex representational models related to infectious diseases, specifically Neglected Tropical Diseases (NTDs), interests many researchers. The current study aims at transforming a complex knowledge body about NTD transmission and the possible processes which occurs due to their development, linking it to mortality information, as an ontology: The Neglected Tropical Disease Ontology (NTDO). From the basic disease transmission model, including arthropod vectors, pathogens, hosts, geographic locations and caused diseases, it was possible to describe an Ontology Design Pattern (ODP) to represent such processes, refined and tested by means of Description Logics (DL) queries. Other result was about the representation of complex processes related to passing from life to death, by specific causes. In the current study, the knowledge about NTD was described as legacy data available in tables, and was expressed as a formal ontology. The ontology highlights complex events with temporal marks and process sequences, from a pathogen transmission to the death of a given person due to a disease. As for usefulness, NTDO can support be used for querying Morbidity and Mortality databases and enabling the study of multiple variables, inherent to morbidity and mortality registries, and creating new knowledge using health related data.

**Keywords:** Ontology, Neglected Tropical Diseases, Ontology Design Pattern, Tabular Representation, Mortality.



## SUMÁRIO

1. Introdução	12
1.1. Objetivos	16
1.1.1. Objetivo Geral	16
1.1.2. Objetivos Específicos	16
1.2. Justificativa	18
2. Estado da Arte	21
2.1. Ontologias	21
2.2. Ontologias Biológicas	27
2.3. Engenharia de Ontologias	31
2.4. Lógica de Descrições - DL	35
2.5. Web Ontology Language - OWL	42
2.6. Padrões de Projetos Ontológicos– PPO	45
3. Doenças Tropicais Negligenciáveis – DTN	48
3.1. O que são as doenças tropicais negligenciáveis?	48
3.2. Quais são e o no que resultam?	49
3.3. Regiões de Ocorrência	51
3.4. Transmissão de DTN por Vetores Artrópodes	53
3.4.1. Processo de Transmissão	53
4. Aspectos Metodológicos	59
4.1. Modelagem Ontológica	59
4.1.1. Propósito	59
4.1.2. Escopo	60
4.1.3. Fontes de Conhecimento	61
4.1.4. Transferência de Conhecimento Tabular para a NTDO	62
4.1.5. Integração com outras Ontologias	64
4.1.6. Desenvolvimento da NTDO	65
4.1.7. Requisitos, Normalização e Documentação	66
4.1.8. Manutenção	66
5. Uma Ontologia para as Doenças Tropicais Negligenciáveis	68
5.1. Representações Tabulares e DTN	69
5.1.1. Formalização do PPO para representação da Transmissão	71
5.1.2. Entidades Geográficas	72
5.1.3. Transmissão do Patógeno	73
5.1.4. Disposições e Manifestações	75
5.1.5. Dependência das Doenças nos Patógeno	76
5.1.6. Axiomas de Inclusão Geral	77
5.1.7. Axiomas de Disjunção	79
5.1.8. Geração de Ontologia baseada em script	79
5.1.9. Pós-processamento da Ontologia	79
5.1.10. Avaliação da Ontologia	80
5.1.11. Questões relacionadas ao desempenho	87
5.2. Modelo de Injúrias e Mortalidade por Causas Específicas	89

5.2.1. Representando Injurias e Mortalidade	89
5.2.2. Desafios Representacionais do Modelo de Mortalidade	92
5.3. Estrutura de Importação e Hierarquia Completa	104
6. Discussão	108
6.1. Considerações sobre o PPO	108
6.2. Considerações sobre o Modelo de Mortalidade	110
6.3. Trabalhos Relacionados	110
6.4. Casos de Uso	112
7. Conclusão	113
8. Referências Bibliográficas	117

## LISTA DE ABREVIATURAS

A.C.	Antes de Cristo
AIDS	Síndrome da Imunodeficiência Adquirida ( <i>Acquired Immunodeficiency Syndrome</i> )
BFO	Ontologia Básica Formal ( <i>Basic Formal Ontology</i> )
BFO-RO	Ontologia Básica Formal - Ontologia de Relações ( <i>Basic Formal Ontology - Relation Ontology</i> )
BioTop	<i>Top-domain ontology for the life Sciences</i>
CQ	Questões de Competência ( <i>Competency Questions</i> )
DAML	Linguagem de Marcação para Agentes DARPA ( <i>Darpa Agent Markup Language</i> )
DL	Lógica de Descrições ( <i>Description Logics</i> )
DO	Ontologia de Doenças ( <i>Disease Ontology</i> )
DOLCE	Ontologia Descritiva para a Linguística e Engenharia Cognitiva ( <i>Descriptive Ontology for Linguistic and Cognitive Engineering</i> )
DTN	Doença Tropical Negligenciável
GCI	Classe de Inclusão Geral ( <i>General Class Inclusion</i> )
GFO	Ontologia Geral Formal ( <i>General Formal Ontology</i> )
GO	Ontologia dos Genes ( <i>Gene Ontology</i> )
HIV	Vírus da Imunodeficiência Humana ( <i>Human Immunodeficiency Virus</i> )
IDH	Índice de Desenvolvimento Humano
IDO	Ontologia para as Doenças Infecciosas ( <i>Infectious Disease Ontology</i> )
IDOMAL	Ontologia para as Doenças Infecciosas - Malária ( <i>Infectious Disease Ontology - Malaria</i> )
KB	Base de Conhecimento ( <i>Knowledge Base</i> )
LC	Leishmaniose Cutânea
LCD	Leishmaniose Cutânea Difusa

LMC	Leishmaniose Mucocutânea
LV	Leishmaniose Visceral
MSF	Médicos Sem Fronteiras ( <i>Medicins Sans Frontieres</i> )
NTDO	<i>Ontologias para as Doenças Tropicais Negligenciáveis (Neglected Tropical Disease Ontology )</i>
OBI	Ontologia para Investigações Biomédicas ( <i>Ontology for Biomedical Investigations</i> )
OBO	Ontologias Biológicas Livres ( <i>Open Biomedical Ontologies</i> )
ODL	<i>Ontology Design Pattern</i>
OIL	<i>Linguagem de Troca de Ontologia (Ontology Interchange Language)</i>
OMS	Organização Mundial de Saúde
ONU	Organização das Nações Unidas
OWL	Linguagem de Ontologias para a Web( <i>Ontology Web Language</i> )
OWL2	Linguagem de Ontologias para a Web v.2 ( <i>Ontology Web Language v.2</i> )
OWL2 EL	Linguagem de Ontologias para a Web v.2 - Linguagem EL ( <i>Ontology Web Language v.2 - EL Language</i> )
OWL2 QL	Linguagem de Ontologias para a Web v.2 - Linguagem de Consulta ( <i>Ontology Web Language v.2 - Query Language</i> )
OWL2 RL	Linguagem de Ontologias para a Web v.2 - Linguagem de Regras ( <i>Ontology Web Language v.2 - Rule Language</i> )
PPO	Padrão de Projeto Ontológico
RDF	Estrutura de Descrição de Recursos ( <i>Resource Description Framework</i> )
RDF Schema	Esquema de Estrutura para Descrição de Recursos ( <i>Resource Description Framework Schema</i> )
RI	Inclusão de Papel (Role Inclusion)
SGBD	sistemas de gerenciamento de bancos de dados
SIM	Sistema de Informação de Mortalidade

SINAN	Sistema de Informação de Agravos de Notificação
SNOMED RT	Nomenclatura Sistematizada da Medicina - Terminologia de Referência (Systematized Nomenclature of Medicine - Reference Terminology)
SNOMED-CT	Nomenclatura Sistematizada da Medicina - Termos Clínicos (Systematized Nomenclature of Medicine-Clinical Terms)
SPARQL	Linguagem SPARQL para consulta em RDF ( <i>SPARQL Query Language for RDF</i> )
SVN	Subversão ( <i>Subversion</i> )
UMLS	Sistema de Linguagem Médica Unificada ( <i>Unified Medical Language System</i> )
W3C	Consórcio <i>World Wide Web</i> ( <i>World Wide Web Consortium</i> )
WS	Web Semântica
XML	Linguagem de Marcação Extensível ( <i>eXtensible Markup Language</i> )

## LISTA DE FIGURAS

FIGURA 1: Hierarquia de Bibliotecas de Ontologias segundo (GUARINO, 1998)	25
FIGURA 2: Hierarquia de Bibliotecas de Ontologias, segundo Freitas (FREITAS, 2003).	25
FIGURA 3: Atividades e Estados para a Engenharia de Ontologias (FERNANDEZ <i>et al.</i> , 1997).	32
FIGURA 4: Relacionamento entre a <i>Tbox</i> e a <i>Abox</i> . Adaptado de (BAADER <i>et al.</i> , 2003).	36
FIGURA 5: Tríade Epidemiológica com a Inclusão do Vetor.	54
FIGURA 6: Processo de Transferência de Conteúdo Tabular para uma Ontologia.	64
FIGURA 7: Estrutura da NTDO.	65
FIGURA 8: Ciclo de vida de um organismo.	89
FIGURA 9: O intervalo de tempo descrito para um <i>gfo:Chronoid</i> e seus limites ( <i>gfo:LeftTimeBoundary</i> e <i>gfo:RightTimeBoundary</i> ) no eixo temporal.	95
FIGURA 10: Modelo Gráfico de um evento de um <i>DeathProcess</i> , <i>InjuryProcess</i> e <i>BiologicalDeathProcess</i> e suas relações de participação ( <b>hasDeathPatient</b> , <b>hasInjuredPatient</b> e <b>hasConvalescentPatient</b> ), de acordo com os participantes dos processos (e.g. <i>Human</i> ).	102
FIGURA 11: Conexões entre a NTDO, BioTop e GFO.	103
FIGURA 12: Modelo de Exportação da NTDO.	104
FIGURA 13: Processos da NTDO	105
FIGURA 14: Objetos Materiais e Imateriais na NTDO	106
FIGURA 15: Disposições e Tempo	107

## LISTA DE TABELAS

TABELA 1: Espécies patogênicas (aos humanos) de <i>Leishmania</i> , seus vetores no Novo Mundo, manifestações clínicas e suas distribuições geográficas, baseado em Sharma & Singh (SHARMA; SINGH, 2008).	56
TABELA 2: Tabela de DTN transmitidas por vetores, por localização geográfica, patógeno e manifestação clínica.	62
TABELA 3: Padrão geral de uma matriz para DTN transmitidas por vetores.	71
TABELA 4: Ontologia simples para realização de testes de raciocínio.	81
TABELA 5: Mapeamentos entre as entidades presentes na BioTop e GFO.	96

# 1. INTRODUÇÃO

**C**om a explosão de informações criada pelo desenvolvimento dos variados campos de pesquisa na biologia e na medicina, tratar uma enorme quantidade de conhecimento tornou-se uma tarefa difícil e complexa.

Muitas aplicações não conseguem tratar a ambiguidade presente em fontes de dados e informação. Muitas vezes esta é representada de diversas formas em sistemas que têm por objetivo executar tarefas semelhantes. Outras vezes, sistemas complementares não conseguem trocar informações correlatas por não apresentarem rigor formal sobre a representação dos dados.

Levando em consideração o conteúdo informacional proveniente da área biomédica este pode ser encontrado em sites, artigos científicos, livros, entre outras formas de publicações, além de bases de dados. A forma tradicional de identificar tais fontes é a partir da utilização de palavras-(termos) chave. Entretanto, as variadas fontes não são guardadas levando em consideração o sentido que emana da informação. Por consequência, cria-se uma dispersão no conhecimento.



Assim, o tratamento das informações de forma que possibilite que máquinas possam “entender” e, assim, relacionar as partes das informações, ganhou um maior destaque a partir da publicação do trabalho de Berners-Lee et al. (BERNERS-LEE *et al.*, 2001) sobre a Web Semântica (WS). A WS é uma extensão da web atual em que as informações são descritas com sentido definido, possibilitando a cooperação de trabalho humano-máquina. A WS é estruturada de tal forma que depende de ontologias para possibilitar a descrição formal e a hierarquização do conhecimento necessário para o desenvolvimento de aplicações.

As ontologias são um modo de representar conhecimento, no âmbito da web semântica e possibilitam a integração de informações provenientes de diversas fontes, sem necessariamente levar em consideração as formas de aquisição e de como serão guardadas. Com o surgimento das ontologias relacionadas à computação, final da década de 70, e o crescimento do interesse, desde meados da década de 1990, é possível representar o conhecimento de um domínio ou abstração. Para a ciência computacional, as ontologias são utilizadas para descrever uma dada conceitualização, provendo aos computadores a habilidade de:

- Realizar inferências lógicas;
- Descobrir fatos dentro de um corpo de informação;
- Verificar relacionamentos corretos e entidades diferentes;
- Definir restrições sobre o uso de bases de conhecimento; entre outras.

A partir do aumento do interesse pelo estudo das ontologias, a alta complexidade evidenciada pelas interações entre entidades é um tópico que vem ganhando destaque, principalmente no âmbito biomédico. Assim, o estudo das doenças negligenciadas (na grande maioria das vezes ocorre em regiões tropicais) vem sendo discutido como tópico de interesse pelo impacto causado por tais doenças.

As Doenças Tropicais Negligenciadas (DTN) são constituídas por um grupo de doenças as quais representam as infecções mais impactantes para a população mundial. Alguns exemplos de DTN são a Doença de Chagas, Febre do Dengue,

Leishmaniose, Filariose Linfática, entre outras. Dentre as DTN, um grupo recebe destaque: as doenças transmitidas por vetores artrópodes, as quais representam o conjunto de doenças que mais resultam em morte de humanos no século XX.

Nesse sentido, o objetivo central desse trabalho será transformar um conjunto de conhecimento complexo sobre as DTN em ontologia, possibilitando o suporte ao estudo de eventos de morbidade e mortalidade na população. A partir do presente trabalho, é possível estudar a dinâmica transmissível dos patógenos e seus vetores nas regiões do país, juntamente com os casos registrados de tais doenças em bancos de dados de morbidade e auxiliar na definição de ações de saúde de forma mais imediata.

Indo além, pode possibilitar o estudo de registros de morbidade e mortalidade conjuntamente. Atualmente, tal procedimento é realizado de maneira *ad hoc* e conta com muito pouca precisão no casamento entre os dados, principalmente por causa das técnicas tradicionais de bancos de dados. Serão, tanto para profissionais que estudam epidemiologia de doenças transmissíveis (médicos, biólogos, enfermeiros, sanitaristas, entre outros) quanto os que desenvolvem sistemas e aplicações relacionados às DTN (profissionais de tecnologias de informação para a saúde), os potenciais utilizadores da ontologia a ser desenvolvida.

Portanto, o estudo de ontologias sobre as DTN poderá prover um melhor conhecimento sobre os eventos de doença as quais subsistem na população. Vários fatores como: *status* epidemiológico para definição de medidas de prevenção e controle, informações sobre as áreas de ocorrência, entre outras informações, são importantes para a descrição do processo saúde-doença em uma população. Assim uma ontologia sobre as DTN pode ser considerada como de grande importância para o conhecimento sobre a dinâmica das doenças transmitidas por vetores e, a partir disso, para o estabelecimento de metas e prioridades de saúde.

Este trabalho foi desenvolvido a partir de uma parceria entre o Centro de Informática (CIn), da Universidade Federal de Pernambuco (UFPE), o Centro de Pesquisas Aggeu Magalhães (CPqAM/Fiocruz-PE), o Departamento de Informática Médica da *Albert-Ludwigs-Universität Freiburg* (Freiburg im Breisgau, Alemanha) e o Instituto de

Informática Médica, Estatística e Documentação da *Graz General Hospital and University Clinics* (Graz, Áustria), resultando assim na publicação de 4 trabalhos (2 em conferências internacionais e 2 em periódico internacional) (FREITAS, FRED; SCHULZ; MEDEIROS, 2009; FREITAS, FREDERICO; SCHULZ; MORAES, 2009a; SANTANA; FERNANDES; *et al.*, 2011; SANTANA *et al.*, 2012).

## 1.1. OBJETIVOS

### 1.1.1. Objetivo Geral

O objetivo do presente estudo é representar conhecimento complexo sobre a transmissão de Doenças Tropicais Negligenciadas e os possíveis processos que ocorrem a partir do desenvolvimento destas doenças em ontologia. O desenvolvimento de tal conjunto de doenças pode culminar, por exemplo, com o falecimento de indivíduos. Tais fatos apresentam diversos desafios representacionais, que também serão atacados e discutidos ao longo do trabalho.

### 1.1.2. Objetivos Específicos

- Definição e representação dos conceitos-chave com base em publicações científicas (livros, periódicos, entre outros);
- Descrição de um padrão para representação de transmissão de patógenos para as DTN, reutilizando a representação de doenças/sinais/sintomas presente na BioTop (modelo de morbidade), descrito em Schulz et al. (SCHULZ *et al.*, 2011), visto que a morbidade na população por tal conjunto de doenças decorre dos processos que ocorrem após a transmissão e sendo este um ponto importante para o bloqueio do ciclo de transmissão.
- Construção de um Padrão de Projeto Ontológico (PPO) (*Ontology Design Patterns - ODP*) para o modelo de transmissão de DTN, o qual seja reutilizável para outras doenças de mesma natureza;
- Validação do Padrão de Projeto Ontológico por meio de Questões de Competência informais (linguagem natural) e formais (consultas em Lógica de Descrições);
- Proposição de um modelo para representação de eventos de morte (modelo de mortalidade), evidenciando as marcas temporais, causas e sequências de processos os quais fazem parte do ciclo de vida dos seres vivos e que apresentam grande relevância epidemiológica.
- Disponibilização de uma ontologia a qual seja útil a profissionais os quais necessitam estudar a dinâmica das DTN (médicos, biólogos, biomédicos, entre outros), e também para auxiliar desenvolvedores de sistemas e

aplicações (e.g. por profissionais de tecnologias de informação) os quais necessitem manipular algum conteúdo relacionado ao conjunto de doenças em consideração.

## 1.2. JUSTIFICATIVA

A demanda pela utilização de uma forma diferenciada para a modelagem de dados e conhecimento é um requisito levantado desde a estabilização dos sistemas de gerenciamento de bancos de dados (SGBD), principalmente quando relacionado a modelagem conceitual. Já a comunidade de engenharia de software, motivada pelas linguagens orientadas a objetos, começou a reconhecer a importância sobre *modelagem de domínio*. Outro fator motivador foi o crescimento e a complexidade dos programas, criando dificuldades para a sua manutenção e a disponibilização para novos usuários (SMITH; WELTY, 2001), entre várias outras.

Voltados para a solução de tais dificuldades, em 2001 houve a introdução do conceito de web semântica, por Tim Berners-Lee *et al.* (BERNERS-LEE *et al.*, 2001), motivados pela possibilidade de tornar o conteúdo (ou parte da web) compreensível por máquinas. Sendo assim, as ontologias foram propostas como um modo de representar conhecimento, no âmbito da web semântica, e possibilitar a integração de informações provenientes de diversas fontes, sem necessariamente levar em consideração as formas de aquisição e de como serão guardadas.

As ontologias também são consideradas um *framework* para representação formal de entidades presentes na biomedicina e várias outras áreas, evidenciando termos e relações (BODENREIDER *et al.*, 2005). Nesse sentido, os estudos sobre ontologias na área biomédica são motivados pela necessidade em gerenciar uma quantidade imensa de dados heterogêneos e semanticamente relacionados, os quais demandam esforços para analisar e estruturar informações, possibilitando ainda a extração do conhecimento que emana de tais fontes (STENZHORN *et al.*, 2007).

Levando em consideração apenas a área biomédica, as ontologias apresentam algumas vantagens quando se trata do estudo de doenças. Alguns exemplos destes benefícios a serem considerados são:

- Auxiliar na aquisição de informação, servindo como base de conhecimento de suporte que provê restrições e deduções ao processo de aquisição;

- Prover acesso inteligente a informações científicas, possibilitando a recuperação de informações extremamente específicas disponíveis em fontes heterogêneas, p.ex. para a criação de consultas em bases de dados com estrutura de dados diferentes, mas com semântica semelhante ou comum;
- Evitar redundâncias de dados e informações, pois as representações ontológicas podem inferir outros dados e informações que não precisam ser armazenados;
- Proporcionar interoperabilidade entre fontes heterogêneas de dados em saúde, por representar a informação sem estar preocupado com a estrutura, p.ex. quando utilizado para mapear dados disponíveis em dois bancos de dados diferentes;
- E, por fim, auxiliar no estudo de doenças e desordens confluindo diversas formas de conhecimento em busca de um denominador comum (HADZIC; CHANG, 2005).

Uma vez que o estudo de doenças e dos fatores que corroboram para o seu estabelecimento nas populações baseado em ontologias se revela importante, é possível que a descrição do conhecimento a cerca das DTN possa trazer reflexos e resultados relevantes. Uma razão para tal reside nesse conjunto de doenças subsistir em condições de pobreza, principalmente em ambientes socioeconomicamente empobrecidos de regiões tropicais (WORLD HEALTH ORGANIZATION, 2010).

Assim, o presente estudo se resguarda nesse cenário, levando as DTN a uma representação ontológica e disponibilizando conhecimento necessário para subsidiar o processo de tomada de decisão em saúde. Pelos benefícios listados para o estudo das ontologias relacionando doenças e pelo impacto social causado pelas DTN, é visível a importância de tal artefato.

Como artifício requerido para subsidiar a luta contra as DTN, o presente estudo traz uma descrição ontológica dos mecanismos de transmissão, utilizando como exemplo

a Leishmaniose, aliado a uma descrição de eventos que levam a morte dos indivíduos por tais doenças. A ontologia criada se chama *Neglected Tropical Disease Ontology* (NTDO, do inglês “Ontologia para as Doenças Tropicais Negligenciáveis”) e será abordada mais detalhadamente nos próximos capítulos.

Portanto, o estudo de tais processos relacionados à saúde pode tornar possível a avaliação automática de múltiplas fontes de dados em saúde, possibilitando também a descoberta de conhecimento situacional (ainda não tratado na literatura), diminuindo o impacto causado pelas DTNs em populações expostas.

Em seguida, serão descritos alguns tópicos pertinentes ao desenvolvimento da ontologia a qual é o objeto do presente trabalho. Tais tópicos contextualizarão, principalmente, as ontologias biológicas, o processo de engenharia de ontologias, entre outros.



## 2. ESTADO DA ARTE

Neste capítulo, são apresentados os principais conceitos e o estado da arte referente aos principais tópicos relacionados ao desenvolvimento do trabalho, como as ontologias, as linguagens de representação, entre outros pontos.

### 2.1. ONTOLOGIAS

A noção de ontologias foi, indiretamente, introduzida por Aristóteles, no século 4 A.C., com a publicação de *Organon* e a introdução da discussão sobre a “natureza” e a “estrutura” da realidade, sendo assim considerada a ciência do *ser pelo ser* (SPEAR, 2006). Aristóteles tentou criar a primeira forma de lógica, criando também as *categorias* como uma taxonomia para os objetos do mundo (FREITAS, 2003). Um dos pontos cruciais evidenciado pelo autor foi a utilização da lógica como algo mecânico para descrever as coisas como elas são pensadas. Entretanto, Aristóteles frisou que a lógica por si só não é suficiente para representar tudo.

O termo *ontologia* foi introduzido apenas em 1613, independentemente, por dois filósofos: Rudolf Göckel e Jacob Lorhard, em *Lexicon philosophicum* e *Ogdoas*

*Scholastica*, respectivamente. Sua popularização aconteceu por volta do ano 1930 por Christian Wolff em *Philosophia prima sive ontologia* (GUIZZARDI, 2005). Segundo a definição, 'ontologia' é utilizada para descrever uma investigação filosófica sobre a existência, englobando os conceitos de *ser* e *o que existe* (CRAIG, 1998).

Nesse sentido, as ontologias, estudadas pela filosofia, são relacionadas ao desenvolvimento de teorias, como causalidade, classificação, identidade, entre outras. Essas também são utilizadas como base de teorias científicas como a descrição de propriedades subatômicas e relações de parte-todo entre uma pessoa e a sociedade. De uma forma geral, elas são descritas por meio de conceitos comuns, como composição, tempo, espaço, processos, além de vários outros (GUIZZARDI, 2005).

Outro marco importante, no início do século 20, foi introdução da ideia de *ontologia formal*, por Edmund Husserl, partindo da comparação com a lógica formal. A distinção descrita por Husserl reside na lógica formal estar relacionada à verdade das interconexões (sentido proposicional), contando com inferência, consistência e validade. Em contrapartida, as ontologias formais lidam com as interconexões entre as coisas, levando em consideração os objetos da realidade (sobre um domínio específico) e suas interconexões (SMITH, 1998).

No final da década de 1970, *sistemas baseados em conhecimento* começaram a ser estudados pela computação juntamente com a incorporação da programação lógica (FREITAS, 2003; HAYES-ROTH *et al.*, 1978), contando com:

- Formalismo de Representação;
- Estratégia de Resolução para o formalismo escolhido (motor de inferência);
- Estratégias de controle e escalonamento de inferências, objetivando guiar e aperfeiçoar inferência.

As ontologias apareceram na computação nesse âmbito, a partir da tentativa de se raciocinar de acordo com o senso comum, i.e. ter muito conhecimento sobre o

mundo de forma a se produzir um 'ser computacional' minimamente inteligente e com um comportamento racional semelhante ao ser humano. Algumas iniciativas dessa natureza foram o trabalho de John Sowa (SOWA, 1999) e a ontologia Cyc (LENAT, DOUGLAS B.; GUHA, 1990). Caso o raciocínio baseado em senso comum fosse alcançado, seria possível realizar inferências e derivar conhecimento a partir dos níveis genéricos apresentados por tais ontologias.

Durante a década de 1990, as ontologias formais ganharam um novo fôlego com estudos realizados pela comunidade computacional. A primeira contribuição para a definição do termo ontologia (computacional), nesse âmbito, foi realizada por Gruber (GRUBER, 1993), descrevendo ontologia como sendo *uma especificação de uma conceitualização*. Entretanto, o conceito descrito por Studer (STUDER, R, 1998), unindo o que foi produzido por Gruber (GRUBER, 1993) e Borst (BORST, 1997), resultou em que *uma ontologia é uma especificação explícita e formal de uma conceitualização compartilhada*, sendo esta a mais utilizada. Em 2001, as ontologias ganharam popularidade com a descrição do que viria a ser a web semântica, por Tim Berners-Lee e colaboradores (BERNERS-LEE *et al.*, 2001).

Tomando como base a definição de Studer (STUDER, R, 1998), uma 'conceitualização' se refere a um modelo abstrato de algum fenômeno natural, evidenciando as principais entidades desse fenômeno. Já 'explícita' é relacionada às entidades e as restrições impostas a estas serem explicitamente definidas, por meio de axiomas, *i.e.* postulados primitivos que são assumidos como verdadeiros (TARSKI, 1994). Ainda, o termo 'formal' é utilizado para descrever que tais artefatos devem ser processados por máquinas (necessitam de uma sintaxe específica, como a *Web Ontology Language* - OWL) e que sejam amplamente reconhecidas e utilizadas com o sentido que foi criado. Para isso, deve ser utilizado um formalismo de representação formal, como a *Lógica de Descrições* (*Description Logics* - DL) (BAADER *et al.*, 2003). Por último, 'compartilhada' significa que a entidade representada e as relações descritas devem demonstrar uma determinada porção de conhecimento aceito por uma comunidade (STUDER, R, 1998).

Do ponto de vista estrutural, as ontologias são organizadas a partir da identificação de *conceitos* (ou classes) e *relações*, descritos como predicados unários e binários, respectivamente. As classes e as relações, também, auxiliam na descrição de axiomas, os quais são construídos para possibilitar descrição do conteúdo de uma classe, como ele as relações são construídas frente a outras classes, e como as relações se comportam entre elas. Já o arcabouço de uma ontologia é constituído de uma hierarquia de classes, organizados taxonomicamente em superclasses e subclasses (STAAB; STUDER, RUDI, 2009).

Um exemplo de classes os quais são utilizados para representar pessoas é *Humano*. A partir desta, é possível representar as ocorrências específicas, como *Filipe*, *Eliane* ou *Mariana*, também conhecidos como instâncias. Classes se relacionam com outras classes a partir de relações, como em “uma pessoa **tem irmão** outra pessoa”; “pessoa **compra** alimento”; entre outras.

Um ponto crucial referente às ontologias são os níveis de abstração empregados na representação. Assim, ontologias podem ser organizadas como uma biblioteca, seguindo o nível de detalhamento apresentado (e a utilização desejada). Guarino (GUARINO, 1998) descreveu um modelo baseado em uma ontologia de **Topo** (com as entidades mais abstratas, como espaço, tempo, processos, entre outros), independente do domínio; ontologia de **domínio**, com as entidades específicas de uma domínio de conhecimento; de **tarefa**, com uma tarefa ou atividade genérica, respeitando o alto nível; e a ontologia de **aplicação**, com entidades dependentes do domínio e de tarefa (Fig. 1).

Alternativamente, Freitas (FREITAS, FRED, 2003) descreveu uma visão para a classificação ontológica, segundo a granularidade e complexidade do objeto (FIG. 2), em que também eram incluídas ontologias *core*.

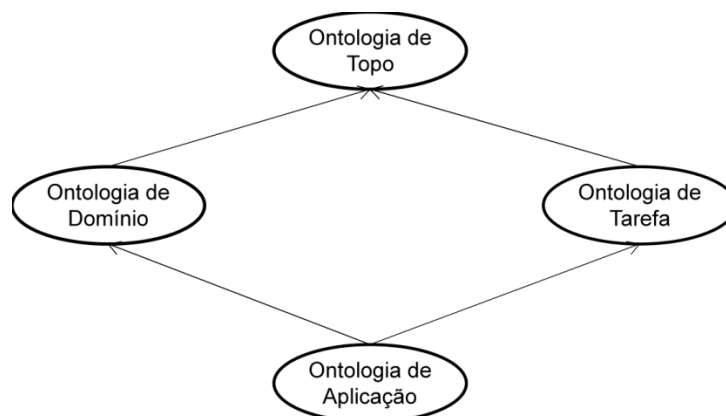


FIGURA 1: Hierarquia de Bibliotecas de Ontologias segundo (GUARINO, 1998).

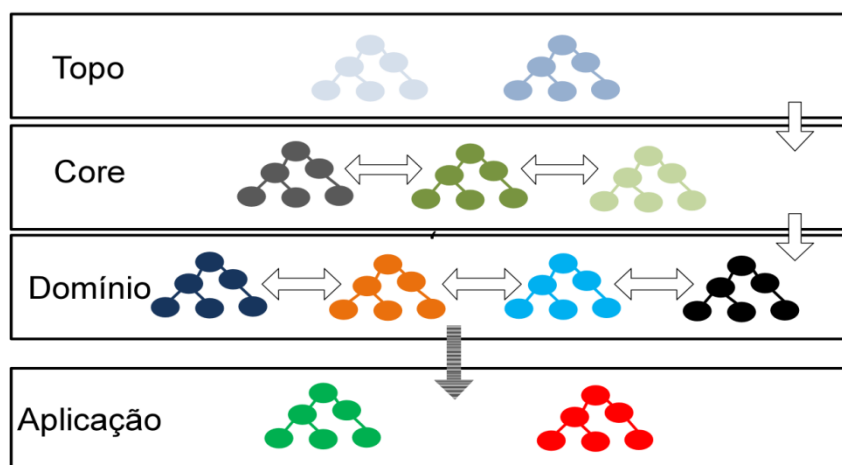


FIGURA 2: Hierarquia de Bibliotecas de Ontologias, segundo Freitas (FREITAS, FRED, 2003).

Na Figura 2, a descrição do **Topo** e o domínio são semelhantes à abordagem descrita por Guarino (GUARINO, 1998). Exemplificando, ontologias equivalentes a este nível são:

- Ontologia Básica Formal (*Basic Formal Ontology*- BFO) (GRENON *et al.*, 2004);
- Ontologia Descritiva para a Linguística e Engenharia Cognitiva (*Descriptive Ontology for Linguistic and Cognitive Engineering* - DOLCE) (GANGEMI *et al.*, 2002);

- Ou, a Ontologia Geral Formal (*General Formal Ontology* - GFO) (HELLER; HERRE, 2004; HERRE *et al.*, 2007).

Entretanto há um nível adicional comumente chamado de **Core** (VALENTE; BREUKER, 1997), ou supra-domínio (*top domain*), o qual provê as classes básicas para modelar qualquer componente de domínio que se refere a uma grande área do conhecimento e possui mapeamentos para ontologias de alto nível. Um exemplo de ontologia *core* é a BioTop (BEISSWANGER *et al.*, 2008), a qual representa o domínio biomédico e é baseada na BFO e na DOLCE. A BioTop pode ser estendida ou mapeada para qualquer ontologia de domínio da mesma natureza, como a *Gene Ontology* (GO) (THE GENE ONTOLOGY CONSORTIUM, 2000) ou a Ontologia para Doenças Infecciosas (*Infectious Disease Ontology* - IDO) (COWELL; SMITH, 2010).

Várias áreas da Ciência da Computação vêm estudando as ontologias, objetivando principalmente o compartilhamento e o reuso de conhecimento. Ontologias são usualmente descritas pela Inteligência Artificial Simbólica para a representação do conhecimento estático de um domínio em um sistema baseado em conhecimento. Alguns exemplos de áreas específicas que estudam e usam as ontologias são: a informática biomédica (SANTANA; SCHÖBER; MEDEIROS; *et al.*, 2011), computação ubíqua (EJIGU *et al.*, 2007), sistemas multi-agentes (RAFI *et al.*, 2009), *e-learning* (AHMED-OUAMER; HAMMACHE, 2010), extração de dados (SEBASTIAN *et al.*, 2009), entre outras.

Como parte do foco do presente trabalho, uma breve introdução sobre as ontologias biológicas é encontrada no tópico a seguir.

## 2.2. ONTOLOGIAS BIOLÓGICAS

A crescente disponibilidade de dados provenientes de pesquisas e recursos intelectuais na área biomédica requer esforços de profissionais biomédicos e informacionais para gerenciar uma enorme quantidade de conteúdo (FREITAS *et al.*, 2009), em sua maioria, semanticamente relacionado. Tal fato foi o gatilho para a proliferação das ontologias biomédicas.

As ontologias biomédicas representam um campo emergente e que ganhou repercussão com o desenvolvimento da *Gene Ontology* (GO) (THE GENE ONTOLOGY CONSORTIUM, 2000), a qual é utilizada para análise de informações de pesquisas em genômica funcional (BODENREIDER *et al.*, 2005).

Assim, as ontologias biológicas são utilizadas (principalmente) para descrever a estrutura e a complexidade do conhecimento do domínio biomédico (RUBIN *et al.*, 2008). Bodenreider *et al.* (BODENREIDER *et al.*, 2005) comentaram que as ontologias biológicas exercem um papel central para a integração de informação sobre diferentes modelos de organismos. O propósito delas é relacionado ao “estudo das entidades da realidade as quais apresentam significância biomédica” (BODENREIDER *et al.*, 2005).

Entretanto, as ontologias biológicas geralmente são desenvolvidas independentemente uma da outra, mesmo quando o conteúdo é semelhante, revelando apenas uma visão parcial do campo (STENZHORN *et al.*, 2007). De tal forma, e como proposto por alguns autores, foram desenvolvidas ontologias *core* com conhecimento generalizado o suficiente para representar grandes áreas específicas (SABOU *et al.*, 2006), e.g. biologia e química. Um esforço que vem sendo mantido, nesse sentido, é relacionado ao desenvolvimento da BioTop, a qual atua como uma guia para a conceitualização de entidades da biomedicina (BEISSWANGER *et al.*, 2008).

O desenvolvimento de ontologias biológicas vem crescendo e os repositórios estão acompanhando. Esforços como o *OBO Foundry* (“The Open Biological and Biomedical Ontologies Foundry,” 2011), o *BioPortal* (“NCBO Bioportal,” 2011), e os

termos clínicos provenientes da Nomenclatura Sistematizada da Medicina (*Systematized Nomenclature of Medicine-Clinical Terms* - SNOMED-CT) (DONNELLY, 2006), entre outras, contêm informações sobre anatomia, fenótipos, química, doenças e muitas outras.

Para exemplificar, o BioPortal, em 2009, tinha 134 ontologias, com 680 mil classes incluídas em seu repositório (NOY *et al.*, 2009). Atualmente, este mesmo repositório conta com 259 ontologias e por volta de 5.3 milhões de termos indexados<sup>1</sup>. Já a SNOMED-CT, derivada da junção da SNOMED RT (*Reference Terminology*) com a terceira versão dos termos clínicos do Reino Unido (UK Clinical Terms v.3) (O'NEIL *et al.*, 1995), em 1999, e implementada em Lógica de Descrições (BAADER *et al.*, 2003), contava com 310.000 termos com uma extensiva rede de propriedades (ELHANAN *et al.*, 2010).

Alguns exemplos de ontologias biológicas serão descritas abaixo:

- *Gene Ontology* (GO) (THE GENE ONTOLOGY CONSORTIUM, 2000)

A *Gene Ontology* (GO) é uma ontologia que foi criada para possibilitar a estruturação de um vocabulário controlado, precisamente definido e comum. Ainda, deve possibilitar a descrição dos papéis dos genes e dos produtos gênicos em qualquer organismo.

A partir do desenvolvimento da GO, vários estudos foram realizados, como o de Bien *et al.* (BIEN *et al.*, 2012) os quais utilizaram essa ontologia para avaliar similaridade semântica entre ela e dados de sequências de proteínas, expressão gênica interações entre proteínas, entre outras. O estudo de Burge *et al.* (BURGE *et al.*, 2012) descreveu a utilização de anotações sobre proteínas, juntamente com a GO, para inferir informações sobre sequências de proteínas não caracterizadas. Outro trabalho, desenvolvido por Chittenden *et al.* (CHITTENDEN *et al.*, 2012) evidenciou o desenvolvimento de um algoritmo para descobrir subclassificação para processos biológicos relacionados aos termos disponíveis na GO.

---

1 Acessado em 8 de junho de 2011



- *Foundational Model of Anatomy* (FMA) (ROSSE; MEJINO, 2003)

A FMA é uma ontologia que foi criada a partir do conteúdo sobre anatomia humana disponível na *Unified Medical Language System* (UMLS) e que engloba as entidades materiais, substâncias e espaços que resultam da expressão de genes estruturais.

Alguns estudos utilizaram a FMA como base. Um exemplo é o trabalho de Travillian *et al.* (TRAVILLIAN *et al.*, 2011), os quais utilizaram a FMA como base para estudar comparativamente a anatomia entre mamíferos. Turner *et al.* (TURNER *et al.*, 2010) apresentaram uma extensão à FMA para acomodar conhecimento sobre neuroanatomia e aplicações em neuro-imagem. Outro trabalho foi desenvolvido por Noy & Rubin (NOY; RUBIN, 2008), com a tradução da FMA na *Ontology Web Language* (OWL), para ser mais bem aproveitada no desenvolvimento de aplicações e estudos relacionadas à ontologias e web-semântica.

- *BioTop* (BEISSWANGER *et al.*, 2008)

A BioTop (*Top-domain Ontology for Life Sciences*) foi concebida sob princípios formais de representação e implementada em OWL DL. O objetivo desta ontologia é relacionado à disponibilização de um nível ontológico o qual possibilite a integração de várias ontologias específicas do domínio das ciências da vida.

Alguns trabalhos os quais utilizaram-se da BioTop foram os de Santana *et al.* (SANTANA *et al.*, 2011; SANTANA; SCHOBER; MEDEIROS; *et al.*, 2011), os quais utilizaram-se do arcabouço da BioTop para a representação de um modelo de mortalidade e um modelo de transmissão de doenças negligenciadas, respectivamente. Outro trabalho baseado na BioTop foi desenvolvido por Schober *et al.* (SCHOBER *et al.*, 2010) o qual criou uma ontologia, a *DebugIT Core Ontology* (DCO), para servir de base na integração de bancos de dados provenientes de diferentes hospitais.

Apesar das mais variadas iniciativas e, por mais que as ontologias biológicas sejam disponibilizadas em um formato padronizado e formalizado, grande parte delas é

criada independentemente e não seguem uma metodologia de engenharia de ontologias que suporte o processo completo de seus ciclos de vida. No próximo tópico será descrito o estado da arte relacionado à criação e manutenção de ontologias baseados em metodologias específicas.

### 2.3. ENGENHARIA DE ONTOLOGIAS

As abordagens de engenharia de ontologias presentes na literatura, comumente, descrevem o processo de criação, evolução e disponibilização de uma ontologia para ser utilizada em uma aplicação ou em um sistema. Algumas metodologias, como a *Methontology* (FERNANDEZ *et al.*, 1997) e a descrita por Uschold & Gruninger (USCHOLD; GRUNINGER, 1996), abordam o desenvolvimento de ontologias. Ainda, outra metodologia, a On-To-Knowledge (STAAB *et al.*, 2001), descreve o ciclo de vida. No presente trabalho, a *Methontology* será utilizada por ser a metodologia mais referenciada na literatura e pela flexibilidade provida ao engenheiro.

A *Methontology*, desenvolvida pelo *Ontology Engineering Group* da *Universidad Politécnica de Madrid*, possibilita o desenvolvimento e criação de ontologias no nível do conhecimento. Isto é possível a partir da identificação de um conjunto de atividades as quais devem ser desenvolvidas para alcançar o produto final. Estas foram demandadas pela ausência de padronização nos processos, ciclo de vida, critérios de modelagem, entre outros (FERNANDEZ *et al.*, 1997).

Assim, para a criação de uma ontologia faz-se necessário à especificação de quais atividades deverão ser realizadas, segundo o planejamento de tarefas:

- Descrição de propósito e escopo;
- Descrição de como o conhecimento para uma ontologia foi adquirido;
- Conceitualização e formalização do objeto;
- Verificação de possíveis integrações (importações de ontologias, entidades similares, entre outros);
- Implementação em uma linguagem formal;
- Avaliação do objeto, documentação;

- E, a descrição de como será realizada a manutenção (FERNANDEZ *et al.*, 1997).

Além disso, há uma necessidade clara sobre o gerenciamento dessas atividades ordenadamente e o nível de detalhamento que cada uma deve ser descrita e realizada. Nesse sentido, a especificação do ciclo de vida da ontologia serve para definir a sequência de passos a serem realizados para a execução de todas as atividades necessárias (FERNANDEZ *et al.*, 1997). Tal sequência de passos é exemplificada na figura 3.

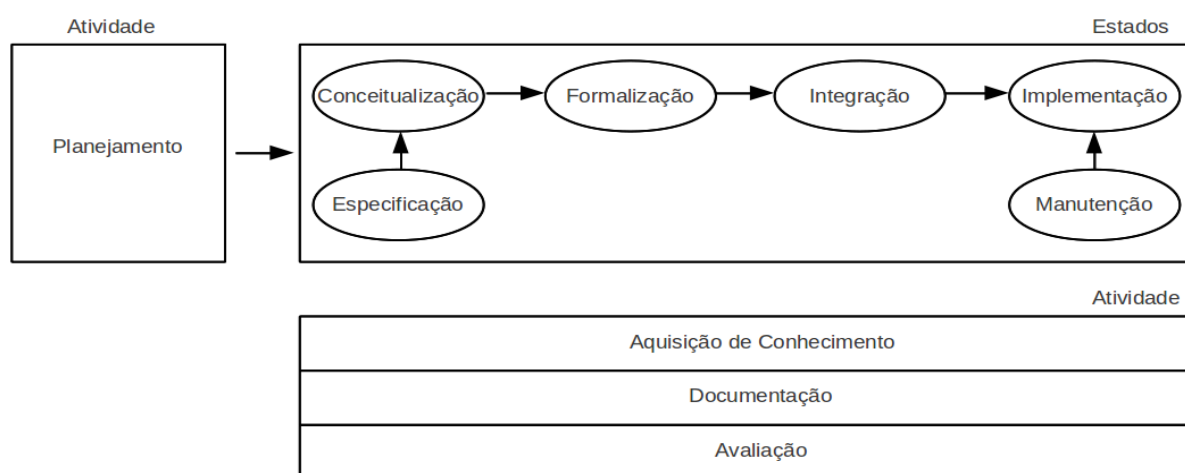


FIGURA 3: Atividades e Estados para a Engenharia de Ontologias(FERNANDEZ *et al.*, 1997).

No ciclo de vida de uma ontologia, após a definição do planejamento, faz-se necessária a especificação do objeto, a conceitualização das entidades, a formalização das descrições, avaliação de possíveis integrações com outras ontologias e, por fim, a implementação do objeto em uma linguagem processável por máquinas. Por último, a manutenção, que inclui a evolução (com enriquecimento da representação) (NOY; KLEIN, MICHEL, 2004), e o versionamento (YILDIZ, 2006)(registro de modificações e alterações realizadas) da ontologia.

Concomitantemente com essas tarefas, a aquisição de conhecimento, a documentação do objeto e a avaliação de tal artefato devem ser realizadas com os demais passos da modelagem para a criação de um produto final (FERNANDEZ *et al.*, 1997).

Assim, após a definição do planejamento e as tarefas que deverão ser realizadas para a construção de uma ontologia, pode ser dado início ao desenvolvimento. Entretanto, a escolha das linguagens de representação e de implementação são dois passos que ocorrem na formalização e implementação, respectivamente. Estes necessitam um estudo mais aprofundado, pois impactarão no que poderá ser representado, na disponibilização e na utilização em sistemas e aplicações.

Um passo não levado em consideração no *Methontology* é a normalização. Uma ontologia normalizada deve apresentar:

- Módulos como árvores de classificação disjuntas;
- E, relações binárias entre módulos distintos devem ser especificados.

Como descrito por Rector (RECTOR, A. L., 2003), com a normalização pode-se conseguir explicitude e modularidade em ontologias, suportando o reuso, manutenção e evolução. Tal abordagem permite uma melhor estruturação do conhecimento por tornar mais conhecimento explícito, ao invés de ser (apenas) ontologicamente limpo, como descrito no OntoClean (GUARINO; WELTY, 2000a, 2000b, 2000c).

Um exemplo de aplicação para a normalização pode ser encontrado no trabalho de Rector (2003), o qual compara a sua abordagem a do OntoClean (GUARINO; WELTY, 2000a, 2000b, 2000c). Este apresenta uma hierarquia básica, e distingue a “borboleta” da “lagarta”, ambos filhos de uma mesma superclasse (*Lepidopteran*). Rector (RECTOR, A. L., 2003) faz a mesma distinção seguindo axiomas, os quais descrevem que uma “lagarta” é um *Lepidopteran* em fase de desenvolvimento jovem, e a borboleta em fase de desenvolvimento adulto.

A representação de Rector aparenta ser mais natural, organizada e próxima da realidade; além de deixar o trabalho de classificação da hierarquia para o raciocinador. Algumas ontologias criadas assim são a BioTop (BEISSWANGER *et al.*, 2008), a Ontologia para Investigações Biomédicas (*Ontology for Biomedical Investigations* – OBI) (COURTOT, MÉLANIE *et al.*, 2008), a NTDO (SANTANA; SCHOBBER; MEDEIROS; *et al.*, 2011), entre outras.

Para que tudo isso seja efetivado, faz-se necessário a escolha de uma linguagem de representação e uma de implementação, a segunda suportando a primeira, para transformar o conhecimento em algo computável. No próximo tópico, serão introduzidos os principais conceitos a cerca da lógica de descrições, formalismo comumente utilizado para representação de ontologias.

## 2.4. LÓGICA DE DESCRIÇÕES - DL

A Lógica de Descrições (*Description Logics* - DL) é um conjunto de linguagens de representação de conhecimento. A DL advém de expressões criadas com conceitos atômicos (predicados unários) e papéis atômicos (predicados binários, ou relações). Uma das principais características sobre a DL é sua semântica baseada em lógica, com uma sintaxe construída sobre um conjunto de construtores Booleanos - conjunção ( $\cap$ ), disjunção ( $\cup$ ) e a negação ( $\neg$ ) -, restrição existencial ( $\exists R.C$ ), restrição de valor ( $\forall R.C$ ), entre outros (BAADER *et al.*, 2003).

Adicionalmente, as descrições de predicados podem ser utilizadas objetivando criar declarações baseadas em DL para uma base de conhecimento (*Knowledge Base* - KB). Comumente, uma KB pode ser dividida em dois blocos: a *Tbox* e a *Abox*. A *Tbox* (com o **T** representando terminológico) engloba o que será representado sobre um dado domínio, como definições de entidades e relacionamentos. A *Tbox* também pode ser chamada de 'definidora', pois as definições das entidades são baseadas em extensões de conceitos primitivos e utilizações de propriedades (ou relações) (BAADER *et al.*, 2008).

Já a *Abox* (**A** de assertiva) é utilizada para descrever situações concretas, pelo instanciamento das entidades descritas na *Tbox*. Adicionalmente, contempla axiomas os quais descrevem as instâncias de um dado conceito e como ele se relaciona com as de outras classes, por meio de propriedades descritas na *Tbox* (BAADER *et al.*, 2008). Na figura 4 há uma breve representação de como a *Tbox* e a *Abox* se comportam em um sistema baseado em conhecimento.

Alguns sistemas baseados em DL provêm aos seus usuários, através do raciocínio automatizado, a possibilidade de inferir conhecimento a partir da descrição, produzindo respostas concretas em tempo finito. Entretanto, tal possibilidade de inferência pode ser limitada pela expressividade da DL utilizada. Assim, existe um balanço entre expressividade e decidibilidade, o qual é refletido na complexidade do que é representado, podendo ser distinguido pelas sublinguagens da DL (BAADER *et al.*, 2003).

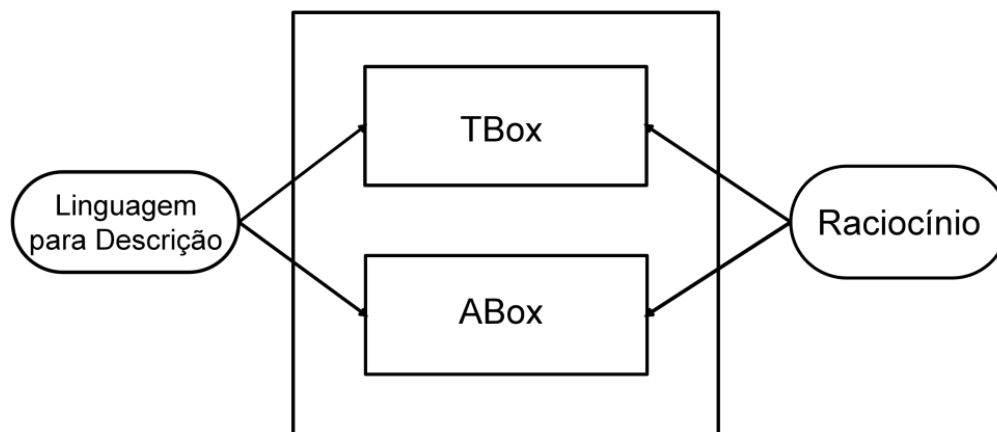


FIGURA 4: Relacionamento entre a *Tbox* e a *Abox*. Adaptado de (BAADER *et al.*, 2003).

A figura acima contempla a *Tbox* e a *Abox* como componentes de algo maior (evidenciado pela caixa maior). Ambos são relacionados entre si pela linguagem para descrição e pela possibilidade de checar o conteúdo da *Abox*, a partir do que foi descrito na *Tbox*, com a utilização de raciocínio.

O raciocínio em DL é considerado uma inferência lógica e pode ser realizado separadamente (*Tbox* ou *Abox*) ou ao mesmo tempo (*Tbox* e *Abox*). Levando em consideração apenas a *Tbox*, o raciocínio pode ser utilizado para verificação de satisfabilidade (coerência), subsunção (uma classe denota ser subclasse de outra), equivalência (duas classes representam, por definição, a mesma coisa) e a disjunção (duas classes não poderão compartilhar instâncias) (BAADER *et al.*, 2003).

Já para a *Abox*, a utilização do raciocinador é relacionada à checagem de consistência entre as instâncias presentes nela e as descrições presentes na *Tbox*. A junção dessas funcionalidades torna possível a comparação e.g. a um banco de dados no qual a ausência de instâncias representa uma informação negativa (inexistente, “mundo fechado”) e para uma ontologia indica ausência de conhecimento (falta de conhecimento, “mundo aberto”) (BAADER *et al.*, 2003), não por isso inviabilizando o raciocínio.

Como já comentado, a DL é um conjunto de linguagens e, por isso, conta com sub-linguagens que podem ser diferenciadas pelos construtores utilizados para representar as classes e as relações. A linguagem mínima de interesse prático é a



$\mathcal{AL}$  (*Attribute Language* – linguagem de atributos), sendo todas as outras uma extensão dessa (BAADER; NUTT, 2007). A sintaxe da,  $\mathcal{AL}$  é a seguinte:

Sendo  $X$  e  $Y$  classes arbitrárias e  $R$  uma relação binária arbitrária, temos:

$X, Y \rightarrow$	$A$   (conceito atômico)
	$\top$   (conceito universal)
	$\perp$   (conceito inferior)
	$\neg A$   (negação atômica)
	$X \sqcap Y$   (interseção)
	$\forall R.X$   (restrição de valor)
	$\exists R.T$   (quantificação existencial limitada)

Exemplificando o que pode ser representado com a  $\mathcal{AL}$ , é possível descrever que, sendo *Pessoa* e *Feminino* conceitos atômicos,  $Pessoa \sqcap Feminino$  e  $Pessoa \sqcap \neg Feminino$  são conceitos descritos com base na  $\mathcal{AL}$ . Assim, pode-se intuitivamente interpretar a primeira afirmação como “as pessoas e que são fêmeas” e, a segunda como, “as pessoas que não são fêmeas” (BAADER *et al.*, 2003). Ainda, adicionando um papel atômico como **temCaráter**, é possível criar conceitos do tipo:  $Pessoa \sqcap \exists \text{temCaráter}.T$ , ou  $Pessoa \sqcap \forall \text{temCaráter}.Feminino$ , os quais denotam “as pessoas de caráter” e “todas as pessoas do sexo feminino tem caráter”, respectivamente.

Partindo para uma definição de semântica formal para os conceitos relacionados à  $\mathcal{AL}$ , serão consideradas as interpretações / as quais consistem em um conjunto não vazio  $\Delta'$  (domínio de interpretação) e uma função de interpretação, e que afirma que para todo conceito atômico  $A$  um conjunto  $A' \subseteq \Delta'$ , e para todo papel atômico  $R$  uma relação binária  $R' \subseteq \Delta' \times \Delta'$ . A função interpretação é extensível para descrições de conceitos pela seguinte fórmula indutiva (BAADER *et al.*, 2003)

$$\begin{aligned}
 \top' &= \Delta' \\
 \perp' &= \emptyset \\
 (\neg A)' &= \Delta' \setminus A' \\
 (X \sqcap Y)' &= X' \cap Y' \\
 (\forall R.C)' &= \{a \in \Delta' \mid \forall b. (a,b) \in R' \rightarrow b \in C'\} \\
 (\exists R.T)' &= \{a \in \Delta' \mid \exists b. (a,b) \in R'\}
 \end{aligned}$$

Ainda, para descrever que dois conceitos são equivalentes ( $C$  é equivalente à  $D$ ), é escrito  $C \equiv D$ , se  $C' = D'$  para todas as interpretações  $I$ .

Incrementos de expressividade podem ser obtidos se outros construtores forem adicionados à  $\mathcal{AL}$ . Para indicar cada um, são adicionadas letras para indicar o tipo de construtor utilizado. A título de exemplo,  $\cup$  é utilizado para a união,  $\mathcal{N}$  para restrição numérica não-qualificada, entre outros. Uma lista pode ser encontrada no quadro 1 (BAADER *et al.*, 2003).

Nome	Sintaxe	Semântica	Símbolo
Superior	$\top$	$\Delta'$	$\mathcal{AL}$
Inferior	$\perp$	$\emptyset$	$\mathcal{AL}$
Interseção	$X \sqcap Y$	$X' \cap Y'$	$\mathcal{AL}$
União	$X \sqcup Y$	$X' \cup Y'$	$\cup$
Negação	$\neg C$	$\Delta' / A'$	$C$
Restrição de Valor	$\forall R.C$	$\{a \in \Delta' \mid \forall b. (a,b) \in R' \rightarrow b \in C'\}$	$\mathcal{AL}$
Quantificação existencial	$\exists R.C$	$\{a \in \Delta' \mid \exists b. (a,b) \in R' \wedge b \in C'\}$	$\mathcal{E}$
Restrição numérica não-qualificada	$\geq n R$ $\leq n R$ $= n R$	$\{a \in \Delta' \mid  \{b \in \Delta' \mid (a,b) \in R'\}  \geq n\}$ $\{a \in \Delta' \mid  \{b \in \Delta' \mid (a,b) \in R'\}  \leq n\}$ $\{a \in \Delta' \mid  \{b \in \Delta' \mid (a,b) \in R'\}  = n\}$	$\mathcal{N}$
Restrição numérica qualificada	$\geq n R.C$ $\leq n R.C$ $= n R.C$	$\{a \in \Delta' \mid  \{b \in \Delta' \mid (a,b) \in R' \wedge b \in C'\}  \geq n\}$ $\{a \in \Delta' \mid  \{b \in \Delta' \mid (a,b) \in R' \wedge b \in C'\}  \leq n\}$ $\{a \in \Delta' \mid  \{b \in \Delta' \mid (a,b) \in R' \wedge b \in C'\}  = n\}$	$\mathcal{Q}$
Mapeamento papel-valor	$R \subseteq S$ $R = S$	$\{a \in \Delta' \mid \forall b. (a,b) \in R' \rightarrow (a,b) \in S'\}$ $\{a \in \Delta' \mid \forall b. (a,b) \in R' \leftrightarrow (a,b) \in S'\}$	
Concordância e Não-concordância	$u_1 \doteq u_2$ $u_1 \not\doteq u_2$	$\{a \in \Delta' \mid \exists b \in \Delta'. u_1'(a) = b = u_2'(a)\}$ $\{a \in \Delta' \mid \exists b_1, b_2 \in \Delta'. u_1'(a) = b_1 \neq b_2 = u_2'(a)\}$	$\mathcal{F}$
Nominal	$I$	$I' \subseteq \Delta'$ com $ I'  = 1$	$\mathcal{O}$

Quadro 1: Expressividade da DL com sua sintaxe e semântica correspondente (BAADER *et al.*, 2003).

A partir de tais construtores, é possível descrever axiomas como

$\text{HomemFeliz} \equiv \text{Humano} \sqcap \neg \text{Fêmea} \sqcap (\exists \text{casado.Doutora})$

$\sqcap (\forall \text{temFilho.}(\text{Doutor} \sqcup \text{Professor}))$

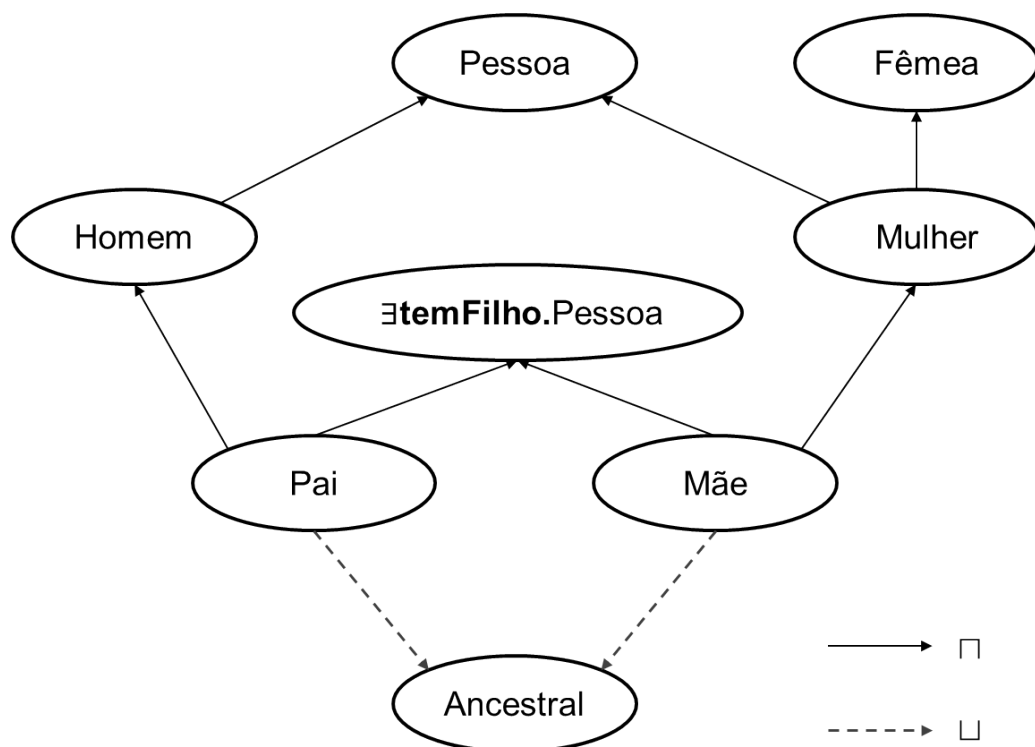
$\sqcap \geq 2 \text{ temFilho} \sqcap \leq 4 \text{ temFilho}$

que significa que um homem feliz, é um humano não-fêmea (macho, então), que é casado com pelo menos uma doutora, tem filhos doutores ou professores em quantidade maior ou igual a dois, ou menor ou igual a 4.

Tendo em mente tais construtores, é possível exemplificar algumas utilizações de raciocínio em DL. Assim, um exemplo de raciocínio por subsunção em *Tbox* é descrito abaixo (baseado em Baader *et al.* (BAADER *et al.*, 2003),

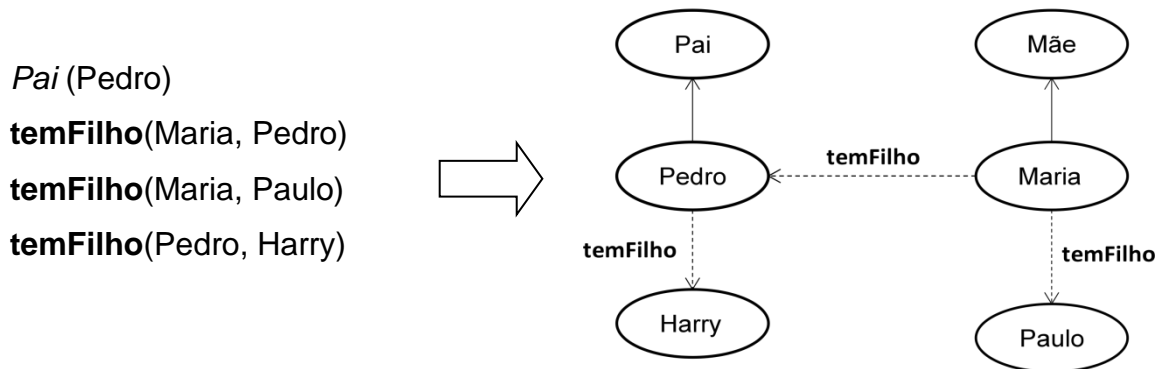
$$\begin{aligned}
 \text{Mulher} &\equiv \text{Pessoa} \sqcap \text{Fêmea} \\
 \text{Homem} &\equiv \text{Pessoa} \sqcap \neg \text{Mulher} \\
 \text{Mãe} &\equiv \text{Mulher} \sqcap \exists \text{temFilho.Pessoa} \\
 \text{Pai} &\equiv \text{Homem} \sqcap \exists \text{temFilho.Pessoa} \\
 \text{Ancestral} &\sqsubseteq \text{Mãe} \sqcup \text{Pai}
 \end{aligned}$$

Analisando a *Tbox* acima, pode-se derivar por subsunção a seguinte hierarquia:

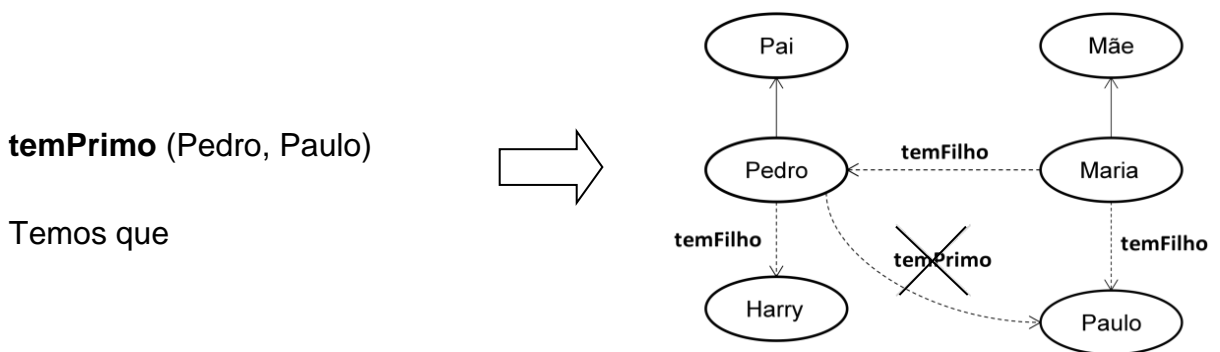


indicando que uma pessoa pode ser homem ou mulher, esta última fêmea. O homem representa os pai e a mulher a mãe, sendo assim descritos a partir da condição de ter pelo menos um filho. Por fim, ambos são descritos como ancestrais.

Tomando como base o exemplo de raciocínio em *Tbox*, uma extensão para a *Abox* (com a descrição de instâncias) e sua visualização gráfica está descrita a seguir.



Seguindo o mesmo raciocínio, é possível construir mentalmente uma relação entre primos (filhos de ancestrais que não são irmãos) e semelhante ao que acontece na realidade, adicionamos a seguinte informação a *Abox*,



Entretanto, a relação **temPrimo** entre Pedro e Paulo, que em uma situação real deve ser do tipo **temIrmão**, seria identificada como inconsistente (marcado com o X), por um raciocinador, caso haja axiomas adicionais que permitam a restrição de afirmações sobre 'irmão' e 'primo'.

De maneira geral, o raciocínio em DL é utilizado para verificar se conceitos não são contraditórios (inconsistentes) e derivar relações. Por meio de algoritmos de raciocínio por subsunção (SCHMIDT-SCHAUSS, 1989), é possível determinar o

relacionamento entre um super-conceito e um sub-conceito, além de computar as instâncias presentes em tais relações. Fazendo uma analogia entre uma base de conhecimento em DL e um banco de dados, temos que uma *Tbox* pode ser comparada ao esquema do banco de dados, enquanto que a *Abox* pode ser comparada com os dados guardados no banco (BAADER *et al.*, 2003).

Apesar disso, a semântica da *Abox* difere da do banco de dados no que concerne à interpretação: no banco, admite-se apenas uma interpretação para o dado (a que está no esquema). Na *Abox*, é possível que uma instância tenha mais do que uma interpretação (uma mesma instância pode ser computada em diversas relações que a classe participe) (BAADER *et al.*, 2003).

Dessa forma, para garantir que o raciocínio seja completo e previsível, a sub-linguagem da DL empregada deve ser no mínimo decidível. O poder de expressividade empregado deve ser restringido (STAAB; STUDER, RUDI, 2009). Alguns exemplos de sublinguagens da DL são a  $\mathcal{ALC}$  ( $\mathcal{AL}$  incluindo negações sobre propriedades); a  $\mathcal{FL}^-$  é equivalente a  $\mathcal{AL}$ , excluindo negação atômica; a  $\mathcal{FL}^0$ , sendo equivalente a  $\mathcal{FL}^-$ , mas não inclui a quantificação existencial limitada; entre outras (BAADER *et al.*, 2003).

A título de exemplo, algumas aplicações baseadas na utilização de DL podem ser encontradas em ontologias para a aviação civil (WANG, H. *et al.*, 2011), ontologias para DTNs (SANTANA; SCHOBBER; MEDEIROS; *et al.*, 2011), combinação de objetivos em agentes inteligentes (WU, 2009), entre muitos outros.

Para o processamento por máquinas de conhecimento representado com base em DL, faz-se necessário à utilização de uma linguagem para implementação. No próximo tópico, a *Ontology Web Language* (OWL), uma recomendação da *World Wide Web Consortium* (W3C) para a representação de ontologias e criada com base na DL, será introduzida.

## 2.5. WEB ONTOLOGY LANGUAGE - OWL

A *Web Ontology Language*, mais conhecida pelo seu acrônimo 'OWL', é uma linguagem utilizada para implementação de ontologias criadas em lógica de descrições. Ela nasceu da junção de outras duas linguagens, a *Linguagem de Marcação para Agentes DARPA* (DAML - *Darpa Agent Markup Language*) (HENDLER; MCGUINNESS, 2000) e da *Linguagem de Permuta para Ontologia* (OIL - *Ontology Interchange Language*) (FENSEL *et al.*, 2000), DAML+OIL. A partir desse ponto, para auxiliar na resolução de problemas decorrentes das linguagens presentes na época, a *World Wide Web Consortium* (W3C) propôs a OWL (MCGUINNESS; HARMELEN, 2004).

A OWL é baseada na habilidade da *eXtensible Markup Language* (XML) em definir anotações personalizadas, e na abordagem de representação de dados flexível do *Resource Description Framework* (RDF). Além disso, a OWL é capaz de representar a estrutura de classes e propriedades, declarando-as com base em subsunção (hierarquias de classe-subclasse) e sendo especificadas com combinações lógicas entre relações e outras classes (axiomas) (HORROCKS, IAN *et al.*, 2003). Indo além e para suportar o raciocínio, a *Ontology Web Language* (OWL) foi concebida para tomar parte de todos esses requisitos (MCGUINNESS; HARMELEN, 2004), e possibilitar de checagem de inconsistências (quando a representação é decidível) e raciocínio por subsunção (ambos exemplificados no tópico 2.4).

Nesse sentido, a OWL foi concebida com três sublinguagens, em ordem crescente de expressividade, relacionadas ao uso final: OWL-Lite, OWL-DL e OWL-Full. A OWL-Lite é mais utilizada quando a sintaxe e a inferência apresenta pouca complexidade (HORROCKS *et al.*, 2003), i.e. necessidades primárias por hierarquias e restrições pouco complexas (MCGUINNESS; HARMELEN, 2004). Já a OWL-DL apresenta como ponto principal o máximo de expressividade, mas mantendo a completude e decidibilidade, semelhante à DL. A última, a OWL-Full é destinada ao máximo de expressividade sem preocupação com completude e decidibilidade (quando a representação tem como principal objetivo ser voltada para a compatibilidade com o RDF e o RDF *Schema*) (HORROCKS *et al.*, 2003).

Alguns benefícios relacionados à modelagem de ontologias com a OWL descritas na literatura são: compartilhamento, evolução, interoperabilidade, detecção de inconsistências, balanço entre expressividade e escalabilidade, facilidade de utilização, compatibilidade com outros padrões (utilizados pela indústria ou na web) e internacionalização (ZHIHONG; MINGTIAN, 2003).

No ano de 2009, a W3C estendeu e revisou a OWL criando novas funcionalidades do ponto de vista sintático e de expressividade, como

- Chaves;
- Cadeias de propriedades;
- *Datatypes* mais ricos;
- Imagem de dados;
- Restrições qualificadas de cardinalidade;
- Assimetria;
- Reflexividade e disjunção de propriedades;
- E, uma melhor capacidade de anotação.

Assim, foi criada a OWL2 e novos perfis de sintaxe foram definidos: OWL2 EL (algoritmos em tempo polinomial, faz alusão à sublinguagem  $\mathcal{EL}^{++}$ ), OWL2 QL (*Query Language* - consultas com conjunções respondidas em *LogSpace*, utilizando tecnologias de bancos de dados relacionais), e OWL2 RL (*Rule Language*, implementação de algoritmos de raciocínio em tempo polinomial) (WORLD WIDE WEB CONSORTIUM, 2009). A OWL2 conta com a expressividade  $\mathcal{SROIQ}$  da DL (STAAB; STUDER, 2009).

Outro aspecto importante relacionado à OWL, desde a primeira versão, foi o desenvolvimento da *Manchester Syntax* (sintaxe Manchester). O desenvolvimento desta objetiva a introdução de pessoas não familiarizadas com a lógica a uma

sintaxe que facilitasse a construção de ontologias (HORRIDGE *et al.*, 2006). A *Manchester Syntax* foi atualizada para a *Manchester OWL2 Syntax* com a criação da OWL2. Tal atualização não acarretou em mudanças significativas (HORRIDGE; PATEL-SCHNEIDER, 2009).

A *Manchester OWL2 Syntax* se apresenta como uma descrição em linguagem natural do sentido dos construtores lógicos, disponibilizados pela DL, para a OWL2. Por exemplo, alguns construtores lógicos como o “ $\exists$ ” e o “ $\forall$ ” são traduzidos para “some” e “only”, respectivamente (HORRIDGE; PATEL-SCHNEIDER, 2009).

Muitos trabalhos utilizaram-se da OWL para a representação de suas respectivas ontologias. Por exemplo, Yan & Zha (2010) apresentaram, em seu trabalho, uma ontologia para gerenciar conhecimento sobre consumidores, demonstrando a capacidade expressiva da OWL em representar tal conhecimento. Thomas *et al.* (THOMAS *et al.*, 2009) desenvolveram uma ontologia em OWL para relacionar entidades físicas, químicas e descrições funcionais de nano partículas para aplicações no diagnóstico do câncer e para fins terapêuticos.

Outro trabalho, desenvolvido por Bouamrane *et al.* (BOUAMRANE *et al.*, 2010), construiu e implementou uma ontologia em OWL para auxiliar na avaliação, como suporte a decisão, em eventos cirúrgicos, e Santana *et al.* (SANTANA *et al.*, 2011; SANTANA; SCHOBER; MEDEIROS; *et al.*, 2011) utilizou a OWL2 para representar padrões de representações ontológicas e eventos complexos relacionados às doenças tropicais negligenciáveis. O próximo tópico introduz os Padrões de Representações Ontológicas, parte importante para o entendimento do presente trabalho.



## **2.6. PADRÕES DE PROJETOS ONTOLÓGICOS – PPO**

### **ONTOLOGY DESIGN PATTERNS – ODP**

Seguindo os processos de Engenharia de Ontologias, pelos quais são descritos o ciclo de vida, o reuso e o processo de modelagem para situações complexas é algo pouco explorado. Algumas razões para o pouco reuso em ontologias pode ser a alta complexidade dos modelos, descrição pobre, ausência de critérios para as definições e as fragilidades inerentes às ferramentas disponíveis (GANGEMI, 2005).

Para tal, os Padrões de Projetos Ontológicos (PPO) (do inglês *Ontology Design Patterns* – ODP) foram propostos e são utilizados para a modelagem de um problema relacionado à descrição de classes de domínio, além das propriedades, levando em consideração também problemas relacionados ao conteúdo da ontologia (GANGEMI, 2005). Ou seja, PPOs são soluções para resolver problemas recorrentes de modelagem ontológica (BLOMQVIST *et al.*, 2009).

De acordo com Egaña *et al.* (EGAÑA *et al.*, 2008), são descritos 3 conjuntos distintos de PPOs: Extensional, Boas Práticas e Modelagem de Domínio. Os PPOs Extensionais visam à extensão dos limites de uma linguagem de representação (EGAÑA *et al.*, 2008). Exemplos deste são os PPOs de:

- Exceção (“Exception ODP,” 2008):
  - Utilizado para classificar as exceções às normas biológicas, como por exemplo, às células sanguíneas humanas que não apresentam núcleo em comparação com todas as outras células do corpo. É fundamentado na utilização de disjunções entre entidades “típicas” e “atípicas”;
- Relações n-árias (“N-ary Relationship,” 2008):
  - Representação da relação entre mais do que dois elementos, em OWL. Para tal, faz-se necessário criar uma nova classe que, por reificação, integra as demais classes que fazem parte da relação n-ária.

Já os PPOs de Boas Práticas são relacionados ao melhoramento das práticas de modelagem ontológica, como criação de módulos ontológicos bem definidos,

ontologias mais eficientes e uma maior manutenção, entre outros (EGAÑA *et al.*, 2008). Alguns exemplos são os de:

- Normalização (“Normalization ODP,” 2008): Aplicado em hierarquias com muitas superclasses para evitar dificuldade de manutenção em ontologias grandes, semântica definida implicitamente. Para tal, faz-se necessário o desmembramento (*untangling*) de uma hierarquia e utilização de restrições para definir as relações classe-subclasse, semelhante ao que é descrito por Rector (RECTOR, A. L., 2003);
- Partição de Valor (“Value Partition ODP,” 2008): Utilizado para descrever os tipos de valores admitidos para um atributo, quando esse pode ser diferentemente descrito para entidades diferentes. Nesse caso, utiliza-se a definição de um determinado atributo em dada superclasse e, nas subclasses fazem-se as especificações necessárias com a definição do valor para o atributo.
- Ontologias *core* (“Upper Level Ontology ODP,” 2008): Utilizado para descrever ontologias *core* que, ideologicamente, são utilizadas para integrar ontologias da mesma natureza. Para tal, a introdução de classes primitivas deve ser realizada por meio de disjunções.

Por fim, os PPOs para Modelagem de Domínio são utilizados para modelar uma parte concreta do domínio, além de suas peculiaridades (EGAÑA *et al.*, 2008). Podem ser encontrados exemplos de tais PPOs no catálogo público de PPOs (“Ontology Design Patterns (ODPs) Public Catalog,” 2008), sendo um desses apresentados posteriormente neste trabalho.

Outra iniciativa para a criação, desenvolvimento e manutenção de PPOs é evidenciada pelo trabalho de Gangemi (GANGEMI, 2005), Presutti & Gangemi (PRESUTTI; GANGEMI, 2008) e Blomqvist *et al.* (BLOMQVIST *et al.*, 2009). A junção desses trabalhos resultou no desenvolvimento de vários padrões de representações, os quais estão dispostos em um portal específico (“Ontology Design Patterns.org (ODP),” 2011).

Alguns exemplos utilização e descrição de PPOs podem ser encontrados em

- Santana et al. (SANTANA; SCHOBBER; MEDEIROS; *et al.*, 2011), com a representação do padrão de transmissão de DTNs, especificamente Leishmaniose e as implicações de tal processo de transmissão (SANTANA; FERNANDES; *et al.*, 2011);
- Damjanovic (DAMJANOVIC, 2009), a qual descreveu um PPO para refletir a semântica e a sintaxe presente na definição de processos de negócio;
- Bobillo et al. (BOBILLO *et al.*, 2007), os quais descreveram um PPO para representar quais partes do conhecimento sobre um dado domínio devem ser considerados em um dado cenário mediante o contexto;
- E, Gangemi et al. (GANGEMI *et al.*, 2004) apresentaram um padrão de projeto para representar processos inflamatórios.

Adicionalmente, e de forma diferente ao que foi descrito por Egaña et al. (EGAÑA *et al.*, 2008), os padrões de representação dispostos no portal de PPOs ("Ontology Design Patterns.org (ODP)," 2011) são divididos em 6 grupos: Conteúdo, re-engenharia, alinhamento, lógicos, arquiteturais e léxico-sintáticos. Do ponto de vista do presente trabalho (e por questões de objetividade), focaremos nos PPOs de conteúdo, pois estes se prestam para solucionar problemas relativos à modelagem de domínio, semelhante aos PPOs de modelagem de domínio descritos por Eganã et al. (EGAÑA *et al.*, 2008).

De maneira geral, com a utilização dos PPOs juntamente à modelagem ontológica, há um possível ganho de qualidade geral na ontologia, principalmente quando o padrão é bem representado e é comprovada, de alguma forma, sua viabilidade. Em Santana et al. (SANTANA; SCHOBBER; MEDEIROS; *et al.*, 2011) foi descrito um PPO para a representação do mecanismo de transmissão para vetores de doenças tropicais negligenciáveis, especificamente, Leishmaniose e modificado, posteriormente, para a representação de Tuberculose (SANTANA; FERNANDES; *et al.*, 2011). Uma descrição mais aprofundada sobre as doenças em questão, parte do tema da pesquisa, será apresentado a seguir.

## **3. DOENÇAS TROPICAIS NEGLIGENCIÁVEIS – DTN**

Neste capítulo, serão tratados os conceitos relativos às Doenças Tropicais Negligenciáveis (DTN) e sua importância para subsidiar o presente trabalho. Este inclui uma breve descrição sobre o impacto causado na população, tanto do ponto de vista da morbidade (quantidade de pessoas doentes por uma determinada doença em relação à população), quanto da mortalidade.

### **3.1. O QUE SÃO AS DOENÇAS TROPICAIS NEGLIGENCIÁVEIS?**

**E**m 2001, a Organização Não-governamental Médicos Sem Fronteiras (MSF) produziu um documento intitulado *Fatal Imbalance* (MEDECINS SANS FRONTIERES (MSF), 2001) o qual propunha uma nova classificação geral para doenças, compartimentalizando-se em três grandes grupos. Estes são o grupo das doenças ‘globais’ (ocorrem em todo o mundo), o das ‘negligenciadas’ (de maior prevalência nos países em desenvolvimento) e o das ‘mais negligenciadas’

(exclusivas de países em desenvolvimento) (MEDECINS SANS FRONTIERES, 2001).

Neste mesmo ano, por iniciativa da Organização Mundial de Saúde (OMS) foi produzido o 'Relatório da Comissão sobre Macroeconomia e Saúde', introduzindo uma classificação similar a dos MSF e dividindo as doenças em Tipo I (equivalente à classificação "doenças globais" proposta pela MSF), Tipo II (negligenciadas da MSF) e Tipo III (mais negligenciadas da MSF) (WORLD HEALTH ORGANIZATION, 2001).

A partir daí, tais denominações vem sendo utilizadas como referência a um conjunto de doenças causadas por agentes infecciosos e parasitários, endêmicas e que mais afetam populações de baixa renda (HOTEZ *et al.*, 2007). Consequentemente, foi obtido um novo olhar no conceito de doenças negligenciadas, antes conhecida de maneira errônea por (apenas) doenças tropicais, por contemplar apenas os fatores geográficos, e não também, os contextos de desenvolvimento político, econômico e social de um país (MOREL, 2006).

### **3.2. QUAIS SÃO E O NO QUE RESULTAM?**

As doenças tropicais negligenciadas representam um conjunto de doenças que causam morbidade e mortalidade (HOTEZ *et al.*, 2009; LINEHAN *et al.*, 2011), e estão associadas a precárias condições de vida, pobreza, iniquidades sanitárias e de saúde. Ainda, auxiliam na permanência do quadro de desigualdade econômico e social, ao representar uma importante barreira no desenvolvimento de um país (DECIT, 2010; HOTEZ *et al.*, 2009). Além disso, compõe um grupo de doenças que recebem de maneira geral, baixo investimento e um reduzido apoio à pesquisa nos países em desenvolvimento (HOTEZ *et al.*, 2009).

Estima-se que grandes partes das pessoas portadoras de uma ou mais DTN vivem com menos de US\$ 2,0 por dia, seja no Brasil ou outros países da América Latina e do Caribe, na África, na Ásia e em algumas localidades dos Estados Unidos e Europa (HOTEZ *et al.*, 2008).

Hoje seu combate é essencial para o cumprimento dos objetivos de desenvolvimento do Milênio, da Organização das Nações Unidas (ONU), os quais

visam a erradicação total da pobreza até o ano de 2015 (ALLOTEY *et al.*, 2010; MAUDLIN *et al.*, 2009). Verifica-se que as DTN causam efeitos adversos na gravidez, na produtividade agrícola, e no desenvolvimento infantil (HOTEZ; AKSOY, 2011; LINDOSO; LINDOSO, 2009). Essas doenças são consideradas como um dos grupos mais importantes de infecções que aprisionam populações carentes a um ciclo de pobreza e desnutrição, promovendo assim, cada vez mais a iniquidade social (HOTEZ *et al.*, 2008, 2007; KARIUKI *et al.*, 2011).

Vale salientar, ainda, a quantidade de mortes causadas pelas doenças negligenciadas. De acordo com Hotez et al. (HOTEZ *et al.*, 2006), as DTNs causam 534 mil mortes, anualmente, sem contar as 5,6 milhões causadas por Malária-Tuberculose-HIV/AIDS, em todo o mundo. Tal conjunto de doenças podem ser tratadas com menos de US\$1 (1 dólar americano) *per capita* por ano. Entretanto, muitas delas apresentam um quadro emergente, como a Leishmaniose, a Febre do Dengue e a Doença de Chagas.

Para a maioria dos casos, as DTN tendem a cronicidade, produzindo sérias deficiências nas populações vulneráveis. Dessa forma, causam incapacidade grave e permanente, com altas taxas de morbidade, podendo levar à morte. Suas taxas de mortalidade consideradas baixas fizeram com que, no passado, eles recebessem pouca prioridade (HOTEZ *et al.*, 2006).

Em crianças, por exemplo, algumas doenças negligenciadas de maior prevalência, tais como ancilostíase, helmintíase e esquistossomose, resultam em desnutrição e dificuldades no crescimento, além de déficits cognitivos e mentais (BETHONY *et al.*, 2006). Em mulheres grávidas, as infecções por esse grupo de doenças resultam em prematuridade, baixo peso da criança ao nascer e aumento da morbimortalidade materna (HOTEZ *et al.*, 2008).

Com o objetivo de alcançar as metas do Milênio e do plano mundial de luta contra as doenças negligenciadas, a OMS aponta temas fundamentais para a redução da prevalência, os quais visam:

- Programar intervenções de saúde que atendam às necessidades e contribuam para o desenvolvimento socioeconômico dos países endêmicos;
- E desenvolver sistemas integrados de vigilância em saúde, importantes para melhoria da qualidade dos dados de saúde e contribuindo dessa forma, na tomada de decisões por parte dos gestores públicos (WORLD HEALTH ORGANIZATION, 2011).

### **3.3. REGIÕES DE OCORRÊNCIA**

As DTN ocorrem em muitos países da América Latina, Ásia e África Sub-Saariana. Adicionalmente, há uma larga distribuição das doenças negligenciadas seguindo as condições climáticas (entre 15°C e 40°C) e a altitude (abaixo de 2.200 metros), em sua maior parte dentro da região tropical e por isso são também conhecidas como Doenças Tropicais Negligenciáveis (DTN). Os indivíduos que vivem nessas regiões (por volta de 2,7 milhões de pessoas) são considerados os mais pobres do mundo e, também, os mais vulneráveis as DTNs (HOTEZ *et al.*, 2007).

Na região da América Latina e do Caribe podem ser encontradas as maiores disparidades de renda entre classes sociais, com 48% de toda a renda alocada no 1/3 mais rico da população e o terço menos favorecido apenas 1,6% (THE WORLD BANK, 2003). Dentre a população com menor potencial financeiro, 1/3 desses vivem em zona rural (praticando agricultura de subsistência, pesca, entre outros) (AULT, 2007). Nessa parcela da população, é peculiar a presença indígena e afrodescendente, com um alto nível de exclusão e iniquidade social, privados de acesso à água para consumo e serviços de saúde (AULT, 2007; RILEY *et al.*, 2007).

Os outros dois terços da população pobre estão distribuídos em favelas, assentamentos, bairros pobres ou em áreas periféricas das cidades. Nestas áreas há uma combinação de água de qualidade duvidosa, baixos níveis sanitários e proliferação de roedores, reservatórios e vetores de doenças (AULT, 2007; RILEY *et al.*, 2007).

Assim como em muitas áreas do mundo, no Brasil, as DTN são mais elevadas em regiões empobrecidas, também sendo relacionadas ao índice de desenvolvimento

humano (IDH). As regiões do país mais afetadas são o Norte e o Nordeste, apresentando altos níveis de prevalência da Doença de Chagas, Leishmaniose, Esquistossomose, Dengue, Oncocercose e a Filariose Linfática. No Brasil, existe uma estimativa na qual 40 milhões de pessoas estão infectadas com uma ou mais DTN (HOTEZ, 2008).

Ainda, entre 2001 e 2008, houve quase 40 mil casos de morte por DTN (LINDOSO; LINDOSO, 2009), descritos no Sistema de Informações de Mortalidade (SIM). Tais doenças figuram também o grupo das Doenças de Notificação Compulsória, do Sistema de Informação de Agravos de Notificação (SINAN) do Ministério da Saúde (MS/Brasil) (BRASIL, 2009).

O conjunto de informações sobre ocorrências de doenças (morbidade, como presente no SINAN) e o informe de mortes, *viz.* SIM, são as principais fontes de informações necessárias e disponíveis para avaliar a saúde da população. Entretanto, as relações e o sentido eminente das fontes de informações de mortalidade e morbidade não são explorados. Nesse sentido, a junção de ambas as fontes permitiria conhecer individualmente, e de forma regionalizada, como se comportam as transmissões das doenças, o que decorre da infecção por tais patógenos, e o agravamento com falecimento de indivíduos.

Nesse âmbito surgiu o presente trabalho visando, por meio da utilização de ontologias, ser mais um braço na luta contra as DTN. O presente trabalho evidencia os processos de transmissão, ponto crucial para bloquear a proliferação de vetores, e a explicitação das principais causas que levam a morte de indivíduos, para servir de suporte na avaliação dos dados e no direcionamento das ações de saúde para a população.



### **3.4. Transmissão de DTN por Vetores Artrópodes**

Dentre as DTN, o conjunto de doenças transmitidas por vetores artrópodes, e.g. Febre do Dengue, Leishmaniose, Doença de Chagas, Tripanossomíase Africana, Filariose Linfática, Febre Amarela, entre outras, resultam em um alto nível de mortalidade em regiões pouco favorecidas (HOTEZ *et al.*, 2009). Dentre todos os casos de doenças transmitidas por vetores que ocorrem na América latina, 67% da leishmaniose visceral, 46% dos casos de leishmaniose cutânea (ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE, 2007), 80,8% dos casos de dengue (PAN AMERICAN HEALTH ORGANIZATION, 2008) e 8% dos casos de filariose linfática ocorrem no Brasil (WORLD HEALTH ORGANIZATION, 2006).

Um dos pontos principais para ações de prevenção e controle em doenças transmitidas por vetores é o processo de transmissão. O bloqueio do mecanismo de transmissão evita que o agente causador de uma doença, de alguma forma, entre em contato com o ser humano e a doença seja desenvolvida (BRASIL, 2009). Em um comentário escrito por TaUIL (TAUIL, 2010), é enfatizado a necessidade do bloqueio da transmissão do mosquito da Febre Amarela, como uma estratégia de controle da doença no Brasil.

Por isso, faz-se necessário conhecer o processo de transmissão das DTN transmitidas por vetores, como um ponto crucial para a definição de ações em saúde frente a tal conjunto de enfermidades. A seguir, há uma breve introdução sobre o processo de transmissão por vetores.

#### **3.4.1. Processo de Transmissão**

Para o entendimento do presente tópico, faz-se necessário à introdução do que é um mecanismo de transmissão de patógenos e quando este é realizado por vetores. O processo de transmissão de patógenos por vetores é o resultado da interação entre um artrópode (p.ex. um mosquito, besouro, entre outros) o qual leva um determinado e potencial patógeno (p.ex. leishmania, vírus da febre do dengue, entre outros), a um hospedeiro (p.ex. ser humano, cão, entre outros) susceptível a desenvolver uma enfermidade, em um dado ambiente (NEVES *et al.*, 2005) (Figura 5).

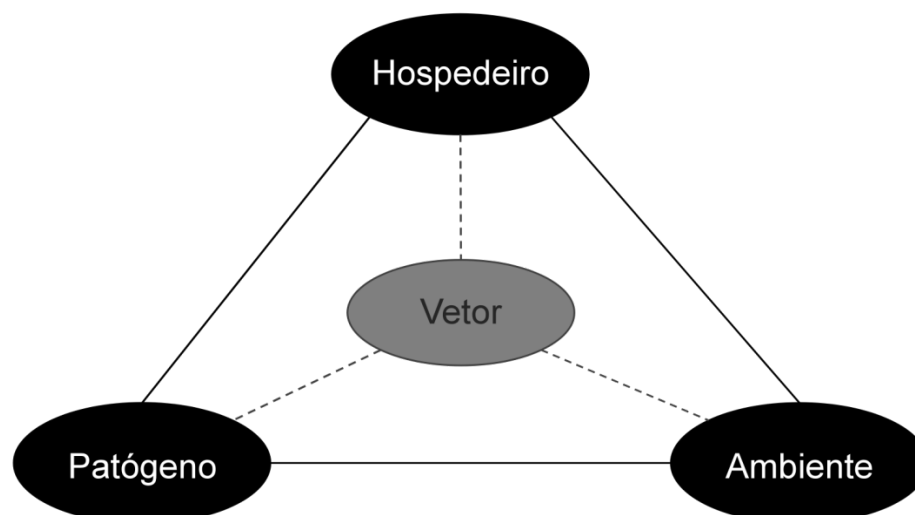


FIGURA 5: Tríade Epidemiológica com a Inclusão do Vetor. Os principais componentes em uma transmissão são o hospedeiro, o patógeno e o ambiente. O vetor é frequentemente relacionado a todos os outros componentes fazendo dele o cerne do processo de transmissão em doenças transmitidas por vetores (NEVES *et al.*, 2005).

Outros importantes participantes, mas nem sempre presente, são os reservatórios (de patógenos). Estes são outros animais, o solo, matéria inanimada, entre outros que proveem um meio para a manutenção do patógeno (NEVES *et al.*, 2005).

Frequentemente, as DTN transmitidas por vetores são associados a um perfil de localidade no qual o hospedeiro, o vetor e o patógeno (além do reservatório) podem compartilhar uma mesma localização geográfica, i.e. habitat, por um determinado tempo (REISEN, 2010). Assim, para a representação de qualquer modelo relacionado a esse conjunto de doenças em questão, faz-se necessário a modelagem de todos os componentes evidenciados na figura 5, já que esse é o princípio básico.

Frequentemente, tal conjunto de interações é descrita na literatura como uma representação tabular, como pode ser encontrado em Piscopo & Azzopardi (PISCOPO; MALLIA AZZOPARDI, 2007), Sharma & Singh (SHARMA; SINGH, 2008) e Mendoza *et al.* (MENDOZA *et al.*, 2009). Na maioria das vezes essas tabelas evidenciam, principalmente, os componentes que estão descritos na figura 5. No presente estudo, utilizaremos a tabela sobre Leishmaniose descrita por Sharma & Singh (SHARMA; SINGH, 2008) (Tabela 1), pela extensão e quantidade de

informação representada, e por causa disso, como exemplo. Antes de apresentar uma tabela sobre a Leishmaniose, faz-se necessário introduzi-la.

A Leishmaniose é uma doença infecciosa causada por protozoários do gênero *Leishmania*, sendo descritas pelo menos 20 espécies patogênicas (PISCOPO; MALLIA AZZOPARDI, 2007), e é estritamente relacionada à virulência da espécie e a resposta imunológica do hospedeiro. Tal doença é endêmica em 88 países, sendo 22 do Novo Mundo e 66 do Velho Mundo (SHARMA; SINGH, 2008).

A infecção pela *Leishmania* pode causar desde uma simples lesão cutânea localizada até a infecção difusa do sistema mononuclear fagocitário. São descritos 4 principais quadros clínicos:

- Leishmaniose Cutânea (LC) ou Tegumentar;
- Leishmaniose Visceral (LV), ou *kala-azar*;
- Leishmaniose Mucocutânea (LMC), ou espundia;
- E, a Leishmaniose Cutânea Difusa (LCD) (SHARMA; SINGH, 2008).

A forma como as *Leishmanias* são transmitidas se dá através da picada de fêmeas de mosquitos do gênero *Phlebotomus* (Velho Mundo) ou *Lutzomyia* (Novo Mundo). Tais vetores vivem em ecossistemas desérticos/semi-áridos e em florestas úmidas, respectivamente (SHARMA; SINGH, 2008). Na tabela 1 há uma descrição extensa sobre os principais locais de transmissão de patógenos da Leishmaniose no Novo Mundo.

TABELA 1: Espécies patogênicas (aos humanos) de *Leishmania*, seus vetores no Novo Mundo, manifestações clínicas e suas distribuições geográficas, baseado em Sharma & Singh (SHARMA; SINGH, 2008).

País	Espécie de <i>Leishmania</i> *	Doença Causada	Mosquito Vetor <sup>%</sup>
Argentina	<i>L. (L.) chagasi</i>	LV	<i>Lu. longipalpis</i>
	<i>L. (V.) braziliensis s.l.</i>	LC	<i>Lu. intermedia</i>
Belize	<i>L. (L.) mexicana</i>	LC	<i>Lu. olmeca olmeca</i>
	<i>L. (V.) braziliensis s.l.</i>	LC	<i>Lu. ovallesi</i>
Bolivia	<i>L. (L.) amazonensis</i> ,	LC, LCD,	<i>Lu. Flaviscutellata</i>
	<i>chagasi</i>	LV	<i>Lu. longipalpis</i>
	<i>L. (V.) braziliensis s.l.</i> ,	LC, LMC	<i>Lu. carrerae carrerae</i>
	<i>yucumensis</i> & <i>llanosmartini</i>		
Brasil	<i>L. (L.) amazonensis</i>	LC, LCD, LMC, VL	<i>Lu. flaviscutellata</i>
	<i>chagasi</i>	VL	<i>Lu. longipalpis</i>
	<i>L. (V.) braziliensis</i>	LC, LMC	<i>Lu. wellcomei</i>
	<i>guyanensis</i>	LC, LMC	<i>Lu. umbralis</i>
	<i>lainsoni</i>	LC	<i>Lu. ubiquitalis</i>
	<i>naiffi</i>	LC	<i>Lu. ayrozai</i>
	<i>shawi</i>	LC	<i>Lu. whitmani</i>
Colombia	<i>L. (L.) amazonensis</i>	LC, LCD	<i>Lu. flaviscutellata</i>
	<i>chagasi</i>	VL	<i>Lu. evansi</i>
	<i>mexicana</i>	LC, LCD	<i>Lu. columbiana</i>
	<i>L. (V.) braziliensis s.l.</i> ,	LC, LMC	<i>Lu. spinicrassa</i>
	<i>colombiensis</i>	LC	<i>Lu. hartmanni</i>
	<i>guyanensis</i>	LC	<i>Lu. umbralis</i>
	<i>panamensis</i>	LC, LMC	<i>Lu. trapidoi</i>
Costa Rica	<i>L. (L.) mexicana</i>	LC	<i>Lu. olmeca olmeca</i>
	<i>L. (V.) braziliensis s.l.</i>	LC, LMC	<i>Lu. trapidoi</i>
	<i>panamensis</i>	LC	<i>Lu. trapidoi</i>
República Dominicana	<i>L. (L.) mexicana-like</i>	LCD	desconhecido

Equador	<i>L. (L.) mexicana</i>	LC	<i>Lu. ayacuchensis</i>
	<i>L. (V.) braziliensis s.l.</i>	LC, LMC	<i>Lu. gomezi</i> & <i>Lu. trapidoi</i>
El Salvador	<i>L. (L.) chagasi</i>	LV,	<i>Lu. longipalpis</i>
	<i>mexicana</i>	LC	<i>Lu. longipalpis</i>
Guiana	<i>L. (L.) amazonenses</i>	LC, LCD	<i>Lu. flaviscutellata</i>
Francesa	<i>L. (V.) braziliensis s.l.</i> ,	LC, LMC	<i>Lu. whitmani</i>
	<i>guyanensis</i>	LC	<i>Lu. umbretilis</i>
	<i>naiffi</i>	LC	<i>Lu. squamiventris</i>
Guadaloupe	<i>L. (L.) chagasi</i>	LV	<i>Lu. longipalpis</i>
Guatemala	<i>L. (L.) chagasi</i>	LV	<i>Lu. olmeca olmeca</i>
	<i>mexicana</i>	LC	<i>Lu. ylephiletor, olmeca &amp; cruciata</i>
	<i>L. (V.) braziliensis s.l.</i>	LC	<i>Lu. ovallesi</i>
Guiana	<i>L. (V.) guyanensis</i>	LC	<i>Lu. umbretilis, anduzei</i>
	<i>Leishmania sp.</i>	LMC	<i>Lu. whitmani</i>
Honduras	<i>L. (L.) chagasi</i>	LV, LC	<i>Lu. olmeca olmeca</i>
	<i>mexicana</i>	LC, LCD	<i>Lu. olmeca olmeca</i>
	<i>L. (V.) braziliensis s.l.</i>	LC, LMC	<i>Lu. ovallesi</i>
	<i>L. (V.) panamensis</i>	LC, LMC	<i>Lu. hartmanni</i>
Martinica	<i>L. (L.) sp.</i>	LC	<i>Lu. whitmani</i>
México	<i>L. (L.) chagasi, mexicana</i>	LV, LC	<i>Lu. longipalpis,</i>
	& outra sp	LCD	<i>Lu. olmeca olmeca</i>
Nicarágua	<i>L. (L.) chagasi</i>	LV	<i>Lu. longipalpis</i>
	<i>L. (V.) braziliensis s.l.</i>	LC, LMC	<i>Lu. ylephiletor</i>
	<i>panamensis</i>	LC, LMC	<i>Lu. panamensis</i>
Panamá	<i>L. (L.) aristedesi</i>	LC	<i>Lu. trapidoi</i>
	<i>L. (V.) braziliensis s.l.</i> ,	LC	<i>Lu. ovallesi</i>
	<i>panamensis</i>	LC	<i>Lu. gomezi</i>
	<i>Leishmania sp.</i>	LMC	<i>Lu. panamanensis</i>

Paraguai	<i>L. (L.) amazonensis</i>	LC, LCD	<i>Lu. flaviscutellata</i>
	<i>chagasi</i>	LV	<i>Lu. longipalpis</i>
Peru	<i>L. (V.) braziliensis s.l.</i>	LC, LMC	<i>Lu. whitmani</i>
	<i>peruviana</i>	LC	<i>Lu. peruensis&amp;verrucarum</i>
Suriname	<i>Leishmania sp.</i>	LC	<i>Lu. flaviscutellata</i>
EUA	<i>L. (L.) mexicana</i>	LC, LCD	<i>Lu. olmeca olmeca</i>
Venezuela	<i>L. (L.) infantum chagasi</i>	LV	<i>Lu. evansi</i>
	<i>garnhami</i>	LC	<i>Lu. youngi</i>
	<i>pifanoi</i>	LC, LCD	<i>Lu. olmeca bicolor</i>
	<i>venezuelensis</i>	LC	<i>Lu. spinicrassa</i>
	<i>L. (V.) braziliensis s.l.</i>	LC, LMC	<i>Lu. umbralis</i>
	<i>colombiensis</i>	LV	<i>Lu. hartmanni</i>
	<i>guyanensis</i>	LC	<i>Lu. ovallesi</i>

\*Na coluna de 'Espécies', a letra 'L.' indica o gênero *Leishmania* e as letras entre parênteses, V. e L. indicam o subgênero da espécie, *Viannia* e *Leishmania*, respectivamente.

%Na coluna 'Mosquito Vetor', a sílaba 'Lu.' indica o gênero do mosquito vetor *Lutzomyia*.

Nessa tabela é possível encontrar os principais países da América Latina, Central e do Norte os quais apresentam alguma distribuição do patógeno da Leishmaniose, p.ex. em *L. (L.) venezuelensis*, associado a possível patologia que pode vir a ser desenvolvida por um hospedeiro susceptível (p.ex. humano). Tal patógeno pode ser transmitido por um vetor, p.ex. *Lu. spinicrassa*, nesse local.

A seguir, serão abordados os aspectos metodológicos que subsidiam o desenvolvimento do presente trabalho.

## 4. ASPECTOS METODOLÓGICOS

No presente capítulo, será apresentada a metodologia empregada desde a concepção da ontologia até a sua manutenção. O processo completo será descrito separadamente para facilitar o entendimento sobre o objeto de estudo. Este segue o Methontology (FERNANDEZ *et al.*, 1997), uma metodologia de engenharia de ontologias a qual descreve todos os passos a serem tomados para a criação de uma ontologia.

### 4.1. MODELAGEM ONTOLÓGICA

A seguir, serão descritos os passos seguidos para a construção do modelo representacional, na forma de ontologia, sobre as DTNs.

#### 4.1.1. Propósito

A presente ontologia, criada e disponibilizada como *NTDO – Neglected Tropical Disease Ontology* (NTDO, 2012) (NTDO, Ontologia para as Doenças Tropicais Negligenciáveis) presta-se à representação de um domínio, as DTN, pouco explorado e com relevância e impacto reconhecido pela literatura e por instituições

de saúde pública nacionais e internacionais. A ontologia descreve detalhadamente o processo de transmissão de doenças e o falecimento de indivíduos que podem vir a ocorrer em a partir da doença transmitida.

Os usuários de tal artefato serão profissionais de saúde (médicos, biólogos, biomédicos, enfermeiros, sanitaristas, entre outros) os quais estão incumbidos de vigiar e avaliar os eventos relacionados à saúde da população. Outro grupo de usuários identificados são os profissionais de Tecnologias de Informação que desenvolvem sistemas e aplicações na área da saúde. Com a NTDO poderão realizar uma análise criteriosa das causas de morbidade e mortalidade para direcionar ações de saúde específicas, baseado na utilização de raciocínio lógico e da representação ontológica dos principais mecanismos que levam um indivíduo a ser infectado por um determinado patógeno até uma possível condição posterior de falecimento.

#### **4.1.2. Escopo**

A ontologia aqui descrita, a NTDO, apresenta uma estrutura hierárquica concebida a partir de entidades provenientes da BioTop (*Top-domain ontology for the life Sciences*), pela especialização dos nós folha e generalização de entidades mais específicas do domínio para as mais gerais da BioTop. Esta ontologia foi criada nos moldes da BFO (GRENON *et al.*, 2004) e das recomendações do OBO *Foundry*. Adicionalmente, é descrita como uma camada (em nível *core*) que se presta para ligar e integrar várias ontologias do domínio das ciências da vida (BEISSWANGER *et al.*, 2008).

Os principais componentes a serem descritos na NTDO são relacionados à representação da transmissão das DTN (como um PPO) e a descrição temporal de processos que são vinculados aos indivíduos. Entre estes está à representação dos processos passíveis de acontecimento a partir da transmissão, como o desenvolvimento e aparecimento de doenças/sinais/sintomas (SCHULZ *et al.*, 2011), além de outros fatores os quais, sendo realizados, podem incorrer no falecimento de um dado indivíduo. No presente estudo, morte significa a ausência de funções cerebrais e cessação de todas as funções biológicas inerentes ao corpo humano



(MILLER; TRUOG, 2010). Tal definição será utilizada como base para descrição do modelo de mortalidade.

Do ponto de vista das classes e relações, estas são agrupadas na ontologia de tal forma que permita a representação da complexidade eminente entre a interação dos seres vivos, ao nível de transmissão de patógenos e possíveis acometimentos patológicos. Além disso, possibilita o estabelecimento de uma sequência de fatos que permitem a visualização de um ciclo de vida, cujo foco seja à propensão à infecção e sua possível realização, culminando com a morte.

#### 4.1.3. Fontes de Conhecimento

As fontes de conhecimento utilizadas, os termos empregados e suas respectivas descrições em linguagem natural são provenientes de publicações relacionadas às DTN (livros, periódicos nacionais/internacionais, textos da OMS e do Ministério da Saúde brasileiro, entre outros). Vale ressaltar que grande parte das entidades descritas na NTDO são provenientes da BioTop e algumas da GFO, ambas ontologias também consideradas fonte de conhecimento. Adicionalmente, outras ontologias e suas respectivas descrições sobre o assunto foram verificadas para possíveis criações de referências.

A título de exemplo, entidades relacionadas ao 'tempo' foram incorporadas e mapeadas a partir da representação de *Chronoid* e *TimeBoundary* provenientes da GFO (HELLER; HERRE, 2004; HERRE *et al.*, 2007) para a NTDO e BioTop, respectivamente. Na GFO, entidades materiais, *viz.* *biotop:MaterialObject* e *gfo:Presential*, tem sua existência definida pela determinação de limites temporais (*gfo:RightTimeBoundary* e *gfo:LeftTimeBoundary*). Adicionalmente, as entidades processuais, *viz.* *biotop:Process* e *gfo:Process*, são projetadas temporalmente sobre *gfo:Chronoid*. Esta última classe, proveniente da GFO, apresenta limites próprios de início e fim. Uma explicação mais detalhada será apresentada no tópico 5.2.

Outra fonte importante de conhecimento para o presente trabalho é a forma como as informações sobre a transmissão de patógenos (nas DTN, especificamente) é apresentada em publicações científicas. Geralmente, estas se utilizam de tabelas

para facilitar o entendimento e apresentar dados para os leitores, de tal forma que evidenciam claramente as interações entre a informação simbólica. Um exemplo de representação tabular pode ser encontrado abaixo (Tabela 2), baseado no trabalho de Sharma & Singh (SHARMA; SINGH, 2008).

TABELA 2: Tabela de DTN Transmitidas por vetores, por localização geográfica, patógeno e manifestação clínica.

Localização Geográfica	Vetor	Patógeno	Manifestação Clínica
Argentina*	<i>Lutzomyia intermedia</i>	<i>Leishmania (V) braziliensis</i>	Leishmaniose Tegumentar
Brasil*	<i>Lutzomyia longipalpis</i>	<i>Leishmania (L) chagasi</i>	Leishmaniose Visceral
América do Sul <sup>\$</sup>	<i>Culex quinquefasciatus</i>	<i>Wuchereria bancrofti</i>	Filariose Linfática
México ao Extremo Sul da América do Sul <sup>#</sup>	<i>Rhodnius prolixus</i> <i>Triatoma infestans</i> <i>Triatoma dimidiata</i>	<i>Trypanosoma cruzi</i>	Doença de Chagas
África <sup>%</sup>	<i>Aedes aegypti</i>	Yellow Fever Virus	Febre Amarela

Dados coletados em:

\*(SHARMA; SINGH, 2008);

<sup>\$</sup>(MENDOZA *et al.*, 2009);

<sup>#</sup>(RASSI JR *et al.*, 2010);

<sup>%</sup>(BARRETT; HIGGS, 2007).

Nesta tabela, podem ser encontrados os principais componentes de um processo de transmissão de patógenos por vetores artrópodes, logicamente o vetor, o patógeno, manifestação clínica vinculada a ele, e o local em que foi identificado. Comparando a tabela 1 com a tabela 2, esta apresenta exemplos de outras doenças as quais podem ser apresentadas similarmente à matriz descrita anteriormente.

#### 4.1.4. Transferência de Conhecimento Tabular para a NTDO

A transferência de conhecimento da representação tabular para a ontologia, de forma que suporte raciocínio concreto e correto é um processo com vários passos e, algumas vezes, iterativo. O procedimento de aquisição de informações tabulares

para descrição do PPO e a representação do modelo de mortalidade pode ser descrito pelo seguinte fluxo:

- Análise ontológica da representação tabular ao nível textual no qual a tabela foi inserida. Primeiramente, são decididas quais entidades são classes (e.g. vetores ou patógenos) e quais são indivíduos (e.g. vetores, patógenos, ou localidades geográficas). Posteriormente, as classes ou categorias superiores são selecionadas e os indivíduos são abstraídos. Referências implícitas a entidades as quais não são explicitamente descritas na tabela (mas estão presentes no texto) são identificadas e ligadas a possíveis categorias ontológicas (quando necessário). Finalmente, demais relações e dependências entre as entidades são identificadas e registradas em linguagem natural. Um ponto importante é a verificação de dependências existenciais, pois estas necessitam ser avaliadas (e.g. uma doença é existencialmente dependente em um patógeno, mas não vice-versa) quando a informação representada for exaustiva, e.g. quando é definido que uma doença é causada somente por três patógenos específicos;
- Formulação de um Padrão de Projeto Ontológico (no presente estudo, apenas para a representação do processo de transmissão baseado em vetores). Utilizando-se a linguagem de representação escolhida, uma ou mais expressões axiomáticas são construídas como protótipo. Um exemplo de tal passo pode ser encontrado em Bobillo et al. (BOBILLO *et al.*, 2007), os quais descrevem, a partir de um caso de uso relacionado à assistência a saúde, um PPO para avaliar o contexto de determinada situação e evidenciar as partes do conhecimento do domínio que são pertinentes para um dado cenário em consideração
- Implementação do padrão de representação, genérico para as doenças transmitidas por vetores, tanto manualmente quanto de maneira semi-automática por um processador de padrão de representação, ou por um conjunto de regras, pelas quais padrões são interpretados e a ontologia desejada é construída segundo o que está contido em tabelas (no presente estudo, apenas para a representação do processo de transmissão baseado em vetores). Um exemplo de tal passo também foi descrito por Peters et al. (PETERS *et al.*, 2009), trabalho este que enfatiza a

representação tabular como fonte para geração automática de classes, relações e axiomas.

- Revisão manual da ontologia expandida semi-automaticamente (no presente estudo, apenas para a representação do processo de transmissão baseado em vetores). Isto inclui a reestruturação manual da ontologia gerada corrigindo-a ou enriquecendo-a (e.g. reconstruindo taxonomias) e finalmente a integração com outras ontologias. Este passo é necessário visto que a geração automática de código é passível de erros de sintaxe, possivelmente inseridos no momento da criação do gerador de código (geralmente um script), como em Santana et al. (SANTANA; SCHOBBER; MEDEIROS; *et al.*, 2011) e Peters et al. (PETERS *et al.*, 2009), ou quando se inclui a informação tabular no gerador..

Abaixo, a figura 6 resume o processo de transferência de conhecimento tabular.

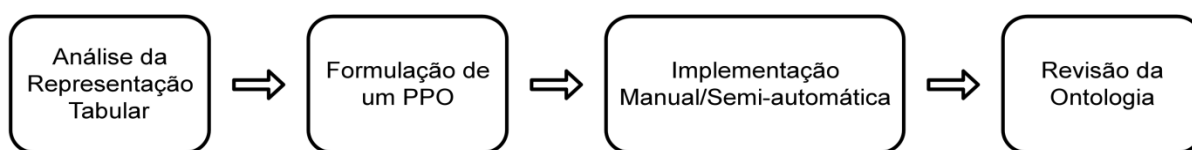


FIGURA 6: Processo de Transferência de Conteúdo Tabular para uma ontologia.

Por fim, é proposto um modelo para a representação de eventos de morte, o qual foi construído a partir da necessidade levantada e pela importância que é dado à morte de indivíduos, do ponto de vista epidemiológico (principalmente). Para tal, faz-se uma descrição passo-a-passo de todas as ações desenvolvidas do ponto de vista ontológico para a construção do modelo. Os fatos a serem descritos em tal modelo evidenciam, principalmente, os possíveis processos os quais podem ocorrer posteriormente à transmissão de patógenos e ao desenvolvimento de alguma enfermidade, que é o falecimento de indivíduos.

#### 4.1.5. Integração com outras Ontologias

A NTDO, para uma representação mais completa do domínio e, para fins de reuso, traz referências completas no corpo da representação de classes e relações presentes em outras ontologias, e.g. *General Formal Ontology* (GFO) (HELLER; HERRE, 2004; HERRE *et al.*, 2007), para representação temporal de processos e

entidades materiais, e da BioTop. Esta representação está descrita, detalhadamente, na seção 5.2.2 e exemplificada na figura 7.

Como já citado, a NTDO importa e estende a BioTop (Beisswanger et al., 2008), usada como fonte primária de classes básicas e relações, seguindo uma abordagem *middle-out*.

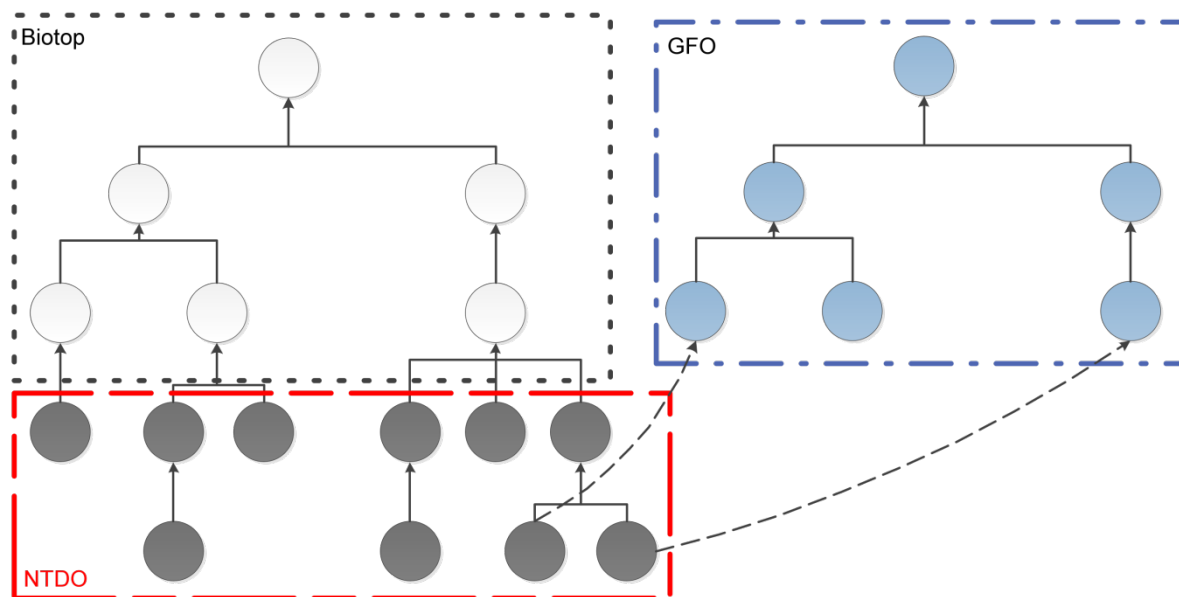


FIGURA 7: Estrutura da NTDO. A NTDO importa e estende a BioTop e apresenta alguns mapeamentos para a GFO.

#### 4.1.6. Desenvolvimento da NTDO

Do ponto de vista da modelagem da ontologia em uma linguagem formal, compreensível e processável por máquina, a NTDO foi concebida sobre uma ótica formal, sendo representada em DL (BAADER *et al.*, 2003) e implementada em OWL2 (WORLD WIDE WEB CONSORTIUM, 2009).

Para tal, foi utilizado o editor de ontologias Protégé v.4.1 (NATIONAL LIBRARY OF MEDICINE *et al.*, 2012), auxiliado pelo motor de raciocínio HermiT (MOTIK *et al.*, 2007) v.1.3.5 para classificação automática de classes por raciocínio de subsunção e verificação estrutural de satisfibilidade na criação de classes e relações durante o processo de modelagem.

#### 4.1.7. Requisitos, Normalização e Documentação

Como requisitos da ontologia, foram utilizados Questões de Competência (*Competency Questions* – CQ) (GRUNINGER; FOX, 1994) para descrição dos principais componentes necessários para a construção do PPO. Ao passo que a ontologia é desenvolvida, segundo as CQ, é necessária à realização de ciclos iterativos para avaliar se o que está representado está em conformidade com o que é requerido. Tal avaliação é realizada segundo consultas em DL e, caso haja necessidade de complementar ou corrigir algo, novos ciclos são iniciados.

Já para a normalização da NTDO, o presente trabalho é baseado no estudo desenvolvido por Rector (RECTOR, A. L., 2003). Este descreve que, a normalização é um processo que tem como principal objetivo explicitude e modularidade do domínio a partir da análise criteriosa sobre a disjunção de classes e estabelecimento de axiomas entre elas. Isto possibilita um melhor reuso, manutenção e evolução da ontologia. Tais requisitos são vitais para um ambiente que sofre constantes modificações e que poderá ser integrado a diferentes aplicações, como a NTDO. Para maiores detalhes sobre normalização de ontologias, consultar o tópico 2.3.

Portanto, a NTDO foi agrupada em pequenos módulos, cada um apresentando um conjunto definido de classes e relações utilizadas para auxiliar na reutilização de suas sub-partes e de outras as quais já estão sendo importadas por outros trabalhos.

A documentação da ontologia consiste no presente documento e em publicações já disponíveis na literatura (SANTANA; FERNANDES; *et al.*, 2011; SANTANA; SCHOBBER; MEDEIROS; *et al.*, 2011), além das que poderão vir a ser publicadas posteriormente. Mais informações podem ser encontradas em <http://www.cin.ufpe.br/~ntdo>.

#### 4.1.8. Manutenção

A manutenção da presente ontologia é feita segundo a utilização de um repositório de subversão (*SVN – subversion*), o qual possibilita uma fácil detecção do processo de trabalho, descritos em arquivos de histórico.

Evoluções do modelo e ampliações de escopo seguirão a mesma sistemática de trabalho descrita aqui, principalmente no que concerne a utilização da NTDO em outros estudos ou aplicações, e que requeira adição de novos módulos e novos ciclos de iteração para refinamento do objeto.

Adicionalmente, no site da NTDO (NTDO, 2012), foi disponibilizada uma aba de contato direto com os desenvolvedores principais, podendo assim, receber questionamentos externos sobre o conteúdo e possíveis indicações de correção ou ampliação de conteúdo.

## ***5. UMA ONTOLOGIA PARA AS DOENÇAS TROPICAIS NEGLIGENCIÁVEIS***

***NEGLECTED TROPICAL DISEASE ONTOLOGY - NTDO***

**N**este tópico será descrito a NTDO, a partir dos processos e demais entidades relacionados à transmissão de patógenos até o início de uma determinada DTN. Como exemplo, será utilizada a representação da Leishmaniose. Em seguida, são apresentados os principais e possíveis processos que acontecem posteriormente ao desenvolvimento de uma determinada doença em um dado indivíduo, o falecimento.



### 5.1. REPRESENTAÇÕES TABULARES E DTN

Levando em consideração o impacto causado pelas DTN, principalmente as transmitidas por vetores, e as representações tabulares apresentadas anteriormente evidenciando processos de transmissão de patógenos por vetores, o desenvolvimento ontologia é iniciado pelo processo de extração de conhecimento, com uma análise ontológica das tabelas. Para relembrar uma das tabelas, abaixo está novamente a tabela 2.

TABELA 2: Tabela de DTN Transmitidas por vetores, por localização geográfica, patógeno e manifestação clínica.

Localização Geográfica	Vetor	Patógeno	Manifestação Clínica
Argentina*	<i>Lutzomyia intermedia</i>	<i>Leishmania (V) braziliensis</i>	Leishmaniose Tegumentar
Brasil*	<i>Lutzomyia longipalpis</i>	<i>Leishmania (L) chagasi</i>	Leishmaniose Visceral
América do Sul <sup>\$</sup>	<i>Culex quinquefasciatus</i>	<i>Wuchereria bancrofti</i>	Filariose Linfática
México ao Extremo Sul da América do Sul <sup>#</sup>	<i>Rhodnius prolixus</i> <i>Triatoma infestans</i> <i>Triatoma dimidiata</i>	<i>Trypanosoma cruzi</i>	Doença de Chagas
África <sup>%</sup>	<i>Aedes aegypti</i>	Yellow Fever Virus	Febre Amarela

Dados coletados em:

\*(SHARMA; SINGH, 2008);

<sup>\$</sup>(MENDOZA *et al.*, 2009);

<sup>#</sup>(RASSI JR *et al.*, 2010);

<sup>%</sup>(BARRETT; HIGGS, 2007).

Na presente tabela, a coluna mais a esquerda apresenta nomes referentes a países individuais (ou regiões), sendo consideradas como instancias de regiões geográficas. A próxima coluna contém termos os quais denotam os vetores artrópodes, sendo esses subclasses da classe *Arthropod*, da BioTop. As células da terceira coluna são relacionadas aos patógenos os quais podem ser subclasse de várias classes da BioTop, e.g. *Bacteria*, *Protist* ou *Virus*, dependendo do patógeno descrito na literatura para a doença a ser descrita. Aqui, como o exemplo escolhido

para representação é a Leishmaniose, as leishmanias fazem parte da classe dos protistas (biotop:*Protist*).

Por fim, as células da última coluna contêm nomes de manifestações (doenças) as quais são subclasses de *PathologicalProcess* na BioTop (SCHULZ *et al.*, 2011). Outras observações são dignas de atenção:

1. Há células com mais de um termo, denotando mais de uma entidade;
2. Nem todas as células em uma coluna contêm classes disjuntas, p.ex. *Leishmania* sp., a qual representa um gênero que também pode denotar o sentido de outras classes, como *Leishmania donovani*;
3. Os indivíduos na primeira coluna são espacialmente relacionados, e.g. a região “México ao Extremo Sul da América do Sul” inclui espacialmente o Brasil, a Argentina, entre outros países.

Analisando o conteúdo que está descrito nas linhas, é possível concluir que cada uma descreve um tipo diferente de padrão de transmissão. Precisamente, cada linha representa uma subclasse da classe ntdo:*Transfer*,

ntdo:*Transfer* subClassOf biotop:*BiologicalProcess*

a qual é subclasse de biotop:*BiologicalProcess*. De acordo com o que foi interpretado, o sentido geral de uma tabela pode ser ou não considerado como uma descrição exaustiva.

Ligando as classes descritas pela reutilização das relações descritas na BioTop (em OWL, *object properties*): cada instância do processo de transmissão tem uma localização (biotop:**hasGeographicLocation**) e um agente (biotop:**hasAgent**), viz. o vetor (segunda coluna); e um participante passivo (biotop:**hasPatient**), o patógeno a ser transmitido. Quanto o processo é finalizado (e somente então) ele é instanciado e nesse momento o patógeno está localizado no hospedeiro.

O hospedeiro aparenta ser um ponto de ligação perdido nessa tabela pelo fato de ser restrito aos *Homo sapiens* (pelo interesse epidemiológico). Mesmo assim ainda é necessário representá-lo. A relação entre o patógeno e o hospedeiro é, acima de

tudo, locacional pelo fato do patógeno estar localizado em um hospedeiro no final do processo de transmissão. Assim, foi usada a relação biotop:**physicallyLocatedIn**.

A relação entre o patógeno (interno ao hospedeiro ou ao vetor) e a manifestação não é perfeita, visto que nem todo processo de transmissão implicar em uma infecção do hospedeiro. Este último fato pode acontecer (não sendo uma relação existencial) depois de um processo de transmissão do patógeno pelo vetor para um hospedeiro.

Por isso, foi necessário incluir a distinção entre disposição e manifestação, descrita por Schulz et al. (SCHULZ *et al.*, 2011). As disposições são relacionadas às manifestações (e.g. doenças) pelas propriedades biotop:**hasRealization** e seu inverso, biotop:**realizationOf**, com o primeiro sendo descrito por meio de restrições de valor ('only', na Manchester Syntax); e o último com restrições existenciais ('some', na Manchester Syntax). O que é típico para as doenças levadas em consideração, assim para a Leishmaniose, é que elas ocorrem em organismos infectados por gêneros (ou espécies) de patógenos específicos.

Após definir como será descrito cada entidade do processo de transmissão para uma tabela, é necessário formalizar os processos de transmissão como um padrão representação, já que as transmissões por vetores se comportam de uma forma semelhante. Este será apresentado no próximo tópico.

### 5.1.1. Formalização do PPO para Representação da Transmissão

A formalização do PPO para a representação do processo de transmissão de doenças baseado em vetor segue a notação apresentada na tabela 3.

TABELA 3: Padrão geral de uma matriz para DTN transmitidas por vetores.

Localização Geográfica	Artrópode (Vetor)	Vertebrado (Hospedeiro)	Protista (Patógeno)	Manifestação (Doença)
$G_{a1} G_{a2} \dots G_{ak}$	$V_{a1} V_{a2} \dots V_{aj}$	$H_{a1} H_{a2} \dots H_{am}$	$P_{a1} P_{a2} \dots P_{al}$	$D_{a1} D_{a2} \dots D_{an}$
$G_{b1} G_{b2} \dots G_{bk}$	$V_{b1} V_{b2} \dots V_{bj}$	$H_{b1} H_{b2} \dots H_{bm}$	$P_{b1} P_{b2} \dots P_{bl}$	$D_{b1} D_{b2} \dots D_{bn}$
...	...	...	...	...
$G_{n1} G_{n2} \dots G_{nk}$	$V_{n1} V_{n2} \dots V_{nj}$	$H_{n1} H_{n2} \dots H_{nm}$	$P_{n1} P_{n2} \dots P_{nl}$	$D_{n1} D_{n2} \dots D_{nn}$

...	...	...	...	...
G <sub>z1</sub> G <sub>z2</sub> ...G <sub>zk</sub>	V <sub>z1</sub> V <sub>z2</sub> ...V <sub>zj</sub>	H <sub>z1</sub> H <sub>z2</sub> ...H <sub>zm</sub>	P <sub>z1</sub> P <sub>z2</sub> ...P <sub>zl</sub>	D <sub>z1</sub> D <sub>z2</sub> ...D <sub>zn</sub>

Na tabela acima, cada linha representa uma transmissão. Comparando com as tabelas da mesma natureza apresentadas até agora, há apenas a inclusão do hospedeiro do processo, o qual está, geralmente, subentendido como humano. Assim,

- A primeira coluna do lado esquerdo apresenta todas as entidades geográficas listadas para uma transmissão, p.ex. G<sub>z1</sub>G<sub>z2</sub>...G<sub>zk</sub> poderia representar Brasil, Argentina ou Guadalupe;
- A segunda coluna, da esquerda para a direita, apresenta os possíveis vetores presentes em uma transmissão, p.ex. V<sub>z1</sub>V<sub>z2</sub>...V<sub>zj</sub> pode representar *Lu. evansi*, *Lu. youngi*, entre outros;
- A coluna central, incluída para descrever os possíveis hospedeiros presentes no processo de transmissão, p.ex. H<sub>b1</sub>H<sub>b2</sub>...H<sub>bm</sub> no presente estudo será apenas biotop:*Human*, pelo que pode ser causado a esses seres;
- A penúltima coluna deverá conter a lista de patógenos ordenados, para uma dada transmissão, p.ex. P<sub>z1</sub>P<sub>z2</sub>...P<sub>zl</sub> pode representar *L. (L.) aristedesi* ou *L. (V.) braziliensis s.l.* no presente estudo;
- E, a coluna mais a direita conterà a lista de manifestações que podem acontecer a partir da transmissão e infecção do hospedeiro, p. ex. tal coluna pode conter valores como Leishmaniose Visceral ou Leishmaniose Mucocutânea em D<sub>z1</sub>D<sub>z2</sub>...D<sub>zn</sub>.

Levando em consideração apenas as entidades geográficas, estas apresentam certas peculiaridades as quais necessitam ser devidamente descritas. Estas estão descritas no próximo tópico.

### 5.1.2. Entidades Geográficas

As entidades geográficas foram incluídas, no presente trabalho, como subclasses do tipo `ntdo:GeographicLocation`, i.e. *BrazilLocation*, *ArgentinaLocation*, entre outros. Reificando as classes de entidades geográficas é possível emular subsunção

taxonômica para representações espaciais (SCHULZ *et al.*, 2005) que, além de facilitar a representação, mantém o raciocínio apenas na *TBox*. Considerando que o indivíduo

‘Brazil’ **part\_of** ‘South America’

como descrito na tabela 2, a classe pode ser descrita como

*ntdo:BrazilLocation* subClassOf *ntdo:SouthAmericaLocation*

Assim, os componentes necessários para a realização de consultas em DL no nosso modelo de domínio são, preferivelmente, descritos como classes e não como instâncias.

O seguinte axioma demonstra o sentido das localizações geográficas ( $G_i$ ) transformadas em classes ( $G\_loc_i$ ).

$G\_loc_i$  equivalentTo

*GeographicLocation* and (**geographicallyLocatedIn** value  $G_i$ )

De fato, as definições completas necessárias para o raciocínio com as instâncias reais podem ser completamente alcançadas por meio de subclasses de *GeographicLocation* ( $G\_Loc_i$ , viz. *ArgentinaLocation*, *GuatemalaLocation*, entre outras). Por causa disso, não há necessidade de criar instâncias do tipo ‘ $G_i$ ’ na ontologia, já que tais classes se referem única e exclusivamente às instâncias de mesmo nome. De antemão, uma representação geográfica completa fugiria do escopo do presente trabalho, já que para o cumprimento dos objetivos a descrição apresentada é suficiente.

Com as peculiaridades geográficas em mente, seguiremos com a representação do processo de transmissão.

### 5.1.3. Transmissão do Patógeno

Como descrito em algumas passagens anteriores, cada linha 'n' presente na tabela é interpretada como um processo de transmissão e, aqui, como uma subclasse de *ntdo:PathogenTransferByVector*,

*PathogenTransferByVector* equivalentTo

*PathogenTransferByVector<sub>a</sub>* or  
*PathogenTransferByVector<sub>b</sub>* or . . . or  
*PathogenTransferByVector<sub>n</sub>* or . . . or  
*PathogenTransferByVector<sub>z</sub>*

Assim, cada *PathogenTransferByVector<sub>a</sub>...PathogenTransferByVector<sub>z</sub>* corresponde aos processos de transmissão individuais, i.e. cada um indica uma linha da tabela e limitados apenas à quantidade de linhas da tabela. A tabela no presente estudo foi considerada como sendo uma descrição exaustiva do domínio, desde que a classe mãe *ntdo:PathogenTransferByVector* represente a disjunção das classes filhas. Assim, cada subclasse *ntdo:PathogenTransferByVector<sub>n</sub>* é completamente definida e carrega consigo um conjunto de restrições de valor.

*ntdo:PathogenTransferByVector<sub>n</sub>* equivalentTo *ntdo:PathogenTransferByVector*

and (biotop:**hasAgent** some ( $V_{n1}$  or  $V_{n2}$  or . . . or  $V_{nj}$ ))  
and (ntdo:**hasGeographicLocation** some ( $G\_loc_{n1}$  or . . . or  $G\_loc_{nk}$ ))  
and (biotop:**hasPatient** some ( $P_{n2}$  or  $P_{n2}$  or . . . or  $P_{nl}$ )  
and (biotop:**physicallyLocatedIn** some ( $H_{n1}$  or  $H_{n2}$  or . . . or  $H_{nm}$ ))))

Neste primeiro axioma de equivalência (descrito por meio de restrições existenciais), é descrito que em uma dada transmissão de patógeno por um vetor, sendo composta por, pelo menos, um agente (nesse caso classes de vetores); a localização geográfica do processo de transmissão; o participante direto da transmissão; o patógeno que está sendo levado pelo vetor; e em quem o patógeno estará localizado no final do processo de transmissão. Assim, é assumido que pelo menos um patógeno estará no corpo do hospedeiro no instanciamento do processo.

Exemplificando tal axioma,

ntdo:*PathogenTranferByVector*<sub>1</sub> equivalentTo ntdo: *PathogenTranferByVector*  
 and (biotop:**hasAgent** some ntdo:*LutzomyiaLongipalpis*)  
 and (ntdo:**hasGeographicLocation** some ntdo:*MexicoLocation*)  
 and (biotop:**hasPatient** some (ntdo:*LeishmaniaChagasi*  
 and (biotop:**physicallyLocatedIn** some biotop:*Human*)))

indica a transmissão da *Leishmania chagasi* (patógeno) para um humano (hospedeiro), por um *Lutzomyia longipalpis* (vetor) no México (local).

Adicionalmente, faz-se necessário à inclusão de axiomas de subclasse

*PathogenTranferByVector*<sub>n</sub> subClassOf *Transfer*  
 and (biotop:**hasAgent** only ( $V_{n1}$  or  $V_{n2}$  or ... or  $V_{nj}$ ))  
 and (ntdo:**hasGeographicLocation** only ( $G_{loc_{n1}}$  or ... or  $G_{loc_{nk}}$ ))  
 and (biotop:**hasPatient** only ( $P_{n1}$  or  $P_{n2}$  or ... or  $P_{ni}$ ))  
 and (biotop:**physicallyLocatedIn** only( $H_{n1}$  or  $H_{n2}$  or ... or  $H_{nm}$ ))  
 and (biotop:**causes** only ( $D_{n1}$  or  $D_{n2}$  or ... or  $D_{nr}$ )))

para adicionar restrições de valor às entidades participantes. Diferentemente do axioma de equivalência, neste há a inclusão não apenas dos pacientes, localidades e agentes, mas também o que o patógeno virá a causar após a transmissão de tais patógenos. Ambas as partes da definição são requeridas para definir logicamente a classe de forma que o raciocinador de DL possa explorá-lo para a resolução de consultas. Em seguida, há a descrição das patologias decorrente do processo transmissivo.

#### 5.1.4. Disposições e Manifestações

Na presente modelagem, foi considerado que seres vivos, especificamente humanos, exibem disposições para realizar um dado processo, relativo principalmente as restrições inerentes ao próprio ser de enfrentar tal processo. Assim, levando em consideração ‘disposições para doenças’, estas expressam o fato de, após uma infecção, há um estado diferenciado do organismo, independente

do modo como a doença irá se apresentar (eventualmente) (SCHULZ *et al.*, 2011). Tal distinção é descrita na BioTop como apresentado abaixo

$D\_disp_i$  equivalentTo biotop:*PathologicalDisposition*  
and (biotop:**hasRealization** only  $D_i$ )

$D_i$  equivalentTo biotop:*PathologicalProcess*  
and (biotop:**realizationOf** some  $D\_disp_i$ )

$D_i$  subClassOf biotop:**realizationOf** only  $D\_disp_i$

indicando que uma dada disposição  $D\_disp_i$  se dá na realização de qualquer tipo de doença  $D_i$ . Já uma doença  $D_i$  é representada como um processo patológico (biotop:*PathologicalProcess*) o qual acontece em decorrência de pelo menos uma disposição ( $D\_disp_i$ ) em desenvolver tal doença, sendo todas as disposições disponíveis para essa doença do mesmo tipo ( $D\_disp_i$ ).

Utilizando a Leishmaniose como exemplo, os mesmos axiomas são descritos assim:

ntdo:*VisceralLeishmaniasisDisposition* equivalentTo biotop:*PathologicalDisposition*  
and (biotop:**hasRealization** only ntdo:*VisceralLeishmaniasis*)

ntdo:*VisceralLeishmaniasis* equivalentTo biotop:*PathologicalProcess*  
and (biotop:**realizationOf** some ntdo:*VisceralLeishmaniasisDisposition*)

ntdo:*VisceralLeishmaniasis* subClassOf  
biotop:**realizationOf** only ntdo:*VisceralLeishmaniasisDisposition*

Complementarmente, a seguir é detalhada a dependência das doenças no patógeno que elas estão relacionadas.

#### 5.1.5. Dependência das doenças nos patógenos

De acordo com a literatura, a Leishmaniose e várias outras doenças causadas por patógenos (não só as DTN), apresentam-se como a consequência da infecção por



um ou mais patógenos semelhantes. Exemplos de doenças amplamente conhecidas e causadas por patógenos são a Síndrome da Imunodeficiência Adquirida (SIDA), causada pelo vírus HIV; a Tuberculose, causada pela bactéria *Mycobacterium tuberculosis*; entre outras. Especificamente sobre a Leishmaniose, esta é causada por espécies patogênicas do gênero *Leishmania*.

Entretanto, 'causalidade' é um conceito complexo e pode haver outros fatores condicionantes para o aparecimento da doença. Estes podem ser considerados também fatores causais, mas ainda não são conhecidos ou não foram listados na literatura. Michalek (MICHALEK, 2009) apresentou, em sua tese de doutorado, as várias abordagens possíveis a serem tomadas acerca da "causalidade".

No presente trabalho, e como descrito por Michalek (MICHALEK, 2009) a causalidade será considerada a partir da união probabilística de todas as versões possíveis de resultados semelhantes de que, unicamente, a presença do patógeno é que pode determinar a causa de uma doença, se esta só pode ser causada por um patógeno específico, no contexto das DTN. Ou seja, se um indivíduo desenvolve leishmaniose ou dengue, ele deve (necessariamente) apresentar o patógeno em algum estágio de vida e estar, imunologicamente, disposto a receber o patógeno (o organismo é incapaz de expulsá-lo) e desenvolver a doença (organismo é incapaz de, no mínimo, neutralizar o patógeno).

Por isso, o seguinte axioma foi incluído

$D_i$  subClassOf biotop:**causedBy** only ( $P_{n1}$  or  $P_{n2}$  or ... or  $P_{nl}$ )

Indicando que toda doença  $D_i$  será causada por um ou mais patógenos descritos em um conjunto definido. Por exemplo, a Leishmaniose Visceral pode ser causada por *Leishmania chagasi*, *Leishmania colombiensis* e/ou *Leishmania mexicana*.

Já estando próximo da finalizar o modelo, ainda faltam alguns pequenos detalhes. Um deles está descrito no próximo tópico.

#### 5.1.6. Axiomas de Inclusão Geral

Até agora, foram descritos o processo de transmissão, onde o patógeno se localiza em cada momento da transmissão, a descrição de um organismo ter a disposição em desenvolver uma determinada doença (se exposto a ela), evidenciando a dependência da doença no patógeno. Entretanto, ainda é possível afirmar que se um hospedeiro porta um determinado patógeno, significa que ele tem pelo menos a disposição de desenvolver a doença causada por ele.

De forma sutil, tal afirmação pode ser criada mentalmente por causa das inúmeras doenças infecciosas que o hospedeiro torna a desenvolver – “*Se o indivíduo tem dentro de si um determinado patógeno, ele tem a doença*”. Entretanto, o hospedeiro com o tempo pode adquirir resistência ao patógeno, de tal forma que a presença dele no hospedeiro e o desenvolvimento da doença dependerá de mais de um fator, como a quantidade de patógenos dentro do hospedeiro, tempo de exposição, resposta imunológica do hospedeiro, entre outros. Isso indica que o fato do indivíduo ter a disposição em adquirir a doença não significa que ele irá desenvolvê-la, mas que pode acontecer e só ocorre no instanciamento dela.

Para tal, foi necessário o Axioma de Inclusão Geral (*General Inclusion Axiom* - GCI) para descrever tal peculiaridade.

$(H_{n1} \text{ or } H_{n2} \text{ or } \dots \text{ or } H_{nm})$  and

biotop:**physicalLocationOf** some  $(P_{n1} \text{ or } P_{n2} \text{ or } \dots \text{ or } P_{nl})$  EquivalentTo

$((\text{biotop:} \text{bearerOf}$  some  $D\_disp_1)$  or

$(\text{biotop:} \text{bearerOf}$  some  $D\_disp_2)$  or ...

$(\text{biotop:} \text{bearerOf}$  some  $D\_disp_i)$

Este axioma indica que: todo hospedeiro que tem localizado em si um ou mais patógenos, significa que ele tem a propensão a desenvolver um determinado processo patológico, indicado pelas ‘disposições’. Exemplificando, no momento que um indivíduo não é capaz de bloquear a entrada do patógeno em seu corpo, ou mesmo é incapaz de expulsar o patógeno, este é descrito como agente causador de uma doença e o organismo não apresenta barreiras à multiplicação do patógeno.

Isso implica em que um indivíduo que estando com uma, e.g. leishmania, em seu corpo poderá vir a desenvolver leishmaniose já que o organismo ofereceu algumas condições para tal. Entretanto, o organismo ainda pode oferecer resistência suficiente para o não-desenvolvimento do processo patológico, mas não para expulsão da leishmania.

A seguir, as últimas considerações sobre o modelo de transmissão.

#### 5.1.7. Axiomas de Disjunção

Por padrão, todas as classes descritas até aqui são disjuntas. Pela especificidade e complexidade do domínio descrito, é necessário que a *TBox* seja a mais fechada o possível para checagem das restrições.

A partir daqui, já é possível montar um script para a geração automática de *TBox*.

#### 5.1.8. Geração de Ontologia Baseada em Script

Após a descrição de todos os detalhes necessários para especificar o domínio da forma mais precisa possível, foi construído um simples gerador automático de *Tbox* para código OWL2 correspondente as classes presente na tabela, incluindo as relações descritas até o momento.

Para tanto, a geração da ontologia foi realizada de acordo com a seguinte sequência de passos:

- Primeiro, o conteúdo da tabela foi copiado em uma planilha no Microsoft Excel.
- As relações descritas anteriormente foram implementadas em um macro VBA (*Visual Basic for Application*), o qual transforma o conteúdo da planilha em uma ontologia em OWL2.

Tal solução foi utilizada pelo fato de ferramentas e formalismos existentes, e.g. (O'CONNOR, 2011) não poder ter sido utilizado ou facilmente adaptado pela presença de múltiplas variáveis por célula e a necessidade de adicionar sufixos (para as classes de localização geográfica e disposição) nos símbolos originais.

O próximo passo, avaliação do material gerado, é descrito no próximo tópico.

#### 5.1.9. Pós-processamento da Ontologia

Após gerar a ontologia, esta foi importada no Protégé para análise da existência de classes irmãs às quais necessitam ser hierarquicamente reestruturadas, seguindo o que é descrito na BioTop. Esse passo foi utilizado em dois casos específicos.

Por exemplo, a entrada “desconhecido” – *unknown* (na coluna dos vetores), na tabela original de Sharma & Singh (SHARMA; SINGH, 2008) – presente na tabela de vetor foi substituído pela classe geral de *Insect*, e as entradas *Leishmania* sp. foram mantidas para representar as classes de espécies de leishmanias que, de um modo geral, ainda não foram taxonomicamente descritas.

Após ter todo o conteúdo da ontologia gerado e as falhas corrigidas, foi possível partir para a avaliação da representação de transmissão, assunto do próximo tópico.

#### 5.1.10. Avaliação da Ontologia

A ontologia foi testada a partir de um conjunto de questões de competência as quais foram formuladas como consultas em DL, descritas com a *Manchester Syntax* para OWL2, e submetidas à interface de consultas em DL (*DL query*) no Protégé v4.1.

A título de exemplo, e para um melhor entendimento, um subconjunto dos dados sobre Leishmaniose (incluindo vetor, patógeno e manifestação), disponível na tabela 1 e baseado no trabalho de Sharma & Singh (SHARMA; SINGH, 2008), foi extraído aleatoriamente e utilizado (descrito na tabela 4). A subdivisão da tabela foi realizada como estratégia para aperfeiçoar a visualização dos resultados e conferência dos dados, já que a tabela completa é bastante extensa.

É importante ter em mente que os processos de transmissão a serem listados não indicam que tais patógenos não possam ser transmitidos por outros tipos de vetores, apenas indicam que até o momento, os mecanismos de transmissão identificados foram dessa natureza e com tais participantes.

Também, como cada linha da tabela indica um processo de transmissão e esses são representados com reificação de entidades, cada linha da tabela foi numerada de 1 a 4 (pelo gerador), i.e. a transmissão em Guadalupe (na primeira linha) será descrito como *PathogenTransferByVector\_1* e assim por diante até o 4 (da 4ª linha).

TABELA 4: Ontologia simples para realização de testes de raciocínio.

Localização Geográfica	Artrópode (Vetor)	Vertebrado (Hospedeiro)	Protista (Patógeno)	Manifestação (Doença)
Guadaloupe	<i>Lu. longipalpis</i>	Humano	<i>L. chagasi</i>	LV
México	<i>Lu. longipalpis</i>	Humano	<i>L. chagasi</i>	LV
	<i>Lu. olmeca olmeca</i>		<i>L. mexicana</i>	LC
			<i>L. sp</i>	LACD
Paraguai	<i>Lu. flaviscutellata</i>	Humano	<i>L. amazonensis</i>	LC
	<i>Lu. whitmani</i>		<i>L. chagasi</i>	LACD
	<i>Lu. whitmani</i>			
Peru	<i>Lu. peruensis</i>	Humano	<i>L. braziliensis</i>	LC
	<i>Lu. verrucarum</i>		<i>L. peruviana</i>	LM

A seguir, são listadas as questões de competência relacionadas à avaliação da representação de transmissão de patógenos, tomando como base a DTN Leishmaniose, e uma breve introdução sobre a forma como foi construída e buscando qual tipo de informação. Todas as questões que serão descritas foram criadas da mesma forma como se pergunta normalmente, ou seja, não levando em consideração a estrutura sintática da DL.

#### ✓ Questão de Competência 1

‘Qual patógeno pode ser transmitido por um dado vetor em uma região geográfica?’

Esta questão foi descrita objetivando descobrir se o modelo descrito até agora respeita um princípio básico: a descrição dos patógenos possíveis para cada região geográfica, seguindo a tabela 4. Adicionalmente, é possível identificar na questão um ‘pode’, partícula vocabular que insere uma possibilidade diante das respostas. Na tentativa de contornar o uso de possibilidade e pelo raciocínio empregado ser de mundo aberto, a questão foi inicialmente negada para descobrir quais entidades “não podem” ser transmitidas, obrigatoriamente.

Tal consulta ainda conta com uma restrição importante, o patógeno deve acompanhar o local de sua transmissão, indicando assim a distribuição espacial das espécies envolvidas na transmissão.

A seguir, o axioma da questão 1.

**DL Query:**

*Protist* and not (**patientIn** some (*PathogenTransferByVector*  
and (**hasGeographicLocation** some *GuadeloupeLocation*)  
and (**hasAgent** only *LutzomyiaLongipalpis*)))

Como exemplo de consulta, foram inicialmente questionados os possíveis protistas que não participam em um processo de transmissão, o qual ocorre em Guadeloupe e tem como vetor a *Lutzomyia longipalpis*.

Em seguida, os resultados da questão negada.

**Resultado:**

*LeishmaniaAmazonensis*, *LeishmaniaBraziliensis*, *LeishmaniaMexicana*,  
*LeishmaniaPeruviana*, *LeishmaniaSp*.

Como resultado, foram listados todos os patógenos que, por definição, **não** participam de um processo de transmissão por *Lutzomyia longipalpis* em Guadeloupe. Então, em poder das entidades a quais **não** satisfazem a questão, é possível encontrar a(s) entidades que são esperadas como resultado.

**Segunda Consulta:**

*Protist* and not (*LeishmaniaAmazonensis* or *LeishmaniaBraziliensis* or  
*LeishmaniaMexicana* or *LeishmaniaPeruviana* or *LeishmaniaSp*)

**2º Resultado:** *LeishmaniaChagasi*, o qual é o resultado esperado.

O resultado encontrado, a partir da negação ao grupo que não satisfaz diretamente a questão de competência 1 é o mesmo que pode ser encontrado na tabela 4 para Guadeloupe.

### ✓ Questão de Competência 2

*‘Uma doença X pode ser transmitida por um vetor Y em uma dada região geográfica Z?’*

A seguinte questão tem por objetivo determinar se uma doença que foi transmitida por um dado vetor em um local foi realmente descrito na ontologia. Isto é, determinar se as descrições limitam o raciocínio para o que está descrito no padrão de transmissão, como um teste de satisfiabilidade.

A seguir, a consulta descrita procura os processos de transmissão que ocorrem no México e tem como participante um protista que causa LMC.

#### DL Query:

```
PathogenTransferByVector
and (hasGeographicLocation some MexicoLocation)
and (hasPatient some (Protist
and (causes some MucocutaneousLeishmaniasis)))
```

**Resultado:** Insatisfatório, o que é o resultado esperado. De acordo com a tabela 4, protistas no México não causam LMC. Isto demonstra que o raciocinador inferiu o que as definições lógicas implicam.

### ✓ Questão de Competência 3

*‘Qual tipo de doença pode ser transmitida em uma dada região geográfica?’*

Tal questão visa extrair todo o conjunto de doenças que são transmitidas em uma localidade. Assim como a questão de competência 1, esta conta com a inclusão de uma possibilidade na sua formulação. Dessa forma, a heurística utilizada para a consulta 3 é similar, i.e. a consulta deve ser negada para conseguir extrair os resultados impossíveis. Por sua vez, estes são negados novamente objetivando o resultado esperado.

A seguir, é descrito quais os processos patológicos causados por protistas que participaram de um processo de transmissão no Peru.



**DL Query:**

*PathologicalProcess* and not (**causedBy** some (*Protist*  
and (**patientIn** some (*PathogenTransferByVector*  
and**hasGeographicLocation** some *PeruLocation*))))

**Resultado:** *AcuteDiffuseCutaneousLeishmaniasis, VisceralLeishmaniasis*

Como na questão de competência 1, lembrando, os resultados encontrados na consulta negada devem ser, novamente, negados para atingir o objetivo da questão. Então:

**Segunda Consulta:**

*PathologicalProcess*  
and not (*AcuteDiffuseCutaneousLeishmaniasis* or  
*VisceralLeishmaniasis*)

**2º Resultado:** *CutaneousLeishmaniasis, MucocutaneousLeishmaniasis*,  
os quais são o resultado esperado para o país Peru.

✓ **Questão de Competência 4**

*‘É possível adquirir a doença X por transmissão com vetor na região Y?’*

Esta questão visa verificar satisfibilidade, como na questão 2, evidenciando a satisfibilidade das relações descritas. Assim, nesta questão procura-se verificar se é possível adquirir uma doença a partir (apenas) da transmissão dela. Mesmo construída como possibilidade, as respostas possíveis para esta questão são: ‘sim’ ou ‘não’.

Então,

**DL Query:**

*PathogenTransferByVector\_1* and (*causes* some *VisceralLeishmaniasis*)

**Resultado:** Insatisfável, de acordo com a descrição da doença. Em um processo de transmissão, por definição não é possível se adquirir a doença diretamente da

transmissão. A classe *PathogenTransferByVector\_1* corresponde à primeira linha da tabela 4 (transmissão em Guadalupe).

### ✓ Questão de Competência 5

*‘Quais vetores podem transmitir um certo patógeno de uma doença X em uma região Y?’*

Esta questão objetiva recuperar do exemplo quais vetores são responsáveis pela transmissão de uma doença em determinado local. Como em questões anteriores, aqui há, novamente a utilização de possibilidade. Isto indica a necessidade de se utilizar dupla negação por causa do raciocínio de mundo aberto.

Para responder tal questão, a primeira consulta exemplo é relacionada à quais artrópodes não são agentes em uma transmissão, no Paraguai, que tem como participante passivo um protista que causa LC e LACD.

#### DL Query:

*Arthropod*

and not (**agentIn** some (PathogenTransferByVector  
and (**hasGeographicLocation** some ParaguayLocation)  
and (**hasPatient** some (*Protist* and **causes** some  
(*CutaneousLeishmaniasis* or  
*AcuteDiffuseCutaneousLeishmaniasis*))))))

**Resultado:** *LutzomyiaOlmecaOlmeca*, *LutzomyiaPeruensis*, *LutzomyiaVerrucarum*, *LutzomyiaWhitmani*.

Como é de se imaginar, tais resultados são contrários ao requerido. Então, devem ser negados novamente.

#### Segunda Consulta:

*Arthropod*

and not (*LutzomyiaOlmecaOlmeca* or *LutzomyiaPeruensis*  
or *LutzomyiaVerrucarum* or *LutzomyiaWhitmani*)

**2º Resultado:** *LutzomyiaFlaviscutellata* e *LutzomyiaLongipalpis*, como esperado.

✓ **Questão de Competência 6**

*‘Poderia um vetor causar diretamente uma certa doença?’*

Esta questão tem por objetivo verificar se um vetor pode causar uma determinada doença diretamente, seguindo as definições. Tal questão é construída de forma semelhante à questão 4, na qual buscou-se verificar quando uma transmissão pode causar uma doença.

**DL Query:**

*Arthropod*

and **causes** some (*AcuteDiffuseCutaneousLeishmaniasis* or  
*VisceralLeishmaniasis* or *CutaneousLeishmaniasis*  
or *MucocutaneousLeishmaniasis*)

Resultado: Insatisfatório, seguindo a descrição da doença. De acordo com interpretação dada à tabela e como as DTNs se manifestam, doenças podem ser causadas pelos protistas (patógenos) e não pelos artrópodes (vetores).

✓ **Questão de Competência 7**

*‘Qual transferência pode ter, no final, um patógeno dentro de um humano?’*

Tal questão nasceu da necessidade em verificar se, seguramente, no instanciamento do processo de transmissão, o patógeno estará necessariamente dentro do hospedeiro. O exemplo criado será com base na primeira linha da tabela. Então, foi formulada a seguinte consulta.

**DL Query:**

*PathogenTransferByVector*

and not ((**hasGeographicLocation** some *GuadeloupeLocation*)  
and (**hasPatient** some (*LeishmaniaChagasi*  
and (**physicallyLocatedIn** only *Human*))))

Tal consulta tenta verificar em quais processos de transmissão que não ocorrem em Guadalupe e que tem como patógeno a *Leishmania chagasi* localizado em um humano. O resultado foi:

**Resultado:** *PathogenTransferByVector\_2*, *PathogenTransferByVector\_3* e *PathogenTransferByVector\_4*.

E então a segunda consulta para achar o resultado esperado.

### Segunda Consulta:

*PathogenTransferByVector*

and not (*PathogenTransferByVector\_2* or *PathogenTransferByVector\_3*  
or *PathogenTransferByVector\_4*)

**2º Resultado:** *PathogenTransferByVector\_1*, que compete à transferência que ocorre em Guadalupe e a primeira linha da tabela 4.

#### 5.1.11. Questões relacionadas ao desempenho

O teste de satisfiabilidade (para a questão de competência 2) levou menos de 1 segundo em um computador com processador Intel Core i7 820QM, 8GB de memória e em um Sistema Operacional Windows 7 64bits. Aumentando artificialmente o tamanho da ontologia de leishmaniose em um fator 10, o mesmo teste levou cerca de 14 minutos, uma vez que o raciocínio para a DL  $\mathcal{SHOIQ}^+$  é realizado em tempo não-determinístico exponencial (NExpTime-completo) (MOTIK *et al.*, 2009). Isto aparenta ser o maior obstáculo para o uso de consultas em DL em ontologias grandes e expressivas.

O desempenho no raciocínio é uma questão conhecida quando uma DL expressiva (disjunção, relações inversas e negações) é utilizada, como no caso da NTDO com atualmente 146 classes, 62 axiomas de equivalência, 175 axiomas de subclasse, 9 axiomas de disjunção e 23 inclusões de classes gerais (GCI – *General Class Inclusion*), utilizando a expressividade  $\mathcal{SHIQ}(\mathcal{ALC} + \mathcal{H} + \mathcal{I} + \mathcal{Q})$ .

No próximo tópico, será apresentada uma representação sobre mortalidade, a qual foca nos processos os quais ocorrem em uma fase posterior à infecção por um patógeno. Esta fase posterior acontece em decorrência de complicações na doença e que podem culminar com o falecimento de indivíduos.

## 5.2. MODELO DE INJÚRIAS E MORTALIDADE POR CAUSAS ESPECÍFICAS

Neste tópico, são descritos os resultados encontrados relacionados à proposição de um modelo de mortalidade, modelo este que se apresenta causalmente relacionado a processos anteriores e que podem levar ao falecimento de um dado indivíduo. Tal representação mostra-se necessária para a descrição da completude dos processos que ocorrem desde a transmissão de um patógeno, com acometimento de um indivíduo por uma doença e o desenrolar dela com um possível falecimento. Além disso, evidenciará vários pontos complexos de modelagem.

### 5.2.1. Representando Injúrias e Mortalidade

A representação de mortalidade a qual será apresentada no presente trabalho é baseada no ciclo de vida descrito abaixo (Fig.8)

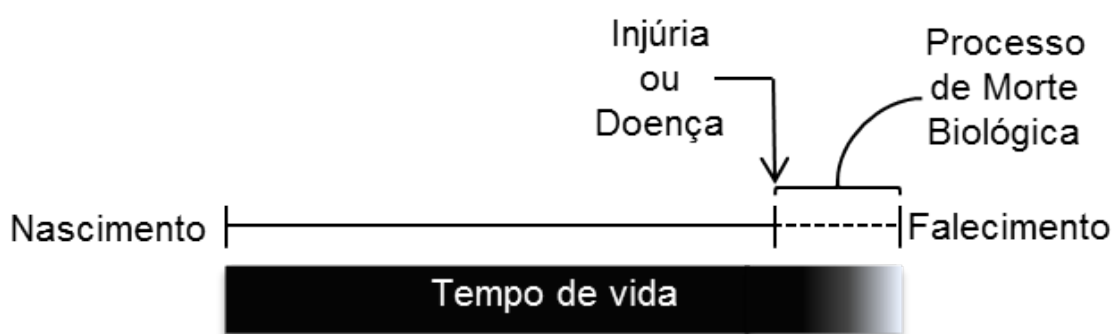


FIGURA 8: Ciclo de Vida de um organismo. Nesta figura estão incluídos os principais processos e os pontos que os mesmos ocorrem.

Tomando o nascimento como o ponto inicial, a vida de um organismo individual continua até sua morte. Entretanto, o tempo de vida pode se sobrepor, no final, com um ou mais processo de morte biológica os quais irão levar ao falecimento. Esta definição é baseada no fato de vários fatores possibilitarem, simultaneamente, a influência sobre o tempo de vida de um organismo, provocando sua morte, como um acidente ou uma doença, ou ambos, e.g. quando acontece um ataque cardíaco em um condutor de veículo decorrendo em um acidente com vítima fatal.

Não sendo apenas uma simplificação introduzida para facilitar a manipulação de dados estatísticos, vários registros de mortalidade possibilitam capturar e guardar apenas uma causa básica de morte. Na presente ontologia, iremos seguir esta

simplificação, pois estender a representação para acomodar mais de uma causa elevaria em muito a complexidade da ontologia. Entretanto, não traria custos computacionais adicionais para a criação de consultas.

Ainda sobre a interpretação da figura 8, em um dado momento, um organismo individual pode adquirir certa doença, e.g. Febre do Dengue, a qual pode causar falecimento prematuro, dependendo do tipo da doença, co-morbididades e condições de saúde. O ponto final da existência de um indivíduo é usualmente relacionado a uma ou mais causas, as quais são determinadas por investigação médica. Em termos clínicos, uma causa de doença é definida em função do estado fisiológico de um indivíduo; isso pode ser relacionado a processos patológicos anteriores, registrados ou não.

Assumindo que há dados disponíveis sobre falecimentos, seria possível descrever e traçar a sequência de causas induzidas e os momentos sobrepostos os quais processos patológicos afetaram a vida do organismo, levando de uma condição de nascido para morto. Processos patológicos irão causar danos ao estado fisiológico do organismo até tal ponto que eles iniciam o processo de morte fisiológica, levando ao falecimento. Esta sequência de processos pode ser evidenciada por registros de mortalidade de indivíduos, quando a causa da morte foi previamente registrada em um sistema de informação de mortalidade, *i.e.* a causa primária da morte já era conhecida.

Para embasar o presente trabalho, a definição de vida utilizada (necessária, mas não necessária e suficiente), foi baseada no estudo de Koshland (KOSHLAND, 2002), o qual identificou as principais características como sendo:

- a) A capacidade inerente de desenvolvimento de acordo com um programa (como descrito no DNA) em condições normais de vida;
- b) A realização de interações cinéticas (como o genoma e o metabolismo);
- c) Capacidade de mutar e possibilitar a seleção natural;

- d) O portador de níveis de sobrevivência é um corpo compartimentalizado em células e/ou órgãos, o qual metaboliza substâncias para gerar energia para adaptação, regeneração e segregação.

Já a definição de 'Nascimento' é baseado na descrição de "nascido vivo", disponibilizada pelo Instituto Brasileiro de Geografia e Estatística (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE), 2011). Isto corresponde à completa expulsão ou extração de um produto gerado pelo corpo materno após a concepção, o qual após ser separado do corpo materno, respire ou exiba outro sinal vital, e.g. batimento cardíaco, contração muscular voluntária, contração do cordão umbilical (sem considerar o cordão cortado ou não), e que a placenta tenha sido ou não expelida. De maneira contrária, "morte" significa um estado de ausência de funções cerebrais e cessação de todas as funções biológicas, inerentes ao corpo humano (MILLER; TRUOG, 2010).

Entretanto, há algumas dificuldades relacionadas a uma representação acurada sobre os processos que levam um indivíduo à morte:

- Complexidade é um ponto, tal qual a natureza causal, o qual pode ser muitas vezes indireto quando há muitos fatores desconhecidos como comorbidades e outras influências as quais convergem em um processo de morte;
- Outro ponto é relacionado a sequencia de processos e o tempo, com uma descrição precisa de 'quando' cada processo acontece, quando começou e como identificar os limites (início e fim).

Não obstante, esta exata informação é, provavelmente, não tão importante se o objetivo do modelo é voltado para dados de mortalidade. Ao invés disso, o que é usualmente conhecido e achado em bases de dados é o conhecimento do que é a sequência de sinais e sintomas típicos de uma doença, por causa das restrições de tempo envolvidas, e.g. durante a tuberculose, a tosse com secreção é seguida de uma infecção pulmonar, podendo ser checado em notificações de morbidade e mortalidade. Para descrever um registro de morte em um banco de dados de



notificação de mortalidade, viz. Sistema de Informação de Mortalidade (SIM) do Ministério da Saúde (MS/Brasil), um médico certifica a causa básica da morte e outras secundárias. A ontologia deve suportar ambas as descrições.

Também, no presente trabalho, é assumida a noção de “processos instantâneos” disponíveis na BioTop como sendo equivalente à ‘evento’, descrito na *General Formal Ontology* (GFO). Isto faz o paciente indica um comportamento que é relacionado, casualmente ou não, a outros processos (HERRE *et al.*, 2007).

A seguir, são apresentados os principais desafios relacionados à representação de entidades de domínio e os axiomas lógicos os quais caracterizam e resolvem tais problemas. Para tal, a representação é dividida em 4 versões que demonstram a abordagem de otimização iterativa e a evolução do modelo para uma proposta final.

#### **5.2.2. Desafios Representacionais do Modelo de Mortalidade**

Vários desafios foram encontrados durante a criação de uma representação coerente para o processo de morte que permitisse preservar a identidade de indivíduos pela adição de restrições de cardinalidades. Além disso, possibilitasse representar a ontologia resultante em uma DL decidível, também não deixando maiores dificuldades para os raciocinadores que irão realizar consultas de usuários. Cada um desses itens é discutido em versões consecutivas até um modelo satisfatório.

A seguir, o primeiro ciclo de iteração para a construção do modelo.

*Versão1*

Inicialmente, uma definição ingênua sobre morte poderia ser:

*DeathEvent* equivalentTo *Event*

and (**hasLocus** some *GeographicLocation*)

and (**hasPatient** some *DeadOrganism*)

(1)

and (**hasProcessualPart** some *BiologicalDeathProcess*)

and (**hasInstant** some *PointInTime*)

indicando que um processo de morte é um processo instantâneo que ocorre em, pelo menos, uma localidade e na qual um organismo morto é um participante. Este axioma também descreve um ou mais processos de morte biológica como parte do processo de morte, e que ocorre em determinado momento no tempo.

Esta definição carrega várias imprecisões do tipo: como preservar a identidade entre o organismo que está vivo e o que morreu, de forma que o organismo vivo não é especificado na descrição axiomática. De acordo com a definição de classe acima, não há garantias que um corpo vivo e um morto sejam idênticos, a partir do momento que os participantes de um *DeathProcess* (classe que indica processo de morte) e de um *BiologicalDeathProcess* (classe que indica processo de morte biológica) podem não ser os mesmos.

Além disso, o axioma não expressa nenhuma restrição de cardinalidade. Isto permite que sejam criadas diferentes interpretações, como a possibilidade de haver mais do que um indivíduo morrendo no mesmo processo de morte.

Simultaneamente, seguir a ideia de que um organismo vivo é eventualmente transformado em um morto causa problemas representacionais adicionais. Primeiro, o alto nível que a seguinte representação importa (no caso a BioTop) tem os organismos restritos para apenas os vivos, necessitando expressões adicionais em classes para referir a organismos mortos (e.g. utilizando a relação **transformationOf**).

Como consequência, um humano morto não é mais um humano, embora possua órgãos humanos, características, entre outros. Além disso, perder a “humanidade” implica na perda de identidade, desde que qualquer classificação de seres vivos seja rígida, *i.e.* uma vez que um indivíduo é uma instância de uma classe rígida, ela deixa de existir como instância quando o ser deixar de existir (GUARINO; WELTY, 2000b). Mesmo que seja assumido que esta descrição seja a de um *phased sortal* (GUARINO; WELTY, 2000b), *i.e.* entidades que mudam de fase (de ‘vivo’ para ‘morto’), não fica claro até quando a identidade será preservada. Em outras palavras, até onde as cinzas de um organismo morto podem ser identificadas como o a pessoa que faleceu e ambos sendo chamados de Hans?

### Versão 2

Uma solução para contornar os desafios representacionais apresentados na versão 1 é não representar entidades que causam confusão, *viz.* *DeadOrganism*. Esta, de fato, não apresenta utilidade para a maioria das aplicações voltadas para a área da saúde. Uma solução, então, consiste em representar organismos vivos baseado na sua existência temporal, limitado por dois pontos no tempo como descrito pela GFO (HERRE *et al.*, 2007). Tal representação emprega a definição de *gfo:Chronoid*, *i.e.* um intervalo não definido como um conjunto de pontos, implicando no tempo representado como um *continuum*, sendo equivalente ao biotop: *TimeInterval*.

biotop: *TimeInterval* equivalentTo *gfo:Chronoid* (2)

Todo *Chronoid* (lê-se cronóide) apresenta dois limites externos, conhecidos como limites de tempo (*gfo:TimeBoundary*) ou pontos no tempo (biotop: *PointInTime*). Na GFO, há dois tipos de limites temporais, indicando os limites esquerdos (início) e direito (fim) de um intervalo de tempo, *i.e.* *gfo:LeftTimeBoundary* e *gfo:RightTimeBoundary*. Por definição, ambos não podem assumir valores semelhantes em um mesmo cronóide (HERRE *et al.*, 2007). Uma representação esquemática pode ser encontrada a seguir (Fig. 9).

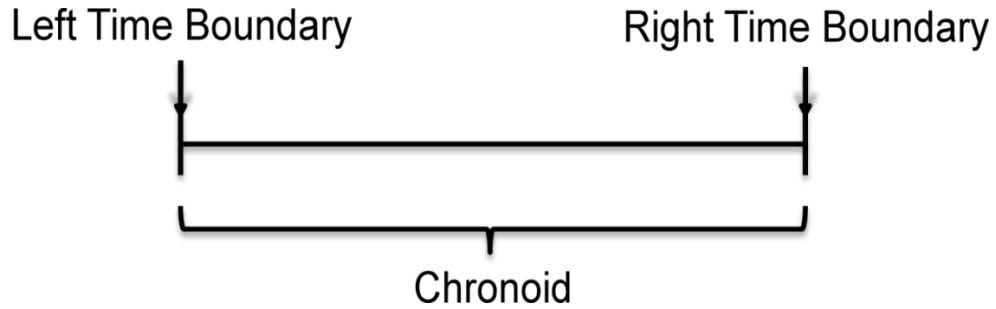


FIGURA 9: O intervalo de tempo descrito para um *gfo:Chronoid* e seus limites (*gfo:LeftTimeBoundary* e *gfo:RightTimeBoundary*) no eixo temporal.

Para tornar mais claro, abaixo seguem as definições de *Chronoid* e seus limites temporais de acordo com a GFO:

*gfo:Chronoid* subclassOf

(*gfo:hasLeftTimeBoundary* exactly 1 *gfo:LeftTimeBoundary*) and (3)

(*gfo:hasRightTimeBoundary* exactly 1 *gfo:RightTimeBoundary*)

e os limites temporais direito e esquerdo são

*gfo:LeftTimeBoundary* equivalentTo *gfo:TimeBoundary* (4)

and (*gfo:leftTimeBoundaryOf* some *gfo:Chronoid*)

*gfo:LeftTimeBoundary* equivalentTo *gfo:TimeBoundary* (5)

and (*gfo:leftTimeBoundaryOf* some *gfo:Chronoid*)

Quando há cronóides em sequência, o limite temporal direito de um processo antecedente deve ser contíguo com o limite temporal esquerdo do subsequente, i.e. um representa o fim e o outro o início, respectivamente. Vale ressaltar ainda que a soma mereológica de *gfo:Chronoids* representa a noção de região temporal (HERRE *et al.*, 2007).

Seguindo a idéia da GFO, entidades que existem inteiramente em um intervalo no tempo são chamadas de *gfo:Presential* (HERRE *et al.*, 2007). De tal forma, a classe *gfo:Presential* engloba entidades materiais que são representados na BioTop como

biotop:*MaterialObject*. Assim, a primeira pode ser considerada como superclasse da segunda.

Seguindo tal premissa, o axioma abaixo deve ser incluído

biotop:*MaterialObject* subClassOf *Presential* (6)

biotop:*LivingOrganism* subClassOf (gfo:**exists\_at** exactly 1 gfo:*TimeBoundary*) (7)

determinando que um organismo vivo existe apenas em um intervalo de tempo.

Por último, entidades processuais são projetadas (gfo:**projectsTo**) em um cronóide (HERRE *et al.*, 2007). Seguindo a mesma ideia de mapeamento entre a GFO e a BioTop, a classe gfo:*Process* pode ser mapeada para a classe biotop:*Process*, já que englobam entidades da mesma natureza. Na tabela abaixo, há uma breve descrição dos principais mapeamentos criados entre a BioTop e a GFO.

TABELA 5: Mapeamentos entre as entidades presentes na BioTop e GFO

Biotop	Tipo de Mapeamento	GFO
<i>MaterialObject</i>	<b>subclassOf</b>	<i>Presential</i>
<i>TimeInterval</i>	<b>equivalentTo</b>	<i>Chronoid</i>
<i>PointInTime</i>	<b>equivalentTo</b>	<i>TimeBoundary</i>
<i>Process</i>	<b>equivalentTo</b>	<i>Process</i>

Finalmente, o *DeathProcess* pode ser modificado para trocar *DeadOrganism* por um *LivingOrganism*, como segue:

*DeathProcess* equivalentTo *InstantaneousProcess*  
 and (**hasLocus** some *GeographicLocation*)  
 and (**hasPatient** some *LivingOrganism*) (8)  
 and (**processualPartOf** some *BiologicalDeathProcess*)  
 and (**hasnstant** some *PointInTime*)

De um lado, os problemas ontológicos relacionados à existência de um *DeadOrganism* estão resolvidos, incluindo o problema de identidade, as instâncias de *LivingOrganism* são formadas em um certo momento no tempo (*gfo:LeftTimeBoundary*) e destruídas em outro (*gfo:RightTimeBoundary*). Do outro lado, por definição a relação *biotop:hasPatient* permite mais de um elemento no conjunto imagem, o que pode levar a interpretações errôneas de que um processo de morte por injúria ou doença acontece para várias pessoas simultaneamente.

Ainda, contém alguns problemas de indentidade:

- a) Um que relaciona os participantes de *DeathProcess* e *BiologicalDeathProcess*;
- b) O conjunto de definições descritas até aqui não incluem o momento da morte ou o fazem igual ao final de um *BiologicalDeathProcess*; e
- c) O mesmo se aplica ao *LivingOrganism*, pelo qual o *RightTimeBoundary* deve coincidir com ambos *DeathProcess* e *BiologicalDeathProcess* resultantes.

De fato, *DeathProcess* é exatamente a última parte (em termos temporais) de um *BiologicalDeathProcess*; isto também é uma questão de coerência desde que o oposto (um *BiologicalDeathProcess* sendo parte de um *DeathProcess*) significa que um processo instantâneo deveria ter como parte um processo projetado em um intervalo no tempo.

### Versão 3

O ultimo problema mencionado na versão 2 pode ser resolvido apenas com a representação de processos instantâneos. Para este propósito, a classe *biotop:InstantaneousProcess* foi utilizada, sendo um processo que ocorre no final de outro processo precedente, de forma a possibilitar uma sequência de processos, conectando o final de um processo com o início do próximo utilizando o operador de concordância ( $\doteq$ ) da DL. Este operador é utilizado em cadeias de propriedades para indicar que as instâncias a serem descritas são conectadas.

É importante notar a diferença entre  $\doteq$  e  $=$ . O primeiro representa a coincidência no valor de duas propriedades, ou, em outras palavras, uma referência a um mesmo objeto (BAADER *et al.*, 2003). O segundo define a regra de formação de uma propriedade, a qual é usualmente baseada em cadeias de propriedades (BAADER *et al.*, 2003). Na presente ontologia, há a necessidade de estabelecer, por exemplo, que certo processo termina exatamente quando outro começa; isto é denotado por uma concordância.

Nesse momento, é necessário incluir a definição de instante para designar o momento que a morte acontece. A condição “instantânea” é definida como um exato ponto no tempo (ntdo:**hasInstant**). Isto pode ser alcançado por fazer instantâneo um processo que ocorre apenas no limite direito de um processo precedente, como descrito abaixo.

ntdo:**hasInstant** = gfo:**projectsTo** o gfo:**hasRightTimeBoundary** (9)

Um ponto importante que precisa ser levantado é a relação entre injúrias (*InjuryProcess*) e *DeathProcess*. Para a descrição de um processo de injúria, é necessário determinar onde isso aconteceu (ntdo:**hasGeographicLocation**), sua causa e o paciente que sofreu a injúria. As causas de injúrias são descritas aqui como sendo exclusivamente causadas por processos de origem não-biológica. Tudo isso está descrito no axioma abaixo:

ntdo:*InjuryProcess* equivalentTo biotop:*InstantaneousProcess*  
 and (ntdo:**hasGeographicLocation** some ntdo:*GeographicLocation*)  
 and (biotop:**causedBy** some (biotop:*ProcessualEntity*  
 and (not biotop:*BiologicalProcessualEntity*)))  
 and (ntdo:**hasInjuredPatient** some biotop:*LivingOrganism*) (10)

Mesmo não sendo o foco, que é sobre mortes causadas por doenças, é importante distinguir processos patológicos, estruturas e disposições (SCHULZ *et al.*, 2011). Desordens são causadas por acidentes, uma lesão ou uma fratura e pode levar à doença. Adicionalmente, desordens seguem injúrias.

A partir do que foi exposto, é possível introduzir uma nova definição de *DeathProcess*:

$\text{ntdo:DeathProcess equivalentTo biotop:InstantaneousProcess}$   
 $\text{and (biotop:hasLocus some ntdo:GeographicLocation)}$   
 $\text{and (ntdo:hasPatient some biotop:LivingOrganism)}$  (11)  
 $\text{and (ntdo:precededBy some biotop:BiologicalDeathProcess)}$   
 $\text{and (ntdo:hasInstant } \doteq \text{ ntdo:precededBy o gfo:hasRightTimeBoundary)}$   
 $\text{and (ntdo:hasDeathPrimaryCause exactly 1 biotop:ProcessualEntity)}$   
 $\text{and (ntdo:hasPatient } \doteq \text{ ntdo:precededBy o ntdo:hasPatient)}$

Este axioma descreve onde a morte aconteceu, qual organismo que faleceu é o participante, e quais processos são a causa primária da morte. As condições de concordância são as mais importantes, pois elas asseguram que a morte ocorre exatamente quando o processo de morte biológica termina (junto com o axioma 8). Ainda, determina que a pessoa falecida é a mesma que participou de um processo que levou a morte, mantendo a identidade do paciente (última condição).

Finalmente, para completar a representação ontológica de mortalidade, a classe *ntdo:BiologicalDeathProcess* foi criada para indicar a existência de um agregado (união de processos ocorrendo em paralelo) de processos não-completamente conhecidos que ocorrem para o organismo que está morrendo e corroboram para a morte. Um processo de morte biológica (da doença para a morte) é uma entidade processual biológica (*biotop:BiologicalProcessualEntity*) a qual é causada por uma injúria (não-biológica) ou um outro processo biológico (mas, claramente não o próprio processo de morte biológica). Tem como participante um organismo convalescente e sua duração é delimitada



$\text{ntdo:BiologicalDeathProcess} \text{ equivalentTo } \text{biotop:BiologicalProcessualEntity}$   
 and (biotop:**causedBy** some  
     ( $\text{ntdo:InjuryProcess}$  or  $\text{biotop:PathologicalProcess}$ )) (12)  
 and (ntdo:**hasPatient** some  $\text{biotop:LivingOrganism}$ )  
 and (gfo:**projectsTo** exactly 1  $\text{gfo:Chronoid}$ )

Este axioma leva em consideração os processos que ocorrem antes do processo de morte e depois de uma injúria ou uma doença. Para a representação de participantes (também descrito em *DeathProcess*), é necessário identificar a existência de um ou mais processos, mesmo sendo imperceptíveis ou indiretamente relacionados.

De um ponto de vista epidemiológico, isto só pode ser completamente definido *a posteriori*, desde que uma causa anterior (doença/injúria) pode apenas ser ligada à causa de morte principal na análise *post mortem* (pela autópsia, por exemplo). Geralmente, isto é identificado na afirmação sobre a causa da morte por um médico o qual estava cuidando do paciente até o momento do falecimento. Na presente ontologia, a partir dos axiomas até aqui descritos, é possível estabelecer a seguinte sequência de fatos para um organismo: doença/injúria → processo de morte biológica → morte.

Os axiomas formulados no presente trabalho mencionam apenas relações causais (e.g. *InjuryProcess* ou *DeathProcess*). Entretanto, a noção de causalidade a qual é necessária para a representação é baseada no observador do processo, *i.e.* o médico que certificou a causa da morte ou identificou a doença. Tomando como exemplo um registro de morte em um banco de dados de notificação de mortalidade, *viz.* SIM, um médico certifica a causa básica da morte e, algumas vezes, secundária, terciária, e assim por diante.

Na NTDO, este fato é suportado pelo *ntdo:BiologicalDeathProcess* desde que este permita a inclusão de mais de uma causa. Adicionalmente, pode ser estendido em *ntdo:DeathProcess*, desde que está sendo levado em consideração (aqui) apenas a causa primária, a causa definidora da morte (e que pode ainda não ser a causa real).

O modelo apresentado resolve o problema de identidade; ainda assim, um problema escondido, não relacionado com a representação, mas com o raciocínio continua: se concordâncias não forem construídas sobre cadeias de propriedades funcionais, a inferência torna-se indecidível (SCHMIDT-SCHAUSS, 1989).

Outro aspecto a ser considerado é que os processos de morte biológica podem ocorrer por causa de injúrias e causas desconhecidas, sem considerar as doenças.

#### Versão 4

O problema de indecidibilidade mencionado acima é relacionado à cardinalidade da relação *biotop:hasPatient*. Para os propósitos do presente trabalho, esta relação deve ser funcional, *i.e.* cada elemento do conjunto domínio deve ser mapeado para apenas um elemento do conjunto imagem. Em cada processo de morte, apenas um ser humano é o participante. Infelizmente, a BioTop não define a relação **hasPatient** como funcional. Por causa disso e para suportar este requisito, foram criadas as seguintes sub propriedades, todas funcionais:

Functional (**hasDeathPatient**, **hasConvalescentPatient**, **hasInjuredPatient**) (13)

para indicar que um processo de morte ou de injúria admite apenas uma instância de participante (*i.e.* *Human*) (Fig. 10). Por exemplo, a relação funcional *ntdo:hasInjuredPatient* é somente empregada em notificações de assistência a saúde, desde que se refere exclusivamente a injúrias sobre uma única pessoa. A propriedade *ntdo:hasConvalescentPatient* é somente empregada em *BiologicalDeathProcess*, ao passo que, analogamente, *ntdo:hasDeathPatient* é utilizada nas definições de *DeathProcess*, como pode ser encontrado a seguir.

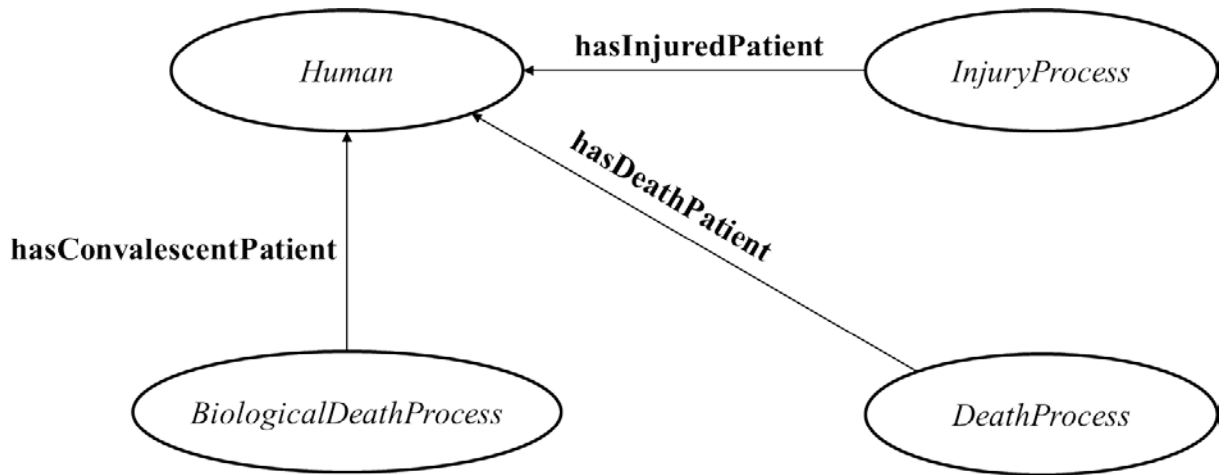


FIGURA 10: Modelo Gráfico de um evento de um *DeathProcess*, *InjuryProcess* e *BiologicalDeathProcess* e suas relações de participação (**hasDeathPatient**, **hasInjuredPatient** e **hasConvalescentPatient**), de acordo com os participantes dos processos (e.g. *Human*).

ntdo:*DeathProcess* equivalentTo biotop:*InstantaneousProcess*  
 and (biotop:**hasLocus** some ntdo:*GeographicLocation*)  
 and (ntdo:**hasDeathPatient** exactly 1 biotop:*LivingOrganism*)  
 and (ntdo:**precededBy** some biotop:*BiologicalDeathProcess*) (14)  
 and (ntdo:**hasInstant**  $\doteq$  ntdo:**precededBy** o gfo:**hasRightTimeBoundary**)  
 and (ntdo:**hasDeathPrimaryCause** exactly 1 biotop:*ProcessualEntity*)  
 and (ntdo:**hasDeathPatient**  $\doteq$  ntdo:**precededBy** o  
 ntdo:**hasConvalescentPatient**)

Esta definição tem a vantagem de enfatizar, explicitamente, o fato de que o participantes do processo de morte coincide com o participante convalescente do *BiologicalDeathProcess*.

Para um melhor entendimento, um modelo esquemático evidenciando as principais classes e relações apresentadas nos axiomas para a representação de morte está descrita na figura 11. Esta descreve o papel transicional de um biotop:*LivingOrganism*, da vida para a morte. Ainda, mostra alguns acordos requeridos para expressar a sequência temporal dos processos.

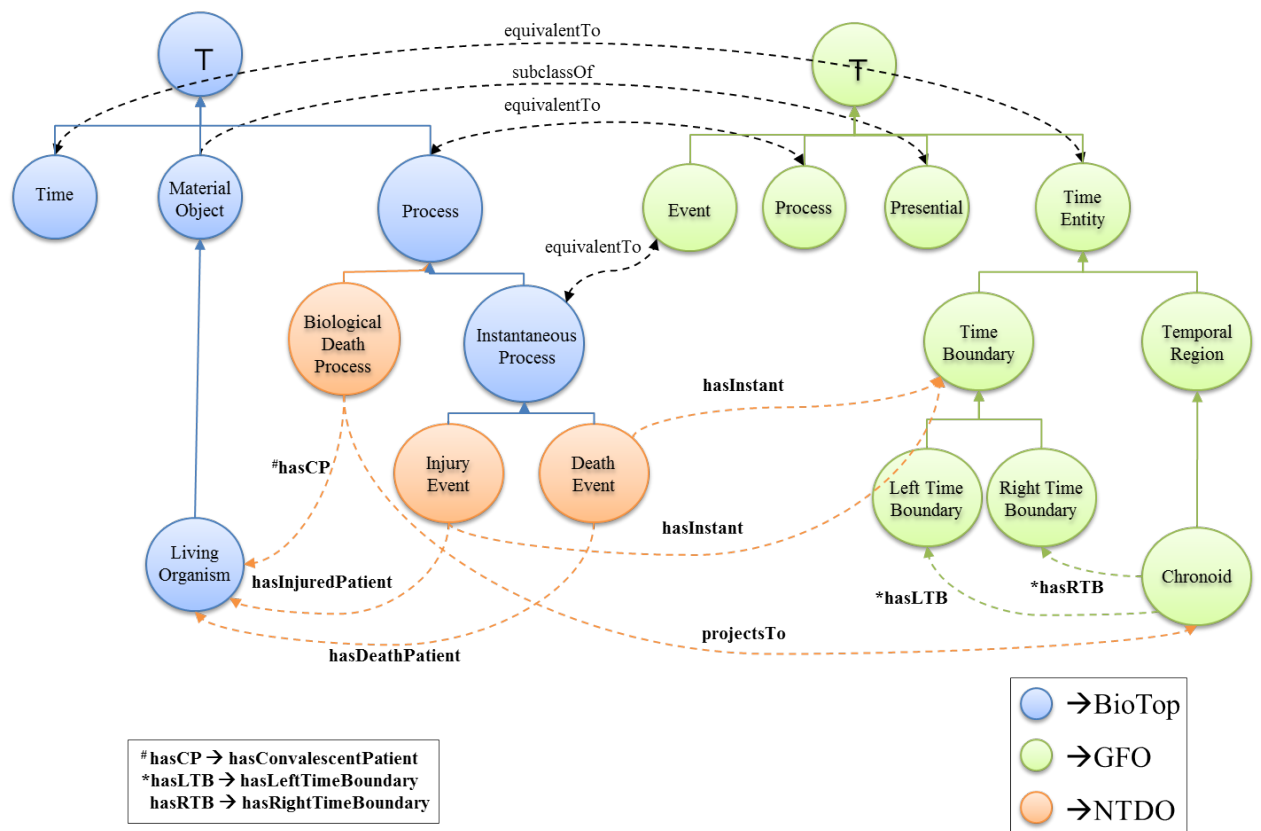


FIGURA 11: Conexões entre a NTDO, BioTop e GFO. Várias classes da NTDO são subclasses da BioTop, ao passo que várias classes da GFO foram importadas para a NTDO.

### 5.3. ESTRUTURA DE IMPORTAÇÃO E HIERARQUIA COMPLETA

A estrutura de importação definida para a NTDO está descrita na figura 12.

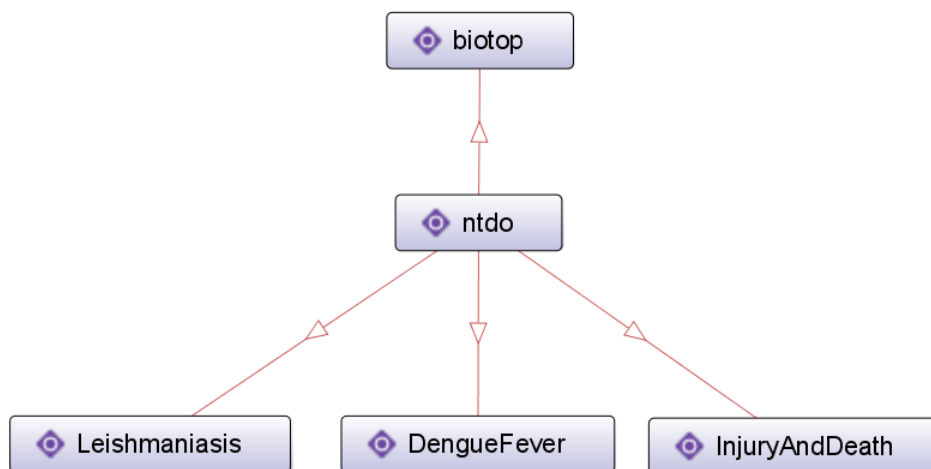


FIGURA 12: Modelo de Exportação da NTDO

A figura 12 apresenta a estrutura de importação da NTDO, com imagem gerada por meio do *plugin* nativo do Protégé v4, OntoGraph. A composição da NTDO, atualmante, é de 3 subconjuntos de ontologias, 2 representando doenças (Leishmaniose e Febre do Dengue) e 1 representando os eventos de injúria e mortalidade (InjuryAndDeath). Tal estrutura foi preservada para facilitar a importação e reutilização de subpartes por outros estudos ou aplicações específicas (Rector, 2003), sem necessariamente ter que recortar o código ou copiar individualmente cada entidade requerida em um novo arquivo OWL.

As classes referentes a representações de Processos (Fig. 13); Objetos Materiais e Imateriais (Fig. 14); e, Disposições e Tempo (Fig 15), estão disponíveis a seguir.

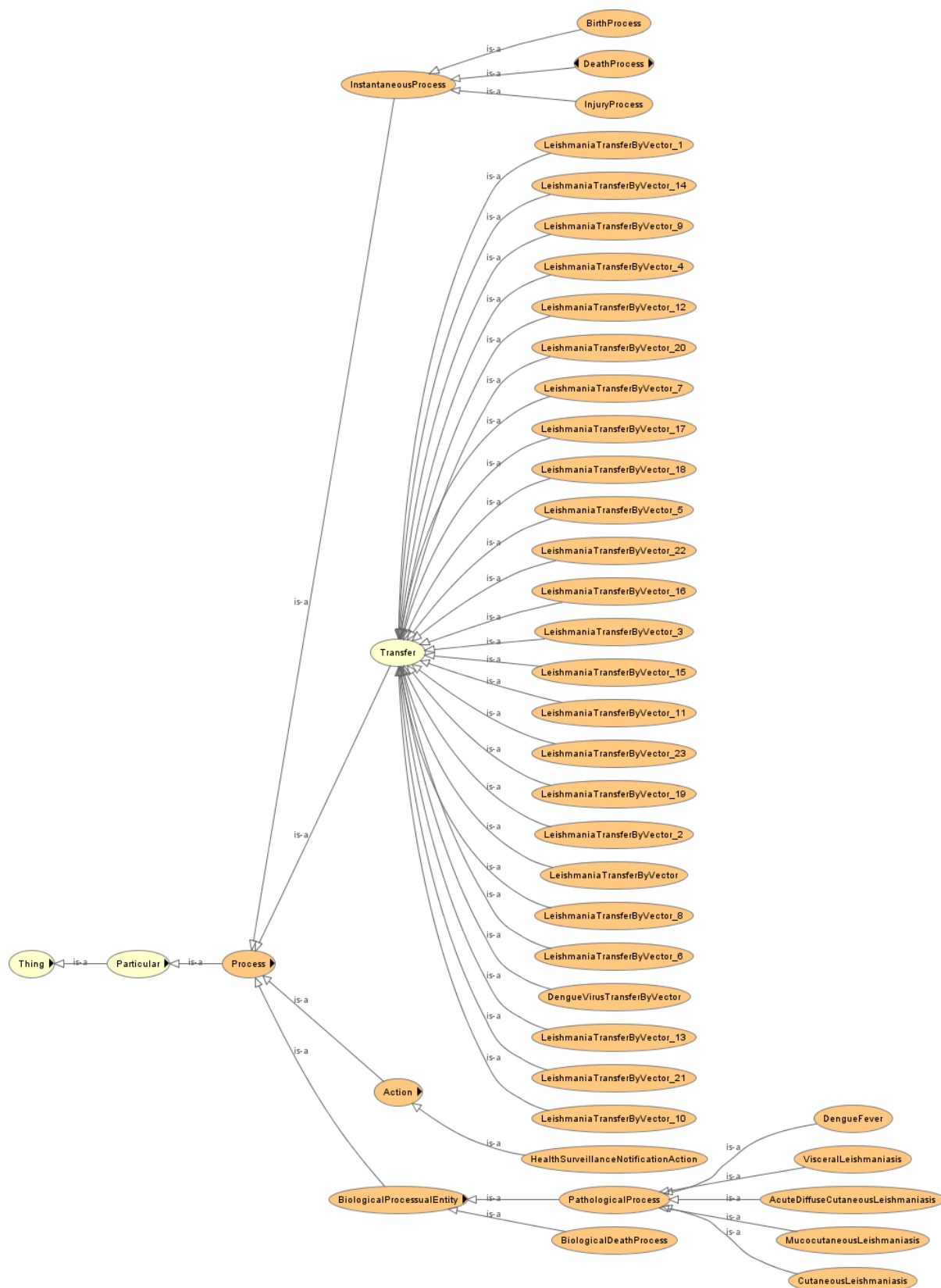


FIGURA 13: Processos na NTDO.

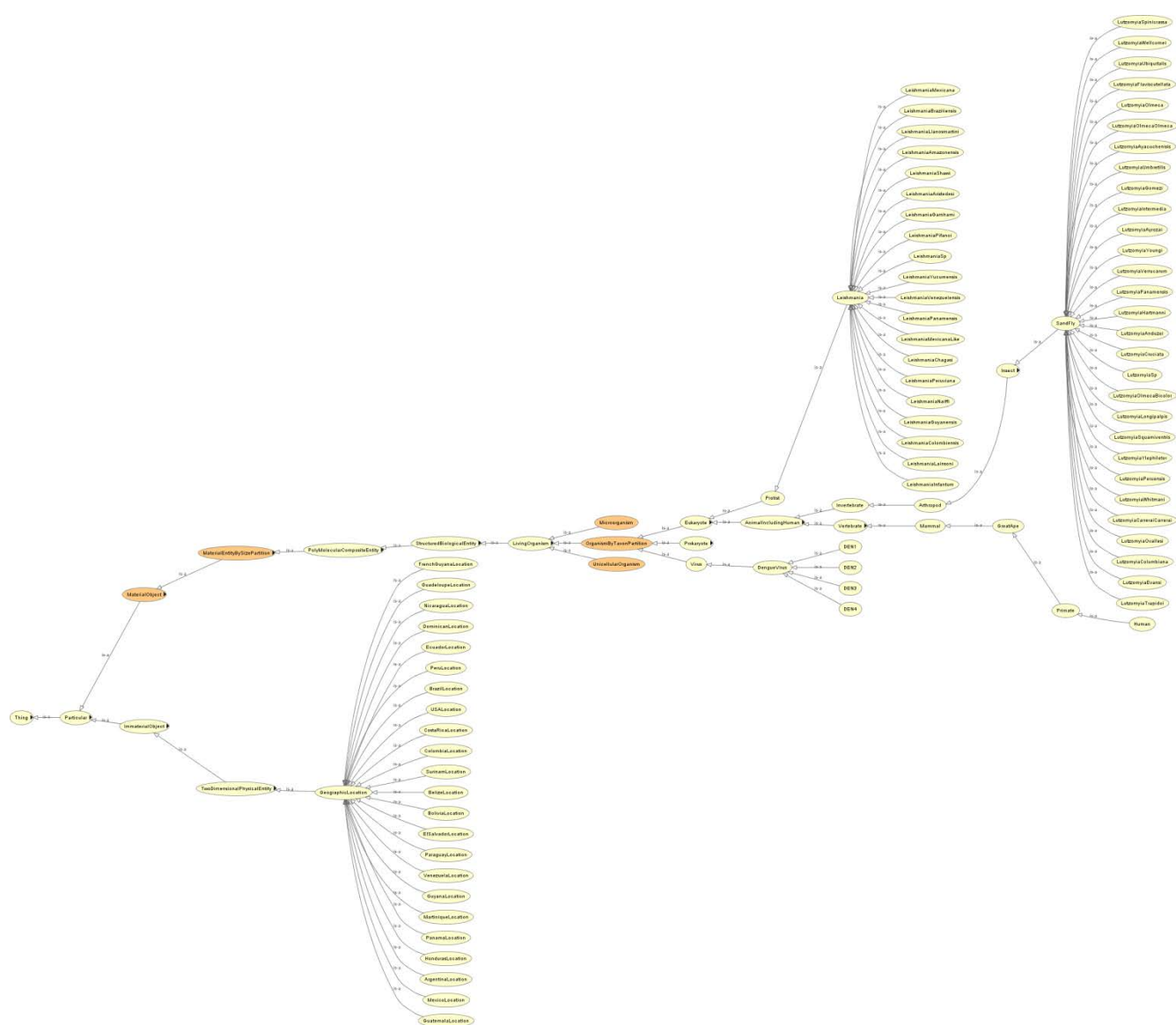


FIGURA 14: Objetos Materiais e Imateriais na NTDO

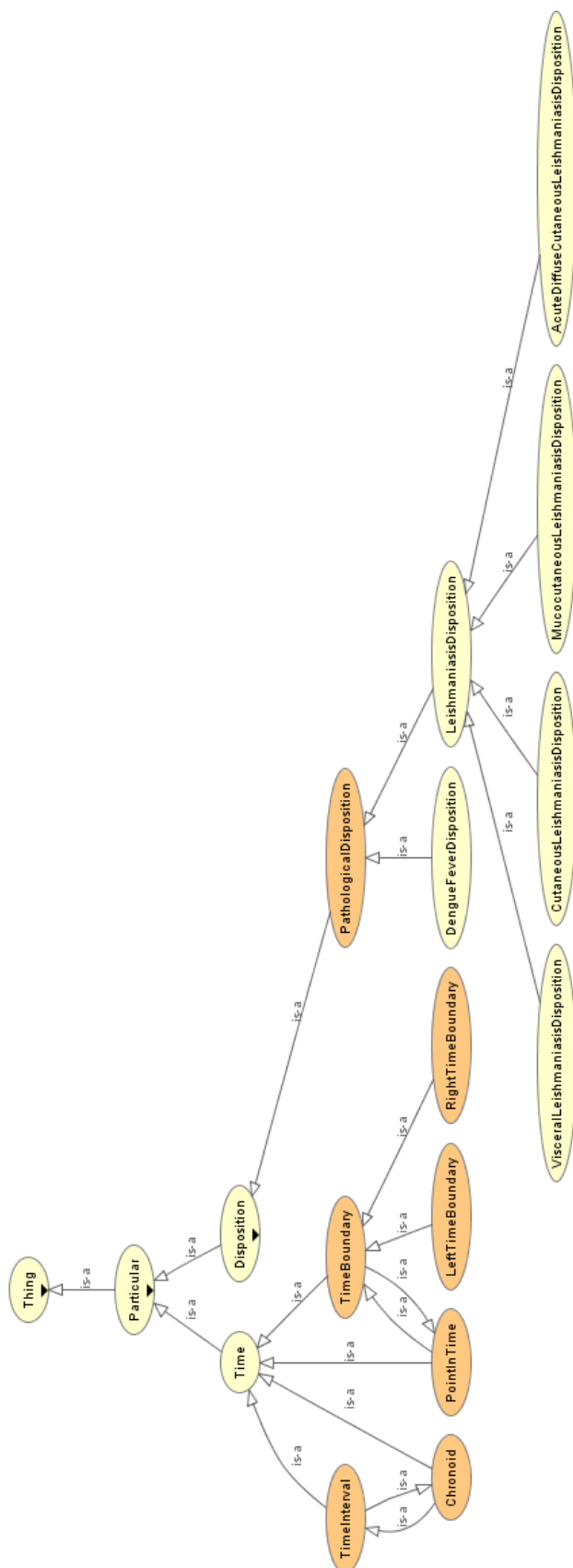


FIGURA 15: Disposições e Tempo



## 6. *DISCUSSÃO*

Neste capítulo, são apresentadas as considerações e os trabalhos relacionados à NTDO. Além disso, são evidenciados os principais achados e no que eles implicam, o que pode ser desenvolvido a partir deles e algumas limitações do presente trabalho.

### 6.1. CONSIDERAÇÕES SOBRE O PPO

O presente trabalho ilustrou e seguiu bases ontológicas fundamentadas pela imposição de distinções categóricas estritas. O padrão formal aqui apresentado para DTN transmitidas por vetor, em particular, reutiliza um conjunto rígido, do ponto de vista filosófico, de relações encontradas na ontologia *core* BioTop.

O presente estudo, indiretamente, pode demonstrar a utilidade e a expressiva superioridade da utilização de consultas em DL ao invés de simples consultas baseadas em SPARQL. Uma vantagem das consultas em DL reside na ausência de variáveis e um modo mais natural de raciocinar sobre hierarquias taxonômicas. Adicionalmente, apresentou-se um PPO de como se construir um modelo

representacional fundamentado, evidenciando os limites da linguagem encontrada e a tentativa de incorporar mais construtores, ou melhor utilizá-los, para uma representação axiomática mais rica.

Para a representação de entidades geográficas, é recomendada a utilização de expressões reificadas, como.

**‘geographicallyLocatedIn** some *BrazilLocation*’

Como resultado, todo o raciocínio pode ser mantido sobre a *Tbox*, i.e. sem fazer referência explícita a particulares (instâncias). Tal relação pode ser encontrada, no presente trabalho, para representar os locais em que ocorrem processos de transmissão de doenças, e também eventos de mortalidade.

A presença de ferramentas que suportem a construção de consultas complexas em DL poderia ser significativamente melhorada, principalmente nos casos em que as ontologias proveem muitas restrições de domínio e imagem.

Outro ponto a ser levado em consideração foi o decréscimo no desempenho do raciocínio. Expressividade focada juntamente com raciocínio em DL é, certamente, um dos critérios mais distintivos do presente trabalho. Quando se compara a NTDO à maioria das ontologias do OBO *Foundry* (SMITH *et al.*, 2007), ou o conjunto de ontologias as quais estão sendo desenvolvidas a partir da *Infectious Disease Ontology* (IDO) (TOPALIS *et al.*, 2011), representando um domínio similar com uma cobertura muito maior, estas apresentam menor expressividade de conteúdo axiomático, com anotações semânticas sendo a *raison d’être* predominante.

A avaliação de casos de uso os quais incluíam raciocínio para recuperação de subclasses e testes de satisfiabilidade, fazendo utilização extensiva dos construtores disponíveis na OWL DL e na ‘Manchester Syntax’, como disjunção, negação e restrições de valor. Ainda, apresentou uma variedade de definições de classes completas em quantidade e complexidade expressiva.

## 6.2. CONSIDERAÇÕES SOBRE O MODELO DE MORTALIDADE

Outro ponto a ser considerado é o modelo de mortalidade aqui apresentado e que demonstra a complexidade em modelar situações encontradas na área biomédica. Poucos trabalhos levam em consideração uma discussão ontológica sobre o tema 'morte'. Um trabalho presente na literatura acerca de representação de mortalidade ontológica foi descrito por Thomasma (THOMASMA, 1984), o qual descreve a ocorrência pontual da mortalidade como sendo apenas uma visão objetiva de finalização da vida.

Partindo do pressuposto que não é completamente possível expressar o que ocorre no momento do falecimento de um indivíduo, então, tal visão não seria completamente aplicável. Corroborando para isso, ainda, está o fato da morte ser identificada, sempre, por um indivíduo externo, não sendo o próprio falecido que atesta a própria morte.

Outro argumento a favor deste tipo de modelagem foi descrito em Schrader (SCHRADER, 1956), o qual aponta que Immanuel Kant relatou o fato de entidades indiscerníveis (como a morte e a vida, entre outras) poderem ser distinguidas pela situação única e particular diante do 'espaço' e do 'tempo', visão essa presente no modelo aqui apresentado.

## 6.3. TRABALHOS RELACIONADOS

Alguns exemplos de ontologias biológicas que apresentam conteúdo da mesma natureza das DTN são os trabalhos de Osborne et al. (OSBORNE *et al.*, 2009), Cowell et al. (COWELL; SMITH, 2010) e Topalis et al. (TOPALIS *et al.*, 2010). O estudo de Osborne et al. (2009) apresenta a *Disease Ontology* (DO), uma ontologia criada, com base no Sistema Linguístico Médico Unificado (*Unified Medical Language System* - UMLS) e fundamentada segundo os preceitos da *OBO Foundry*, para anotar informações sobre doenças e relacioná-las com sequências genéticas para auxiliar a descoberta de doenças.

Já o trabalho de Cowell et al. (COWELL; SMITH, 2010) descreve a IDO para representar, através de anotações, peculiaridades pertinentes a doenças de cunho

infeccioso, como Tuberculose, Febre do Dengue, Influenza, entre outros. Outro exemplo é o trabalho de Topalis et al. (TOPALIS *et al.*, 2010), o qual descreve a extensão da IDO para representar entidades relacionadas à Malária em uma ontologia chamada IDO Malária (IDOMAL), também baseada nas recomendações da OBO *Foundry*.

Um ponto em comum entre os trabalhos de Osborne et al. (OSBORNE *et al.*, 2009), Cowell et al. (COWELL; SMITH, 2010) e Topalis et al. (TOPALIS *et al.*, 2010), além de serem desenvolvidas sob a ótica da OBO *Foundry*, é o foco apenas em anotações e a utilização dessas para recuperar informações. Comparando-os com o presente trabalho, o conteúdo apresentado aqui além de seguir a rigorosidade filosófica empregada pela OBO *Foundry* e disponível na BioTop, também apresenta um enfoque para a semântica formal, com utilização de raciocínio e prova.

Tal junção da semântica formal (utilizada por engenheiros de conhecimento) com a abordagem científica da modelagem ontológica (utilizada geralmente por um especialista de domínio), além de embasamento filosófico, deve perfazer o alicerce da modelagem ontológica e, possivelmente, contribuir para uma melhor ontologia (GOBLE; WROE, 2004). Além disso, boa parte do conteúdo das ontologias provenientes da OBO é descrito de maneira informal e ambígua, no caso em linguagem natural, dificultando o reuso (GOLBREICH *et al.*, 2007).

Outro ponto a ser avaliado no presente trabalho é a criação de uma ontologia a partir de uma representação tabular. Alguns trabalhos foram descritos, como o de O'Connor et al. (O'CONNOR *et al.*, 2011) que propuseram uma solução genérica baseada em uma extensão da OWL *Manchester Syntax* a qual permitia a utilização de células provenientes de planilhas Excel para descrição de PPOs. Enquanto que esses autores criaram um método para 'ontologizar' uma informação tabular pré-existente por permitir a criação de classes e propriedades entre outras possibilidades sofisticadas; Bowers et al. (BOWERS *et al.*, 2010) descreveram uma abordagem baseado em planilha-para-OWL na qual planilhas são populadas com o objetivo de facilitar a construção ontológica. Uma abordagem similar foi desenvolvida por Peters et al. (PETERS *et al.*, 2009). Tais soluções podem ser mais amenas para

biólogos, médicos, entre outros, comparado com OWL. Entretanto, não se prestam para reutilização de dados legados, tal como extraídos de tabelas existentes.

Entretanto, uma importante limitação de transferência de conteúdo tabular em uma ontologia é o tipo de conteúdo. Ao passo que os conteúdos numéricos podem ser representados por *datatypes* em OWL (rudimentarmente suportado por raciocinadores), associações probabilísticas ou expressões *default* estendem o escopo do que é possível expressar em uma ontologia criada em OWL2, assim como a DL não é apropriada para processar tal tipo de conhecimento (FREITAS, FREDERICO; SCHULZ; MORAES, 2009b). Outro ponto importante a ser considerado, é a presença (no trabalho) de alguns construtores adicionais, os quais não estão descritos no âmbito da OWL2 e a OWL2 e a Manchester Syntax ainda não possibilitam a inserção de axiomas complexos que incluem instâncias.

#### **6.4. CASOS DE USO**

Por fim, vale ressaltar os potenciais casos de uso disponíveis para a NTDO. Um possível caso de uso, e em desenvolvimento, é a utilização da NTDO como meio de integração entre bancos de dados com informações epidemiológicas, i.e. morbidade e mortalidade, como o SINAN e o SIM. Além disso, e a partir da integração, pode ser utilizada para a construção de consultas inteligentes em sistemas de suporte a decisão. Outra possível utilização da presente ontologia é como conhecimento de domínio para teste de verificação de algoritmos em *chatbots*, que também se encontra em desenvolvimento. Ainda, pode ser utilizada como base de conhecimento para recuperação de informações textuais sobre o domínio, entre tantas outras possíveis utilizações.

## 7. CONCLUSÃO

No presente estudo, foram investigadas 3 questões.

- ✓ Primeiro, como conhecimento complexo acerca da transmissão de DTNs pode ser expresso de forma que garanta respostas confiáveis às CQs relevantes, formuladas por epidemiologistas;
- ✓ Segundo, como informações legadas contidas em tabelas, principalmente proveniente de publicações científicas, podem ser transformadas em uma ontologia formal, seguindo princípios filosóficos fundamentados, no caso dos processos de transmissão, e uma modelagem ontológica formalmente acurada;
- ✓ E, terceiro, como representar eventos complexos os quais apresentam marcações temporais, envolvendo causalidade e sequencia explícita e objetiva para a representação de entidades processuais que levam a eventos de mortalidade.

De acordo com os objetivos e o que foi exposto, foi possível achar resultados satisfatórios para as três questões. Entretanto, não é completamente possível representar todo o conteúdo do modelo de mortalidade em OWL2, visto que esta

não apresenta o operador de concordância. Mesmo assim, foi possível descrever conhecimento complexo de domínio, repleto de minúcias e pré-condições, satisfatoriamente com DL.

Levando em consideração apenas o terceiro ponto, a modelagem criada parece ter gerado uma representação satisfatória para a descrição dos processos relacionados à morte. Foi construído um modelo representacional evidenciando situações de mortalidade. Ainda, várias dificuldades foram encontradas, principalmente no que tange aos principais termos aqui representados, *viz.* morte, vida, nascimento, entre outros, e suas principais definições servindo como base para a modelagem.

Este terceiro ponto pode, ainda, servir como uma descrição passo-a-passo de como podem ser construídas representações ontológicas, pois pontua claramente os problemas encontrados e as possíveis soluções.

Adicionalmente, e de forma indireta aos objetivos do trabalho, foi desenvolvida com sucesso uma ferramenta baseada em PPO a qual deve ser individualmente construída para cada tabela, *i.e.* adaptada para o conteúdo a ser trabalhado. Tal ferramenta deve evidenciar as peculiaridades principais existentes no conteúdo simbólico disposto tabularmente e que partilha conhecimento do domínio.

Entretanto, algumas limitações foram encontradas. Uma delas é o conhecido problema de escalabilidade relacionado ao raciocínio em DL para utilização de ontologias, que por causa da complexidade computacional inerente, limita o desempenho. Outro ponto a ser levado em consideração se refere à construção de consultas às quais não são satisfatoriamente suportadas por ferramentas amigáveis ao usuário, como na interface de consulta do D2RQ (“D2RQ,” 2011) que utilizam a Linguagem de Consulta para RDF SPARQL (*SPARQL Query Language for RDF*) restringindo as possibilidades representacionais por fazê-los escolher um perfil otimizado de OWL2 para a aplicação desejada. Tal escolha pode levar à escolha pela restrição de expressividade, ou escalabilidade em ontologias extensas.

Uma limitação referente ao uso de tabelas como fonte de conhecimento para criação de ontologias é que seu conteúdo nem sempre pode conter conhecimento em nível

ontológico, em um senso mais estrito. Ainda, tabelas podem conter valores numéricos, os quais podem ser representados, mas são rudimentarmente suportados pela DL, possibilitando apenas a introdução do dado (sem raciocínio), e pela OWL2.

Uma última limitação, que não é relacionada à modelagem, mas à própria OWL2 é a falta de um construtor da DL, o operador de concordância ( $\equiv$ ). A presença de tal construtor permitiria a asserção de valores e a igualdade entre indivíduos em nível axiomático. Isto possibilitaria a definição completa de identidade de um indivíduo, como por exemplo, antes e *post mortem*. Para contornar tal problemática, outros formalismos representacionais, e.g. para conhecimento probabilístico precisam ser mais bem explorados.

De um modo geral, é possível, a partir do presente trabalho, descrever todos os processos relacionados às DTNs, em vida, como a transmissão de patógenos de um indivíduo para outro, o desencadeamento de doenças, desenvolvimento de infecção por tal patógeno e, em alguns casos com a evolução do quadro clínico de um paciente, representar o quadro causal relacionado a um possível falecimento.

Assim, a ontologia aqui descrita pode ser de potencial serventia em Sistemas de Suporte à Decisão, especificamente, para Saúde Pública, pois permitiria um estudo mais aprofundado da situação epidemiológica de determinada área seguindo um índice cronológico. A partir do conhecimento descrito sobre as DTN na NTDO, mais conhecimento pode ser coletado a partir de informações textuais por recuperação de informações do domínio em fontes textuais e, ainda, servir como casos de testes comparativos para *chatterbots* aprendizes, os quais coletam informações de domínio e, a partir disso, geram novas ontologias.

Por fim, o conjunto de conhecimento descrito na NTDO, e que pode ser gerado a partir dela, pode permitir a construção de consultas inteligentes em bancos de dados de Morbidade e Mortalidade. Assim, as representações semânticas aqui apresentadas podem permitir uma inovação no que concerne a vigilância de casos de doenças relacionados a infecções por doenças, principalmente negligenciadas.



Isto possibilitará o estudo de um amplo conjunto de variáveis, inerentes aos registros de morbidade e mortalidade, com formação de novo conhecimento sobre dados guardados separadamente e em muitos casos relacionados. Mais estudos são necessários para o desenvolvimento de aplicações baseadas na NTDO que possam auxiliar no estudo e avaliação do perfil epidemiológico e auxiliar na definição de ações de saúde para uma população.

## 8. REFERÊNCIAS

### BIBLIOGRÁFICAS

AHMED-OUAMER, R.; HAMMACHE, A. **Ontology-based information retrieval for e-Learning of computer science**. 2010 International Conference on Machine and Web Intelligence. **Anais...** [S.l.]: IEEE. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5648113>>. Acesso em: 23 abr. 2011. , out 2010

ALLOTEY, P.; REIDPATH, D. D.; POKHREL, S. Social sciences research in neglected tropical diseases 1: the ongoing neglect in the neglected tropical diseases. **Health research policy and systems / BioMed Central**, v. 8, n. 1, p. 32, jan 2010.

AULT, S. K. Pan American Health Organization's Regional Strategic Framework for addressing neglected diseases in neglected populations in Latin America and the Caribbean. **Memórias do Instituto Oswaldo Cruz**, v. 102, p. 99-107, out 2007.

BAADER, F.; HORROCKS, IAN; SATTTLER, U. Description Logics. In: HARMELEN, F. VAN; LIFSCHITZ, V.; PORTER, B. (Eds.). **Handbook of Knowledge Representation**. 1. ed. [S.l.]: Elsevier Inc., 2008. p. 135-179.

BAADER, F.; MCGUINNESS, D. L.; NARDI, D.; PATEL-SCHNEIDER, P. **The Description Logic Handbook: Theory, implementation, and applications**. 1. ed. Cambridge: Cambridge University Press, 2003.

BAADER, F.; NUTT, W. Basic Description Logics. In: BAADER, F.; CALVANESE, D.; MCGUINNESS, D. L.; NARDI, D.; PATEL-SCHNEIDER, P. (Eds.). **THE DESCRIPTION LOGIC HANDBOOK: Theory , implementation , and applications**. 2. ed. Cambridge: Cambridge University Press, 2007. p. 47-104.

BARRETT, A. D. T.; HIGGS, S. Yellow fever: a disease that has yet to be conquered. **Annual review of entomology**, v. 52, p. 209-29, jan 2007.

BEISSWANGER, E.; SCHULZ, S.; STENZHORN, H.; HAHN, U. BIOTOP : An Upper Domain Ontology for the Life Sciences. **Applied Ontology**, v. 3, n. 4, p. 205-212, 2008.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web. **Scientific American**, v. 284, n. 5, p. 34-43, maio 2001.

BETHONY, J.; BROOKER, S.; ALBONICO, M. *et al.* Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. **Lancet**, v. 367, n. 9521, p. 1521-32, 6 maio 2006.

BIEN, S. J.; PARK, C. H.; SHIM, H. J. *et al.* Bi-directional semantic similarity for gene ontology to optimize biological and clinical analyses. **Journal of the American Medical Informatics Association : JAMIA**, 28 fev 2012.

BLOMQVIST, E.; GANGEMI, A.; PRESUTTI, V. **Experiments on pattern-based ontology design**. Proceedings of the fifth international conference on Knowledge capture - K-CAP '09. **Anais...** New York, New York, USA: ACM Press. Disponível em: <<http://portal.acm.org/citation.cfm?doid=1597735.1597743>>. Acesso em: 19 dez. 2011. , 2009

BOBILLO, F.; DELGADO, M.; GÓMEZ-ROMERO, J. An Ontology Design Pattern for Representing Relevance in OWL. **Lecture Notes in Computer Science**, v. 4825/2007, p. 72-85, 2007.

BODENREIDER, O.; MITCHELL, J. A.; MCCRAY, A. T. **Biomedical ontologies**. Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing. **Anais...** [S.l: s.n.]. , jan 2005

BORST, W. N. **Construction of engineering ontologies for knowledge sharing and reuse**. [S.l.]: University of Twente, 1997.

BOUAMRANE, M.-M.; RECTOR, A.; HURRELL, M. Using OWL ontologies for adaptive patient information modelling and preoperative clinical decision support. **Knowledge and Information Systems**, 22 out 2010.

BOWERS, S.; MADIN, J. S.; SCHILDHAUER, M. P. Owlifier: Creating OWL-DL ontologies from simple spreadsheet-based knowledge descriptions ☆. **Ecological Informatics**, v. 5, n. 1, p. 19-25, jan 2010.

BRASIL, M. DA S. **Guia de Vigilância Epidemiológica**. 7. ed. Brasília: Ed. Ministério da Saúde, 2009. p. 816

BURGE, S.; KELLY, E.; LONSDALE, D. *et al.* Manual GO annotation of predictive protein signatures: the InterPro approach to GO curation. **Database : the journal of biological databases and curation**, v. 2012, p. bar068, jan 2012.

CHITTENDEN, T. W.; HOWE, E. A; TAYLOR, J. M. *et al.* nEASE: a Method for Gene Ontology Sub-Classification of High Throughput Gene Expression Data. **Bioinformatics (Oxford, England)**, v. 28, n. 5, p. 726-728, 13 jan 2012.

COURTOT, MÉLANIE; BUG, W.; GIBSON, F. *et al.* **The OWL of Biomedical Investigations**. Proceedings of the Fifth OWLED Workshop on OWL: Experiences and Directions. **Anais...** Karlsruhe, Germany: [s.n.] , 2008

COWELL, L. G.; SMITH, B. Infectious Disease Ontology. In: SINTCHENKO, V. (Ed.). **INFECTIOUS DISEASE INFORMATICS**. New York, New York, USA: Springer, 2010. p. 373-395.

CRAIG, E. Ontology. In: CRAIG, E. (Ed.). **Routledge Encyclopedia of Philosophy**. 7. ed. London: Routledge, 1998. .

**D2RQ**. Disponível em: <<http://d2rq.com/>>.

DAMJANOVIC, V. **Ontology design patterns for the semantic business processes**. Proceedings of the 4th International Workshop on Semantic Business Process Management - SBPM '09. **Anais...** New York, New York, USA: ACM Press. Disponível em: <<http://portal.acm.org/citation.cfm?doid=1944968.1944977>>. , 2009

DECIT. Doenças negligenciadas: estratégias do Ministério da Saúde. **Revista de Saúde Pública**, v. 44, n. 1, p. 200-202, fev 2010.

DONNELLY, K. SNOMED-CT: The advanced terminology and coding system for eHealth. **Studies in health technology and informatics**, v. 121, p. 279-90, jan 2006.

EGAÑA, M.; RECTOR, A.; STEVENS, R.; ANTEZANA, E. Applying Ontology Design Patterns in Bio-ontologies. **KNOWLEDGE ENGINEERING: PRACTICE AND PATTERNS: Lecture Notes in Computer Science**, v. 5268/2008, p. 7-16, 2008.

EJIGU, D.; SCUTURICI, M.; BRUNIE, L. **An Ontology-Based Approach to Context Modeling and Reasoning in Pervasive Computing**. Fifth Annual IEEE

International Conference on Pervasive Computing and Communications Workshops (PerComW'07). **Anais...** [S.l.]: Ieee. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4144786>>. , mar 2007

ELHANAN, G.; PERL, Y.; GELLER, J. A Survey of Direct Users and Uses of SNOMED CT: 2010 Status. **AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium**, v. 2010, p. 207-11, jan 2010.

**Exception** **ODP**. Disponível em: <<http://odps.sourceforge.net/odp/html/Exception.html>>.

FENSEL, D.; HORROCKS, I; HARMELEN, F. V. *et al.* **OIL in a Nutshell**. (R. Dieng, Ed.)Proceedings of the 12th European Workshop on Knowledge Acquisition, Modeling, and Management (EKAW'00). **Anais...** [S.l.]: Springer Verlag. , 2000

FERNANDEZ, M.; GOMEZ-PEREZ, A.; JURISTO, N. METHONTOLOGY : From Ontological Art Towards Ontological Engineering. **AAAI Technical Report**, v. SS-97-06, p. 33-40, 1997.

FREITAS, FRED. **Ontologias e a Web Semântica**. (R. Vieira & F. Osório, Eds.)Anais do XXIII Congresso da Sociedade Brasileira de Computação: Jornada de Mini-Cursos em Inteligência Artificial. **Anais...** Campinas: SBC. , 2003

FREITAS, FRED; SCHULZ, S.; MEDEIROS, Z. **Neglected Tropical Diseases - A Challenge to Biomedical Ontology Engineering**. (H. Herre, J. Kelso, & S. Schulz, Eds.)Ontologien in Biomedizin und Lebenswissenschaften (OBML) 2009. **Anais...** Leipzig: [s.n.]. Disponível em: <<https://wiki.imise.uni-leipzig.de/Gruppen/OBML/Workshops/2009?v=ayy>>. , 2009

FREITAS, FREDERICO; SCHULZ, S.; MORAES, E. Survey of current terminologies and ontologies in biology and medicine. **RECIIS**, v. 3, n. 1, p. 7-18, mar 2009a.

FREITAS, FREDERICO; SCHULZ, S.; MORAES, E. Pesquisa de terminologias e ontologias atuais em biologia e medicina. **RECIIS**, v. 3, n. 1, p. 8-20, 11 mar 2009b.

GANGEMI, A. **Ontology Design Patterns for Semantic Web Content**. (Y. Gil, E. Motta, V. R. Benjamins, & M. A. Musen, Eds.)4th International Semantic Web

Conference, ISWC 2005. Lecture Notes in Computer Science. **Anais...** Galway, Ireland: Springer-Verlag LNCS. , 2005

GANGEMI, A.; CATENACCI, C.; BATTAGLIA, M. Inflammation ontology design pattern: an exercise in building a core biomedical ontology with descriptions and situations. **Studies in health technology and informatics**, v. 102, p. 64-80, jan 2004.

GANGEMI, A.; GUARINO, N.; MASOLO, C.; OLTRAMARI, A.; SCHNEIDER, L. Sweetening Ontologies with DOLCE. **Lecture Notes in Computer Science**, v. 2473/2002, p. 223-233, 2002.

GOBLE, C.; WROE, C. The Montagues and the Capulets. **Comparative and functional genomics**, v. 5, n. 8, p. 623-32, jan 2004.

GOLBREICH, C.; HORRIDGE, M.; HORROCKS, IAN; MOTIK, B.; SHEARER, R. **OBO and OWL : Leveraging Semantic Web Technologies for the Life Sciences**. Proceedings of the 6th international The semantic web and 2nd Asian conference on Asian semantic web conference ISWC'07/ASWC'07. **Anais...** [S.l.]: Springer-Verlag Berlin, Heidelberg. , 2007

GRENON, P.; SMITH, B.; GOLDBERG, L. Biodynamic ontology: applying BFO in the biomedical domain. **Studies in health technology and informatics**, v. 102, n. ii, p. 20-38, jan 2004.

GRUBER, T. A translation approach to portable ontology specifications. **Knowledge Acquisition**, v. 5, n. 2, p. 199-220, jun 1993.

GRUNINGER, M.; FOX, M. S. **The Role of Competency Questions in Enterprise Engineering**. Proceedings of the IFIP WG5.7 Workshop on Benchmarking - Theory and Practice. **Anais...** Trondheim, Norway: [s.n.]. , 1994

GUARINO, N. **Formal Ontology and Information Systems**. Formal Ontology in Information Systems. Proceedings of FOIS'98. **Anais...** [S.l: s.n.]. , 1998

GUARINO, N.; WELTY, C. **Identity , Unity , and Individuality : Towards a Formal Toolkit for Ontological Analysis**. (W. Horn, Ed.)Proceedings of ECAI-2000: The European Conference on Artificial Intelligence. **Anais...** Amsterdam: IOS Press. , 2000a

GUARINO, N.; WELTY, C. **A Formal Ontology of Properties**. (R. Dieng & O. Corby, Eds.) Proceedings of EKAW-2000: The 12th International Conference on Knowledge Engineering and Knowledge Management. **Anais...** Menlo Park: AAAI Press. , 2000b

GUARINO, N.; WELTY, C. **Ontological Analysis of Taxonomic Relationships**. (A. Lander & V. Storey, Eds.) Proceedings of ER-2000: The International Conference on Conceptual Modelling. **Anais...** [S.l.]: Springer-Verlag LNCS. , 2000c

GUIZZARDI, G. **Ontological Foundations for Structural Conceptual Models**. Telematics ed. Enschede, The Netherlands: Instituut Telematics, 2005. p. 441

HADZIC, M.; CHANG, E. Ontology-Based Support for Human Disease Study. **Proceedings of the 38th Annual Hawaii International Conference on System Sciences**, v. 00, n. C, p. 143a-143a, 2005.

HAYES-ROTH, F.; WATERMAN, D. A.; LENAT, D. B. Principles of Pattern-Directed Inference Systems. **Pattern-Directed Inference Systems**. New York, NY, USA: Academic Press, 1978. .

HELLER, B.; HERRE, H. Ontological Categories in GOL. **Axiomathes**, v. 14, n. 1, p. 57-76, 2004.

HENDLER, J.; MCGUINNESS, D. L. The Darpa Agent Markup Language. **IEEE Intelligent Systems**, v. 15, n. 6, p. 67-73, 2000.

HERRE, H.; HELLER, B.; BUREK, P. *et al.* **General Formal Ontology (GFO): A Foundational Ontology Integrating Objects and Processes. Part I: Basic Principles**. . Leipzig: [s.n.]. Disponível em: <<http://www.onto-med.de/ontologies/gfo/>>. , 2007

HORRIDGE, M.; DRUMMOND, N.; GOODWIN, J.; RECTOR, A.; WANG, H. H. **The Manchester OWL Syntax**. Proc. of the 2006 OWL Experiences and Directions Workshop (OWL-ED2006). **Anais...** Athens, Georgia: [s.n.]. , 2006

HORRIDGE, M.; PATEL-SCHNEIDER, P. **OWL 2 Web Ontology Language: Manchester Syntax**. . [S.l.: s.n.]. Disponível em: <<http://www.w3.org/TR/owl2-manchester-syntax/>>. , 2009

HORROCKS, IAN; PATEL-SCHNEIDER, P.; HARMELEN, F. VAN. From SHIQ and RDF to OWL: the making of a Web Ontology Language. **Journal of Web Semantics**, v. 1, p. 7-26, dez 2003.

HOTEZ, P. J. The giant anteater in the room: Brazil's neglected tropical diseases problem. **PLoS neglected tropical diseases**, v. 2, n. 1, p. e177, jan 2008.

HOTEZ, P. J.; AKSOY, S. An interfaith dialogue on the neglected tropical diseases. **PLoS neglected tropical diseases**, v. 5, n. 12, p. e1240, dez 2011.

HOTEZ, P. J.; BOTTAZZI, M. E.; FRANCO-PAREDES, C.; AULT, S. K.; PERIAGO, M. R. The neglected tropical diseases of Latin America and the Caribbean: a review of disease burden and distribution and a roadmap for control and elimination. **PLoS neglected tropical diseases**, v. 2, n. 9, p. e300, jan 2008.

HOTEZ, P. J.; FENWICK, A.; SAVIOLI, L.; MOLYNEUX, D. H. Rescuing the bottom billion through control of neglected tropical diseases. **The Lancet**, v. 373, n. 9674, p. 1570-1575, maio 2009.

HOTEZ, P. J.; MOLYNEUX, D. H.; FENWICK, A. *et al.* Control of neglected tropical diseases. **The New England journal of medicine**, v. 357, n. 10, p. 1018-27, set 2007.

HOTEZ, P. J.; MOLYNEUX, D. H.; FENWICK, A. *et al.* Incorporating a rapid-impact package for neglected tropical diseases with programs for HIV/AIDS, tuberculosis, and malaria. **PLoS medicine**, v. 3, n. 5, p. e102, jan 2006.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Nascido Vivo**. Disponível em: <[http://www.ibge.gov.br/home/estatistica/populacao/registrocivil/nascido\\_vivo.shtm](http://www.ibge.gov.br/home/estatistica/populacao/registrocivil/nascido_vivo.shtm)>. Acesso em: 5 maio. 2011.

KARIUKI, T.; PHILLIPS, R.; NJENGA, S. *et al.* Research and capacity building for control of neglected tropical diseases: the need for a different approach. **PLoS neglected tropical diseases**, v. 5, n. 5, p. e1020, jan 2011.

KOSHLAND, D. E. The seven pillars of life. **Science (New York, N.Y.)**, v. 295, n. 5563, p. 2215-6, 22 mar 2002.



LENAT, DOUGLAS B.; GUHA, R. V. **Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project**. 1. ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1990. p. 391

LINDOSO, J. A. L.; LINDOSO, A. A. B. P. Neglected tropical diseases in Brazil. **Revista do Instituto de Medicina Tropical de São Paulo**, v. 51, n. 5, p. 247-253, out 2009.

LINEHAN, M.; HANSON, C.; WEAVER, A. *et al.* Integrated implementation of programs targeting neglected tropical diseases through preventive chemotherapy: proving the feasibility at national scale. **The American journal of tropical medicine and hygiene**, v. 84, n. 1, p. 5-14, jan 2011.

MAUDLIN, I.; EISLER, M. C.; WELBURN, S. C. Neglected and endemic zoonoses. **Philosophical transactions of the Royal Society of London. Series B, Biological sciences**, v. 364, n. 1530, p. 2777-87, 27 set 2009.

MCGUINNESS, D. L.; HARMELEN, F. VAN. **OWL Web Ontology Language: Overview**. Disponível em: <<http://www.w3.org/TR/owl-features/>>.

MEDECINS SANS FRONTIERES (MSF). **Fatal Imbalance The Crisis in Research and Development for Drugs for Neglected Diseases**. Disponível em: <<http://www.msf.org/source/access/Fatal/index.htm>>. Acesso em: 4 dez. 2011.

MENDOZA, N.; LI, A.; GILL, A.; TYRING, S. Filariasis: diagnosis and treatment. **Dermatologic therapy**, v. 22, n. 6, p. 475-90, 2009.

MICHALEK, H. **A Formal Ontological Approach to Causality embedded in the Top-Level Ontology of GFO ( General Formal Ontology )**. [S.l.]: University of Leipzig, 2009.

MILLER, F. G.; TRUOG, R. D. Decapitation and the definition of death. **Journal of medical ethics**, v. 36, n. 10, p. 632-4, out 2010.

MOREL, C. M. [Innovation in health and neglected diseases]. **Cadernos de saúde pública / Ministério da Saúde, Fundação Oswaldo Cruz, Escola Nacional de Saúde Pública**, v. 22, n. 8, p. 1522-3, ago 2006.

MOTIK, B.; SHEARER, R.; HORROCKS, IAN. **Optimized Reasoning in Description Logics using Hypertableaux**. Proceedings of the 21st international

conference on Automated Deduction. **Anais...** Berlin, Heidelberg: Springer Verlag. , 2007

MOTIK, B.; SHEARER, R.; HORROCKS, IAN. Hypertableau Reasoning for Description Logics. **Journal of Artificial Intelligence Research**, v. 36, p. 165-228, 2009.

**N-ary Relationship.** Disponível em: <[http://odps.sourceforge.net/odp/html/Nary\\_Relationship.html](http://odps.sourceforge.net/odp/html/Nary_Relationship.html)>.

NATIONAL LIBRARY OF MEDICINE; THE NATIONAL CENTER FOR BIOMEDICAL ONTOLOGY; BIOMEDICAL; STANFORD CENTER FOR BIOMEDICAL INFORMATICS RESEARCH. **Protégé Ontology Editor 4.1.** Disponível em: <<http://protege.stanford.edu/>>.

**NCBO Bioportal.** Disponível em: <<http://bioportal.bioontology.org/>>.

NEVES, D. P.; MELO, A. L.; LINARDI, P. M.; VITOR, R. W. A. **Parasitologia Humana**. 11. ed. São Paulo: Atheneu, 2005. p. 494

NOY, N. F.; KLEIN, MICHEL. Ontology Evolution: Not the Same as Schema Evolution. **Knowledge and Information Systems**, v. 6, n. 4, p. 428-440, 8 mar 2004.

NOY, N. F.; RUBIN, D. L. Translating the Foundational Model of Anatomy into OWL. **Web semantics (Online)**, v. 6, n. 2, p. 133-136, jan 2008.

NOY, N. F.; SHAH, N. H.; WHETZEL, P. L. *et al.* BioPortal: ontologies and integrated data resources at the click of a mouse. **Nucl. Acids Res.**, v. 37, n. Web Server issue, p. W170-3, 1 jul 2009.

NTDO. **Neglected Tropical Disease Ontology.** Disponível em: <<http://www.cin.ufpe.br/~ntdo/>>.

**Normalization ODP.** Disponível em: <<http://odps.sourceforge.net/odp/html/Normalisation.html>>.

ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE. **Saúde nas Américas.** Publicação ed. Washignton: Organizaçao Pan-Americana de Saúde, 2007.

OSBORNE, J. D.; FLATOW, J.; HOLKO, M. *et al.* Annotating the human genome with Disease Ontology. **BMC genomics**, v. 10 Suppl 1, p. S6, jan 2009.

**Ontology Design Patterns (ODPs) Public Catalog.** Disponível em: <<http://www.gong.manchester.ac.uk/odp/html/index.html>>.

**Ontology Design Patterns.org (ODP).** Disponível em: <[http://ontologydesignpatterns.org/wiki/Main\\_Page](http://ontologydesignpatterns.org/wiki/Main_Page)>.

O'CONNOR, M. J.; DAS, A. K. A Method for Representing and Querying Temporal Information in OWL. **Communications in Computer and Information Science**, v. 127, p. 97-110, 2011.

O'NEIL, M.; PAYNE, C.; READ, J. Read Codes Version 3: a user led terminology. **Methods of information in medicine**, v. 34, n. 1-2, p. 187-92, mar 1995.

PAN AMERICAN HEALTH ORGANIZATION. **2008: Number of Reported Cases of Dengue & Dengue Hemorrhagic Fever (DHF), Region of the Americas (by country and subregion).** Disponível em: <<http://www.paho.org/english/ad/dpc/cd/dengue-cases-2008.htm>>. Acesso em: 21 mar. 2011.

PETERS, B.; RUTTENBERG, A.; GREENBAUM, J. *et al.* Overcoming the Ontology Enrichment Bottleneck with Quick Term Templates. **Nature Precedings**, 16 nov 2009.

PISCOPO, T. V.; MALLIA AZZOPARDI, C. Leishmaniasis. **Postgraduate Medical Journal**, v. 83, n. 976, p. 649-657, fev 2007.

PRESUTTI, V.; GANGEMI, A. Content Ontology Design Patterns as Practical Building Blocks for Web Ontologies. **Lecture Notes in Computer Science**, v. 5231/2008, p. 128-141, 2008.

RAFI, M.; QURESHI, H.; KHATOON, H. **Ontology Maintenance via Multi-agents.** 2009 Fifth International Joint Conference on INC, IMS and IDC. **Anais...** [S.l.]: Ieee. , 2009

RASSI JR, A.; RASSI, A.; MARIN-NETO, J. A. Chagas disease. **The Lancet**, v. 375, n. 9723, p. 1388-1402, abr 2010.

RECTOR, A. L. **Modularisation of domain ontologies implemented in description logics and related formalisms including OWL.** Proceedings of the international conference on Knowledge capture - K-CAP '03. **Anais...** New York, New York, USA: ACM Press. , 2003

REISEN, W. K. Landscape epidemiology of vector-borne diseases. **Annual review of entomology**, v. 55, p. 461-83, jan 2010.

RILEY, L. W.; KO, A. I.; UNGER, A.; REIS, M. G. Slum health: diseases of neglected populations. **BMC international health and human rights**, v. 7, p. 2, jan 2007.

ROSSE, C.; MEJINO, JOSÉ L V. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. **Journal of biomedical informatics**, v. 36, n. 6, p. 478-500, dez 2003.

RUBIN, D. L.; SHAH, N. H.; NOY, N. F. Biomedical ontologies: a functional perspective. **Briefings in bioinformatics**, v. 9, n. 1, p. 75-90, jan 2008.

SABOU, M.; D'AQUIN, M.; MOTTA, E. **Using the Semantic Web as Background Knowledge for Ontology Mapping**. (P. Shvaiko, J. Euzenat, Natalya Fridman Noy, *et al.*, Eds.) Proceedings of the 1st International Workshop on Ontology Matching (OM-2006). **Anais...** Athens, Georgia, USA: [s.n.], 2006

SANTANA, F.; FERNANDES, R.; SCHOBER, D. *et al.* **Representing “casualties” for epidemiological data processing**. (H. Herre & R. Hoehndorf, Eds.) OBML 2011 Workshop Proceedings. **Anais...** Berlin, Heidelberg: Universität Leipzig, Medizinische Fakultät. Disponível em: <<https://wiki.imise.uni-leipzig.de/Gruppen/OBML/Workshops/2011>>. , 2011

SANTANA, F.; FERNANDES, R.; SCHOBER, D. *et al.* Towards an ontological representation of morbidity and mortality in Description Logics. **Journal of biomedical semantics**, 2012.

SANTANA, F.; SCHOBER, D.; MEDEIROS, Z.; FREITAS, FRED; SCHULZ, S. Ontology patterns for tabular representations of biomedical knowledge on neglected tropical diseases. **Bioinformatics**, v. 27, n. 13, p. i349-i356, 2011.

SCHMIDT-SCHAUSS, M. **Computational aspects of an order-sorted logic with term declarations**. Lecture no ed. Berlin; New York: Springer Verlag, 1989. p. 171

SCHOBER, D.; BOEKER, M.; BULLENKAMP, J. *et al.* The DebugIT core ontology: semantic integration of antibiotics resistance patterns. **Studies in health technology and informatics**, v. 160, n. Pt 2, p. 1060-4, jan 2010.

SCHRADER, G. A. Heidegger's Ontology of Human Existence. **The Review of Metaphysics**, v. 10, n. 1, p. 35-56, 1956.

SCHULZ, S.; DAUMKE, P.; SMITH, B.; HAHN, U. **How to distinguish parthood from location in bioontologies**. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium. **Anais...** [S.l.: s.n.]. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19833604>>. , jan 2005

SCHULZ, S.; SPACKMAN, K.; JAMES, A.; COCOS, C.; BOEKER, M. Scalable representations of diseases in biomedical ontologies. **Journal of biomedical semantics**, v. 2 Suppl 2, n. 1, p. S6, jan 2011.

SEBASTIAN, Y.; LOH, B. C. S.; THEN, P. H. H. **A Paradigm Shift: Combined Literature and Ontology-Driven Data Mining for Discovering Novel Relations in Biomedical Domain**. 2009 IEEE International Conference on Data Mining Workshops. **Anais...** [S.l.]: IEEE. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5360528>>. Acesso em: 23 abr. 2011. , dez 2009

SHARMA, U.; SINGH, S. Insect vectors of Leishmania: distribution, physiology and their control. **Journal of vector borne diseases**, v. 45, n. 4, p. 255-72, dez 2008.

SMITH, B. The Basic Tools of Formal Ontology. In: GUARINO, N. (Ed.). **Formal Ontology in Information Systems**. Amsterdam, Oxford, Tokyo, Washington, DC: IOS Press, 1998. p. 19–28.

SMITH, B.; ASHBURNER, M.; ROSSE, C. *et al.* The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. **Nature biotechnology**, v. 25, n. 11, p. 1251-5, nov 2007.

SMITH, B.; WELTY, C. **Ontology: Towards a New Synthesis**. Proceedings of the international conference on Formal Ontology in Information Systems - FOIS '01. **Anais...** Ogunquit, Maine, USA: ACM Press. , 2001

SOWA, J. F. **Knowledge Representation: Logical, Philosophical, and Computational Foundations**. 1. ed. [S.l.]: Brooks/Cole, 1999. p. 594

SPEAR, A. D. Ontology for the Twenty First Century: An Introduction with Recommendations by. p. 1-132, 2006.

STAAB, S.; STUDER, RUDI. **Handbook on Ontologies**. Second Edition. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. p. 811

STAAB, S.; STUDER, RUDI; SCHNURR, H.-P.; SURE, Y. Knowledge processes and ontologies. **IEEE Intelligent Systems**, v. 16, n. 1, p. 26-34, jan 2001.

STENZHORN, H.; BEISSWANGER, E.; SCHULZ, S. Towards a top-domain ontology for linking biomedical ontologies. **Studies in health technology and informatics**, v. 129, n. Pt 2, p. 1225-9, jan 2007.

STUDER, R. Knowledge engineering: Principles and methods. **Data & Knowledge Engineering**, v. 25, n. 1-2, p. 161-197, mar 1998.

TARSKI, A. **Introduction to Logic and to the Methodology of Deductive Sciences**. 4. ed. New York, New York, USA: Oxford University Press, 1994. p. 294

TAUIL, P. L. Aspectos críticos do controle da febre amarela no Brasil. **Revista de Saúde Pública**, v. 44, n. 3, p. 555-558, jun 2010.

THE GENE ONTOLOGY CONSORTIUM. Gene Ontology: tool for the unification of biology. **Nature Genetics**, v. 25, n. may, p. 25-29, 2000.

THE WORLD BANK. **Inequality in Latin America & the Caribbean: Breaking with History?** Disponível em: <<http://go.worldbank.org/TFJHCL2B30>>. Acesso em: 17 mar. 2011.

THOMAS, D. G.; PAPPU, R. V.; BAKER, N. A. Ontologies for cancer nanotechnology research. **Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference**, v. 2009, p. 4158-61, jan 2009.

THOMASMA, D. C. The comatose patient, the ontology of death, and the decision to stop treatment. **Theoretical Medicine**, v. 5, n. 2, p. 181-196, jun 1984.

TOPALIS, P.; DIALYNAS, E.; MITRAKA, E. *et al.* A set of ontologies to drive tools for the control of vector-borne diseases. **Journal of biomedical informatics**, v. 44, n. 1, p. 42-7, fev 2011.

TOPALIS, P.; MITRAKA, E.; BUJILA, I. *et al.* IDOMAL: an ontology for malaria. **Malaria journal**, v. 9, p. 230, jan 2010.

TRAVILLIAN, R. S.; DIATCHKA, K.; JUDGE, T. K.; WILAMOWSKA, K.; SHAPIRO, L. G. An ontology-based comparative anatomy information system. **Artificial intelligence in medicine**, v. 51, n. 1, p. 1-15, jan 2011.

TURNER, J. A; MEJINO, JOSE L V; BRINKLEY, J. F. *et al.* Application of neuroanatomical ontologies for neuroimaging data annotation. **Frontiers in neuroinformatics**, v. 4, n. June, p. 1-12, jan 2010.

**The Open Biological and Biomedical Ontologies Foundry.** Disponível em: <<http://www.obofoundry.org/>>.

USCHOLD, M.; GRUNINGER, M. Ontologies: Principles, methods and applications. **Knowledge Engineering Review**, v. 11, n. 2, p. 93-136, 1996.

**Upper Level Ontology ODP.** Disponível em: <[http://odps.sourceforge.net/odp/html/Upper\\_Level\\_Ontology.html](http://odps.sourceforge.net/odp/html/Upper_Level_Ontology.html)>.

VALENTE, A.; BREUKER, J. **Ontological Engineering with Principled Core Ontologies.** (A. Farquhar & M. Gruninger, Eds.)AAAI Technical Report SS-97-06. **Anais...** [S.I.]: AAAI Press. , 1997

**Value Partition ODP.** Disponível em: <[http://www.gong.manchester.ac.uk/odp/html/Value\\_Partition.html](http://www.gong.manchester.ac.uk/odp/html/Value_Partition.html)>.

WANG, H.; GAO, S.; WANG, J. **The Application Research of Description Logic in Civil Aviation Domain Ontology.** 2011 International Conference on Management and Service Science. **Anais...** [S.I.]: IEEE. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5998918>>. Acesso em: 3 jan. 2012. , ago 2011

WORLD HEALTH ORGANIZATION. **Macroeconomics and Health: Investing in Health for Economic Development.** World Health. Geneva: [s.n.], 2001

WORLD HEALTH ORGANIZATION. Global Programme to Eliminate Lymphatic Filariasis. **Weekly epidemiological record / Health Section of the Secretariat of the League of Nations**, v. 81, n. 22, p. 221-32, jun 2006.

WORLD HEALTH ORGANIZATION. **First WHO report on neglected tropical diseases 2010: Working to overcome the global impact of neglected tropical diseases.** World Health. Geneva: [s.n.], 2010

WORLD HEALTH ORGANIZATION. **World Health Statistics**. . Geneva: [s.n.], 2011

WORLD WIDE WEB CONSORTIUM (W3C). **OWL 2 Web Ontology: Language Document Overview**. Disponível em: <<http://www.w3.org/TR/2009/REC-owl2-overview-20091027/>>. Acesso em: 21 nov. 2011.

WU, X. **Goals combination in intelligent agents based description logics**. 2009 ISECS International Colloquium on Computing, Communication, Control, and Management. **Anais...** [S.l.]: Ieee. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5268013>>. , ago 2009

YILDIZ, B. **Ontology Evolution and Versioning**. . Vienna: Vienna University of Technology. Disponível em: <[http://publik.tuwien.ac.at/files/pub-inf\\_4603.pdf](http://publik.tuwien.ac.at/files/pub-inf_4603.pdf)>. , 2006

ZHIHONG, Z.; MINGTIAN, Z. **Web Ontology Language OWL and its description logic foundation**. Proceedings of the 8th International Scientific and Practical Conference of Students, Post-graduates and Young Scientists. Modern Technique and Technologies. MTT'2002 (Cat. No.02EX550). **Anais...** [S.l.]: IEEE. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1236278>>. Acesso em: 1 jun. 2011. , 2003.