



Programa de Pós-Graduação em

Computação Aplicada

Mestrado/Doutorado Acadêmico

Diego Pinheiro da Silva

Extração Semiautomática de Redes Bayesianas a partir de
Ontologias com Base Em Sumarização

São Leopoldo, 2019

Diego Pinheiro da Silva

**EXTRAÇÃO SEMIAUTOMÁTICA DE REDES BAYESIANAS A
PARTIR DE ONTOLOGIAS COM BASE EM SUMARIZAÇÃO**

Dissertação apresentada como requisito parcial
para a obtenção do título de Mestre, pelo
Programa de Pós-Graduação em Computação
Aplicada da Universidade do Vale do Rio dos
Sinos – UNISINOS

Orientador: Prof. Dr. Sandro José Rigo.

Co-orientadora: Prof.^a Dr.^a Marta Rosecler Bez.

São Leopoldo

2019

S586e

Silva, Diego Pinheiro da.

Extração semiautomática de redes bayesianas a partir de ontologias com base em sumarização / Diego Pinheiro da Silva – 2019.

13 f. : il. ; 30 cm.

Dissertação (mestre) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Computação Aplicada, 2019.

“Orientador: Prof. Dr. Sandro José Rigo Co-orientadora: Prof^a. Dr^a. Marta Rosecler Bez .”

1. Ontologia. 2. Redes bayesianas. 3. Extração de dados. 4. Sumarização de ontologias. I. Título.

CDU 004

Dados Internacionais de Catalogação na Publicação (CIP)
(Bibliotecária: Amanda Schuster – CRB 10/2517)

Diego Pinheiro da Silva

Extração Semiautomática de Redes Bayesianas a partir de Ontologias com Base Em Sumarização

Dissertação/Tese apresentada à Universidade do Vale do Rio dos Sinos – Unisinos, como requisito parcial para obtenção do título de Mestre/Doutor em Computação Aplicada.

Aprovado em 14 de março de 2019

BANCA EXAMINADORA

Prof. Dr. Sandro José Rigo - Orientador

Prof. Dr. Marcelo Ladeira - UNB

Prof. Dr. Cristiano André da Costa - UNISINOS

Prof. Dr. Sandro José Rigo - Orientador

Visto e permitida a impressão

São Leopoldo,

Prof. Dr. Rodrigo da Rosa Righi
Coordenador PPG em Computação Aplicada

AGRADECIMENTOS À CAPES

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Dedico esta pesquisa a minha família, professores, colegas e amigos, por todo apoio e motivação que mantiveram meus sonhos vivo nestes dois anos de mestrado.

AGRADECIMENTOS

Agradeço primeiramente aos meus pais e irmãos, heróis que me deram amor, incentivo e um apoio incondicional.

Ao Guilherme Frozza, por todo auxílio e companheirismo que me manteve firme nessa caminhada e que continua ao meu lado em todos os momentos.

Ao Sandro Rigo, por todo ensinamento, por me guiar em todo caminho e mostrar que eu posso mais, exemplo de pessoa para mim.

A Marta Bez, por todo o seu exemplo, amor e amizade que me inspira a ser melhor, que eu possa um dia ser metade da pessoa que és.

A Blanda Mello, por juntos termos superado todos os obstáculos, um apoiando ao outro, sem perder a alegria e a inspiração para fazer ciência.

Aos amigos e colegas, em especial aos que integram o grupo de pesquisa “Computação Aplicada”, dos quais fizeram parte da minha formação e que sempre vão estar presentes em minha vida.

A todos os envolvidos, muito obrigado.

RESUMO

Ontologias são modelos de representação de conhecimento facilmente interpretados tanto por seres humanos como por computadores. Redes Bayesianas são modelos de representação de conhecimento que incorporam o tratamento de incerteza. A sumarização de ontologias tem como proposta facilitar e melhorar o entendimento de uma ontologia, com o intuito de delimitar os conceitos mais importantes dentro o conjunto de conceitos representados. Existe uma grande dificuldade na construção de Redes Bayesianas, segundo a literatura. A maioria dos modelos conhecidos envolve uma extensa interação manual. Por outro lado, existe uma disponibilidade crescente de ontologias que descrevem o conhecimento de diversas áreas. Essas ontologias podem ser utilizadas como fontes para a criação de Redes Bayesianas, através de modelos diversos. Este trabalho apresenta um novo modelo para a extração semiautomática de Redes Bayesianas a partir de Ontologias. O diferencial desse trabalho localiza-se na abordagem com análise de relevância dos aspectos semânticos das ontologias, presentes em um algoritmo de conversão desenvolvido, bem como a estruturação do conhecimento necessário a ser convertido com uso de um método de sumarização de ontologias, o que é considerado como uma abordagem inédita na área. O texto descreve o embasamento teórico e os trabalhos relacionados, bem como as hipóteses formuladas, o modelo proposto, o protótipo desenvolvido e o experimento de avaliação de resultados. O modelo foi implementado em um caso real de geração de Redes Bayesianas para casos clínicos e foi integrado a um editor interativo de Redes Bayesianas. Realizou-se três experimentos diferentes, através de análises comparativas, com especialistas na área da saúde. Os resultados indicam boas possibilidades na geração das redes bayesianas, sendo efetivos quando comparados com resultados manuais e com avanços em relação ao estado da arte. O modelo foi bem recebido e considerado útil por especialistas.

Palavras-Chave: Ontologia, Redes Bayesianas, Extração de Dados, Sumarização de Ontologias.

ABSTRACT

Ontologies are models of representation of knowledge easily interpreted by both humans and computers. Bayesian networks are models of knowledge that work with uncertain reasoning, providing a way to treat uncertainty. The ontology summarization aims to facilitate and improve the understanding of an ontology, in order to restrict the knowledge of the domain to the most important concepts. There is a great difficulty in the construction and / or generation of Bayesian Networks, according to the literature. Most known models involve extensive manual interaction. On the other hand, there is a growing availability of ontologies that describe the knowledge of several areas. These ontologies can be applied as sources for the creation of Bayesian networks, through several models. This work presents a new model for the semiautomatic extraction of Bayesian Networks from ontologies. The differential of this work is the analysis of the relevance of the semantic aspects of the ontologies present in a developed conversion algorithm, as well as the structuring of the knowledge needed to be converted, where in this work we use the ontology summarization. The literature presents no approach that resembles this one. The text describes the theoretical basis and related works, as well as the formulated hypotheses, the developed model and the preliminary evaluation experiment. The model was implemented in a real case of Bayesian network generation for clinical cases and was integrated to a Bayesian Network editor. Three different experiments were carried out, through comparative analyzes, with specialists and questionnaires. The results indicate good possibilities in the generation of Bayesian networks, being effective when compared with manual results and advances in relation to the state of the art. The model was well received and endorsed by experts.

Keywords: Ontology, Bayesian Network, Data Extraction, Ontology Summarization.

LISTA DE FIGURAS

Figura 1: Classificação da pesquisa.....	28
Figura 2: Cadastro do protocolo na ferramenta StArt.	38
Figura 3: Validação realizada pela ferramenta ToS.	39
Figura 4: Gráfico comparativo de resultados entre as bases selecionadas.	40
Figura 5: Resultados inseridos na ferramenta StArt.....	41
Figura 6: Cadastro de extração de dados na ferramenta StArt.	43
Figura 7: Fluxograma dos resultados das fases de seleção.....	44
Figura 8: Título e ano dos artigos selecionados.	44
Figura 9: Comparativo do ano de publicação dos artigos selecionados.....	45
Figura 10: Comparativo da base de dados dos artigos selecionados.....	45
Figura 11: Revistas nas quais os artigos selecionados foram publicados.....	46
Figura 12: Comparativo de técnicas utilizadas nos artigos.	49
Figura 13: Comparativo de metodologias usadas.....	51
Figura 14: Comparativo de tipos de extração.....	54
Figura 15: Gráfico comparativo de tipos de validação.....	56
Figura 16: Gráfico comparativo de linguagens utilizadas.....	58
Figura 17: Comparativo de ferramentas utilizadas.....	60
Figura 18: Visão geral do modelo de extração de RBs a partir de ontologias.	64
Figura 19: Fragmento de código OWL	67
Figura 20: Fragmento de código JSON	68
Figura 21: Ontologia de Diagnóstico de Gripe.....	72
Figura 22: Rede Bayesiana de diagnóstico de gripe.....	73
Figura 23: Web Interface do Editor Bayesiano	75
Figura 24: RB de náusea gerada manualmente por especialistas.....	80
Figura 25: Ontologia de diagnóstico de náusea.....	81
Figura 26 - RB de náusea gerada pelo módulo de extração.	82
Figura 27: Fragmento de Ontologia de Eliminação Urinária Prejudicada.	83
Figura 28: Rede Bayesiana Original de Eliminação Urinária Prejudicada.	84
Figura 29: Rede Bayesiana de Eliminação Urinária Prejudicada gerada pelo protótipo.....	84

LISTA DE TABELAS

Tabela 1: Perguntas de pesquisa adotadas	35
Tabela 2: Comparação dos resultados nas bases de dados.	39
Tabela 3: Técnicas utilizadas e seus autores.	48
Tabela 4: Metodologias usadas e seus autores.	50
Tabela 5: Ligação entre os temas propostos e seus autores.....	51
Tabela 6: Protótipos desenvolvidos e seus autores.....	54
Tabela 7: Formas de validação e seus autores.	55
Tabela 8: Artigos aplicados na área da saúde e seus autores.	56
Tabela 9: Linguagens de programação utilizadas e seus autores.	57
Tabela 10: Resultados positivos e seus autores.	58
Tabela 11: Ferramentas utilizadas e seus autores.....	59
Tabela 12: Comparativo das abordagens utilizadas em cada publicação.	61
Tabela 13: Comparativo da RB gerada da ontologia com a do especialista.....	82
Tabela 14: Comparativo das redes desenvolvidas na validação.....	76
Tabela 15: Questionário para especialistas em saúde.....	77
Tabela 16: Questionário para especialistas em computação.	78
Tabela 17: Comparativo dos nodos	85

LISTA DE ABREVIATURAS

RBs	Redes Bayesianas
PIPICA	Programa Interdisciplinar de Pós-Graduação em Computação Aplicada
UNISINOS	Universidade do Vale do Rio dos Sinos
PLN	Processamento de Linguagem Natural
DICOM	<i>Digital Imaging and Communications in Medicine</i>
HL7	<i>Health Level Seven International</i>
IA	Inteligência Artificial
W3C	<i>World Wide Web Consortium</i>
ToS	<i>My Tree of Science</i>
JSON	<i>JavaScript Object Notation</i>
NET	<i>Netviz Project File</i>
OWL	<i>Ontology Web Language</i>
PF	Perguntas Focalizadas
PE	Pergunta Estática
PG	Perguntas Gerais

SUMÁRIO

1. INTRODUÇÃO	25
1.1. QUESTÃO DE PESQUISA.....	27
1.2. OBJETIVOS.....	27
1.2.1. Objetivo geral	27
1.2.2. Objetivos específicos	27
1.3. METODOLOGIA.....	27
1.4. ORGANIZAÇÃO DO DOCUMENTO.....	28
2. REFERÊNCIAL TEÓRICO	29
2.1. ONTOLOGIAS	29
2.2. SUMARIZAÇÃO DE ONTOLOGIAS	31
2.3. REDES BAYESIANAS.....	32
3. TRABALHOS RELACIONADOS	35
3.1. ELABORAÇÃO DAS QUESTÕES DE PESQUISA DA REVISÃO SISTEMÁTICA.....	35
3.2. CONSTRUÇÃO DO PROCESSO DE PESQUISA	36
3.3. DEFINIÇÃO DOS CRITÉRIOS DE INCLUSÃO/EXCLUSÃO	37
3.4. FERRAMENTAS DE APOIO	37
3.5. FASES DE SELEÇÃO.....	39
3.5.1. Fase 1 - Critérios de inclusão/exclusão;	40
3.5.2. Fase 2 – Validação ToS	41
3.5.3. Fase 3 – Título, palavras-chave e resumo.....	42
3.5.4. Fase 4 - Introdução e conclusão.....	42
3.5.5. Fase 5 - Leitura integral	42
3.6. RESULTADOS DA REVISÃO SISTEMÁTICA.....	43
3.6.1. Perguntas respondidas.....	46
3.6.2. Análise dos artigos.....	48
3.7. LACUNAS E DESAFIOS	60
4. MODELO PROPOSTO	63
4.1. VISÃO GERAL DO MODELO PROPOSTO.....	63
4.1.1. Componente do Editor de Redes Bayesianas	65
4.1.2. Componente de Comunicação	65
4.1.3. Componente de Serviços	65
4.1.4. Componente de Conversão	65
4.1.5. Componente de Recursos.....	66
4.1.6. Componente de Fonte de Dados	66
4.2. DETALHAMENTO DO PROCESSO DE ANÁLISE DE ONTOLOGIA E GERAÇÃO DE RB	68
4.2.1. Importação do OWL	68
4.2.2. Análise de contexto.....	69
4.2.3. Análise da estrutura	70
4.2.4. Extração dos dados	71
4.3. EXEMPLO DE APLICAÇÃO DO MODELO	71
5. IMPLEMENTAÇÃO E AVALIAÇÃO.....	75
5.1. PROTÓTIPO	75
5.2. EXPERIMENTOS	76

5.2.1.	Experimento com especialistas da saúde e computação utilizando o protótipo 1.0 ..	76
5.2.2.	Experimento com ontologias da web usando o protótipo 2.0.....	79
5.2.3.	Caso de estudo de enfermagem com protótipo 2.0.....	83
6.	CONCLUSÃO.....	87
6.1.	CONTRIBUIÇÕES	88
6.2.	TRABALHOS FUTUROS	88
	REFERÊNCIAS	89
	APÊNDICES.....	96
	APÊNDICE A - PROTOCOLO DA REVISÃO SISTEMÁTICA.....	97
	APÊNDICE B – DICIONÁRIO CRIADO PARA ANÁLISE DE CONTEXTO	100
	APÊNDICE C – EXEMPLO DE ONTOLOGIA EXTRAÍDA (NÁUSEA).....	101

1. INTRODUÇÃO

As Ontologias e as Redes Bayesianas (RBs) são duas abordagens para representação de conhecimento que vêm sendo utilizadas amplamente e possuem aspectos particulares que as tornam vantajosas em determinadas situações, determinando a sua preferência em determinadas áreas e aplicações. Nos últimos anos, um número crescente de aplicações utiliza ontologias, em diversas áreas. Uma destas áreas é a saúde, com registros eletrônicos de saúde ou suporte a diagnósticos (ZANNETI & BONACIN, 2014), (URBANSKY, 2018). O conhecimento do campo médico pode ser bem representado através do poder expressivo das ontologias (ZANNATA et al., 2013). Como evidência dessas iniciativas, atualmente, existem mais de 500 ontologias biomédicas registradas no BioPortal (KALET et al., 2017). As Redes Bayesianas (PINHEIRO et al., 2015), (NOLLETO, BARROS & PINHEIRO, 2017), por sua vez, consistem em artefatos que fornecem representações robustas do conhecimento probabilístico e são particularmente adequadas para aplicações de tomada de decisão sob a incerteza observada nos diversos domínios de conhecimento, tais como na saúde (STAJDUHAR, DALBELO-BASIC & BOGUNOVIC, 2011), (HELWANGER et al., 2017).

No entanto, a construção de RBs ainda é um problema a ser abordado (FENZ, 2012), (PINHEIRO et al., 2015), (KALET et al., 2017). A criação de estruturas de representação de conhecimento de RBs é uma atividade dispendiosa e desafiadora. Há um alto ganho de tempo, trabalho e investimento ao se reutilizar as bases de conhecimento existentes (ZANNETI & BONACIN, 2014), (FENZ, 2012), (CHAN et al., 2018). A construção automática ou semiautomática de RBs pode ser suportada por modelos que aproveitam a ampla disponibilidade de ontologias. Isso pode ser feito já que as ontologias definem os termos e propriedades fundamentais que compõem o vocabulário de uma área, bem como, as regras para relacionar termos e propriedades em algum vocabulário (KALET et al., 2017). As ontologias têm um alto poder expressivo, mas, em geral, não fornecem um tratamento apropriado da incerteza, situação esta que é melhor gerenciada pelas RBs (RODRIGUES, BEZ & FLORES, 2013), (CHANG, HUNG & JUANG, 2013), (PINHEIRO et al., 2015), (FENZ, 2012). Devido aos aspectos complementares dos formalismos, tanto das Ontologias como das RBs, a pesquisa sobre o seu uso conjunto pode ser interessante para fomentar a melhor aplicação destes recursos (PINHEIRO et al., 2015), (RODRIGUES, BEZ & FLORES, 2013).

Com base nesta motivação apresentada, observa-se um conjunto de iniciativas existentes para a utilização de Ontologias como base para a geração automática ou semiautomática de Redes Bayesianas. O estudo destes trabalhos permite observar que grande parte das técnicas empregadas necessita tratar das questões originadas pela falta de padronização de componentes das ontologias. Neste sentido, observa-se que alguns autores construíram sua própria ontologia para permitir um melhor tratamento dos elementos da ontologia, já tendo em vista a abordagem a ser empregada para a extração da RB, como a inserção de campos que representavam nodos, inserções de valores de probabilidades, ligação entre os nodos, dentre outros (AGARWAL, VERMA & MALLIK, 2016), (MOK & MIN, 2015). Em outras abordagens, em casos de extração com uso de ferramentas interativas, observa-se o uso de procedimentos de mapeamento das classes das ontologias a elementos específicos de RBs (LIN et al., 2014), (CHANG, HUNG & JUANG, 2013). Nestes trabalhos, percebe-se a identificação, por exemplo, de que as classes das ontologias são consideradas como equivalentes a nodos de RBs e que as relações existentes nas ontologias são consideradas como nodos com sua ligação. Este tipo de abordagem torna necessário que um especialista confirme e ajuste as operações conforme o processo de mapeamento é realizado.

Os aspectos hierárquicos existentes nas ontologias também foram explorados como elementos para a geração de RBs. Autores que utilizaram o nível hierárquico da ontologia para a extração e empregaram um algoritmo genérico de mapeamento em geral adotam uma abordagem que identifica as classes como os nodos-pai da rede, as propriedades que representam a ligação de um nodo-pai à um nodo-filho, e as instâncias, que são os nodos-filhos da rede (MESSAOUD et al., 2015). Há também outra técnica de mapeamento semelhante, que identifica apenas as classes e subclasses como os nodos da rede, fazendo com que a classe principal seja o nodo-pai, a subclasse seu nodo filho e assim por diante (KOVALENKO & ANTOSHCHUK, 2015).

A literatura apresenta uma série de dificuldades existentes nas iniciativas conhecidas. A maioria das abordagens em matéria de construção de base ontológica para RB não fornecem um modelo para usar ontologias existentes sem extensões específicas para a construção da RB. Na maioria das vezes, é necessário instalar complementos e baixar programas, dificultando o uso para usuários leigos, além da grande necessidade de conhecimento e intervenção do especialista que está construindo a RB, tanto na parte de software, como no conhecimento ontológico. Abordagens que não exigem essas extensões não suportam a incorporação automática de resultados e Tabelas de probabilidade condicionais. Não há nenhum modelo que seja capaz de gerar resultados concretos sem o auxílio de outras extensões e as Tabelas probabilísticas ainda são um grande desafio a ser atendido, visto que as ontologias não possuem apenas uma inferência lógica. Abordagens baseadas em ontologia existentes necessitam de modificações substanciais para permitir uma conversão automatizada, sendo que os experimentos analisados mostram a necessidade da intervenção do especialista em quase todo o processo de conversão, exigindo um alto nível de conhecimento e trabalho. Dadas as vantagens e limitações de ambos os formalismos, a investigação sobre a sua utilização complementar torna-se valiosa.

Buscando identificar com maior precisão os dados existentes, a técnica de sumarização de ontologias consiste em um método que busca melhorar o entendimento de uma ontologia, gerando um resumo da mesma (TROULLINO et al., 2015). As abordagens de sumarização de ontologias utilizam recursos para identificar quais são as partes mais importantes de uma ontologia e produzir uma nova versão resumida com os pontos principais. Tal geração de resumos possibilita a utilização das informações mais relevantes de uma ontologia sem que haja a necessidade de um conhecimento prévio da mesma (SOUZA, DURAN & VIEIRA, 2014).

Neste contexto, este trabalho apresenta um novo modelo para a extração semiautomática de RBs de ontologia. Este modelo permite usar o conhecimento descrito no formato de ontologia. Também oferece apoio no trabalho de especialistas em RB, o que pode ser um recurso essencial, pois a construção manual das RBs é demorada e trabalhosa, bem como exige uma grande quantidade de conhecimento específico. Além disso, o modelo proposto também fornece suporte ao uso integrado de ambos os formalismos, promovendo precisão na avaliação dos resultados da RB em eventuais diagnósticos.

Uma das contribuições do modelo apresentado é a utilização de um conjunto flexível de regras para avaliar o contexto em ontologias e mapear os conceitos e as relações necessárias para uma estrutura de RB, sem restrições quanto à estrutura da ontologia. Outra contribuição consiste em propor o uso de um método de sumarização de ontologias para filtrar os dados mais relevantes a serem utilizados em uma RB. A única intervenção manual necessária é o ajuste das probabilidades finais em cada nó da RB. Essa abordagem não é observada na literatura. A abordagem adotada permite que o conhecimento convertido em RB seja utilizado em diversos contextos. Para avaliar os resultados, foi utilizado um simulador de casos clínicos (BEZ et al., 2018), (ANTONELI et al., 2018), que utiliza o modelo de conhecimento para gerar as probabilidades de seus casos.

1.1. Questão de pesquisa

Portanto, a questão de pesquisa adotada neste trabalho foi definida da seguinte forma:

Como utilizar ontologias para construir a estrutura das Redes Bayesianas, preservando as restrições semânticas da ontologia e minimizando o trabalho do especialista?

1.2. Objetivos

1.2.1. Objetivo geral

O objetivo geral deste trabalho é o desenvolvimento de um modelo que utilize os dados de ontologias para a geração de RBs, respeitando a semântica original e desta forma auxiliando o especialista da área de domínio tratada.

1.2.2. Objetivos específicos

Buscando atender o objetivo geral, foram delineados os seguintes objetivos específicos:

1. Realizar estudos sobre Ontologias e Redes Bayesianas.
2. Avaliar trabalhos correlatos disponíveis na literatura quanto aos detalhes da sua abordagem de conversão de ontologias em Redes Bayesianas.
3. Analisar ontologias existentes na internet para identificar aspectos relevantes ao trabalho de conversão.
4. Analisar as necessidades para conversão do conhecimento das ontologias para a estrutura das redes bayesianas.
5. Descrever um modelo para realização da conversão semi-automaticamente.
6. Desenvolver um protótipo para avaliação do modelo.
7. Avaliar os resultados de experimentos com especialistas de um domínio.

1.3. Metodologia

A Figura 1 apresenta um resumo da metodologia que foi adotada neste trabalho, caracterizando os atributos pertinentes a esta pesquisa em vermelho.

Caracteriza-se o trabalho como de natureza aplicada, pois visa o desenvolvimento de um modelo que pode solucionar problemas concretos encontrados na literatura e com especialistas quando se trata de criar Redes Bayesianas.

A forma de abordagem será qualitativa, pois através da análise de especialistas e da literatura, buscar-se-á uma ampla compreensão de como é gerada uma Rede Bayesiana e de conhecimento sobre ontologias. Além disso, a medida em que a análise ocorrer, a mesma deverá ser validada junto a especialistas da área.

Os objetivos deste trabalho podem ser caracterizados como uma pesquisa do tipo exploratória, que visa o estudo de determinada área de ação, suas premissas e necessidades para que, a partir disso, possa ser desenvolvido o modelo que converta os dados de ontologias para a geração de Redes Bayesianas.

Figura 1: Classificação da pesquisa



Fonte: Adaptado de Bez (2013).

Quanto aos procedimentos técnicos, considera-se o trabalho como bibliográfico e experimental (WAZLAWICK, 2014). O procedimento bibliográfico será realizado através de pesquisas em trabalhos acadêmicos, livros, artigos e publicações referente a Redes Bayesianas, Ontologias, formas de conversão de dados e técnicas de inteligência artificial. O levantamento bibliográfico servirá como base para a definição da proposta de modelo que possa auxiliar os especialistas na geração de Redes Bayesianas. Será analisado o modo de construção, dados, semântica, ligações e contextos inseridos nas Redes Bayesianas.

Caracteriza-se o trabalho desenvolvido como experimental (KÖCHE, 2011), sendo realizada a análise e proposto o desenvolvimento de um modelo que visa solucionar o problema prático específico, observando como a tecnologia pode auxiliar na conversão do conhecimento e o suporte que prestará para os especialistas quanto ao conhecimento de domínio.

1.4. Organização do documento

No Capítulo 2 são apresentados os conceitos fundamentais sobre ontologias e sumarização de ontologias. Já o Capítulo 3 descreve os conceitos gerais sobre RBs. Uma revisão sistemática sobre sistemas de geração de Redes Bayesianas a partir de ontologias é descrita no Capítulo 4. O modelo proposto neste trabalho é apresentado no Capítulo 5. No Capítulo 6, é exibida a implementação do modelo em um editor bayesiano e a avaliação preliminar dos resultados, finalizando com as conclusões no Capítulo 7.

2. REFERÊNCIAL TEÓRICO

2.1. Ontologias

As ontologias são modelos de representação que relatam um domínio formalmente, determinando seus conceitos bem como relacionamentos, restrições e regras relativas. Utiliza-se de uma estrutura de dados abstrata que descreve o aspecto semântico de um conjunto de informações (GRUBER, 1993), (GRUBER, 2002). Esta estrutura é facilmente processada por computadores, possibilitando assim sua reutilização e compartilhamento (ZANATTA ET AL., 2013). Segundo Urbanskyy (2018), as ontologias permitem que o conhecimento tácito seja explicitado, formalizando uma visão relevante do mundo e tornando este modelo passível de processamento e interpretação por parte dos computadores.

Em geral, o objetivo de uma ontologia é representar o conhecimento através de variáveis e relações relevantes entre elas, fornecendo assim o entendimento comum para um domínio (KALET et al., 2017), (CONG, 2014). Tal entendimento facilita a comunicação precisa e eficaz das informações, o que, por sua vez, leva a outros benefícios, como a interoperabilidade, reutilização e compartilhamento (ARSENE, DUMITRACHE & MIHU, 2011). A representação é feita através de uma base de conhecimento, que se torna disponível para aplicações que necessitam usar e/ou compartilhar o conhecimento de um domínio. Uma ontologia fornece a capacidade de se modificar, de estabelecer hierarquias e relacionamentos entre entidades e, por fim, realizar inferências lógicas. Estas características têm se provado útil em vários campos (KALET et al., 2017).

As ontologias vêm sendo aplicadas em diferentes áreas da computação, dentre elas, a Integração de Dados, Inteligência Artificial (IA) e a Web Semântica. Visando tais áreas, as ontologias têm sido usadas como uma maneira de uniformizar o compartilhamento de informações em diferentes tipos de aplicações, relatando um domínio de interesse, seja ele de forma a representar uma classificação de conceitos e/ou axiomas de um teorema ou um esquema de banco de dados (SOUZA, DURAN & VIEIRA, 2014). A principal funcionalidade de uma ontologia é elevar a compreensão do conhecimento de um determinado domínio, usando um método de conceituação formal, facilmente compartilhável. Tal representação de conhecimento é formalmente especificado de maneira explícita para facilitar a interpretação e compreensão tanto de computadores como de pessoas (YAN et al., 2013). Através desta abordagem existe a possibilidade de utilização desta representação do conhecimento do mundo real por parte dos computadores, a fim de lidar com fatos de interesse (ZANATTA et al., 2013).

Na ciência da computação, as ontologias são consideradas um complemento para artefatos de software e também como linguagem para modelagem de recursos específicos. As mesmas tornaram-se a ferramenta de escolha das comunidades de IA para representar o conhecimento de uma área específica. de modo a facilitar o compartilhamento e reutilização das informações (SOUZA et al., 2014), (URBANSKY, 2018), (PINHEIRO et al., 2018). Existem hoje inúmeras ferramentas disponíveis para ajudar no desenvolvimento de uma ontologia, tais como Protégé, Ontop, OntoDebug, entre outras. Como o uso de ontologias tem crescido, torna-se mais fácil reutilizar outras ontologias para construção ou ampliação em algum caso específico de necessidade de um domínio do conhecimento. Isso ajuda a evitar o problema de silos de conhecimento (KALET et al., 2017).

Dentre as vantagens de seu uso na computação, destaca-se a possibilidade dada aos programadores de reaproveitar o conhecimento, podendo realizar alterações e implementar extensões para sua utilização. A criação de estruturas de conhecimento é uma das atividades mais longas, caras e complexas de um sistema especialista. Há um alto ganho de tempo, trabalho

e investimento quando reutilizada uma ontologia, visto que existe uma ampla quantidade de ontologias disponíveis para uso, reutilização e integração, podendo ser complementadas de acordo com os conceitos dos domínios (ZANETTI et al. 2013), (PINHEIRO et al., 2018).

A construção de ontologias envolve peculiaridades arquiteturais e metodológicas que a diferenciam da análise e desenvolvimento de sistemas de informação tradicionais. Dessa forma, no intuito de sistematizar essa atividade, foram propostas diversas metodologias específicas para a elaboração de ontologias. Entretanto, nenhuma das metodologias existentes é totalmente madura, principalmente se comparadas com metodologias de engenharia de software, não havendo, assim, um padrão para sua construção (URBANSKY, 2018).

O mais recente desenvolvimento em linguagens de ontologias é o OWL (*Web Ontology Language*), aprovado pelo *World Wide Web Consortium (W3C)* para promover a visão da Web Semântica, tornando-se a linguagem de uso padrão (ARSENE, DUMITRACHE & MIHU, 2011). O W3C inaugurou um ambiente para pesquisa e troca de informações, principalmente considerando-se a quantidade de dados que armazena. No entanto, estas informações são, em sua maioria, armazenadas sem qualquer critério de organização, padronização ou classificação quanto ao significado que carregam. Isso dificulta o processamento dos dados por meio de mecanismos inteligentes e automatizados (ARSENE, DUMITRACHE & MIHU, 2011).

Por esta razão, tem-se hoje a necessidade de transformar a Web atual em uma Web Semântica na qual a informação é provida tanto para o consumo humano quanto para o processamento de máquina. No entanto, para que a Web Semântica funcione, componentes de software devem ter acesso às informações e também as regras de inferência e significado destes dados. Nesse sentido, a forma de se representar a informação no contexto da Web Semântica é de vital importância nos dias atuais (LIMA et al., 2015).

Devido a existência de diversas aplicações de ontologias, é importante enfatizar as vantagens dessa utilização (URBANSKY, 2018), tais como pode ser observado nos pontos a seguir:

- Comunicação e colaboração entre pessoas: ontologias reduzem os conflitos conceituais e terminológicos na organização, uma vez que definem uma conceituação unificada da mesma.
- Formalização: devido a natureza formal de uma ontologia, há uma eliminação de contradições e inconsistências
- Interoperabilidade entre sistemas: ontologias podem ser utilizadas na tradução de diferentes representações de bases de dados, fornecendo uma conceituação única a partir da qual outras conceituações se normalizam.
- Representação de conhecimento e reuso: ontologias formam um vocabulário de consenso e representam o conhecimento de forma explícita no seu mais alto nível de abstração, possuindo um elevado potencial de reuso.

Uma ontologia representa, explicitamente, as classes de entidades de um domínio de aplicação, suas propriedades, suas relações, os papéis que pode desempenhar, como eles são decompostos em partes, os eventos e os processos que estão envolvidos. O conhecimento pode ser extraído de uma ontologia usando o raciocínio lógico, que explora tanto as relações entre classes (conceitos) e os fatos armazenados (as instâncias das classes) (ANDREA & FRANCO, 2012). A utilização das ontologias permite o acesso automático a bases de conhecimento de modo que, em conjunto com motores de inferência, podem ser usados para inferir novos conhecimentos a partir de fatos já conhecidos (FENZ, 2012).

Nos últimos anos, diversas aplicações foram desenvolvidas com a utilização de ontologias, inclusive em áreas relacionadas à saúde, tais como, informações armazenadas em registros eletrônicos de saúde e conhecimento medicinal (ZANATTA et al., 2013), (URBANSKY, 2018). Atualmente, há mais de 500 ontologias biomédicas registradas no BioPortal. Em padrões de sintaxe e semânticas internacionais, uma ontologia tem ampla aplicabilidade em uma variedade de contextos (KALET et al., 2017). O conhecimento medicinal é estruturado nas ontologias mediante a sua complexidade e o formalismo usado para a representação de conceitos, propriedades e relações. Esse formalismo possui um amplo poder expressivo, sendo altamente adequado para organizar grandes conjuntos de conhecimento (ZANATTA et al., 2013).

Quanto ao uso de ontologias na saúde, existe um paralelo com os avanços na área que podem ser feitos. Existe uma transição do formato de armazenamento do histórico de saúde dos pacientes, normalmente registrado manualmente em prontuários armazenados em grandes arquivos físicos localizados geralmente nos subsolos de hospitais e clínicas. Atualmente a maioria das instituições de saúde mantém alguma forma de digitalização de dados e observa-se o desenvolvimento de prontuários eletrônicos. Com isso, dois protocolos envolvendo informações médicas foram criados como forma de padronizar os dados e garantir a melhor interoperabilidade entre os estabelecimentos de saúde: *Digital Imaging and Communications in Medicine* (DICOM) e o *Health Level Seven International* (HL7). Tais padrões, embora importantes, são extremamente limitados e não resolvem o problema da semântica da informação (KALET et al., 2017). Por outro lado, esta semântica está sendo representada em formatos de ontologias, gerando um cenário favorável à novas aplicações na área da saúde (GRUBER, 2002).

2.2. Sumarização de Ontologias

A sumarização de ontologias pode ser definida como um método de análise do conhecimento expresso em uma ontologia para gerar uma versão resumida da mesma, que seja mais adequada para um determinado usuário ou tarefa (CARDOSO, PARDO & NUNES, 2013). Em geral, o método de sumarização consiste em “extrair uma parte saliente da ontologia para permitir uma investigação precisa e eficaz”, conforme afirma Souza, Duran & Vieira (2014). Na área de Processamento de Linguagem Natural a definição do termo “sumarização” apresenta características como: construir um sumário de um único ou vários documentos; preservar importantes e fundamentais informações, mantendo o sentido. Tais características são genéricas e fundamentais para sumarizar qualquer fonte de dados de grande volume. A segunda característica é fundamental para distinguir a sumarização de ontologia de outros métodos semelhantes, dos quais reduzem a complexidade ou o tamanho das ontologias (ZACARIAS, 2016).

A sumarização tem como proposta facilitar e melhorar o entendimento de uma ontologia. No intuito de ajudar no entendimento das ontologias, a sumarização busca restringir o conhecimento do domínio para os conceitos mais importantes. Tais conceitos destacados podem considerar tanto a opinião do usuário como a sua relevância na ontologia como um todo (SOUSA, PIRES & SALGADO, 2013). Este processo utiliza procedimentos semelhantes que são encontrados na geração de resumos de textos, tendo como a principal diferença o tipo de informação a ser resumida. Estes procedimentos que constituem a geração de um resumo envolvem etapas gerais tais como, em uma primeira etapa, identificar quais são as partes importantes do conteúdo, avaliando a relevância da informação em relação a todo o conteúdo ou relacionando a opinião do usuário. Uma segunda etapa envolve usar as partes importantes

que foram identificadas para gerar uma versão resumida do conteúdo, permanecendo com a sua coerência e sentido em relação as informações originais. (ZACARIAS & FELIPPO, 2015)

A aplicação da sumarização de ontologias pode ser composta por um sistema de compartilhamento de dados com base no conhecimento semântico, que usa de ontologias para representar dados. Ela também pode ser aplicada em softwares para auxiliar na visualização, navegação e desenvolvimento de uma ontologia, pois a redução da complexidade da ontologia facilita o entendimento e as atividades do mesmo (TROULLINOU et al., 2015). Uma ontologia pode conter entidades distintas, tais como classes, propriedades e instâncias, os quais geram termos primitivos e constituem os elementos básicos de uma ontologia. A modelagem estrutural desenvolvida na ontologia para representar um conhecimento é baseada principalmente nos conceitos e propriedades (GRUBIŠIĆ, STANKOV & PERAIĆ, 2013).

Um método para sumarização de ontologias consiste em identificar um trecho na ontologia que contenha os conceitos mais relevantes para gerar um resumo da mesma (QUEIROZ-SOUSA ET AL. 2013). Em função disso, o conceito estabelecido consiste em produzir uma outra ontologia por meio de uma réplica da original, que deve conter os conceitos mais relevantes da ontologia, sempre respeitando seus relacionamentos e propriedades semânticas (QUEIROZ-SOUSA ET AL. 2013).

No próximo Capítulo, é apresentado um referencial teórico sobre Redes Bayesianas.

2.3. Redes Bayesianas

De um modo geral, RBs podem ser consideradas modelos de representação de conhecimento incerto, baseados no Teorema de Bayes (PINHEIRO et al., 2015). Elas compõem um modelo probabilístico gráfico que representa um conjunto de variáveis aleatórias e suas dependências condicionais através de um gráfico acíclico orientado. Tais redes são amplamente utilizados como modelo gráfico para a representação do conhecimento probabilístico sob incerteza (FENZ, 2012), (ARSENE, DUMITRACHE & MIHU, 2015).

As RBs são usadas principalmente para apresentar a incerteza de vários domínios de uma forma natural e facilmente compreensível (STAJDUHAR, DALBELO-BASIC & BOGUNOVIC, 2011). Elas baseiam-se no princípio de que, como grande parte das variáveis de um domínio é condicionalmente independente, não é necessário calcular todas as suas probabilidades conjuntas, sendo possível ignorar ramificações irrelevantes para a consulta que se está fazendo (PEARL, 2014). Com isso, podem ser modelados dois tipos de informações: estrutura (aspecto qualitativo) e informação numérica (probabilidades - aspecto quantitativo) (ARSENE, DUMITRACHE & MIHU, 2015).

A principal vantagem do raciocínio probabilístico sobre o raciocínio lógico é o fato de que agentes podem tomar decisões racionais mesmo quando não existe informação suficiente para provar que uma ação funcionará (HELWANGER et al., 2017). Os autores Stajduhar, Dalbello-Basic & Bogunovic (2011) relatam que o processo de diagnóstico (independente se for médico, de reparo automobilístico ou qualquer outra área) quase sempre envolve incerteza. Nesses casos, o especialista precisa identificar a causa do problema a partir de sinais e sintomas, ou seja, dos efeitos da causa observada.

As redes possuem nodos que representam variáveis aleatórias no sentido bayesiano, podendo ser quantidades observáveis, variáveis latentes, parâmetros desconhecidos ou hipóteses. As ligações entre os nodos representam dependências condicionais. Nodos que não estão ligados representam variáveis que são condicionalmente independentes um do outro. Cada nó está associado com uma função de probabilidade que recebe como entrada um conjunto

particular de valores para as variáveis controladoras, calculando assim, a probabilidade da variável representada pelo nodo (FENZ, 2012).

As RBs são utilizadas em várias aplicações do mundo real, tais como diagnósticos, análise de *spam*, recuperação de informações e PLN (FENZ, 2012), (BEZ et al., 2018), (JUNQUEIRA & FERNANDES, 2018). Elas tornaram-se poderosas representações de conhecimento probabilístico. Como tal, são particularmente bem adequadas para aplicações de raciocínio sob incerteza em domínios médicos. RBs têm sido utilizados na decisão médica durante várias décadas. Enquanto elas podem ser mais comumente conhecidas por seu papel no raciocínio diagnóstico, usos recentes são observados nos campos de análise de relações entre dados genômicos e câncer, meta-análise de dados biomédicos, modelagem e sistemas de apoio à decisão clínica (KALET et al., 2017).

O diagnóstico médico pode ser definido como o processo de identificação de um conjunto de hipóteses que modela o domínio do problema e encontra aquele com maior probabilidade de combinar o estado atual. A incerteza surge da incapacidade de avaliar o grau de verdade de uma hipótese devido a informações pouco confiáveis, incompletas ou inconsistentes (ARSENE, DUMITRACHE & MIHU, 2011). Por estas razões, as RBs são excelentes ferramentas para a representação do conhecimento na medicina. Este tipo de representação corresponde ao raciocínio humano sobre causalidade e incerteza. Além disso, eles podem ser aprendidos a partir de dados. Por esta razão, RBs estão sendo utilizadas de usos em biomedicina e saúde para o diagnóstico, tratamento, prognóstico e descoberta de interações funcionais (STAJDUHAR, DALBELO-BASIC & BOGUNOVIC, 2011).

A dificuldade de utilizar RBs na medicina surge de inúmeras razões, incluindo a falta de conhecimento definitiva, conflitantes teorias e conhecimento especializado que não é facilmente ou amplamente compartilhado. Estes obstáculos podem levar à confusão entre as variáveis que precisam ser consideradas e as conexões entre elas, tornando-se, assim, uma difícil tarefa de se estabelecer a topologia de uma RB. A topologia é o conjunto de variáveis, seus estados e a ligação que diz respeito a duas variáveis. A direção da ligação representa causalidade ou dependência, e as relações são quantificadas por Tabelas de probabilidades condicionais (KALET et al., 2017).

A Tabela de probabilidade condicional é usada para expressar como os estados potenciais dos pais do nó afetam a probabilidade posterior do nó considerado. De um modo geral, dois modelos diferentes ou uma combinação de ambos são usadas para construir uma RB: a partir de uma rede de dados existente ou através de especialistas que inserem domínios de conhecimento complexos. Ambas as abordagens têm suas deficiências específicas. Enquanto a abordagem baseada no especialista de domínio é demorada e propensa a erros, a construção automatizada através de redes de dados existentes enfrenta um problema de polarização (FENZ, 2012).

Para a construção de RBs, é necessário que haja a identificação das variáveis significativas dentro do domínio, juntamente com os seus valores possíveis, e a identificação das relações entre variáveis e na representação gráfica da estrutura (ARSENE, DUMITRACHE & MIHU, 2015). Estas variáveis podem ser de diferentes tipos, de acordo com o problema em questão, sendo elas: Booleanas (“verdadeiro” e “falso”), Numéricas (tamanho, peso, velocidade, entre outras) ou Categóricas (categorias, como cor, letras, entre outras) (HELWANGER, 2016).

Mediante a isto, surgem alguns desafios na construção de RBs, como identificar quais variáveis são relevantes para o domínio considerado (nodo), identificar as relações entre as variáveis identificadas (ligações) e criar uma Tabela de probabilidade condicional para cada

variável, valores que, normalmente, são inseridos a partir de conhecimento tácito (FENZ, 2012).

As classes, propriedades e instâncias de uma ontologia são amplamente utilizados para capturar e representar um determinado domínio de conhecimento. Elas são uma solução potencial para apoiar a construção de RBs, pois as mesmas definem os termos e as propriedades básicas que compreendem o vocabulário de uma área, bem como, as regras para a combinação de termos e propriedades para definir as extensões para o vocabulário (FENZ, 2012). É importante ressaltar que os conceitos de ontologias e RBs apresentam diferenças, mas também compartilham algumas semelhanças. Enquanto ontologias descrevem conceitos de domínio e suas relações, as RBs descrevem as distribuições de probabilidade e suas relações. As ontologias descrevem conceitos como classes, enquanto as RBs usam nodos e ligações para representar um determinado domínio. Pode-se usar a semântica formalmente definida nas ontologias com as classes e propriedades para construir, automaticamente a estrutura gráfica (nodos e ligações) das RBs. Desta forma, pesquisas mostraram que os domínios descritos nas ontologias são valiosos na construção de RBs, principalmente por representar causalidades nestes domínios (FENZ, 2012).

Após o aporte teórico sobre Ontologias, Sumarização de Ontologias e RBs, o próximo Capítulo apresenta os trabalhos correlatos.

3. TRABALHOS RELACIONADOS

Na busca de trabalhos correlatos, foi realizada uma revisão sistemática que, após um processo de filtragem de 1.093 publicações, analisou 33 trabalhos relacionados com a conversão de ontologias em RBs. Esta revisão sistemática sobre extração de dados de ontologias e geração de RBs foi desenvolvida baseando-se no protocolo criado por Medeiros (2016), que utilizou uma mescla do protocolo da pesquisadora Elisabete Kitchenham (KITCHENHAM, 2007), referente à área da computação, e o protocolo de Recomendação PRISMA (PRISMA, 2015), direcionado para à área da saúde. O Protocolo completo desta revisão sistemática está disponível no [Apêndice A](#) deste trabalho. Esta abordagem permite que sejam levados em consideração aspectos do processo de escolha e análise dos artigos, ampliando a assertividade dos resultados, quando comparados com os resultados obtidos a partir de revisões diretas de literatura (COOPER, 2016).

O tema principal desta revisão é a extração de dados de ontologias para conversão em RBs, tendo sido abordados, complementarmente, assuntos de interesse referentes ao tema principal, tais como as aplicações destes recursos na área da saúde e também o conjunto de técnicas de extração de dados e PLN disponíveis para estas finalidades. O protocolo foi desenvolvido com a ajuda de especialistas da área de computação ligados a área da saúde e a área de Inteligência Artificial. Os itens a seguir descrevem o trabalho desenvolvido em cada etapa do protocolo adotado.

3.1. Elaboração das questões de pesquisa da revisão sistemática

As questões de pesquisa da revisão sistemática levaram este estudo a descobrir trabalhos que poderiam estar ligados à extração de dados de ontologias para conversão em RBs, em contextos que preveem o uso destes recursos em atividades ligadas à saúde. Para este trabalho, foram definidas cinco Perguntas Gerais (PG), três Perguntas Focalizadas (PF) e uma Pergunta Estatística (PE). O propósito das PGs é entender de que forma os trabalhos foram aplicados em relação ao assunto. O objetivo das PFs é identificar quais as tecnologias usadas para construir seus modelos. Por último, o propósito associado com as PEs é encontrar dados estatísticos sobre o estudo realizado. Estas questões são apresentadas na Tabela 1.

Tabela 1: Perguntas de pesquisa adotadas

Referência	Questão
<i>Perguntas Gerais</i>	
PG1	Quais foram as técnicas mais utilizadas?
PG2	Que tipos de metodologias foram usadas?
PG3	Qual a ligação entre os temas propostos?
PG4	Que formas de validação foram utilizadas?

PG5 Os artigos foram aplicados à área da saúde?

Perguntas Focalizadas

PF1 Quais as linguagens de programação que foram utilizadas?

PF2 Quais foram as ferramentas utilizadas?

PF3 O artigo apresenta algum protótipo desenvolvido?

Pergunta Estática

PE1 A proposta teve resultados positivos?

Fonte: Elaborado pelo autor.

3.2. Construção do processo de pesquisa

Através do processo definido pelo protocolo, foram escolhidas as bases de dados para aplicar a sequência de pesquisa, a obtenção dos resultados e a identificação dos principais termos. Para a área da computação, foram utilizadas as seguintes bases de dados: *Web Of Science*, que possui o acesso a periódicos em diversas áreas do conhecimento (WEB OF SCIENCE, 2017); *ACM Digital Library*, que é uma base de dados bibliográficos com foco na área da computação (ACM, 2017); *IEEEExplore*, que fornece acesso a algumas das publicações mais citadas no mundo em Ciência da Computação, Eletrônica e Engenharia Elétrica (IEEEXPLORE, 2017). Para a área da saúde, foi utilizada a base de dados *PubMed*, que conta com publicações biomédicas do *MEDLINE* (PUBMED, 2017).

Esta escolha de bases de dados a consultar levou em conta os requisitos definidos para este trabalho, no que se refere a identificação de forma ampla e consistente de material bibliográfico sobre o tema principal de pesquisa, com a utilização de fontes complementares integrando tanto a área de computação como a área de saúde. A *string* de busca foi criada conforme as palavras-chave definidas, tendo como elementos obrigatórios da pesquisa o uso de Ontologias e RBs. A *String* de busca é apresentada na seguinte sequência: ((“*ontology*”) AND (“*bayesian network*”) OR (“*health*”) OR (“*natural language processing*” OR “*NLP*”) OR (“*data extraction*”)).

Na primeira etapa, foi executada a *string* de busca nas bases de dados selecionadas. Todas as publicações encontradas foram exportadas no formato *BibTex* para serem cadastradas na ferramenta *StArt* (HERNANDES et al., 2010). Após o registro, foram validados os critérios de inclusão e exclusão. Os artigos da base *Web Of Science* passaram pela validação do software *My Tree of Science* (ToS) (ROBLEDO-GIRALDO; DUQUE-MÉNDEZ; ZULUAGA-GIRALDO, 2013). Avante, foram realizadas as leituras de todos os títulos, palavras-chave e resumos, passando, em seguida, para uma validação de introdução e conclusão e, na fase final, a leitura integral dos artigos selecionados. As fases do processo de seleção de artigos são exibidas a seguir:

Fase 1 - Validar os critérios de inclusão e exclusão;

Fase 2 - Validar as publicações na ferramenta ToS (conforme as bases compatíveis).

Fase 3 - Leitura do título, palavras-chave e resumo;

Fase 4 - Leitura da introdução e conclusão;

Fase 5 - Leitura integral dos artigos e validação das respostas para as perguntas.

3.3. Definição dos critérios de inclusão/exclusão

Os estudos foram filtrados para selecionar os artigos relevantes. Para isso, alguns critérios de inclusão/exclusão (CIE) foram definidos, conforme indicado a seguir:

CIE1: O ano de publicação do artigo deve estar dentro do período de 2011 e 2017 (6 anos anteriores a data da pesquisa);

CIE2: Ser um artigo científico publicado;

CIE3: Deve estar escrito em inglês;

CIE4: O artigo deve apresentar uma forma de validação;

CIE5: A publicação deve estar disponível na íntegra na internet ou disponível através de convênios das instituições de ensino.

CIE6: Ser validado pelo software *ToS* (ROBLEDO-GIRALDO; DUQUE-MÉNDEZ; ZULUAGA-GIRALDO, 2013).

3.4. Ferramentas de apoio

Para a classificação dos artigos, foi utilizado o software *StArt*, desenvolvido para apoiar todo o processo de uma revisão sistemática. Por meio de uma árvore hierárquica, ele disponibiliza funcionalidades que apoiam a execução de cada fase. Inicialmente, é cadastrado o protocolo criado ([Apêndice A](#)). Após, o software disponibiliza funcionalidades para apoiar as etapas de condução, seleção e extração das informações. Para a importação dos dados na ferramenta, as publicações devem ser exportadas no formato *BibTex* (HERNANDES et al. 2010). A Figura 2 ilustra o cadastro do protocolo na ferramenta *StArt*.

Figura 2: Cadastro do protocolo na ferramenta StArt.

Protocol

Objective:* ?

Está revisão sistemática busca encontrar os artigos que trabalham com ontologias e redes Bayesianas, buscando encontrar técnicas e validações de inteligência artificial, como processamento de linguagem natural e extração de dados, visando principalmente verificar na literatura modos de conversão de ontologias para redes bayesianas aplicadas a saúde.

* This field must be filled in

Main question:* ?

Validar extração de dados de ontologias para conversão em redes bayesianas

Population: Pesquisadores, professores, desenvolvedores e profissionais da computação.

Intervention: Verificar as tecnologias que são utilizadas nas pesquisas.

Control: Não será utilizado.

Results: Gerar um embasamento teórico para a dissertação, assim como a publicação de artigos científicos.

Application: Esta revisão sistemática tem como foco pesquisadores, professores e alunos da área da computação aplicada, pois trará uma visão abrangente de tecnologias e metodologias adotadas até o momento para desenvolvimento de pesquisas.

* This field must be filled in

Keywords and Synonyms* ?

Keywords:

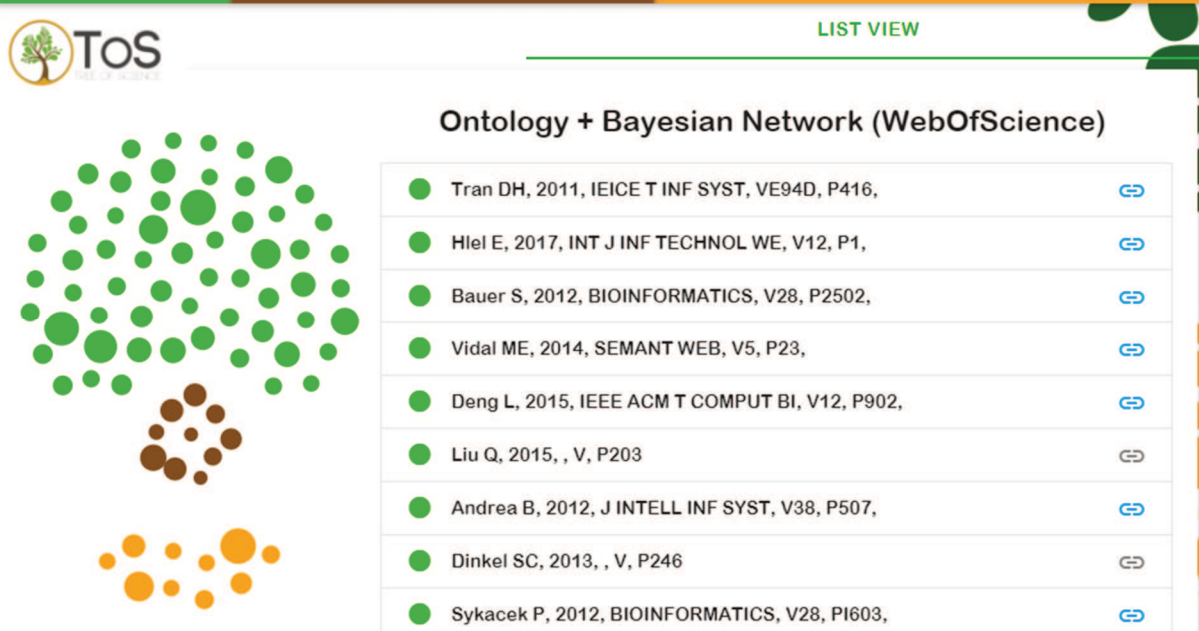
Ontology
Bayesian network

Fonte: Elaborado pelo autor.

Para a análise dos resultados da base de dados da *Web Of Science*, utilizou-se o software *My Tree of Science (ToS)*, proposto pelo grupo de pesquisa GAIA, da Universidad Nacional de Colombia, Sede Manizales (ROBLEDO-GIRALDO; DUQUE-MÉNDEZ; ZULUAGA-GIRALDO, 2013). *ToS* é uma ferramenta web para a seleção de artigos científicos que utiliza os princípios da teoria da rede para identificar os artigos e autores mais relevantes em um determinado assunto. Os artigos são relacionados às suas referências e formam um gráfico onde pode-se calcular os graus de saída e entrada para cada artigo.

Como resultado, ele fornece uma lista ordenada dos artigos relevantes de pesquisa em formato de árvore, onde a raiz (representada em laranja) são artigos clássicos na literatura (não utilizados neste trabalho por não preencher o requisito de tempo de publicação), o tronco (representado em marrom) são os artigos que servem como a estrutura da pesquisa e as folhas (representadas em verde) são as publicações com novas tendências sobre o assunto pesquisado. Na Figura 3 é ilustrada a representação da ferramenta com a *string* de busca apresentada no protocolo. Em seguida, serão apresentados os resultados das buscas realizadas nas bases de dados, as fases de seleção e os critérios de inclusão e exclusão de cada artigo.

Figura 3: Validação realizada pela ferramenta ToS.



The screenshot displays the ToS interface. On the left is a logo with a tree and the text 'ToS'. Below it is a decorative graphic of green and orange dots. The main content area is titled 'Ontology + Bayesian Network (WebOfScience)' and shows a 'LIST VIEW' of search results. The results are as follows:

Author	Year	Journal	Volume	Page	Link
Tran DH	2011	IEICE T INF SYST	VE94D	P416	Link
Hiel E	2017	INT J INF TECHNOL WE	V12	P1	Link
Bauer S	2012	BIOINFORMATICS	V28	P2502	Link
Vidal ME	2014	SEMANT WEB	V5	P23	Link
Deng L	2015	IEEE ACM T COMPUT BI	V12	P902	Link
Liu Q	2015	, , V		P203	Link
Andrea B	2012	J INTELL INF SYST	V38	P507	Link
Dinkel SC	2013	, , V		P246	Link
Sykacek P	2012	BIOINFORMATICS	V28	PI603	Link

Fonte: Elaborado pelo autor.

3.5. Fases De Seleção

A inserção da *string* de busca nas bases de dados foi realizada no dia 17 de agosto de 2017. Somando os resultados das quatro plataformas propostas nesta pesquisa, foram encontradas 1.093 publicações relacionadas a ontologias e RBs. Após os filtros de período e idioma, restaram 631 artigos a serem analisados, que são de interesse desta pesquisa. A Tabela 2 ilustra os resultados encontrados de acordo com a base e a inserção das palavras-chave.

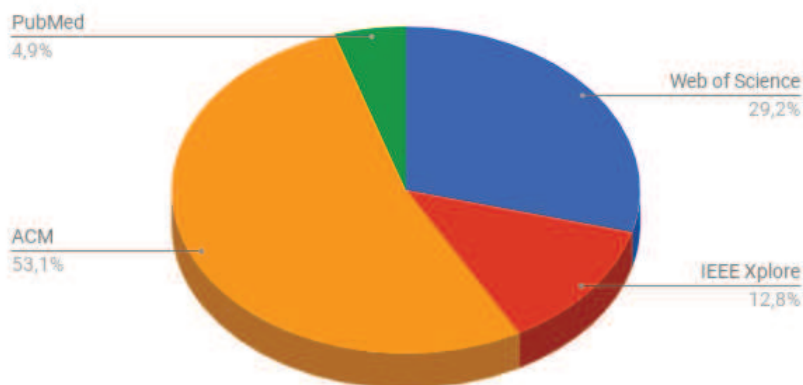
Tabela 2: Comparação dos resultados nas bases de dados.

	<i>Ontology+ BN</i>	<i>Ontology+BN+ NLP</i>	<i>Ontology +BN+Health</i>	<i>Ontology+ BN+Data Extraction</i>	<i>Ontology+ BN+NLP+ Health</i>
<i>Web of Science</i>	184	2	8	3	0
<i>IEEE Xplore</i>	81	5	8	19	1
<i>ACM</i>	335	6	23	0	0
<i>PubMed</i>	31	0	4	0	0
	631	13	43	22	1

Fonte: Elaborado pelo autor.

Para representar a diferença de resultados entre as bases selecionadas, a Figura 4 ilustra um gráfico comparativo da quantidade de publicações encontradas. A seguir, serão apresentados os resultados de acordo com cada fase de seleção.

Figura 4: Gráfico comparativo de resultados entre as bases selecionadas.



Fonte: Elaborado pelo autor.

3.5.1. Fase 1 - Critérios de inclusão/exclusão;

Como os motores de busca disponibilizam opções de filtros, antes da exportação dos resultados, foram aplicados os critérios de exclusão referentes ao ano de publicação, artigo científico publicado e o idioma, excluindo do retorno da busca todos os artigos publicados antes do ano de 2011 e que não estejam escritos em inglês. Os resultados foram exportados no formato *BibTex* e inseridos na ferramenta *StArt*, como mostra a Figura 5.

Figura 5: Resultados inseridos na ferramenta StArt.

General information

String: ("ontology") AND ("bayesian network") OR ("health") OR ("natural language processing" OR "NLP") OR ("data extraction")

Search machine: PubMed Number of papers: 30 Date of the search: 08/23/2017

Observations: Import Reference File
BIBTEX MEDLINE

ID Paper Title Author Status/Selection Status/Extraction Priority Reading Score Remove ALL duplicated papers

ID Paper	Title	Author	Year	Status/Selection	Status/Extraction	Reading Priority	Score
378	Understanding multicellular function and disease with human ...	Greene CS ; Krishnan A ; Wong ...	2015	Unclassified	Unclassified	Low	0
377	SWARMS Ontology: A Common Information Model for the Co...	Li X ; Bilbao S ; Martin-Wanton T...	2017	Unclassified	Unclassified	Low	38
376	Statistical algorithms for ontology-based annotation of scient...	Chakrabarti C ; Jones TB ; Luge...	2014	Unclassified	Unclassified	Low	17
375	Signaling network prediction by the Ontology Fingerprint enh...	Qin T ; Tsol LC ; Sims KJ ; Lu X ; ...	2012	Unclassified	Unclassified	Low	20
374	Role of SPI-1 secreted effectors in acute bovine response to...	Lawhon SD ; Khare S ; Rossetti ...	2011	Unclassified	Unclassified	Low	6
373	Reconstruction of biological networks by incorporating prior k...	Pei B ; Shin DG	2012	Unclassified	Unclassified	Low	14
372	Quantitative utilization of prior biological knowledge in the Ba...	Gao S ; Wang X	2011	Unclassified	Unclassified	Low	11
370	Prioritizing breast cancer subtype related miRNAs using miRN...	Hua L ; Zhou P ; Li L ; Liu H ; Ya...	2013	Unclassified	Unclassified	Low	3
369	Predicting gene ontology annotations of orphan GWAS gene...	Kuppuswamy U ; Ananthasubra...	2014	Unclassified	Unclassified	Low	11
368	Ontology driven decision support for the diagnosis of mild co...	Zhang X ; Hu B ; Ma X ; Moore P...	2014	Unclassified	Unclassified	Low	14
367	Network-based auto-probit modeling for protein function pre...	Jiang X ; Gold D ; Kolaczyk ED	2011	Unclassified	Unclassified	Low	3
366	Multi-comparative systems biology analysis reveals time-cour...	Adams LG ; Khare S ; Lawhon S...	2011	Unclassified	Unclassified	Low	9
365	Integrative transcriptomics and proteomics analysis of longis...	Liu S ; Han W ; Jiang S ; Zhao C...	2016	Unclassified	Unclassified	Low	0
363	Information-theoretic evaluation of predicted ontological ann...	Clark WT ; Radivojac P	2013	Unclassified	Unclassified	Low	9
362	Improving protein complex prediction by reconstructing a hig...	Taghipour S ; Zarrineh P ; Ganjt...	2017	Unclassified	Unclassified	Low	3
361	Identification of Novel Pathways in Plant Lectin-Induced Can...	Shi Z ; Sun R ; Yu T ; Liu R ; Che...	2016	Unclassified	Unclassified	Low	0
360	Identification of key genes associated with the effect of estr...	Zhang ST ; Zuo C ; Li WN ; Fu X...	2016	Unclassified	Unclassified	Low	3
359	Heterogeneous data integration by tree-augmented naive B...	Lin X ; Chen XW	2013	Unclassified	Unclassified	Low	6
358	Gene Ontology consistent protein function prediction: the FA...	Kourmpetis YA ; van Dijk ADJ ; T...	2013	Unclassified	Unclassified	Low	14
357	Finding pathway-modulating genes from a novel Ontology Fl...	Qin T ; Matmati N ; Tsol LC ; Mo...	2014	Unclassified	Unclassified	Low	11
356	Evaluating the effect of unbalanced data in biomedical docu...	Izaz R ; Pavon R ; Rehring-lato...	2011	Unclassified	Unclassified	Low	6

Path:

Name:

Size:

Fonte: Elaborado pelo autor.

Após a inserção das publicações no software, automaticamente foram excluídos 6 artigos detectados como duplicados entre as bases selecionadas. Foi realizada uma revisão manual e detectou-se mais 67 artigos duplicados. Neste momento, os mesmos já estavam sendo classificados de acordo com os critérios de inclusão e exclusão previamente definidos no protocolo. Decidiu-se que alguns critérios seriam realizados nas próximas fases, a fim de otimizar o tempo da análise, tais como a validação da ferramenta *ToS*, que será realizada na Fase 2 e a verificação da disponibilidade dos artigos na íntegra e as formas de validações, que serão usadas na Fase 4. A ferramenta *Start* permite a classificação do artigo, se ele deve ser aceito ou rejeitado para a próxima fase, e ainda informa por qual critério de exclusão ou inclusão foi aceito ou rejeitado. Após a validação desta etapa, 558 artigos foram aceitos para a Fase 2.

3.5.2. Fase 2 – Validação ToS

Para a realização da segunda fase, foram exportadas as publicações da base de dados *Web Of Sciences* no formato de “texto sem formato” e importadas na ferramenta *ToS*. Após a validação do software, foram considerados relevantes 69 artigos dos 184 encontrados na base. 115 artigos foram rejeitados nessa fase pelo critério de exclusão “não validado pelo *ToS*”. As publicações consideradas como estrutura desta área de pesquisa (tronco) são aquelas que dão suporte a pesquisa, como: Hashimoto et al. (2011); Lin et al. (2014); Arsene, Dumitrache & Mihiu (2011); Grubišić, Stankov & Peraić (2013); Fenz (2012).

Os artigos que apresentam as tendências atuais (folhas) envolvendo ontologias e RBs são: Qinghong (2012); Sayed & Lohse (2014), dentre outros.

As bases de dados *IEEE Xplore*, *ACM* e *PubMED* não foram compatíveis com o software *ToS*. Sendo assim, as suas publicações passaram automaticamente para a Fase 3. Após a validação, 443 artigos foram aceitos para a Fase 3.

3.5.3. Fase 3 – Título, palavras-chave e resumo

Na fase 3, realizou-se a leitura dos títulos, palavras-chave e resumo, analisando a presença dos termos “*ontology*” e “*bayesian network*”. Nesta etapa, foram adicionados dois critérios de exclusão que não foram previstos/esperados na criação do protocolo:

- Não se enquadrar na abordagem adotada (devido a quantidade de artigos que, embora utilizem tecnologias semelhantes, não estejam dentro da abordagem de ontologias e RBs);
- Não possuir relação entre os dois assuntos obrigatórios da pesquisa (uma grande quantidade de publicações tratava ou apenas de ontologias ou somente de RBs, desviando do objetivo principal deste trabalho).

Após a leitura dos resumos e a validação dos novos critérios de exclusão, 93 publicações foram aceitas para a quarta fase.

3.5.4. Fase 4 - Introdução e conclusão

Na quarta fase, foram lidas as introduções e conclusões dos artigos aprovados na fase anterior, buscando obter uma ideia mais clara sobre o foco dos artigos que possam ter sido aceitos sem tanta exatidão, bem como, analisar a qualidade dos resultados obtidos. Verificou-se que 60 artigos não estavam apresentando a proposta buscada neste trabalho e foram excluídos. Todos os artigos selecionados foram encontrados na íntegra e possuem alguma forma de validação, não havendo nenhuma exclusão para estas validações. Após a leitura e a análise, 33 publicações foram aprovadas para a última fase, onde foi realizada a leitura completa dos mesmos.

3.5.5. Fase 5 - Leitura integral

A fase 5 consistiu na leitura completa dos 33 artigos que restaram, visando extrair todas as informações das publicações encontradas e organizá-las para que, assim, possam ser respondidas as perguntas pré-definidas no protocolo. A ferramenta *StArt* disponibiliza uma aba de extração de dados dos artigos, onde foram cadastradas as perguntas e respostas propostas, classificando conforme cada publicação e as informações que se buscava extrair, como mostra a Figura 6.

Figura 6: Cadastro de extração de dados na ferramenta StArt.

770 - Ontology based disease diagnosis system with probabilistic inference

Study Data Selection Data **Data Extraction Form** Similar Studies

As publicações apresentam algum tipo de validação? Sim

Quais foram as técnicas mais utilizadas?

Que tipos de metodologias foram usadas?

Quais autores foram a base para a fundamentação teórica?

Qual a ligação entre os temas propostos?

Teve algum protótipo desenvolvido? Sim

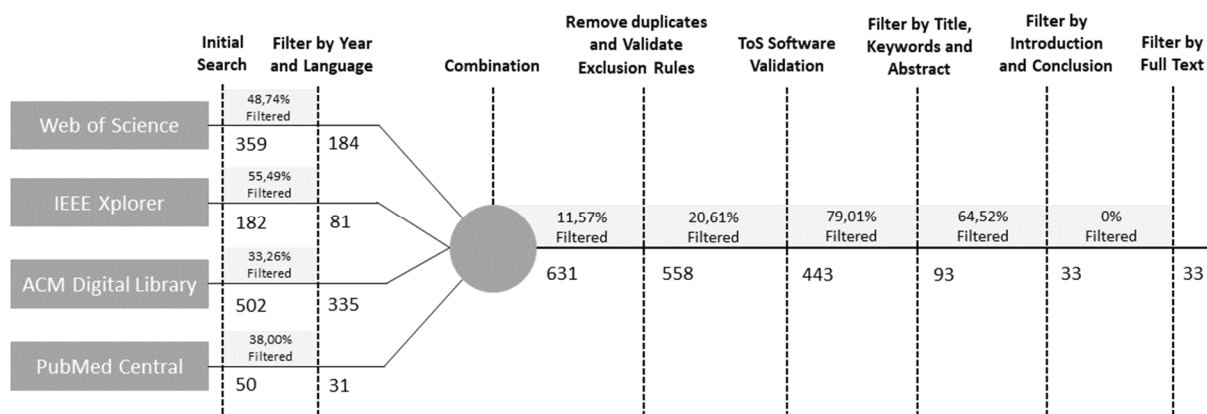
Fonte: Elaborado pelo autor.

Após a leitura das publicações e a extração dos dados concluída, iniciou-se a análise dos resultados obtidos. Alguns destes resultados foram simplificados no formato de gráficos e Tabelas. A apresentação dos resultados será exibida no próximo item.

3.6. Resultados da Revisão Sistemática

Mediante a conclusão da leitura e a extração das informações de 33 publicações selecionadas para a etapa final, é possível responder as perguntas de pesquisa levantadas no protocolo desta revisão sistemática. Para exemplificar o processo ocorrido na Seção anterior, a Figura 7 ilustra um fluxograma geral dos resultados das classificações dos artigos em cada fase de seleção.

Figura 7: Fluxograma dos resultados das fases de seleção.



Fonte: Elaborado pelo autor.

A Tabela ilustrada na Figura 8 mostra os artigos considerados como relevantes para esta revisão sistemática e o seu ano de publicação.

Figura 8: Tabela com o título e ano dos artigos selecionados.

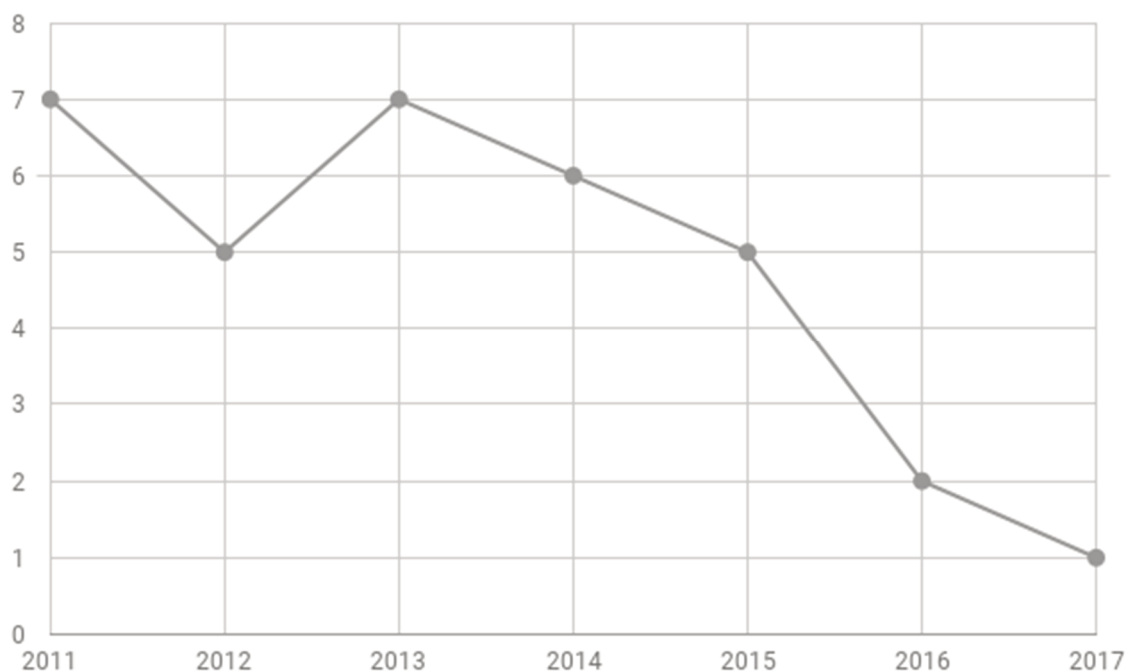
Título	Ano
A framework for proactive assistance: Summary	2014
A framework of ontology-based knowledge information processing for change detection in remote sensing data	2011
A method of constructing user interest model based on semantics	2012
A novel approach to power transformer fault diagnosis based on ontology and Bayesian network	2014
A URREF interpretation of Bayesian network information fusion	2014
About the studying on the uncertainty of ontology on the BayesianNetwork	2014
Active learning of causal Bayesian networks using ontologies: A case study	2013
An ontology-based approach for constructing Bayesian networks	2012
Autonomie decision making based on bayesian networks and ontologies	2016
Bayesian information extraction network for Medline abstract	2013
Bayesian knowledge-driven ontologies: Intuitive uncertainty reasoning for semantic networks	2011
Depression Diagnosis Based on Ontologies and Bayesian Networks	2013
Developing Bayesian networks from a dependency-layered ontology: A proof-of-concept in radiation oncology	2017
Development of a novel asset management system for power transformers based on ontology	2013
Event recognition using a semantic-probabilistic network	2015
Expert system for medicine diagnosis using software agents	2015
Generating Bayesian networks from medical ontologies	2013
Human action recognition using a semantic-probabilistic network	2015
Integration of Ontologies and Bayesian Networks for Maritime Situation Awareness	2012
Medicine expert system dynamic Bayesian Network and ontology based	2011
Mining Bayesian networks out of ontologies	2012
Multiviews ontologies based reasoning for medical diagnosis in VDS	2011
Ontologies and Bayesian Networks in Medical Diagnosis	2011
Ontologies for probabilistic situation assessment in the maritime domain	2013
Ontology based approach to Bayesian student model design	2013
Ontology based Bayesian software process improvement	2014
Ontology based disease diagnosis system with probabilistic inference	2016
Ontology based EMR for decision making in health care using SNOMED CT	2012
Ontology-based context-aware model by applying Bayesian network	2015
Ontology-driven generation of Bayesian diagnostic models for assemblysystems	2014
SemCaDo: A serendipitous strategy for causal discovery and ontologyevolution	2015
The Development of Ontology Information System Based on Bayesian Networkand Learning	2011
Towards Automatic Generation of Ontology-Based Antipattern Bayesian Network Models	2011

Fonte: Elaborado pelo autor.

Para uma análise detalhada, gerou-se alguns gráficos comparativos com relação ao ano das publicações selecionadas e as bases encontradas. Um detalhe importante para ressaltar é que o ano de 2017 é referente somente as publicações realizadas até o mês de agosto, data em

que foi realizada a busca desta revisão. A Figura 9 apresenta o gráfico comparativo do ano das publicações.

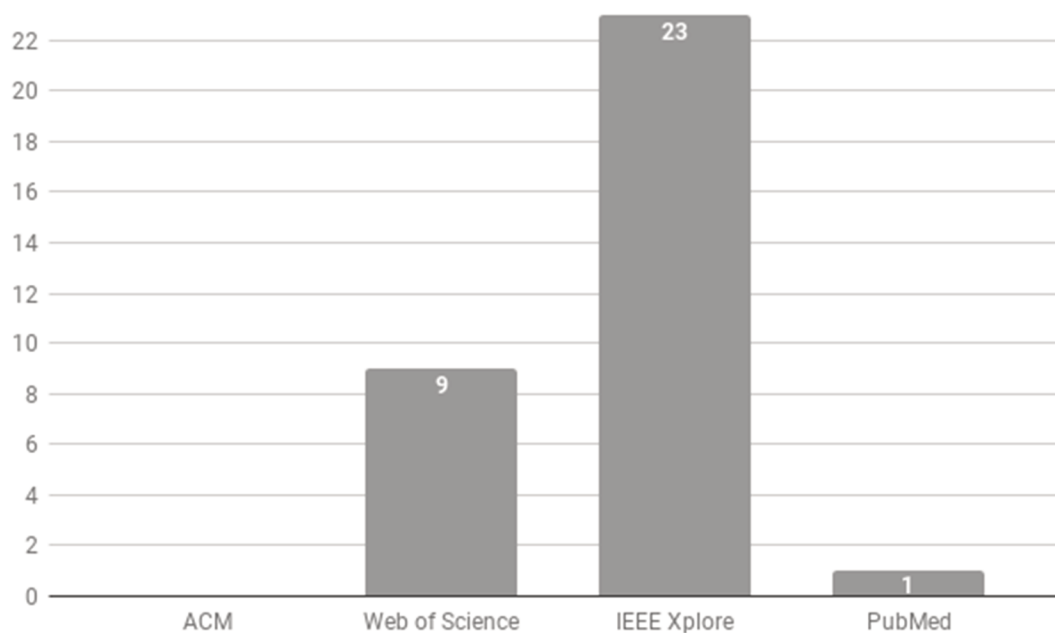
Figura 9: Comparativo do ano de publicação dos artigos selecionados.



Fonte: Elaborado pelo autor.

Através do gráfico é possível verificar que o estudo de extração de dados de ontologias e RBs obteve destaque nos anos de 2011 e 2013, tendo atualmente diminuído o número de trabalhos relacionados ao tema, visto que houve poucas publicações no período de 2016 ao atual. A Figura 10 ilustra o gráfico comparativo das bases de dados.

Figura 10: Comparativo da base de dados dos artigos selecionados.



Fonte: Elaborado pelo autor.

A base de dados da *IEEE* obteve grande êxito quanto as publicações desta pesquisa, seguindo da base *Web Of Science*. Vale destacar que a maioria dos artigos duplicados ocorreram entre estas duas bases, fazendo com que a *IEEE* obtivesse mais relevância na seleção pela quantidade de informações que oferecia. Mas, de forma geral, pode-se dizer que ambas as bases tiveram uma proporção semelhante quanto ao êxito dos resultados. Já as bases *PubMed* e *ACM* não tiveram muita relevância para este estudo, tendo, no final, uma seleção de um a nenhum artigo. Caso haja futuras revisões, pode-se considerar a hipótese de substituí-las.

Também foi possível listar as revistas onde os artigos selecionados foram publicados, visando pesquisar a relevância da revista para a possibilidade de uma publicação deste trabalho futuramente. Verificou-se que a revista *Expert Systems With Applications* possui a maior quantidade de publicações sobre o tema da pesquisa. A Figura 11 ilustra uma Tabela com todas as revistas onde foram publicados os artigos.

Figura 11: Revistas das quais os artigos selecionados foram publicados.

Revista
International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)
International Conference on Information Fusion (FUSION)
India International Conference on Information Processing (IICIP)
Hawaii International Conference on System Sciences
International Conference on Computer Science Education (ICCSE)
Advances In Multimedia, Software Engineering And Computing
Applied Science, Materials Science And Information Technologies Inindustry
Computing Colombian Conference
Data & Knowledge Engineering
Expert Systems With Applications
IEEE International Conference on Semantic Computing
IEEE International Conference on Systems, Man, and Cybernetics
IEEE International Geoscience and Remote Sensing Symposium
IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness
IEEE International Symposium on Intelligent Systems and Informatics
IEEE PES Asia-Pacific Power and Energy Engineering Conference
Information Technologies in Innovation Business Conference
International Conference on Emerging Trends in Networks and Computer Communications
International Conference on Recent Trends in Information Technology
International Conference on Software Engineering and Applications
International Joint Conference on Neural Networks
International Journal Of Advanced Manufacturing Technology
Journal Of Intelligent Information Systems
Knowledge-based Systems
Medical physics
Ninth International Conference on Software Engineering Research, Management and Applications
World Congress on Computer and Information Technology

Fonte: Elaborado pelo autor.

3.6.1. Perguntas respondidas

Através dos resultados da revisão sistemática, foi possível responder às seguintes perguntas:

1. Quais foram as técnicas mais utilizadas?
 - Criação de uma ontologia e geração da RB manualmente (13 artigos).

- Geração da RB de uma ontologia existente utilizando um *plugin* (6 artigos).
 - Criação de uma ontologia e geração da RB conforme o nível hierárquico da ontologia (5 artigos).
 - Geração da RB manualmente de uma ontologia existente (3 artigos).
 - Criação de uma ontologia e geração da RB utilizando um *plugin* (2 artigos).
 - Criação de uma ontologia e geração da RB utilizando um software (2 artigos).
 - Geração da RB de uma ontologia existente utilizando um software (1 artigo).
 - Avaliação de técnicas existentes (1 artigo).
2. Que tipos de metodologias foram usadas?
 - Criação de ontologias e RBs para aplicar em um software (17 artigos).
 - Criação de RBs de uma ontologia existente para aplicar em um software (7 artigos).
 - Criação de ontologias e RBs para experimentação (7 artigos).
 - Criação de RBs de uma ontologia existente para experimentação (1 artigo).
 - Análise de técnicas existentes (1 artigo).
 3. Qual a ligação entre os temas propostos?
 - Extração de uma RB manual de uma ontologia (18 artigos).
 - Extração semiautomática de RB de uma ontologia (11 artigos).
 - Extração automática de RB de uma ontologia (3 artigos).
 - Análise de dados (1 artigo).
 4. Teve algum protótipo desenvolvido?
 - Sim (29 artigos).
 - Não informado (4 artigos).
 5. Que formas de validação foram utilizadas?
 - Professores (11 artigos).
 - Especialistas (10 artigos).
 - Empresa (6 artigos).
 - Usuários (5 artigos).
 - Questionários (1 artigo).
 6. Os artigos foram aplicados à área da saúde?
 - Não (23 artigos).
 - Sim (10 artigos).
 7. A proposta teve resultados positivos?
 - Sim (27 artigos).
 - Não (6 artigos).
 8. Quais foram as ferramentas utilizadas?
 - Não informado (13 artigos).
 - Protege (11 artigos).
 - Netica (6 artigos).
 - Bayesian Genie (2 artigos).
 - Jess Reasoner (1 artigo).
 - RBTab (1 artigo).
 - RHugin (1 artigo).
 - Bayesian Domain Network (1 artigo).

- Norsys (1 artigo).
- Graph API (1 artigo).
- Probt API (1 artigo).
- Owl API (1 artigo).
- Bayes OWL (1 artigo).
- Onto Bayes (1 artigo).

3.6.2. Análise dos artigos

Conforme os dados obtidos nos artigos, foi possível responder a todas as perguntas pré-estabelecidas nessa revisão sistemática. Essa Seção apresentará uma análise aprofundada de cada questão, buscando apresentar os resultados, causas e observações sobre os artigos selecionados para a fase final.

3.6.2.1. Técnicas utilizadas

A maneira utilizada para saber como foram realizadas as extrações de dados de ontologias para conversão em RBs foi analisando as técnicas utilizadas em cada trabalho. A Tabela 3 demonstra quais são as técnicas e os autores que as utilizaram.

Tabela 3: Técnicas utilizadas e seus autores.

Técnica	Artigos
Desenvolvimento de Ontologias e RB manualmente	Pilato (2012), Bucci, Sandrucci & Vicario (2011), Bibi et al. (2014), Agarwal, Verma & Mallik (2016), Julina & Thenmozhi (2012), Mok & Min (2015), Andrea & Franco (2012), Grubišić, Stankov & Peraić (2013), Ding (2011), Hashimoto et al. (2011), Qinghong (2012), Armand, Filliat & Ibañez-Guzman (2014), Cong (2014).
Geração de RB via Plugin	Settas, Cerone & Fenz (2011), Kalet et al. (2017), Fenz (2012), Arsene, Dumitrache & Mihiu (2015), Arsene, Dumitrache & Mihiu (2011), Aguilar (2016).
Geração de RB via nível hierárquico da Ontologia	Kovalenko & Antoshchuk (2015), Messaoud et al. (2015), Kovalenko, Antoshchuk & Hodovychenko (2015), Mannai (2013), Fischer & Beyerer (2013).
Desenvolvimento de RBs manualmente	Santos, Jurmain (2011), Fujita, Kurematsu, Hakura (2011), Sayed & Lohse (2014).
Desenvolvimento de Ontologia e geração de RB via Plugin	Lin et al. (2014), Chang, Hung & Juang (2013).
Desenvolvimento de Ontologia e geração de RB via Software	Rodrigues, Bez & Flores (2013). Yan et al. (2013).

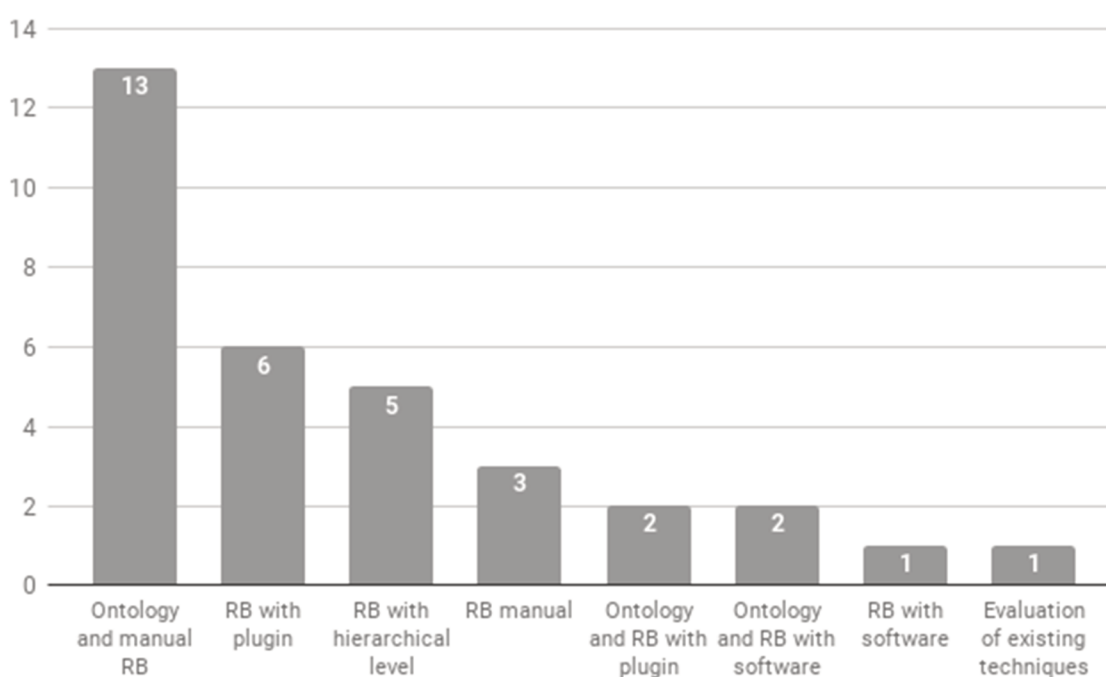
Geração de RB via Software Pavlin & Laskey (2014).

Avaliação de técnicas existentes Messaoud, Leray & Amor (2013).

Fonte: Elaborado pelo autor.

Em relação às técnicas mais utilizadas, verificou-se que 39,4% dos autores criaram suas próprias ontologias e RBs manualmente e 9,1% extraíram uma RB manualmente de uma ontologia já existente para aplicar em suas propostas, das quais suas soluções não são relevantes, pois é justamente o problema que este trabalho busca resolver. Um número de 18,2% dos artigos extraíram uma RB de uma ontologia já existente e 6,1% criaram suas próprias ontologias e extraíram uma RB utilizando um *plugin* vinculado ao software Protégé, como o Nética, RBTab e outros. Porém, seus resultados, embora positivos, não foram satisfatórios, visto que os *plugins* realizam a conversão buscando um padrão hierárquico da ontologia, fazendo com que haja muito trabalho manual e seja necessária a interferência de um especialista. Um total de 15,2% dos artigos indica que criaram uma ontologia e extraíram uma RB conforme o nível hierárquico da mesma, conforme a criação de classes e subclasses, porém, esse modelo não é eficaz, visto que as ontologias existentes não seguem esse padrão. 6,1% criaram uma ontologia e extraíram uma RB e 3% geraram uma RB de uma ontologia já existente utilizando um software ou algoritmo criado pelos autores, mas que geraram resultados negativos, ou por não atingir o objetivo final, ou por criar um sistema viciado que só funciona para a ontologia criada. Por fim, 3,0% dos artigos apenas apresentaram uma avaliação de técnicas existentes na literatura. A Figura 12 apresenta um gráfico com esses resultados.

Figura 12: Comparativo de técnicas utilizadas nos artigos.



Fonte: Elaborado pelo autor.

3.6.2.2. Metodologias usadas

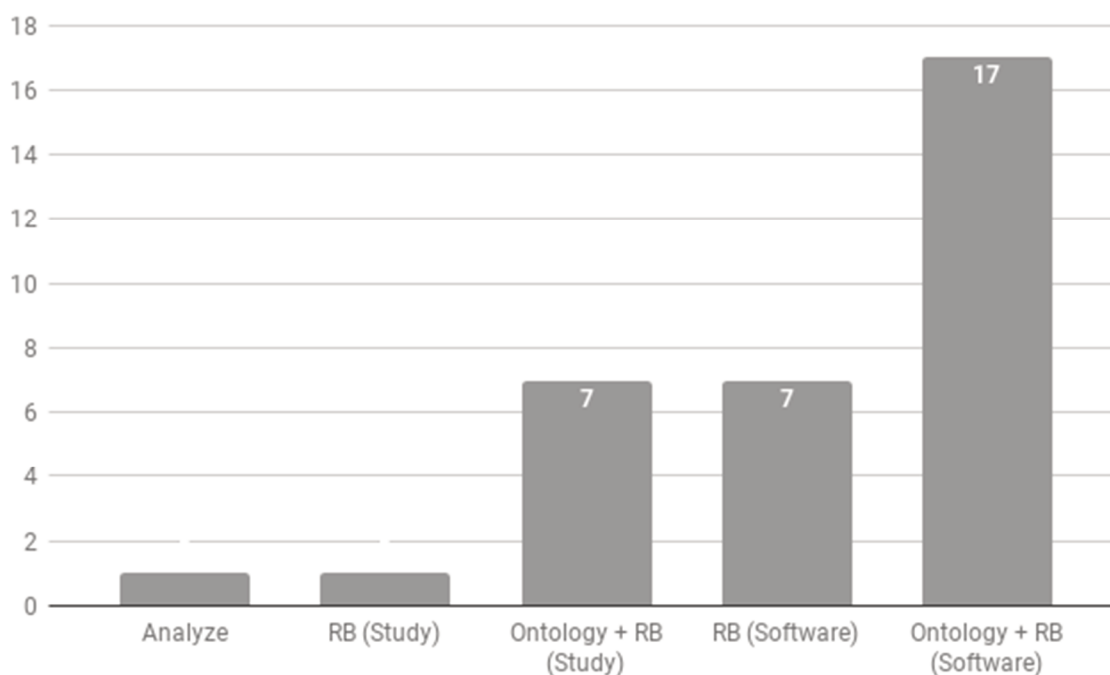
Buscou-se analisar as metodologias utilizadas pelos autores, visando identificar a aplicação da qual foram executadas as ontologias e/ou RBs. A Tabela 4 demonstra quais são as metodologias usadas e os autores que as utilizaram.

Tabela 4: Metodologias usadas e seus autores.

Metodologia	Artigos
Aplicação de Ontologia e RB em um software	Kovalenko & Antoshchuk (2015), Pilato (2012), Fujita, Kurematsu, Hakura (2011), Fischer & Beyerer (2013), Bibi et al. (2014), Julina & Thenmozhi (2012), Mok & Min (2015), Grubišić, Stankov & Peraić (2013), Sayed & Lohse (2014), Hashimoto et al. (2011), Qinghong (2012), Rodrigues, Bez & Flores (2013), Kovalenko, Antoshchuk & Hodovychenko (2015), Chang, Hung & Juang (2013), Lin et al. (2014), Fenz (2012), Santos, Jurmain (2011).
Aplicação de RB em um software	Settas, Cerone & Fenz (2011), Arsene, Dumitrache & Mihiu (2015), Arsene, Dumitrache & Mihiu (2011), Aguilar (2016), Pavlin & Laskey (2014), Kalet et al. (2017), Cong (2014).
Ontologia e RB para experimentação	Bucci, Sandrucci & Vicario (2011), Agarwal, Verma & Mallik (2016), Ding (2011), Mannai (2013), Yan et al. (2013), Armand, Filliat & Ibañez-Guzman (2014). Messaoud et al. (2015).
RB para experimentação	Andrea & Franco (2012).
Análise de técnicas	Messaoud, Leray & Amor (2013).

Fonte: Elaborado pelo autor.

Nas metodologias usadas, 44,8% dos autores aplicaram as ontologias e RBs e 20,7% aplicaram somente as RBs a um software, onde a maioria se encaixa na categoria de simuladores. 17,2% dos artigos utilizaram ontologias e RBs e 10,3% apenas as RBs a fim de experimentação, visando buscar técnicas e algoritmos eficientes. 6,9% dos trabalhos apenas analisaram as técnicas existentes na literatura. A Figura 13 ilustra um gráfico com esses resultados.

Figura 13: Comparativo de metodologias usadas.

Fonte: Elaborado pelo autor.

3.6.2.3. Ligação com os temas propostos

A maneira utilizada para saber como foi realizada a extração de dados de ontologias para conversão em RBs foi analisando as técnicas utilizadas em cada trabalho. A Tabela 5 demonstra quais são as técnicas e os autores que as utilizaram.

Tabela 5: Ligação entre os temas propostos e seus autores.

Ligação entre os temas propostos	Artigos
Geração manual de RB a partir de ontologias	Pilato (2012), Fujita, Kurematsu, Hakura (2011), Bucci, Sandrucci & Vicario (2011), Bibi et al. (2014), Agarwal, Verma & Mallik (2016), Julina & Thenmozhi (2012), Mok & Min (2015), Andrea & Franco (2012), Grubišić, Stankov & Peraić (2013), Ding (2011), Hashimoto et al. (2011), Qinghong (2012), Armand, Filliat & Ibañez-Guzman (2014), Mannai (2013), Santos, Jurmain (2011), Fischer & Beyerer (2013), Sayed & Lohse (2014). Cong (2014).
Geração semiautomática de	Settas, Cerone & Fenz (2011), Kalet et al. (2017), Fenz (2012), Arsene, Dumitrache & Mihi (2015), Arsene, Dumitrache & Mihi (2011), Yan

RB a partir de ontologias	et al. (2013), Pavlin & Laskey (2014), Aguilar (2016), Rodrigues, Bez & Flores (2013), Lin et al. (2014), Chang, Hung & Juang (2013).
Geração automática de RB a partir de ontologias	Messaoud et al. (2015), Kovalenko & Antoshchuk (2015), Kovalenko, Antoshchuk & Hodovychenko (2015).
Análise de dados	Messaoud, Leray & Amor (2013).

Fonte: Elaborado pelo autor.

Na ligação com os temas propostos, buscou-se identificar qual o tipo de extração de RB de uma ontologia que os autores realizaram. Foram detectados três tipos de técnicas de extração: manual, semiautomática e automática.

Um total de 54,5% dos artigos apresenta uma extração de forma manual, visualizando a ontologia e criando a RB conforme seu conhecimento e experiência. Na primeira etapa, é desenvolvida uma ontologia que armazena o conhecimento de um determinado domínio. Alguns autores também utilizaram um modelo ontológico pronto disponível na web. Dessas ontologias, foram extraídos os conceitos e as relações centrais, gerando alternativas de representações de conhecimento das propriedades das classes, utilizando também o restante do modelo ontológico como informações extras de contexto. A partir das representações citadas, inicia-se a criação da estrutura da RB. Também há necessidade de detalhamento para gerar as variáveis, a fim de construir valores de probabilidade que sejam viáveis e corretos. As metodologias analisadas nos trabalhos utilizam fórmulas matemáticas para a obtenção dos valores probabilísticos ou não citam a obtenção de valores de probabilidades. Os resultados, embora eficientes, são demorados e necessitam de um especialista que possua um alto conhecimento na construção de RBs e do domínio expressado na ontologia, problema que este trabalho busca resolver.

Um percentual de 33,3% dos autores realizou uma extração semiautomática, geralmente utilizando um *plugin* ou software, onde, durante e após a extração dos dados, deve-se realizar ajustes e adequações para chegar ao resultado final correto. Os autores buscaram, através desta técnica, desenvolver automaticamente alguns conceitos da RB com base em uma ontologia, necessitando da interação de um usuário para completar os conceitos ausentes ou deixando-os inacabados para uma futura validação humana da RB. Os conceitos que não são obtidos através da técnica diferem-se de valores qualitativos e quantitativos, onde não há um padrão com que os *plugins*/softwares puderam trabalhar para o tratamento.

Alguns autores buscaram extrair a RB utilizando uma ontologia extra com os conceitos registrados para a geração da mesma, como nodos, variáveis e outros, mas não citam sobre a criação de variáveis probabilísticas. Outra questão é que a abordagem necessita de um tratamento da ontologia que se busca extrair a RB, tornando um pouco mais restrita a sua utilização. Isso mostra o grau de importância do conhecimento de um usuário necessário para a realização da tarefa quando utilizado o modelo ontológico para a construção da RB.

Outro caso é de autores que propuseram uma técnica em que, primeiramente, é necessário um usuário para discernir quais são os conceitos que estarão presentes nos nodos da RB. Após, o mesmo deve identificar as relações entre os conceitos dos nodos e o sentido do arco. Em seguida, o usuário deve identificar os valores das variáveis probabilísticas, visto que, normalmente, as ontologias não modelam essa informação. Mediante todos esses ajustes, os autores utilizaram uma extensão/*plugin* para ser gerada a RB. Este processo facilitou a construção de RBs, mas a técnica possui limitações relacionadas a forma semiautomática do

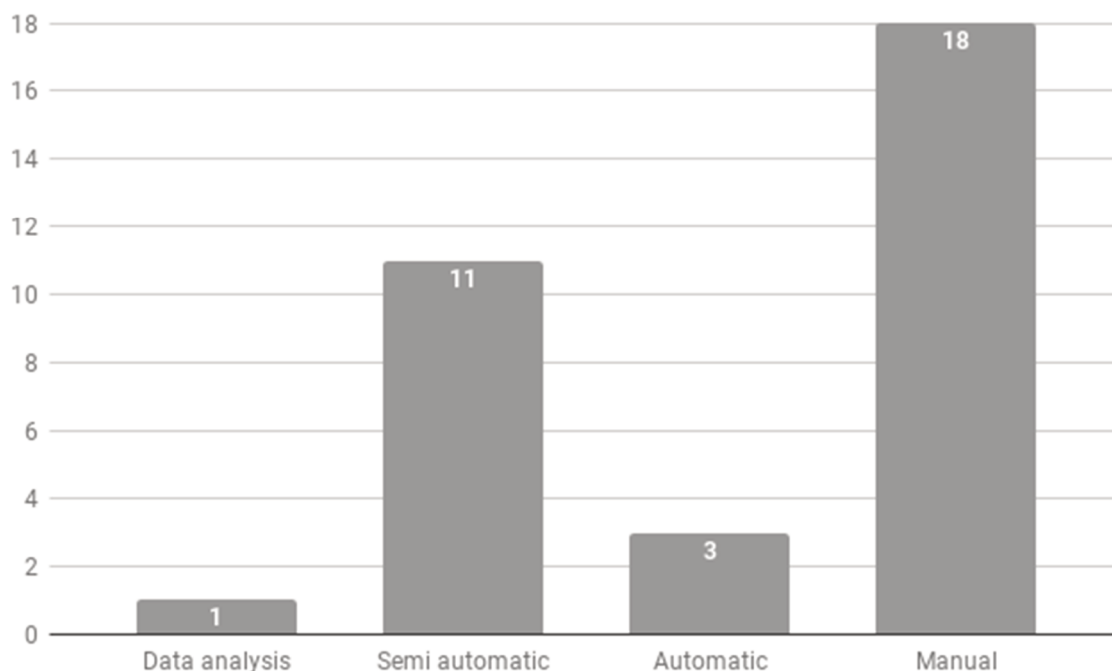
processo, pois exige que o usuário trabalhe em praticamente todo o processo, além de ser necessário o conhecimento para a definição dos conceitos e das funções matemáticas para a geração dos valores das variáveis probabilísticas.

Apenas 9,1% dos trabalhos apresentaram uma extração de RB de forma automática, sem a necessidade de alterações de um usuário ou especialista. Essa forma foi possível visto que todos os trabalhos utilizaram o conceito de hierarquia da ontologia, cujos resultados foram negativos ou funcionam apenas na ontologia criada pelos autores. Para atingir a extração completa da ontologia, esta técnica exige cenários limitados, tanto nas ontologias, utilizando conceitos específicos que normalmente não são utilizados como padrão nas ontologias existentes ou construindo um modelo ontológico para extrair o seu modelo, como pela geração da RB, que acaba em uma estrutura pré-definida.

Em uma abordagem, os autores utilizam o OntoBayes (YANG & CALMET, 2005), um protótipo que integra conceitos de RBs na ontologia OWL, possibilitando assim, a inserção das variáveis de probabilidades no modelo ontológico e sua construção, classificando a hierarquia em classes, propriedades e instâncias. Com isso, pode-se desenvolver um modelo ontológico com grandes domínios de representação de conhecimento e uma extração de RB automaticamente através de um algoritmo. Porém, esse processo é limitado apenas a modelos ontológicos desenvolvidos especificamente para a construção de RBs e utilizando o *plugin*, o que não é o caso da maioria das ontologias disponibilizadas na web.

Outro tipo de abordagem descrita pelos autores é a utilização da estrutura hierárquica da ontologia (classes, propriedades e instâncias), onde a ontologia é construída considerando que as classes serão o nó pai (topo da RB), as propriedades, a ligação entre os nós e, as instâncias, os nós filhos (final da RB). Esse modelo se restringe a uma ontologia criada especificamente para a extração de RBs, visto que nenhuma ontologia já desenvolvida segue essa regra e, normalmente, são desenvolvidas de diferentes formas. Além disso, os valores das variáveis de probabilidade são inseridos manualmente, mediante não estar presente no domínio da ontologia. À extração levando em conta a hierarquia de classes e subclasses segue o mesmo conceito que o anterior.

Além da identificação dos tipos de extração, verificou-se que 3% dos artigos apenas realizaram uma análise de dados. Os resultados dessa Seção são ilustrados na Figura 14.

Figura 14: Comparativo de tipos de extração.

Fonte: Elaborado pelo autor.

3.6.2.4. Protótipos desenvolvidos

Identificou-se os autores que criaram e utilizaram algum tipo de protótipo para realizar a sua pesquisa. A Tabela 6 mostra se foram desenvolvidos protótipos e os autores que os apresentaram.

Tabela 6: Protótipos desenvolvidos e seus autores.

Protótipo Artigos	
Sim	Kovalenko & Antoshchuk (2015), Pilato (2012), Fujita, Kurematsu, Hakura (2011), Bucci, Sandrucci & Vicario (2011), Fischer & Beyerer (2013), Bibi et al. (2014), Agarwal, Verma & Mallik (2016), Julina & Thenmozhi (2012), Mok & Min (2015), Settas, Cerone & Fenz (2011), Kalet et al. (2017), Fenz (2012), Arsene, Dumitrache & Mihu (2015), Arsene, Dumitrache & Mihu (2011), Andrea & Franco (2012), Grubišić, Stankov & Peraić (2013), Sayed & Lohse (2014), Messaoud et al. (2015), Qinghong (2012), Lin et al. (2014), Aguilar (2016), Santos, Jurmain (2011), Yan et al. (2013), Rodrigues, Bez & Flores (2013), Armand, Filliat & Ibañez-Guzman (2014), Kovalenko, Antoshchuk & Hodovychenko (2015), Chang, Hung & Juang (2013), Pavlin & Laskey (2014), Cong (2014).
Não	Ding (2011), Hashimoto et al. (2011), Messaoud, Leray & Amor (2013), Mannai (2013).

Fonte: Elaborado pelo autor.

Em relação ao desenvolvimento de protótipos, verificou-se que 97,9% dos autores criaram um protótipo para seu experimento, seguindo de 12,1% dos trabalhos que não construíram ou não ilustraram um protótipo.

3.6.2.5. Formas de validação

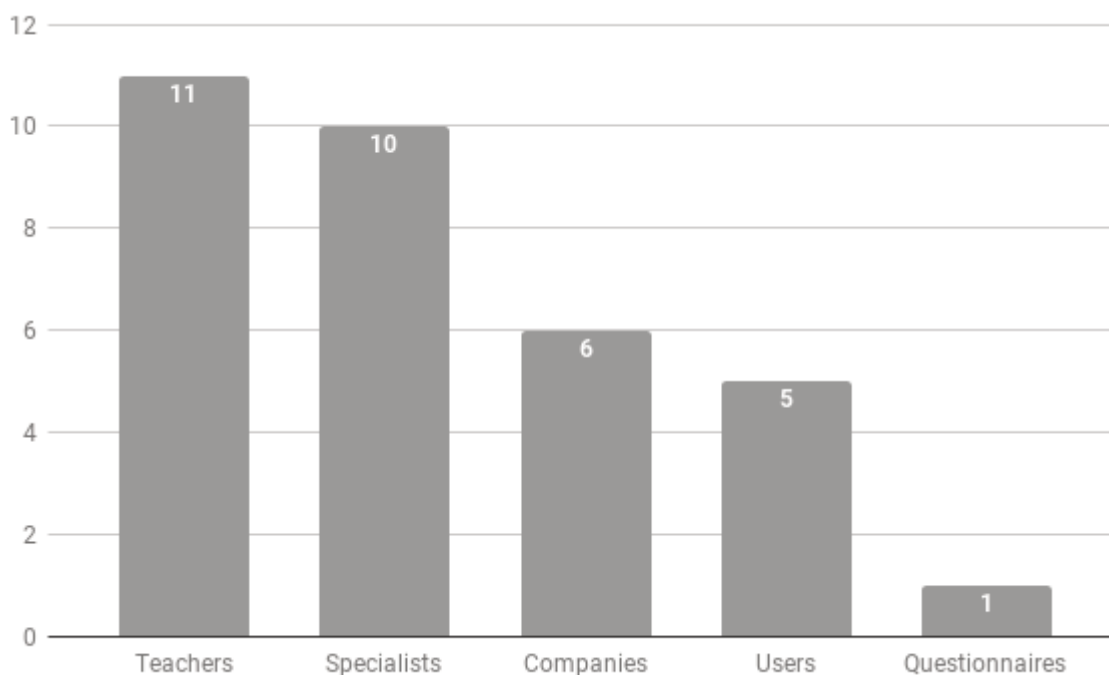
Também buscou-se identificar as formas de validação utilizadas pelos autores. A Tabela 7 demonstra quais são estas formas de validação e os autores que as utilizaram.

Tabela 7: Formas de validação e seus autores.

Validação com	Artigos
Professores	Kovalenko & Antoshchuk (2015), Cong (2014), Andrea & Franco (2012), Ding (2011), Lin et al. (2014), Messaoud, Leray & Amor (2013), Mannai (2013), Yan et al. (2013), Kovalenko, Antoshchuk & Hodovychenko (2015), Chang, Hung & Juang (2013), Grubišić, Stankov & Peraić (2013).
Especialistas	Bucci, Sandrucci & Vicario (2011), Fischer & Beyerer (2013), Kalet et al. (2017), Fenz (2012), Arsene, Dumitrache & Mihiu (2015), Arsene, Dumitrache & Mihiu (2011), Sayed & Lohse (2014), Messaoud et al. (2015), Rodrigues, Bez & Flores (2013) Pavlin & Laskey (2014).
Empresas	Pilato (2012), Bibi et al. (2014), Mok & Min (2015), Settas, Cerone & Fenz (2011), Hashimoto et al. (2011), Aguilar (2016).
Usuários	Agarwal, Verma & Mallik (2016), Julina & Thenmozhi (2012), Qinghong (2012), Armand, Filliat & Ibañez-Guzman (2014) Santos, Jurmain (2011).
Questionários	Fujita, Kurematsu, Hakura (2011).

Fonte: Elaborado pelo autor.

Os autores utilizaram diferentes formas para validar seus experimentos. 33,3% dos trabalhos contaram com a validação de professores e pesquisadores universitários. Já 30,3% contaram com especialistas para realizar a validação, sendo, na maioria dos casos, os que envolvem saúde. 18,2% dos artigos aplicaram seus experimentos em empresas com segmento na área de pesquisa do autor. 15,2% dos trabalhos validaram com usuários. Apenas 3% dos autores utilizaram questionários como forma de validação. Os resultados são exibidos no gráfico da Figura 15.

Figura 15: Gráfico comparativo de tipos de validação.

Fonte: Elaborado pelo autor.

3.6.2.6. Aplicação na área da saúde

Como este trabalho busca a implementação de um protótipo na área de Redes Bayesianas para casos clínicos, foi pesquisada a aplicação de ontologias e RBs na área de saúde. Foram identificados os autores que aplicaram sua pesquisa nesta área. A Tabela 8 mostra quais artigos foram aplicados na área da saúde e os autores que aplicaram.

Tabela 8: Artigos aplicados na área da saúde e seus autores.

Aplicado Artigos	
Sim	Fujita, Kurematsu, Hakura (2011), Bucci, Sandrucci & Vicario (2011), Agarwal, Verma & Mallik (2016), Julina & Thenmozhi (2012), Kalet et al. (2017), Arsene, Dumitrache & Mihiu (2015), Arsene, Dumitrache & Mihiu (2011), Mannai (2013), Rodrigues, Bez & Flores (2013), Chang, Hung & Juang (2013).
Não	Kovalenko & Antoshchuk (2015), Pilato (2012), Fischer & Beyerer (2013), Bibi et al. (2014), Mok & Min (2015), Settas, Cerone & Fenz (2011), Cong (2014), Fenz (2012), Andrea & Franco (2012), Grubišić, Stankov & Peraić (2013), Sayed & Lohse (2014), Messaoud et al. (2015), Ding (2011), Hashimoto et al. (2011), Qinghong (2012), Lin et al. (2014), Aguilar (2016), Santos, Jurmain (2011), Yan et al. (2013), Armand, Filliat & Ibañez-Guzman (2014), Kovalenko, Antoshchuk & Hodovychenko (2015), Pavlin & Laskey (2014), Messaoud, Leray & Amor (2013).

Fonte: Elaborado pelo autor.

Quanto a aplicação na área da saúde, foi verificado que 30,3% dos autores aplicaram ontologias e RBs nessa área. 69,7% dos trabalhos aplicaram os experimentos em diversas áreas, como eletrônica, educação e outros.

3.6.2.7. Linguagem utilizadas

A fim de analisar os protótipos criados e a maneira como reagiram a conversão, buscou-se identificar quais foram as linguagens de programação utilizadas para o desenvolvimento dos mesmos. A Tabela 9 mostra quais foram as linguagens de programação e os autores que as utilizaram.

Tabela 9: Linguagens de programação utilizadas e seus autores.

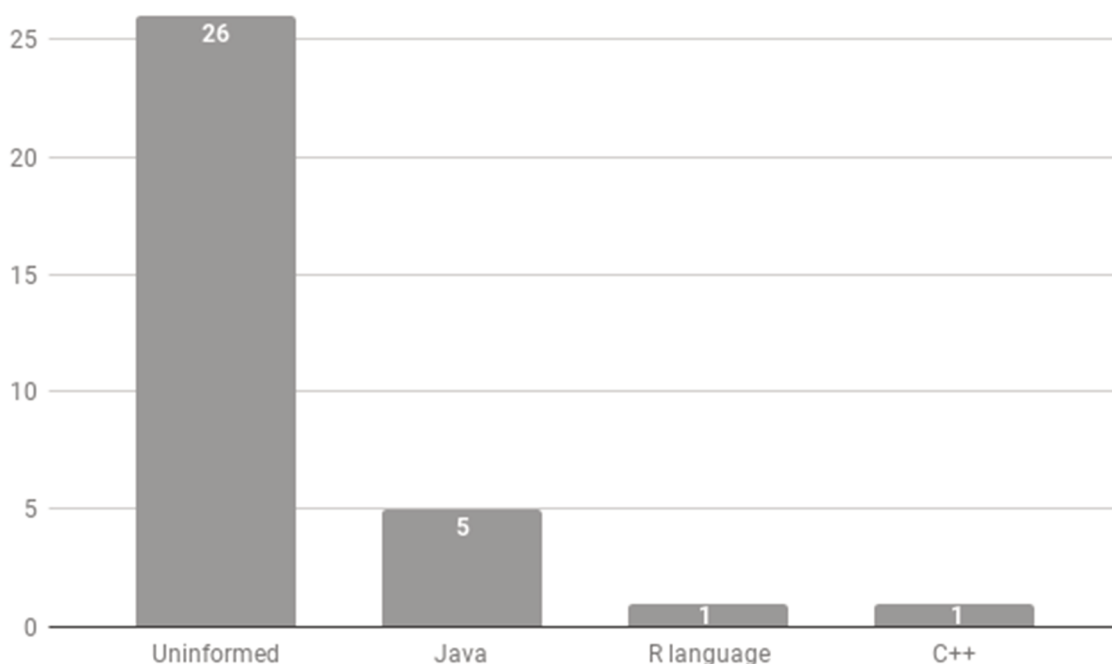
Linguagem	Artigos
Não informado	Bucci, Sandrucci & Vicario (2011), Fischer & Beyerer (2013), Bibi et al. (2014), Agarwal, Verma & Mallik (2016), Julina & Thenmozhi (2012), Mok & Min (2015), Settas, Cerone & Fenz (2011), Cong (2014), Arsene, Dumitrache & Mihiu (2015), Arsene, Dumitrache & Mihiu (2011), Andrea & Franco (2012), Grubišić, Stankov & Peraić (2013), Sayed & Lohse (2014), Ding (2011), Hashimoto et al. (2011), Qinghong (2012), Messaoud, Leray & Amor (2013), Aguilar (2016), Mannai (2013), Santos, Jurmain (2011), Yan et al. (2013), Armand, Filliat & Ibañez-Guzman (2014), Kovalenko, Antoshchuk & Hodovychenko (2015), Chang, Hung & Juang (2013), Pavlin & Laskey (2014), Messaoud et al. (2015).
Java	Pilato (2012), Fujita, Kurematsu, Hakura (2011), Fenz (2012), Lin et al. (2014), Rodrigues, Bez & Flores (2013).
R	Kalet et al. (2017).
C++	Kovalenko & Antoshchuk (2015).

Fonte: Elaborado pelo autor.

78,8% dos autores não informaram em seus trabalhos a utilização de alguma linguagem. 15,2% utilizaram Java, visto o fato que alguns *plugins* necessitavam dessa linguagem para serem executados. 3,0% utilizaram a linguagem R e outros 3,0% citaram que utilizam a linguagem C++. O gráfico da Figura 16 esboça esses resultados.

Além das linguagens de programação, notou-se também as linguagens de descrição de ontologias que foram usadas. 48,8% dos autores citaram a utilização da OWL como linguagem utilizada, visto que é o formato mais popular para a construção de uma ontologia. 15,1% dos artigos citam a utilização de RDF. 12,1% dos autores utilizaram a linguagem SPARQL, principalmente para realizar consultas na ontologia para a extração dos dados. 24% não citaram alguma linguagem de descrição de ontologias.

Figura 16: Gráfico comparativo de linguagens utilizadas.



Fonte: Elaborado pelo autor.

3.6.2.8. Resultados obtidos

Foi analisado se os autores consideraram seus resultados positivos ou realizada tal análise caso não fosse identificada essa informação. A Tabela 10 mostra quais são os autores que apresentaram resultados positivos.

Tabela 10: Resultados positivos e seus autores.

Positivo Artigos	
Sim	Kovalenko & Antoshchuk (2015), Pilato (2012), Fujita, , Kurematsu, Hakura (2011), Bucci, Sandrucci & Vicario (2011), Fischer & Beyerer (2013), Bibi et al. (2014), Agarwal, Verma & Mallik (2016), Julina & Thenmozhi (2012), Mok & Min (2015), Settas, Cerone & Fenz (2011), Kalet et al. (2017), Fenz (2012), Arsene, Dumitrache & Mihiu (2015), Arsene, Dumitrache & Mihiu (2011), Andrea & Franco (2012), Grubišić, Stankov & Peraić (2013), Sayed & Lohse (2014), Messaoud et al. (2015), Qinghong (2012), Lin et al. (2014), Messaoud, Leray & Amor (2013), Aguilar (2016), Santos, Jurmain (2011), Armand, Filliat & Ibañez-Guzman (2014)., Kovalenko, Antoshchuk & Hodovychenko (2015), Chang, Hung & Juang (2013), Pavlin & Laskey (2014).
Não	Cong (2014), Ding (2011), Hashimoto et al. (2011), Mannai (2013), Yan et al. (2013), Rodrigues, Bez & Flores (2013).

Fonte: Elaborado pelo autor.

Sobre os resultados obtidos, verificou-se que 81,8% dos autores consideraram os seus resultados como positivos ou esperados. Ocorre esse valor principalmente pela consideração do objetivo da proposta dos trabalhos, e não ao resultado esperado do que se busca nessa revisão. Um percentual de 18,2% dos autores consideraram seus resultados negativos e/ou não esperados.

3.6.2.9. Ferramentas utilizadas

Identificou-se também quais foram as ferramentas utilizadas para a construção de ontologias ou RBs, *plugins* utilizados ou APIs. A Tabela 11 mostra quais são as ferramentas e os autores que as utilizaram.

Tabela 11: Ferramentas utilizadas e seus autores.

Ferramenta	Artigos
Não informado	Kovalenko & Antoshchuk (2015), Pilato (2012), Bucci, Sandrucci & Vicario (2011), Fischer & Beyerer (2013), Bibi et al. (2014), Andrea & Franco (2012), Ding (2011), Hashimoto et al. (2011), Qinghong (2012), Messaoud, Leray & Amor (2013), Mannai (2013), Santos, Jurmain (2011), Kovalenko, Antoshchuk & Hodovychenko (2015).
Protégé	Agarwal, Verma & Mallik (2016), Julina & Thenmozhi (2012), Mok & Min (2015), Fenz (2012), Arsene, Dumitrache & Mihiu (2015), Arsene, Dumitrache & Mihiu (2011), Rodrigues, Bez & Flores (2013), Armand, Filliat & Ibañez-Guzman (2014), Chang, Hung & Juang (2013) Yan et al. (2013), Lin et al. (2014).
Nética	Fujita, Kurematsu, Hakura (2011), Settas, Cerone & Fenz (2011), Fenz (2012), Arsene, Dumitrache & Mihiu (2015), Lin et al. (2014), Pavlin & Laskey (2014).
Bayesian Genie	Grubišić, Stankov & Peraić (2013), Sayed & Lohse (2014).
Jess Reasoner	Fujita, Kurematsu, Hakura (2011).
RBTab	Settas, Cerone & Fenz (2011).
RHugin	Kalet et al. (2017).
Bayesian Domain Network	Kalet et al. (2017).
Norsys	Cong (2014).
Graph API	Messaoud et al. (2015).
Probt API	Messaoud et al. (2015).
Owl API	Lin et al. (2014).

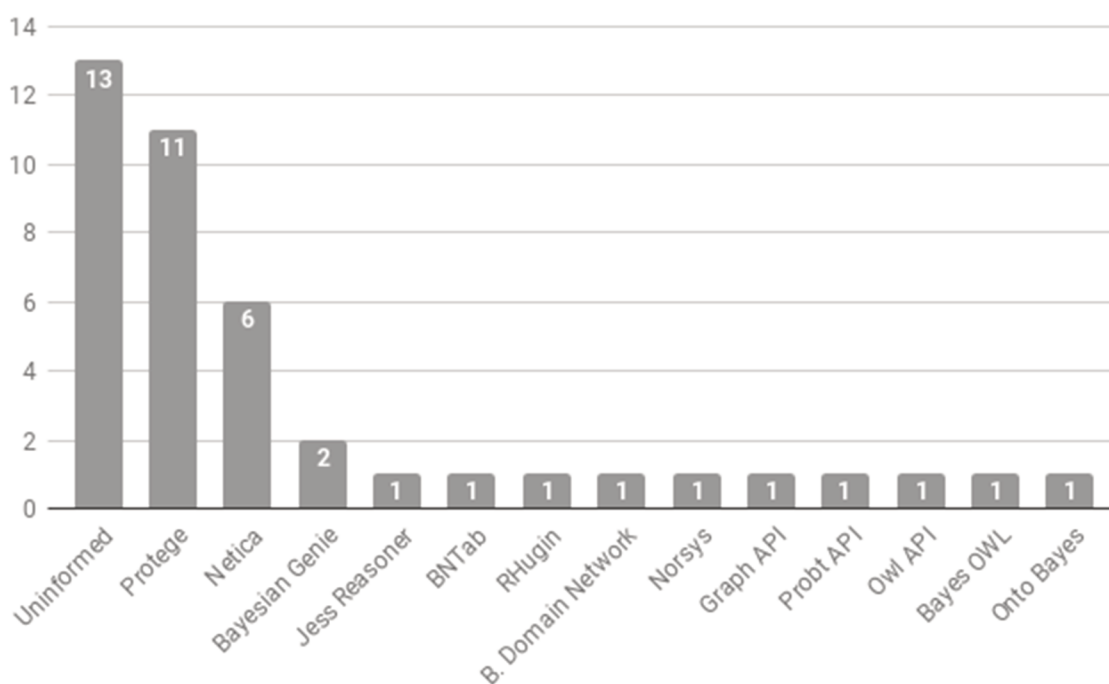
Bayes OWL Aguilar (2016).

Onto Bayes Chang, Hung & Juang (2013).

Fonte: Elaborado pelo autor.

Verificou-se, também, quais foram as ferramentas usadas pelos autores para o desenvolvimento de suas pesquisas. Em 31% dos artigos não foram informadas quais ferramentas foram utilizadas. Um total de 26,2% informou que utilizaram o software Protégé para o desenvolvimento das ontologias. 14,3% dos trabalhos utilizaram o plugin Nética para gerar a RB de uma ontologia no Protégé. 4,8% optaram pelo *plugin* Bayesian Genie. Ferramentas e recursos como Jess Reasoner, RBTAB, RHugin, Bayesian Domain Network, Norsys, Graph API, Probt API, Owl API, Bayes Owl e Onto Bayes estiveram presentes apenas nos trabalhos do qual foram identificados, representando cada uma 2,4%. Na Figura 17 é ilustrado o gráfico com esses resultados.

Figura 17: Comparativo de ferramentas utilizadas.



Fonte: Elaborado pelo autor.

3.7. Lacunas e Desafios

Para a construção de uma análise crítica, construiu-se uma Tabela da Seção 3.6.1, considerada de extrema relevância para esta pesquisa, para que seja possível aprofundar a estratégia abordada por cada trabalho selecionado. A Tabela 12 ilustra o comparativo das abordagens utilizadas em cada publicação.

Tabela 12: Comparativo das abordagens utilizadas em cada publicação.

	D O	O W	R B M	F M V	R B S P	R B A H	A D
A framework for proactive assistance: Summary	X		X	X			
A framework of ontology-based knowledge information processing for change detection in remote sensing data	X		X	X			
A method of constructing user interest model based on semantics	X		X	X			
A novel approach to power transformer fault diagnosis based on ontology and Bayesian network	X			X	X		
A URREF interpretation of Bayesian network information fusion		X			X		
About the studying on the uncertainty of ontology on the Bayesian Network		X	X	X			
Active learning of causal Bayesian networks using ontologies: A case study							X
An ontology-based approach for constructing Bayesian networks		X		X	X		
Autonomie decision making based on bayesian networks and ontologies		X		X	X		
Bayesian information extraction network for Medline abstract	X		X				
Bayesian knowledge-driven ontologies: Intuitive uncertainty reasoning for semantic networks		X	X	X			
Depression Diagnosis Based on Ontologies and Bayesian Networks	X			X	X		
Developing Bayesian networks from a dependency-layered ontology: A proof-of-concept in radiation oncology.		X			X		
Development of a novel asset management system for power transformers based on ontology	X			X	X		
Event recognition using a semantic-probabilistic network	X			X		X	
Expert system for medicine diagnosis using software agents		X			X		
Generating Bayesian networks from medical ontologies	X				X		
Human action recognition using a semantic-probabilistic network	X			X		X	
Integration of Ontologies and Bayesian Networks for Maritime Situation Awareness	X		X				
Medicine expert system dynamic Bayesian Network and ontology based		X			X		
Mining Bayesian networks out of ontologies	X		X	X			
Multiviews ontologies based reasoning for medical diagnosis in VDS		X	X	X			
Ontologies and Bayesian Networks in Medical Diagnosis	X		X				
Ontologies for probabilistic situation assessment in the maritime domain	X		X	X			
Ontology based approach to Bayesian student model design	X		X	X			
Ontology based Bayesian software process improvement	X		X				
Ontology based disease diagnosis system with probabilistic inference	X		X				
Ontology based EMR for decision making in healthcare using SNOMED CT	X		X				
Ontology-based context-aware model by applying Bayesian network	X		X				
Ontology-driven generation of Bayesian diagnostic models for assembly systems	X		X	X			
SemCaDo: A serendipitous strategy for causal discovery and ontology evolution	X			X			
The Development of Ontology Information System Based on Bayesian Network and Learning	X		X	X			
Towards Automatic Generation of Ontology-Based Antipattern Bayesian Network Models		X			X		

Legenda utilizada para as siglas: DO- Desenvolvido uma Ontologia; OW- Ontologia da Web; RBM- RB Manual; FMV- Fórmula Matemática para Variáveis; RBSP- RB Semiautomática com plugin; RBAH- RB Automática com Hierarquia; AD- Análise de Dados.

Fonte: Elaborado pelo autor.

Após a análise da Tabela 12, foi possível identificar questões importantes para esse trabalho, sendo estas avaliadas individualmente em relação a ontologias e RBs. No que se refere a ontologias, identificou-se informações relevantes, como as dificuldades de construir uma RB a partir de uma ontologia, técnicas de mapeamento e possíveis relações ontologias que podem ser mapeadas para elementos de RBs, bem como, identificar as melhores formas de aplicação para atingir o objetivo deste trabalho. Essas informações são descritas a seguir.

Mediante à análise realizada no Capítulo anterior (em especial no 4.6), verificou-se que, dentre as principais dificuldades de extrair uma ontologia para gerar RBs, destaca-se o fato de não haver um padrão para a construção de uma ontologia. Sendo assim, existem inúmeras ontologias disponibilizadas na web, porém construídas de forma diferente, do ponto de vista de sua estrutura. Outra dificuldade observada refere-se ao formalismo utilizado para a construção de uma ontologia, que não possui o mesmo objetivo de uma RB, mesmo que ambos sejam representações de conhecimento. As ontologias possuem uma natureza estática, representando o conhecimento de como é constituído o mundo. As RBs contêm atributos que podem ser dinâmicos, modelando o conhecimento incerto, à fim de gerar probabilidades de ocorrência de eventos.

Grande parte das técnicas de mapeamento analisadas envolvem questões de padronização de campos. Para os autores que construíram sua própria ontologia (AGARWAL, VERMA & MALLIK, 2016), (MOK & MIN, 2015), o desenvolvimento foi feito pensando na extração da RB, como à inserção de campos que representavam nodos, inserções de probabilidade, ligação entre os nodos, dentre outros. Nos casos de extração com ferramentas interativas (LIN et al., 2014), (CHANG, HUNG & JUANG, 2013), observa-se a abordagem de mapeamento das classes, identificando que as mesmas são nodos e suas relações são nodos com sua ligação, abordagem na qual se faz necessário que o especialista confirme e ajuste as operações conforme o andar do processo.

Os trabalhos que utilizaram o nível hierárquico da ontologia para a extração e empregaram um algoritmo genérico de mapeamento (MESSAOUD et al., 2015), em geral adotam uma abordagem que identifica as classes como os nodos-pai da rede, as propriedades que representam a ligação de um nodo-pai a um nodo-filho, e as instâncias, que são os nodos-filhos da rede. Há também outra técnica de mapeamento semelhante (KOVALENKO & ANTOSHCHUK, 2015), que identifica apenas as classes e subclasses como os nodos da rede, fazendo com que a classe principal seja o nodo-pai, a subclasse seu nodo filho e, assim por diante.

Conforme o objetivo deste trabalho, considera-se que nenhuma das técnicas de mapeamento são eficazes para essa proposta, descartando a forma de mapeamento hierárquico, visto que normalmente as ontologias não são construídas dessa maneira. Também não será usada a forma de padronização de campos, pois busca-se realizar uma extração de ontologias já disponíveis na web, que não possuem necessariamente esse padrão. Por fim, não serão adotadas as abordagens interativas, que exigem muito esforço e conhecimento específico do especialista.

O próximo Capítulo apresenta o modelo proposto neste trabalho.

4. MODELO PROPOSTO

Após o estudo dos trabalhos correlatos, foram extraídas as informações consideradas importantes para o desenvolvimento do modelo. Destaca-se neste Capítulo o modelo proposto e seus diferenciais. Considera-se importante destacar diferenciais em dois aspectos. Um deles relacionado com o ponto de vista do usuário do sistema e outro com a abordagem desenvolvida.

O diferencial do modelo desenvolvido em relação ao ponto de vista do usuário é a facilidade de uso, uma vez que a técnica não exige conhecimento adicional do usuário ou qualquer informação direta sobre a estrutura da RB a ser criada. O modelo é baseado na própria informação da ontologia e no uso pretendido para a RB. Portanto, o uso dessa técnica proposta é menos complexo, considerando que não é necessário planejar a estrutura da RB durante a concepção da ontologia ou mesmo durante o processo de extração. A única tarefa que resta e que requer intervenção do usuário é o ajuste das probabilidades na RB.

A contribuição do modelo do ponto de vista da abordagem desenvolvida está centrada em dois elementos. O primeiro é o uso de um **conjunto de regras de conversão**, criadas a partir de uma análise das ontologias. Com este conjunto de regras de conversão é possível realizar a extração de maneira genérica para qualquer ontologia, não se limitando a apenas um tipo de estrutura ou abordagem de construção específica, como observado na maioria dos trabalhos relacionados. O segundo elemento consiste em um **método de sumarização** de ontologias, que permite que a abordagem de regras de conversão seja potencializada, ampliando a precisão do resultado.

Esta contribuição é inédita, pois a utilização destas duas técnicas juntas não foi explorada na literatura identificada nos trabalhos estudados. O diferencial busca respeitar as regras semânticas das ontologias na etapa da conversão, bem como descobrir quais são os dados (conhecimento) realmente necessários para a Rede Bayesiana e a forma como a mesma deve ser montada.

O ponto principal considerado neste modelo é que algumas relações ontológicas podem ser automaticamente mapeadas em RBs, tendo sido identificadas em nossa pesquisa bibliográfica. As classes da ontologia possuem uma grande possibilidade de associação semântica com os nodos da RB, embora não necessariamente seja uma regra diretamente aplicável a todos os casos, pois existem classes genéricas e/ou específicas que não terão utilidade na RB, bem como suas ligações terão que ser analisadas com cuidado, pois existem ontologias que trabalham apenas com classes, outras com propriedades e instâncias.

Como não existe um padrão de construção de ontologias amplamente adotado, todo o contexto da ontologia deve ser analisado para gerar a RB. O problema principal a ser tratado neste caso relaciona-se com a correta compreensão e utilização da informação de um modelo (a ontologia) para o desenvolvimento de outro (a RB). Sendo assim, avalia-se que a solução do problema não está sujeita a estratégias sintáticas. Portanto, adotamos a sumarização de ontologias como nosso procedimento principal, no qual são criadas regras para avaliar se o elemento da ontologia analisada apresenta melhor adequação como um nodo, uma ligação, um nodo-filho, estado da rede ou até mesmo alguma informação sem relevância. Em relação às RBs, foram indicados os itens necessários para a sua construção, o formato de arquivo que as RBs utilizam, como são geradas as suas probabilidades e quais suas fontes de dados.

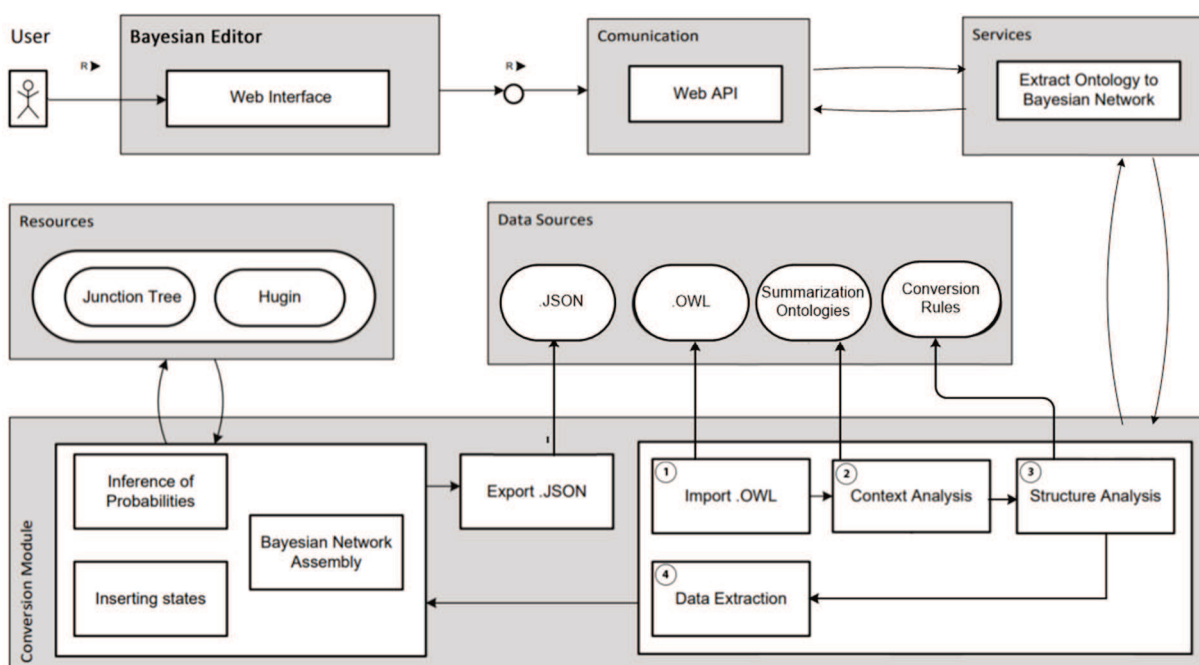
Através da análise, consulta com especialistas e estudo na literatura, foi possível desenvolver o modelo proposto neste trabalho. Essas informações são descritas a seguir.

4.1. Visão geral do modelo proposto

A Figura 18 ilustra todos os componentes envolvidos, destacando que a conexão do editor bayesiano com o restante dos componentes do modelo é flexível e que, portanto, o modelo pode ser aplicado em conjunto com outros editores.

O modelo prevê esta flexibilidade a partir de uma interface flexível que permite disponibilizar a editores de RBs um serviço dedicado ao tratamento de uma ontologia para a sua transformação em RB. Este componente pode ser observado no módulo “Serviços” da Figura 18. Este serviço de Extração é suportado por componentes específicos, descritos adiante, com a interação com fontes de dados para representar tanto as ontologias consultadas quanto os RBs gerados. O modelo prevê a possibilidade de que vários recursos adicionais também possam ser incorporados, quando necessário, para executar processos específicos.

Figura 18: Visão geral do modelo de extração de RBs a partir de ontologias.



Fonte: Elaborado pelo autor.

O modelo prevê o uso dos recursos pelos usuários a partir de um editor de redes bayesianas. Este aspecto é flexível, sendo que o modelo permite o seu uso por diversos editores, provendo uma interface de solicitação de serviços. Desta forma, para iniciar um suposto uso do modelo, o usuário acessa um Editor Bayesiano, o qual solicita a importação de uma ontologia através de um serviço web. Ao receber a solicitação, o modelo conduz a busca, através da Web API, de acordo com o tipo de serviço selecionado. O serviço de extração de dados da ontologia e geração de uma rede bayesiana, ao ser chamado, implica na execução do módulo de conversão, que contém os componentes do algoritmo e exportação (detalhado na Seção 4.2).

O módulo utiliza como fontes de dados os recursos semânticos dos arquivos no formato OWL, o método de sumarização de ontologias desenvolvido e as regras de conversão (melhor explicado na Seção 4.2) e, para exportação, atribui os dados ao formato JSON. Para o auxílio da montagem gráfica e das inferências probabilísticas da RB, são utilizados os recursos de *Junction Tree* (ROSENBERG, 2017).

4.1.1. Componente do Editor de Redes Bayesianas

Considera-se que uma maneira adequada para fomentar o uso do modelo de forma flexível é prever o seu uso com a ampla gama de editores bayesianos que se encontram disponíveis. De modo complementar, considera-se que uma interface web permite que o modelo possa ser usado de modo facilitado por editores de redes bayesianas, promovendo uma vantagem quanto a sua utilidade e disponibilidade.

O componente destinado a representar o editor de redes bayesianas presente na estrutura do modelo proposto possui como objetivo estabelecer um formato flexível de acesso para os diversos tipos de editores bayesianos existentes. Desta forma, os recursos do modelo podem ser empregados de forma mais ampla, tendo em vista que podem ser acessados por diferentes editores.

4.1.2. Componente de Comunicação

O componente de comunicação, através da Web API, tem como função possibilitar o acesso ao tipo de serviço selecionado no editor bayesiano. Este componente é dedicado a facilitar a chamada de serviços disponíveis, tais como a criação de uma RB nova, importação de uma RB existente, importação de uma ontologia, entre outros. Para manter a flexibilidade e o isolamento dos editores utilizados, optou-se por usar uma API Web, implementando ligações para os serviços de extração de ontologias para geração de redes bayesianas. O componente de comunicação também é responsável por enviar os resultados da conversão.

A Web API, além de *open source*, é uma arquitetura considerada leve e adequada para dispositivos que atuam com uma largura de banda limitada, tais como os dispositivos móveis, por exemplo. Sendo assim, ela apresenta uma excelente performance quanto a sua comunicação entre a interface web e os serviços solicitados, apoiando na ampliação da flexibilidade de uso dos recursos do modelo.

4.1.3. Componente de Serviços

O componente de serviços guarda as funções disponíveis no editor bayesiano, tais como: criação de uma nova rede bayesiana, criação de uma rede bayesiana multi-seccionada, importação de uma rede bayesiana descrita no formato JSON (formato padrão do editor) ou importação de uma rede bayesiana .NET (formato disponibilizado em outros editores bayesianos), etc.

Para a inserção do modelo, criou-se um novo serviço de importação de ontologia, o qual é chamado pela Web API. Este serviço executa o componente de conversão, descrito na Seção 4.1.4. A proposta deste componente é atuar para padronizar a funcionalidade como um serviço do editor de redes bayesianas, fazendo com que o usuário se habitue a ora criar uma nova rede bayesiana, ora adquirir/encontrar uma ontologia com as informações que necessita, realizar a conversão e trabalhar com a mesma no editor bayesiano.

4.1.4. Componente de Conversão

O componente de conversão é o componente que implementa todo o processo de extração de dados da ontologia para a construção de uma rede bayesiana. O processo divide-se em duas etapas: Extração de dados e montagem da Rede Bayesiana. Na primeira etapa, o sistema importa a ontologia em padrão OWL e realiza a varredura do seu código, para identificar todos os seus componentes (classes, subclasses, instâncias, propriedades, etc). Após identificados, o sistema aplica sobre os componentes da ontologia uma análise de contexto através do método de sumarização de ontologias, a partir do qual o modelo identifica quais são os elementos chave para a construção da rede e quais poderão ser descartados. Um exemplo de componente que pode ser descartado seriam as classes descritivas, utilizadas em algumas ontologias para fins de documentação. Em seguida, o sistema realiza uma análise da estrutura da ontologia, aplicando as regras de conversão (descritas na Seção 4.2), a fim de descobrir quais componentes estão conectados entre si. Por fim, o sistema coleta todos os dados relevantes da extração e inicia a segunda etapa para a construção da rede bayesiana.

Na segunda etapa, o modelo inicia com a montagem da rede bayesiana, criando os nodos e ligações conforme os dados recebidos da etapa anterior. Após, o sistema insere os estados “Sim” e “Não” a cada nodo, bem como suas probabilidades de “50% sim” e “50% não”. Esta configuração foi definida como padrão para esta pesquisa, mas pode ser alterada posteriormente pelo usuário. Feito isso, o modelo infere as probabilidades a todos os nodos através do algoritmo de *junction tree*, aplicando assim as regras de Bayes para que, caso um estado seja editado, tal probabilidade propague para os demais. Concluindo esse processo, o sistema monta os dados em um arquivo no formato JSON e exibe a rede bayesiana graficamente na Web Interface. Todos os detalhes deste módulo serão apresentados na Seção 4.2.

Para a realização da conversão dos dados neste módulo, o modelo utiliza de recursos e fontes de dados, que serão descritos a seguir.

4.1.5. Componente de Recursos

O modelo utiliza um recurso importante para a montagem e inferência da rede bayesiana: O algoritmo de *junction tree*. Para mostrar todos os resultados em tempo real, diversas inferências precisam ser feitas. O algoritmo *junction tree* é indicado para essa situação, pois é dividido em duas partes: a compilação da *junction tree* e a inferência na mesma. A inferência na *junction tree* é bastante eficiente e a mesma *junction tree* construída para uma rede pode ser reutilizada para todas as inferências nela (BEAL, 2003).

Optou-se pela utilização deste recurso visto que o mesmo demonstra grande eficiência na construção e na inferência da rede bayesiana. Prevê-se a flexibilização para a utilização de outros recursos, tais como bancos de dados ou estrutura de grafos, de modo que estes possam ser agregados tanto em uma extração eficaz, como na flexibilização do conteúdo e montagem da rede bayesiana.

4.1.6. Componente de Fonte de Dados

Como Fonte de Dados, o sistema utiliza tanto formatos já existentes como novos que foram desenvolvidos para este trabalho. As Regras de Conversão e o método de sumarização de ontologias foram criados pelo autor através de análises e consulta com especialistas

(conforme detalhado na Seção 4.2). Como fonte de dados das ontologias, optou-se em tratar apenas do formato OWL, definido pela W3C como o formato padrão para a criação de ontologias (ARSENE, DUMITRACHE & MIHU, 2011). Através deste formato é possível detectar todos os componentes presentes na ontologia. Um fragmento de código OWL é ilustrado na Figura 19.

Figura 19: Fragmento de código OWL

```
<?xml version="1.0"?>
<Ontology xmlns="http://www.w3.org/2002/07/owl#"
  xml:base="http://www.semanticweb.org/diegopinheiro/ontologies/2018/2/OntologiaNausea"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xml="http://www.w3.org/XML/1998/namespace"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  ontologyIRI="http://www.semanticweb.org/diegopinheiro/ontologies/2018/2/OntologiaNausea">
  <Prefix name="" IRI="http://www.semanticweb.org/diegopinheiro/ontologies/2018/2/OntologiaNausea#"/>
  <Prefix name="owl" IRI="http://www.w3.org/2002/07/owl#"/>
  <Prefix name="rdf" IRI="http://www.w3.org/1999/02/22-rdf-syntax-ns#"/>
  <Prefix name="xml" IRI="http://www.w3.org/XML/1998/namespace"/>
  <Prefix name="xsd" IRI="http://www.w3.org/2001/XMLSchema#"/>
  <Prefix name="rdfs" IRI="http://www.w3.org/2000/01/rdf-schema#"/>
  <Declaration>
    <ObjectProperty IRI="#Trata_com"/>
  </Declaration>
  <Declaration>
    <Class IRI="#Gastric_Irritation"/>
  </Declaration>
  <EquivalentClasses>
    <Class IRI="#Antiemetic_drugs"/>
    <ObjectSomeValuesFrom>
      <ObjectProperty IRI="#Ligado_a"/>
      <Class IRI="#Nausea"/>
    </ObjectSomeValuesFrom>
  </EquivalentClasses>
  <SubClassOf>
    <Class IRI="#Skeletal_Musculature"/>
    <Class IRI="#Symptoms"/>
  </SubClassOf>
  <SubObjectPropertyOf>
    <ObjectProperty IRI="#Treatment_of"/>
    <ObjectProperty IRI="#Diagnostia_a"/>
  </SubObjectPropertyOf>
</Ontology>
```

Fonte: Elaborado pelo autor.

As Redes Bayesianas geradas devem utilizar o formato JSON, visto que é o formato aceito por diversos editores de redes bayesianas. Por exemplo, o formato pode ser observado no trabalho descrito por Helwanger (HELWANGER et al., 2017). Um exemplo deste formato está descrito a seguir. O arquivo tem uma estrutura simples e personalizada para o uso no editor. Após a montagem de Rede, o modelo gera o código no formato apresentado. Um fragmento de código JSON utilizado no editor mencionado como exemplo é ilustrado na Figura 20.

Figura 20: Fragmento de código JSON

```

{
  "version": 2,
  "network": {
    "name": "Diagnóstico de Náusea",
    "height": 1000,
    "width": 2000,
    "selectedNodes": [],
    "beliefs": {},
    "propertiesPanelVisible": false,
    "kind": "BN",
    "id": "e5ba01a2-5b41-4604-84c3-2b1a76dea662",
    "subnetworks": [],
    "linkages": {},
    "inferenceEnabled": true
  },
  "nodes": [
    {
      "id": "Nausea",
      "states": [
        "Sim",
        "Não"
      ],
      "parents": [
        "Definitive Features",
        "Gastric Irritation",
        "Treatment Sequela"
      ],
      "cpt": [
        {
          "when": {
            "Definitive Features": "Sim",
            "Treatment Sequela": "Sim",
            "Gastric Irritation": "Sim"
          },
          "then": {
            "Sim": 0.29,
            "Não": 0.71
          }
        }
      ]
    }
  ],
}

```

Fonte: Elaborado pelo autor.

4.2. Detalhamento do processo de análise de ontologia e geração de RB

Para a descrição do processo de análise de ontologia e geração de RB, devem ser referidos os componentes centrais para esta tarefa, que são identificados na Figura 18 e numerados como passos de 1 a 4 no Módulo de Conversão. Eles serão explicados individualmente abaixo.

4.2.1. Importação do OWL

O item “1. Importar OWL” consiste no primeiro passo operacional, pois o modelo permite que o usuário selecione um arquivo contendo uma ontologia de formato OWL a ser convertido em RB. O sistema faz uma validação verificando se a ontologia está no formato

correto e se não há erros. Caso haja algum problema, o sistema informa ao usuário a causa e como proceder. Se estiver tudo correto, o passo 2 é realizado, descrito posteriormente.

4.2.2. Análise de contexto

O objetivo desta etapa é analisar o contexto dos dados presentes na ontologia, pois, para realizar uma conversão e extração eficaz, é necessário identificar quais são os pontos principais, como eles estão conectados e, principalmente, respeitar as regras e estrutura semântica das ontologias, para que, assim, uma RB possa ser construída. Uma simples troca de parâmetros não produz resultados eficazes, comprovado já nos trabalhos correlatos estudados.

Foi estudado e analisado quais são os tipos de nodos essenciais para a construção de uma RB e quais são descartáveis. Detectou-se que a RB trata diretamente com os dados em questão, diferente das ontologias, que possuem muitos dados que são classificadores. Exemplos de dados classificadores são: Diagnósticos; Doenças; Causas; Sintomas. Estes dados servem para indicar quais são os tipos de conteúdo presentes dentro de suas subclasses, o que não é utilizado na RB. Exemplos de dados em questão são: Febre; Vômito; Cefaleia; Dipirona. Estes dados indicam diretamente a questão do que estão ligados.

Inicialmente, através de um estudo com os especialistas, foi desenvolvido um dicionário para analisar esses contextos presentes nas ontologias e identificar sua relevância para a conversão em RB. Este dicionário (disponível no [Apêndice B](#)) permitiu analisar o contexto que a ontologia descreve através de uma avaliação dos termos de acordo com a sua possibilidade de uso na rede bayesiana e excluindo aqueles que não faziam sentido ou que não eram necessários.

Para auxiliar a conduzir a avaliação preliminar do modelo, uma primeira versão de um protótipo no qual foi utilizado este dicionário foi desenvolvida e observaram-se resultados que foram considerados adequados no caso de estudo em questão. Porém, essa abordagem necessita de uma gama de estudo da área de domínio em que será abordada a conversão e também um extenso trabalho manual, aspecto que o modelo proposto busca evitar. Buscou-se então novas abordagens que pudessem analisar melhor estes pontos relevantes dentro da ontologia. Após uma gama de estudos, optou-se por criar um método de sumarização de ontologias, visto que esta abordagem complementar os aspectos do modelo já estabelecidos, gerando a possibilidade de atendimento de um número maior de áreas de domínio sem necessidade de estudo manual preliminar. A descrição do método é descrita a seguir.

4.2.2.1. Método de Sumarização de Ontologias

No item “2. Análise de Contexto” o passo 2 é realizado. Nessa fase, o sistema passa pelo código da estrutura ontológica, analisando todos os seus componentes e o contexto com que estão inseridos. Com o objetivo de ajudar no entendimento de grandes ontologias, reduzir o processamento e remover o processo manual da criação de um dicionário, optou-se pela utilização de técnicas de sumarização de uma ontologia para restringir o conhecimento aos conceitos mais importantes de um determinado domínio. Com isso, definiu-se uma abordagem para sumarizar qualquer ontologia, destacando os pontos importantes em um subconjunto de conceitos que apresenta de forma resumida o conhecimento da ontologia original.

Uma etapa de sumarização de ontologias foi desenvolvida, que busca à indicação de conceitos mais importantes para garantir à precisão dos dados necessários em uma Rede Bayesiana. Para a elaboração da técnica, utilizou-se o algoritmo *Broaden Relevant Paths* (BRP) (FORNAHL, HASSINK & MENZEL, 2015) para a produção de grafos conexos com os nodos mais relevantes e duas medidas de *Centrality Measure* e *Closeness Measure* (BLOCH, MATTHEW & TEBALDI, 2017), (SOUZA, DURAN & VIEIRA, 2014) para avaliar a relevância dos conceitos na ontologia representada nos grafos.

Para realizar os procedimentos necessários, realizou-se um estudo nas áreas da ciência cognitiva, que usa a noção de categoria natural para a identificação de conceitos; topologia de rede, para analisar conceitos que são ricamente caracterizados com propriedades e relacionamentos taxonômicos; estatísticas lexicais, para analisar conceitos que são susceptíveis de serem mais familiares ao domínio; e, por fim, teoria dos grafos, para à utilização de medidas de centralidade na identificação de conceitos centrais das ontologias representadas em grafos.

A medida de relevância é a chave para a sumarização de ontologias, pois é responsável por analisar e definir a relevância dos conceitos. Um estudo de teoria dos grafos (RUOHONEN, 2013) e análise de rede foram os pontos chaves para o desenvolvimento de avaliação dos elementos das ontologias. As medidas de centralidade mais utilizadas são: *Degree Centrality*, *Betweenness Centrality*, *Eigenvector Centrality* e *Closeness Centrality*. Tais medidas determinam qual a importância relativa de um vértice dentro de um grafo, assim, o valor de centralidade do vértice. Após o estudo das técnicas, decidiu-se por utilizar duas medidas para calcular a relevância dos conceitos, uma baseada em *Degree Centrality*, com atribuição de pesos nos tipos de relacionamentos e outra baseada em *Closeness Centrality*. Tais medidas foram selecionadas visto a eficácia apresentada em trabalhos correlatos.

A medida de relevância *Degree Centrality* considera a quantidade e o tipo de relacionamento entre os conceitos da ontologia. Os tipos de relacionamentos considerados são “*standard*”, “*is-à*”, “*part-of*” e “*same-as*” e relacionamentos “*user-defined*”, como “*HasItems*” e “*authorOf*”. Esta medida também pode conter o *Weighted In-Degree*, que considera a direção do relacionamento, contabilizando os relacionamentos, apontando os conceitos e indicando o subconceito ou range de uma propriedade no relacionamento. Para este estudo, utilizou-se as medidas de “*standard*” e “*user-defined*”, considerando pesos nos direcionamentos nos tipos dos relacionamentos.

A *Closeness Centrality* representa relevância aos conceitos que estão próximos aos conceitos mais importantes, fazendo com que o valor da medida de um conceito seja diretamente proporcional a quantidade de conceitos com grande valor de relevância que estão próximos dele. O foco dessa medida é distinguir os conceitos que tem o mesmo valor de relevância, levando em consideração a distância para os conceitos mais importantes. Também, destaca-se os conceitos que têm maior probabilidade de relacionar-se com conceitos relevantes.

Para realizar a sumarização, é necessário identificar os conceitos principais e selecionar estes conceitos. Para identificar, são utilizadas as medidas de relevância, enquanto que a seleção é aplicada pelo algoritmo *Broaden Relevant Paths* (BRP). O algoritmo BRP tem como função identificar um caminho em um grafo que contenha os vértices de maior valor de relevância no grafo.

No método proposto neste trabalho, após a importação da ontologia descrita na Seção 4.2.1, utiliza-se uma notação em grafo orientado com conexões rotuladas para representar a ontologia e aplicar as medidas de relevância nos seus conceitos. Cada conceito tem uma atribuição de um valor máximo de relevância equivalente. Sendo assim, os conceitos próximos ao conceito também terão um aumento na relevância em função da medida *Closeness Centrality*. Após, inicia-se o algoritmo BRP, onde os conceitos são representados por vértices e os relacionamentos por arestas. Após, é realizada uma execução para achar um caminho que atenda as especificações dos parâmetros e contenha os conceitos mais relevantes, excluindo os que não são considerados relevantes. Feito isso, o método avança para à análise da estrutura da montagem da Rede Bayesiana, descrito no próximo item.

4.2.3. Análise da estrutura

No item “3. Análise da estrutura”, é executada a parte principal do modelo. Após a identificação do contexto das classes, o sistema analisa a estrutura da ontologia. Um conjunto de regras foi elaborado de acordo com uma análise realizada em ontologias de exemplo e para os trabalhos relacionados. Esse conjunto de regras fornece a orientação necessária ao modelo para analisar o contexto da ontologia e identificar os elementos a serem gerados como componentes da RB. O aspecto principal dessas regras pode ser resumido da seguinte maneira:

- a) *Considerar as classes pai e filhas da ontologia como nodos e realizar sua ligação hierarquicamente.*
- b) *Analisar as propriedades de objetos presentes entre as classes e inserir uma ligação entre eles, caso possua a classe sintática “and” ou “or”, ou nenhuma.*
- c) *Caso a classe sintática entre as classes e propriedades seja “not”, não inserir a ligação.*
- d) *Analisa a característica “Equivalent to”. As classes inseridas na mesma são consideradas como as classes pai da em questão (por se tratar de uma classe equivalente a), cumprindo as mesmas regras anteriores.*
- e) *Caso as classes estejam como “Disjoint”, não haverá ligação equivalente ou por propriedade nas classes pai e filhos, por se tratar de que “disjoint” é uma regra de apenas um ou o outro.*
- f) *Não inserir na RB classes soltas que não possuem nenhuma ligação (por se tratar de uma regra de RB).*
- g) *Caso haja classes repetidas e com mesma classe pai, esta será convertida em apenas um nodo com vários pais.*

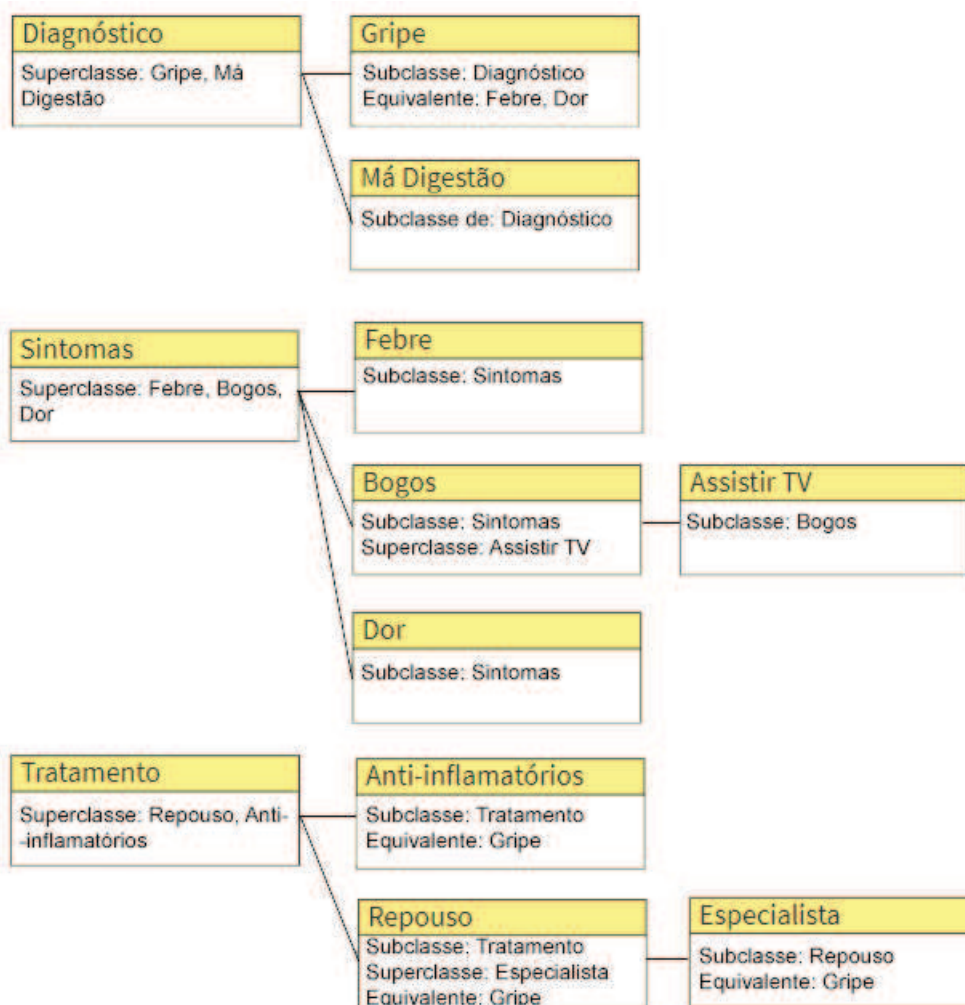
4.2.4. Extração dos dados

Como etapa final, ocorre a extração de dados, executada pelo componente “Extraindo Dados”, após todos os dados detectados e a estrutura analisada, o sistema converte os dados no formato JSON e gera a estrutura para que possa ser lida pelo editor bayesiano. Após, o sistema insere os estados de “sim” e “não” para cada nodo e a probabilidade proporcional em cada nodo (50,00% Sim, 50,00% Não), a qual o especialista pode ajustar a qualquer momento. No próximo passo é realizado o cálculo de Bayes e a inferência das probabilidades. Após a montagem da rede com os elementos extraídos, um algoritmo baseado na *Junction Tree* (ROSENBERG, 2017) aplica a fórmula de Nilsson (1998) para gerar as variáveis probabilísticas dos estados proporcionais, inferindo-os na rede. Por fim, o modelo ilustra graficamente a RB na tela, onde o usuário pode mover os nodos para adaptar ao seu estilo de visualização.

4.3. Exemplo de aplicação do modelo

Para ilustrar o funcionamento do modelo e como atua o modelo de conversão, criou-se uma ontologia fictícia de diagnóstico de gripe. Tal ontologia é ilustrada na Figura 21.

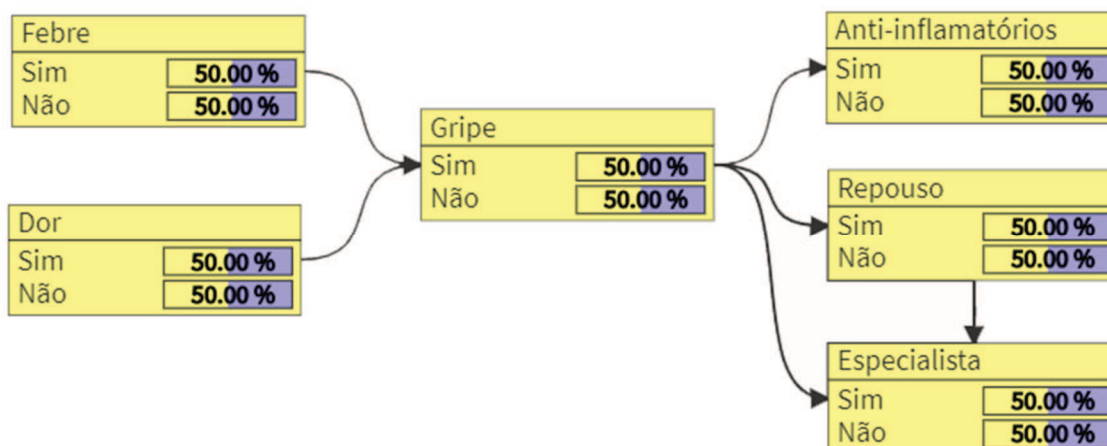
Figura 21: Ontologia de Diagnóstico de Gripe



Fonte: Elaborado pelo autor.

Quando importada a ontologia e analisada a estrutura, o código aplica o método de sumarização de ontologias. Na ontologia ilustrada na Figura 21, o método exclui as classes “diagnóstico”, “sintomas” e “tratamento”, pois as detecta como classes descritivas, onde servem apenas para informar o título de conteúdo de suas subclasses e, assim, não contendo valor para uma RB. Da mesma forma, o modelo identifica que a classe “bogos” e todas as suas subclasses se tratam de questões que não tem relevância, excluindo-as da conversão. O resultado pode ser visto na Figura 22.

Figura 22: Rede Bayesiana de diagnóstico de gripe



Fonte: Elaborado pelo autor.

Após excluir todas as classes que não tem relevância na RB, o modelo aplica as regras referentes a análise da estrutura (Seção 4.2.3), afim de realizar as ligações entre os nodos. Segundo a regra a), o sistema realiza a ligação entre as classes “Repouso” e “Especialista”, visto que as mesmas são classes pai e filha. Ao aplicar a regra “d”, o sistema faz a ligação das equivalências, ligando as classes “Febre” e “Dor” na classe “Gripe”, e a classe Gripe ligando as classes “Anti-inflamatórios”, “Repouso” e “Especialista”. Ao aplicar a regra “f”, o sistema exclui a classe “Má digestão” por não haver ligações na mesma.

Ao finalizar, o sistema monta a estrutura de RB, inserindo os estados “Sim” e “Não” em cada nodo, as probabilidades proporcionalmente (50,00% Sim, 50,00% Não), aplica o cálculo de Bayes e realiza a inferência das probabilidades. O próximo Capítulo apresenta a implementação e a avaliação do modelo proposto.

5. IMPLEMENTAÇÃO E AVALIAÇÃO

Neste Capítulo são descritos o protótipo desenvolvido para permitir a avaliação do modelo proposto e os procedimentos de avaliação realizados.

5.1. Protótipo

Com o objetivo de permitir a avaliação do modelo proposto em um ambiente de uso real de RB, um protótipo do modelo foi desenvolvido e aplicado como caso de estudo no software Bayes Editor (HELWANGER et al., 2017). A escolha deste software deu-se por permitir o acesso a um contexto real de uso de editor de Redes Bayesianas e também por possibilitar o contato com uma equipe de especialistas que utiliza o editor para criação de RBs na área da saúde.

O software Bayes Editor é utilizado pelos especialistas na construção de RBs que estão inseridos no projeto Health Simulator (BEZ et al., 2018). O editor bayesiano é *open source* disponibiliza uma web interface, da qual permite ao usuário a criação e/ou edição de redes bayesianas. A ilustração da web interface é ilustrado na Figura 23.

Figura 23: Web Interface do Editor Bayesiano



Fonte: Elaborado pelo autor.

Para a integração deste editor com o protótipo desenvolvido para avaliar o modelo, criou-se um módulo de conversão de ontologias para Redes Bayesianas. O protótipo dos componentes do modelo foi desenvolvido em Javascript e C#, favorecendo que o funcionamento se adapte as funcionalidades do editor. As conversões dos dados, aplicação do algoritmo de conversão são realizadas com a mesma linguagem de programação. O método de sumarização de ontologias foi desenvolvido em Java, visto a compatibilidade quanto aos arquivos de ontologias e a utilização dos grafos. A seguir, serão ilustrados os experimentados realizados com o protótipo.

5.2. Experimentos

Para a validação do modelo, foram criados dois protótipos. O primeiro, utilizava o dicionário de contexto para a extração dos dados relevantes da ontologia (como explicado na Seção 4.2.2), sendo chamado de protótipo 1.0 neste trabalho. Após o avanço nos estudos, foi desenvolvido um novo estudo substituindo o dicionário de contexto pelo método de sumarização de ontologias, tornando-se como a principal solução para essa proposta. Este será identificado como protótipo 2.0 nesta pesquisa.

Este trabalho dividiu a análise dos resultados em três experimentos. No primeiro experimento, é realizada uma validação com especialistas da área da saúde e computação utilizando a conversão manual e o protótipo 1.0 (utilizando o dicionário de contexto). No segundo, foram geradas RBs com o protótipo 2.0 de 12 ontologias disponíveis na web. Uma comparação com uma RBs já existente é realizada. No terceiro experimento, é realizado um caso de estudo com especialistas em enfermagem, utilizando também uma construção manual e o protótipo 2.0 (usando sumarização de ontologias). O experimento 1 e 2 tiveram dois questionários avaliativos diferentes para relatar a experiência dos usuários.

Através desses experimentos, foi possível identificar a veracidade do modelo como um método de auxílio aos especialistas, buscando comparar desde a construção manual realizada pelos mesmos, a sua efetividade quanto a uma conversão e a opinião dos mesmos. Será visto que o modelo gera Redes Bayesianas através de Ontologias de maneiras similares ou melhores às dos especialistas, tanto em questão de acerto, como de trabalho e tempo.

A assertividade dos nodos em relação ao especialista, as ligações corretas e o tempo de execução são critérios de avaliação que consideramos importantes na maioria dos experimentos realizados. Quanto a extração de ontologias da web, o critério de avaliação é a confirmação do especialista quanto ao sentido dos resultados.

5.2.1. Experimento com especialistas da saúde e computação utilizando o protótipo 1.0

Com o protótipo 1.0 utilizando o dicionário de contexto desenvolvido, foi realizado um segundo experimento envolvendo sete especialistas em saúde e cinco especialistas na área de computação. Esses participantes são especialistas na construção de RBs.

Todos os integrantes participaram coletivamente de uma introdução sobre ontologias, seus conceitos técnicos e funcionamento. Após o estudo teórico, cada perito recebeu um fragmento de ontologia de diagnóstico de Cefaleia, desenvolvida por Rodrigues (2012), com base no qual construíram manualmente uma RB conforme os dados que visualizaram na mesma. Após a conclusão, a ontologia foi inserida no módulo de extração desenvolvido nesta pesquisa para comparar os resultados.

Uma comparação do número de nodos do resultado da extração com as redes construídas pelos especialistas é mostrada na Tabela 14. A Tabela indica o número de nodos de cada categoria identificados pelo especialista usado como referência (coluna E), pelo modelo desenvolvido (coluna M) e pelos participantes do experimento (colunas 1 até 12).

Tabela 13: Comparativo das redes desenvolvidas na validação

Nodos	E	M	1	2	3	4	5	6	7	8	9	10	11	12
-------	---	---	---	---	---	---	---	---	---	---	---	----	----	----

Condutas	20	20	8	13	11	20	27	21	6	8	11	19	23	17
Tratamentos	5	5	1	0	6	0	6	0	6	0	3	4	0	3
Diagnósticos	5	5	3	6	3	5	7	1	6	3	5	5	0	5
Bogus	0	0	0	0	2	0	4	0	4	0	0	3	4	0

Legenda: E=Especialista médico criador de uma RB de Cefaleia (FONSECA, 2013, M=Modelo desenvolvido neste trabalho, Colunas do 1 ao 7=Especialistas em saúde, Colunas do 8 ao 12= Especialistas em computação.

Fonte: Elaborado pelo autor.

Foi possível verificar, através desta comparação, a dificuldade relacionada ao modelo de análise manual. Mesmo com os dados disponíveis, os especialistas levaram mais de duas horas para realizar uma construção manual da RB, e seus resultados, comparados aos de um especialista em cefaleia, são imprecisos e incompletos. O modelo mostrou-se eficaz e preciso quando comparado aos resultados do especialista médico. Observa-se na Tabela 14 que o modelo identificou todos os nodos identificados pelo especialista.

Os especialistas responderam a um questionário de acordo com sua área de especialização. As perguntas e o resultado do questionário para especialistas em saúde é mostrado na Tabela 15.

Tabela 14: Questionário para especialistas em saúde

Perguntas	DT	D	NDC	C	CT
Conheço algum software que realiza este tipo de tarefa.	43%	14,3%	14,3%	14,3%	14,3%
O uso do gerador semiautomático de RBs pode facilitar a formalização no computador do conhecimento da área da saúde.	0%	0%	14,3%	14,3%	71,5%
O uso do gerador semiautomático de RBs pode viabilizar o desenvolvimento de novos aplicativos na área de saúde.	0%	0%	0%	14,3%	85,8%
O modelo proposto é útil para a área da saúde.	0%	0%	0%	14,3%	85,8%
O uso da ontologia como base para geração de RBs pode simplificar o entendimento e a visualização das informações da área da saúde.	0%	0%	14,3%	14,3%	71,5%
Se necessitar de RBs e a informação estiver em ontologias, usarei o gerador semiautomático.	0%	0%	0%	28,6%	71,5%

A geração de RBs, a partir de ontologias, no Bayes Editor, é fácil de usar.	0%	0%	14,3%	14,3%	71,5%
---	----	----	-------	-------	-------

Legenda: DT= Discordo Totalmente, D=Discordo, NDC=Nem Concordo nem Discordo, C=Concordo, CT=Concordo Totalmente.

Fonte: Elaborado pelo autor.

Através deste questionário, é possível identificar que 57,3% dos especialistas em saúde não conhecem algum software que faça uma extração semiautomática de ontologias para geração de RBs, 14,3% nem discordam nem concordam, 28,6% indicam conhecer.

Um percentual de 85,8% dos usuários afirmou que o uso do módulo pode facilitar a formalização do conhecimento na área da saúde e 14,3% nem discordam nem concordam com isso. 100% dos participantes concordam que o gerador pode viabilizar o desenvolvimento de novos aplicativos para a área da saúde. Da mesma forma, 100% dos especialistas em saúde concordam que o modelo proposto é útil para a sua área de atuação.

Já um percentual de 85,8% afirmou que o uso de ontologias como base simplificaram o entendimento de informações da área da saúde. 14,3% nem discordaram nem concordaram com a afirmação. 100% afirmam totalmente que, se houver a necessidade de construir RBs através de ontologias, irão utilizar o gerador. 85,8% dos especialistas em saúde concordam que a extração semiautomática no módulo foi fácil de usar e 14,3% nem concordam nem discordam com esta afirmação.

Quanto aos especialistas na área da computação, o resultado do questionário é exibido na Tabela 16.

Tabela 15: Questionário para especialistas em computação.

PERGUNTAS	DT	D	NDC	C	CT
Conheço algum software que realiza este tipo de tarefa.	100%	0%	0%	0%	0%
A ferramenta proposta pode simplificar a construção de RBs com base em ontologias.	0%	0%	0%	20%	80%
O módulo para geração semiautomático de RBs é de fácil uso.	0%	0%	0%	0%	100%
A ferramenta apresentou erros na sua execução.	80%	0%	0%	0%	20%
A ferramenta apresenta uma interface clara e objetiva.	0%	0%	20%	0%	80%
A ferramenta é útil.	0%	0%	0%	20%	80%
Ao desenvolver um software para a área da saúde, se necessário, usaria o módulo desenvolvido.	0%	0%	0%	20%	80%

Legenda: DT= Discordo Totalmente, D=Discordo, NDC=Nem Concordo nem Discordo, C=Concordo, CT=Concordo Totalmente.

Fonte: Elaborado pelo autor.

Conforme a Tabela 16, é possível afirmar que 100% dos especialistas em computação não conhecem outras ferramentas que façam uma extração semiautomática de ontologias para geração de RBs.

A totalidade, ou seja, um percentual de 100% concordou que o módulo pode simplificar a construção de RBs com base em ontologias, da mesma forma que 100% concordaram que o módulo de extração é de fácil uso. Sobre erros no software, 80% discordaram haver erros, 20% afirmaram que houve um erro durante o processo. Ao analisar o erro, foi visto que o mesmo ocorreu pelo fato de uma oscilação provisória na internet, já que o editor é *online*.

Uma parte significativa, com um percentual de 80% afirmou que o módulo apresenta uma interface clara e objetiva, 20% nem concordam nem discordam dessa afirmação. 100% dos especialistas em computação concordam que a ferramenta é útil e 20% concordam com a afirmação. Igualmente, 100% dos usuários concordaram que utilizariam o módulo de extração de RBs, se necessário, quando desenvolverem um software para a área da saúde.

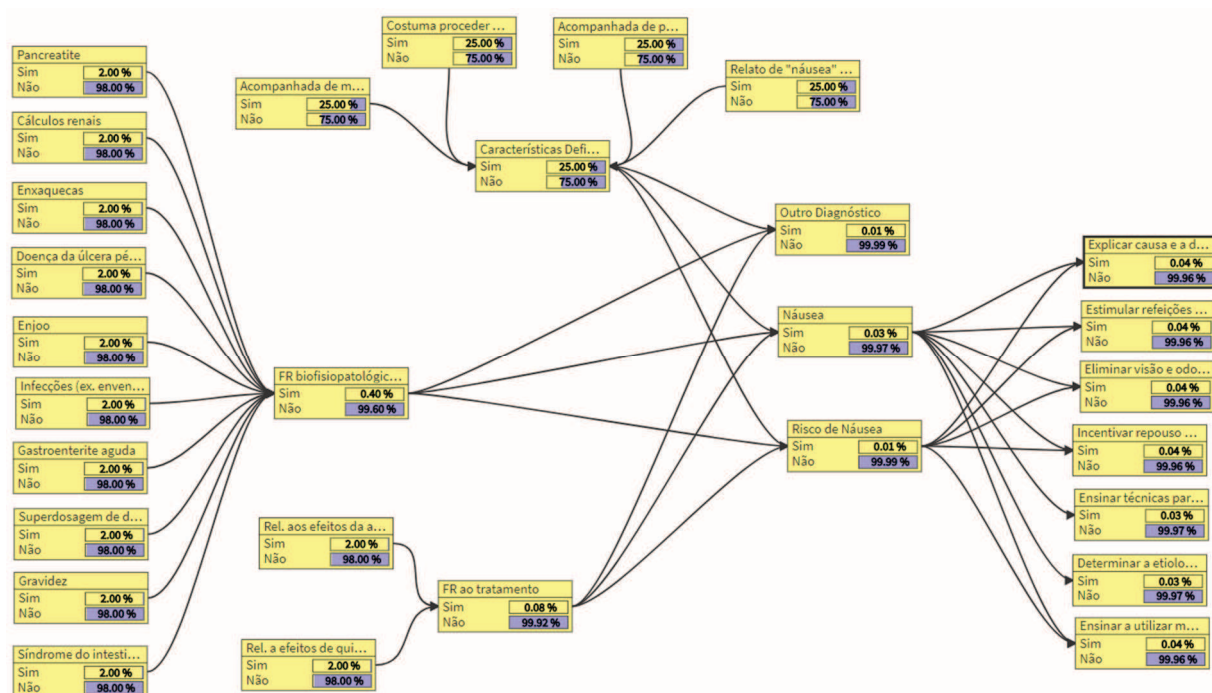
Através desses dados, é possível identificar que o modelo apresentou bons resultados e está funcional quanto a sua estrutura computacional. Também, que o modelo obteve uma boa receptividade dos especialistas na área de saúde. A percepção de utilidade é um ponto forte do modelo, especialmente porque o protótipo está disponibilizado de forma *online*, otimizando ainda mais o trabalho de um especialista, pois ele pode ser acessado a qualquer hora ou local, exigindo apenas uma conexão com a internet.

5.2.2. Experimento com ontologias da web usando o protótipo 2.0

No primeiro experimento, foram analisadas sete ontologias clínicas obtidas em repositórios web (BioPortal e OpenClinical) (BioPortal, 2018), (Open Clinical, 2018) e cinco ontologias construídas por especialistas em saúde. As Ontologias retiradas das bases são: “Alzheimer”, “Intolerância à Atividade”, “Constipação”, “Gripe”, “Meningite fúngica”, “Candidíase” e “Onicomicose”. Já as ontologias construídas por especialistas são: “Náusea” (Exemplo no Apêndice C), “Cefaleia”, “Eliminação Urinária Prejudicada”, “Integridade da Pele Prejudicada” e “Diarreia”. Essas ontologias foram escolhidas por sua capacidade de representar os principais cenários necessários à nossa avaliação. As doze ontologias clínicas foram convertidas em RBs pelo modelo desenvolvido. Os especialistas domínio avaliaram todas as RBs convertidas e confirmaram que os resultados fazem sentido e que a conversão pode ser considerada um sucesso.

Para ilustrar detalhes do procedimento, foi escolhido um dos doze experimentos, por questão de espaço, do qual será esboçado a seguir. O experimento escolhido utiliza uma ontologia que representa o diagnóstico de náusea desenvolvido pelos enfermeiros participantes do projeto Health Simulator, utilizado para gerar uma RB através do protótipo. A ontologia foi desenvolvida com base em um RB de náusea criado manualmente por essas enfermeiras (Figura 24). A RB foi utilizada como referência para comparar com a RB semiautomática gerada pelo modelo e para analisar os resultados. A Figura 25 ilustra um fragmento da ontologia usada no experimento.

Figura 24: RB de náusea gerada manualmente por especialistas.



Fonte: Elaborado pelo autor.

Figura 25: Ontologia de diagnóstico de náusea.

The screenshot shows a Semantic Web browser interface. On the left, a class hierarchy is displayed under 'owl:Thing'. The hierarchy includes:

- Diagnóstico
 - Nausea
 - Outro_Diagnóstico
 - Risco_de_Náusea
- Sintomas
 - Anestesia
 - Características_Definidoras
 - Cálculos_renais
 - Diarreia
 - Enjoo
 - Enxaquecas
 - Estômago
 - Gastroenterite_aguda
 - Gravidez
 - Infecções
 - Intestino_irritável
 - Irritação_Gástrica
 - Musculatura_esquelética
 - Pancreatite
 - Quimioterapia
 - Sequela_Tratamento
 - Superdosagem_de_drogas
 - Vômito
 - Úlcera_péptica
- Tratamento
 - Eliminar_odor_desagradável
 - Ensinar_técnicas
 - Etiologia
 - Explicar_causa_e_duração
 - Medicamentos_antieméticos** (highlighted)
 - Refeições_pequenas
 - Repouso

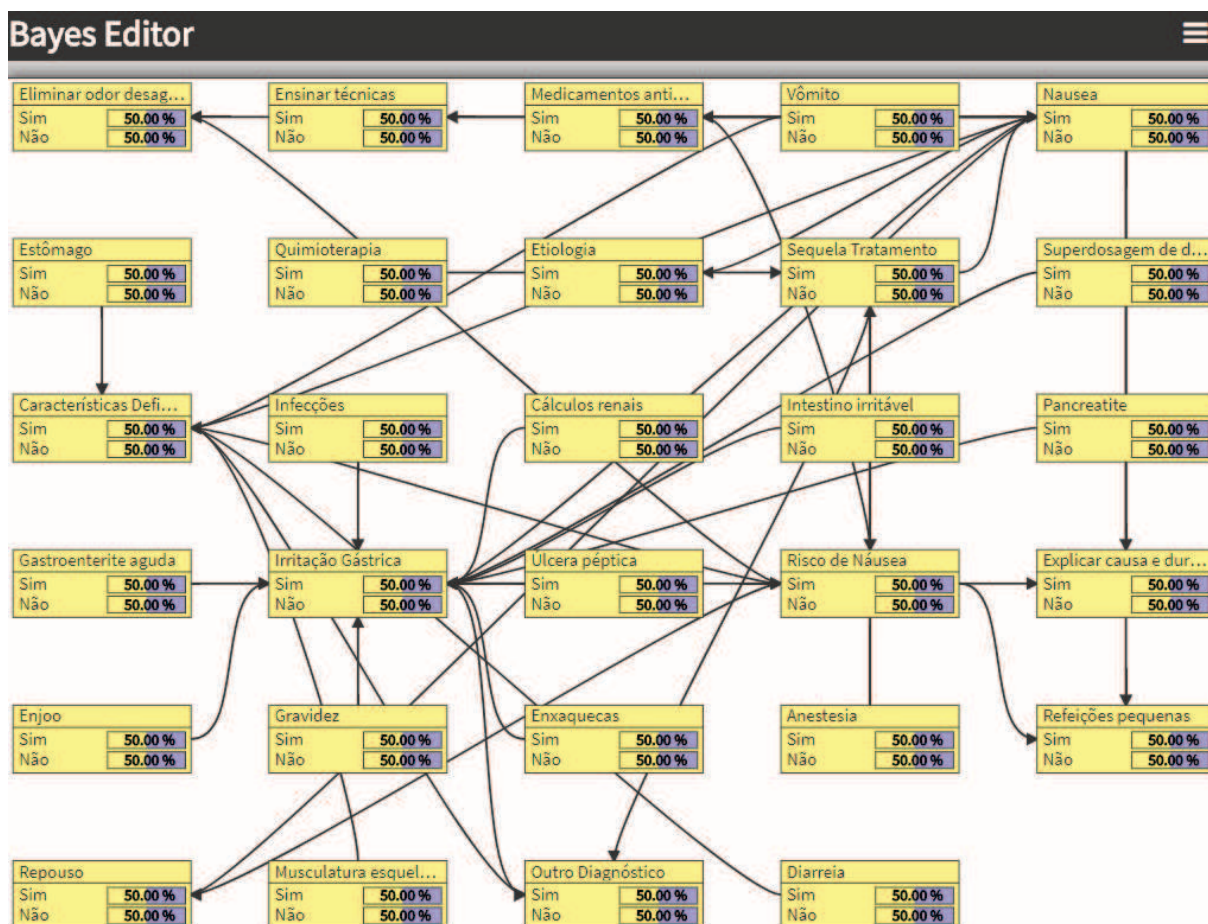
The right pane shows details for the class 'Medicamentos_antieméticos'. It includes a 'Description' section and a list of relationships:

- Equivalent To: Tratamento_de_some (Risco_de_Náusea and Nausea)
- SubClass Of: Tratamento
- General class axioms
- SubClass Of (Anonymous Ancestor)
- Instances
- Target for Key
- Disjoint With
- Disjoint Union Of

Fonte: Elaborado pelo autor.

Nesta ontologia, o modelo identificou com sucesso (100%) os principais nodos de um caso clínico (sinais/sintomas, diagnósticos e tratamentos), bem como suas conexões em cada nodo para ilustrar a lógica correta. Estas afirmações podem ser comprovadas comparando o a RB gerada (Figura 26) com a RB original construída pelo especialista (Figura 24). Embora a visualização esteja diferente, os nodos da RB estão corretos e relacionados positivamente com a ontologia extraída semi-automaticamente.

Figura 26 - RB de náusea gerada pelo módulo de extração.



Fonte: Elaborado pelo autor.

A Tabela 13 ilustra o comparativo do experimento apresentado com a RB construída pelo especialista. Nas colunas são exibidos os números de nós definidos pelos especialistas e pelo modelo.

Tabela 16: Comparativo da RB gerada da ontologia com a do especialista

Nodos	Especialista	Modelo
Condutas	19	19
Tratamentos	7	7
Diagnósticos	3	3
Bogus	0	0
Outros Nodos	0	0
Ligações Corretas	39	39
Ligações Incorretas	0	0

Fonte: Elaborado pelo autor.

5.2.3. Caso de estudo de enfermagem com protótipo 2.0

Para a validação dos resultados fornecidos pelo protótipo utilizando sumarização de ontologias, foi realizado um terceiro experimento envolvendo dez especialistas na área da enfermagem. Esses participantes são especialistas na construção de RBs, tendo entre 22 à 40 anos, dentre elas professoras, enfermeiras ou estudantes de enfermagem (graduação e doutorado).

Todos os integrantes participaram individualmente de uma introdução sobre ontologias, seus conceitos técnicos e funcionamento. Após o estudo teórico, cada perito recebeu um fragmento de ontologia de diagnóstico de Eliminação Urinária Prejudicada (Figura 27) desenvolvida por uma enfermeira, onde, durante duas semanas, construíram manualmente uma RB conforme os dados da mesma.

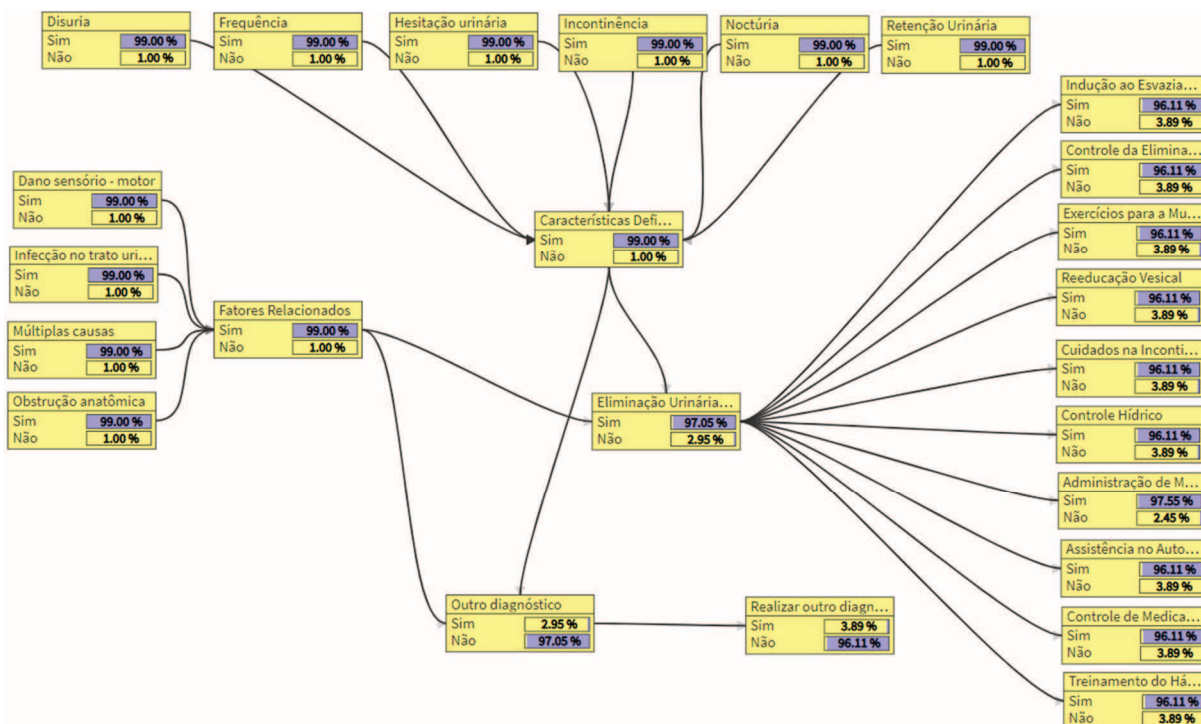
Figura 27: Fragmento de Ontologia de Eliminação Urinária Prejudicada.

The image shows a screenshot of an ontology editor interface. On the left, there is a 'Class Hierarchy' tree under the root 'owl:Thing'. The tree is expanded to show several levels of classes, including 'Investigações', 'Conduitas', 'Sintomas', 'Diagnósticos', and 'Tratamentos'. The class 'Assistência_no_Autocuidado:_Uso_do_Vaso_Sanitário' is highlighted in blue. On the right, the detailed view for this class is shown. It includes the IRI (http://www.semanticweb.org/diegopinheiro/ontologies/2018/9/urinaria#Assistência_no_Autocuidado:_Uso_do_Vaso_Sanitário), an 'Annotations' section with input fields for property and value, a 'Parents' section showing 'Tratamentos', and a 'Relationships' section showing a relationship 'Tratamento_de' with the class 'Eliminação_Urinária_Prejudicada'.

Fonte: Elaborado pelo autor.

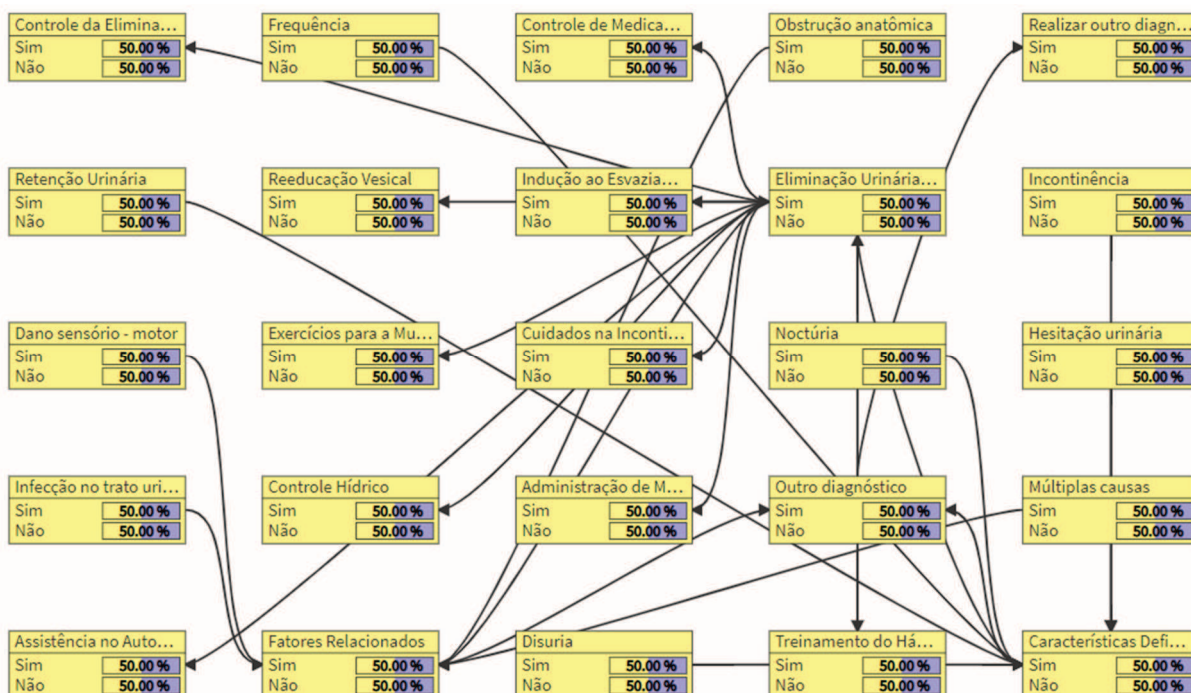
A mesma enfermeira que desenvolveu a ontologia também construiu uma Rede Bayesiana (Figura 28) de Diagnóstico de Eliminação Urinária Prejudicada, sendo assim, possível comparar a semelhança dos dados tanto do sistema desenvolvido como das demais enfermeiras. Após a conclusão, a ontologia foi inserida no módulo de extração desenvolvido nesta pesquisa para comparar os resultados (Figura 29).

Figura 28: Rede Bayesiana Original de Eliminação Urinária Prejudicada.



Fonte: Elaborado pelo autor.

Figura 29: Rede Bayesiana de Eliminação Urinária Prejudicada gerada pelo protótipo.



Fonte: Elaborado pelo autor.

Uma comparação do número de nós e ligações do resultado da extração com as redes construídas pelos especialistas é mostrada na Tabela 17. A Tabela indica o número de nodos de cada categoria identificados pelo especialista usado como referência (coluna E), pelo modelo desenvolvido (coluna M) e pelos participantes do experimento (colunas 1 até 10).

Tabela 17: Comparativo dos nodos

	E	M	1	2	3	4	5	6	7	8	9	10
Sinais	7	7	7	7	6	7	7	7	6	7	5	7
Sintomas	5	5	5	4	4	5	5	5	4	5	3	5
Diagnóstico	3	3	3	3	0	3	2	2	3	3	2	3
Tratamentos	10	10	10	9	8	10	9	8	10	10	7	10
Bogus	0	0	3	3	0	0	3	0	1	0	0	4
Outras	0	0	5	0	2	0	3	0	3	0	3	3
Ligações iguais E	25*	25	18	23	19	23	17	21	23	20	12	15
Ligações diferente E	0*	0	5	7	3	2	7	1	3	7	11	9

Legenda: Coluna E: Especialista em Enfermagem criador da Rede Bayesiana; Coluna M: Resultados do Modelo; Colunas 1 ao 10: Resultados dos demais especialistas em enfermagem.

Fonte: Elaborado pelo autor.

A comparação mostra que há precisão do modelo para análise manual. Mesmo com os dados disponíveis, os especialistas levaram até cinco horas para realizar uma construção manual da BN, e seus resultados, comparados aos do especialista em infecção urinária prejudicada, são imprecisos e incompletos. Também, foi coletado o grau de dificuldade que as enfermeiras enfrentam para a construção dessa rede, sendo relatado a um nível de 7 à 8 (sendo 1 muito fácil e 10 muito difícil). Isso mostra que, mesmo com a experiência, a construção manual de Redes Bayesianas segue sendo um processo difícil de ser realizado.

O modelo provou ser eficaz e preciso quando se trata de seus resultados. O algoritmo obteve um resultado semelhante aos resultados do especialista da enfermagem, extraindo corretamente as informações de condutas, diagnósticos e tratamentos, excluindo falsos e outros elementos inadequados para o caso. Em comparação com outros especialistas, o algoritmo é efetivo em termos de tempo e acurácia dos resultados, uma vez que algumas construções manuais foram incompletas e/ou incorretas de acordo com o caso clínico em questão, bem como o tempo utilizado para sua construção.

Foram elencados pontos, pelas enfermeiras, aspectos positivos e negativos tanto do processo manual como do processo semi-automatizado descrito neste trabalho. Dentre os pontos positivos do processo manual, destaca-se os resultados e a sua aplicabilidade e o uso da lógica e do conhecimento do conteúdo, bem como a visão ampla da doença em si. Quanto aos negativos, foram destacados pontos como o cálculo dos percentuais, o grande tempo gasto com a construção e especialização na doença, a dificuldade na construção da RB e a compreensão dos fatores que estão relacionados.

Quanto à avaliação do método semiautomático, como pontos positivos, destacou-se o tempo levado e a otimização do tempo das enfermeiras, a maior assertividade na construção da rede, a agilidade que trouxe ao processo e a facilidade de colocar em prática os diagnósticos de enfermagem. Como pontos negativos, em geral, as enfermeiras destacaram que não veem pontos negativos em usar a rede criada automaticamente por programas de computador. Uma enfermeira destacou o medo de que possa haver a alimentação de dados incompletos ou errados.

O protótipo utilizando sumarização de ontologias, quanto aos experimentos realizados, obteve uma qualidade semelhante ao do dicionário manual, ganhando principalmente quanto a não inserção manual. O método se torna superior tratando-se de ontologias maiores e de outras áreas de conhecimento de domínio. Destaca-se que, manter o dicionário desenvolvido anteriormente deixa o sistema menos flexível e exige um estudo maior, o que se busca evitar. Este trabalho solucionou este problema modificando o protótipo com o processo de sumarização de ontologias.

Com a execução dos três experimentos, é possível identificar que o modelo desenvolvido respeitando a semântica das ontologias ao convertê-las em Redes Bayesianas e, com isso, auxilia os especialistas da área mapeada na ontologia. O modelo foi bem recebido pelos especialistas, visto pelos comentários e as avaliações do mesmo. Nas comparações de nodos e ligações, é possível identificar que as Redes Bayesianas que foram geradas através de Ontologias são similares ou melhores às dos especialistas, tanto em questão de acerto, como de trabalho e tempo. Sendo assim, considera-se que este trabalho atingiu seu objetivo geral.

6. CONCLUSÃO

Este estudo apresentou a análise, desenvolvimento e validação de um novo modelo de extração semiautomática de RBs utilizando sumarização de ontologias. Foram apresentadas a revisão bibliográfica dos aspectos de embasamento teórico, uma pesquisa de trabalhos relacionados o método de sumarização de ontologias, o modelo desenvolvido e os experimentos de avaliação.

Este modelo foi implementado como caso de estudo no Bayes Editor (HELWANGER et al., 2017) e os dados gerados são usados como base de conhecimento no projeto de pesquisa Health Simulator (BEZ, et al.,2018). O modelo foi avaliado por especialistas na área de computação e saúde e, em um caso específico, da enfermagem, por meio de uma construção comparativa e de um questionário. Através da avaliação realizada, pode-se observar o potencial e importância deste trabalho para a área da saúde e qualquer outra, bem como a receptividade positiva dos especialistas na área de saúde.

Com este modelo, será possível reduzir o esforço destes especialistas que, na prática, tem um extenso trabalho para construir RBs manualmente, mesmo com os dados já presentes em uma base. O modelo não exige que os usuários tenham nenhum conhecimento da RB adicional, conforme exigido em grande parte dos trabalhos relacionados, e o fato de estar *online* o torna prático e acessível. Além disso, é compatível com ontologias, exigindo apenas que elas sejam descritas no formato de estrutura padrão atual. O modelo gera uma RB estruturada de acordo com os dados e estrutura da ontologia e, no final, o especialista deve inserir apenas a probabilidade de cada ação/nodo da rede.

Com isso, foi possível identificar que há uma gama de possibilidades a ser trabalhada para uma extração eficaz de RBs através de um modelo ontológico. Os modelos existentes nas publicações selecionadas são limitados, normalmente dependendo de um extenso trabalho manual, conhecimento de um especialista, construção de uma ontologia específica para o modelo que se espera chegar ou não, obtendo um resultado completo ou satisfatório.

Desta forma, a questão de pesquisa foi atendida, considerando-se que o modelo desenvolvido utiliza ontologias para a geração de RBs, preserva as restrições semânticas das ontologias e minimiza o trabalho dos especialistas. Quanto ao objetivo geral, considerado como desenvolver um modelo que converta os dados de ontologias para a geração de RBs, respeitando a semântica, auxiliando o especialista da área mapeada na ontologia, este foi concluído. No que tange aos objetivos específicos, estes serão apresentados e justificados sobre o atendimento na sequência:

1. Estudar sobre Ontologias e Redes Bayesianas: este foi atendido através da revisão bibliográfica realizada e apresentada nos Capítulos dois e três.
2. Verificar de que forma os trabalhos correlatos na literatura realizam a conversão de ontologias para Redes Bayesianas: este objetivo foi atendido através da revisão sistemática apresentada no Capítulo quatro.
3. Analisar ontologias existentes na internet: este objetivo foi atendido e apresentado junto ao modelo no Capítulo cinco. Este objetivo foi importante para a construção das regras de conversão.

4. Averiguar uma forma de converter o conhecimento das ontologias para a estrutura bayesiana: este objetivo foi atendido através das análises desenvolvidas no decorrer desta pesquisa.
5. Estudar técnicas e algoritmos para realizar a conversão semi-automaticamente: este objetivo foi atendido e o resultado do estudo foi apresentado no Capítulo cinco.
6. Desenvolver um protótipo para validação do modelo: este objetivo foi atendido, tendo sido apresentado no Capítulo seis.
7. Validar os resultados de experimentos com especialistas de um domínio: Este objetivo foi concluído, tendo sido apresentados os resultados dos experimentos no Capítulo seis.

6.1. Contribuições

Ao longo do período, foi desenvolvido um estudo sobre Redes Bayesianas e Ontologias, que resultou no artigo “Exploração do uso de bases de conhecimento e processamento de linguagem natural em um simulador de casos clínicos”, publicado nos anais do evento *Computer on the Beach* do ano de 2018 (PINHEIRO et al., 2018).

A utilização do protótipo como geração de conhecimento de um simulador de casos clínicos gerou um artigo denominado “Inteligência Artificial Aplicada a um Simulador na Área da Saúde”, apresentado no Congresso Brasileiro de Informática na Saúde de 2018 e publicado na revista *Journal of Health Informatics* (PINHEIRO et al., 2018).

Como demais contribuições, destaca-se os protótipos desenvolvidos com base na literatura, análises e consultas com especialistas. De forma pontual, destaca-se o método de sumarização de ontologias desenvolvido neste projeto. O protótipo desenvolvido está disponível para uso pela comunidade, como forma de estimular os estudos nesta área.

6.2. Trabalhos futuros

Como trabalhos futuros, novas formas de análise contextual, diagnóstico de sintaxes de ontologias e outras técnicas de inteligência artificial serão analisadas para aumentar a precisão do modelo. um estudo aprofundado sobre a geração de variáveis e pesos poderia ser realizado para complementar o modelo. um estudo aprofundado sobre a teoria dos grafos pode ser realizado a fim de otimizar a visualização final da rede bayesiana gerada pelo modelo. também será buscada a adaptação a outros formatos de arquivos e versões do Protégé.

REFERÊNCIAS

- ACM. **ACM digital library**. 2017. Disponível em: <<http://dl.acm.org/>>. Acessado em: 29/05/2017.
- AGARWAL, Prema; VERMA, Richa; MALLIK, Anupama. **Ontology based disease diagnosis system with probabilistic inference**. In: 2016 1st India International Conference on Information Processing (IICIP). IEEE, 2016. p. 1-5.
- AGUILAR, J. **Autonomic decision making based on bayesian networks and ontologies**. International Joint Conference on Neural Networks (IJCNN), 2016. p. 3825–3832.
- ANDREA, B.; FRANCO, T. **Mining Bayesian networks out of ontologies**. Journal of Intelligent Information Systems, 2012. v. 38, n. 2, p. 507–532.
- ANTONELI, Gabriela; ANTUNES, Michele; BARROS, Paulo Ricardo Muniz; RIGO, Sandro; CARVALHO, Marie Jane Soares; BEZ, Marta Rosecler. **O uso de simuladores no auxílio do ensino-aprendizagem na Enfermagem**. EDUCAÇÃO & LINGUAGEM, v. 21, p. 25-42, 2018.
- ARMAND, A.; FILLIAT, D.; IBAÑEZ-GUZMAN, J. **A framework for proactive assistance: Summary**. Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics, 2014. v. 2014–January, p. 684–687.
- ARSENE, O.; DUMITRACHE, I.; MIHU, I. **Expert system for medicine diagnosis using software agents**. Expert Systems with Applications, 2015. v. 42, n. 4, p. 1825–1834. Disponível em: <<http://dx.doi.org/10.1016/j.eswa.2014.10.026>>.
- ARSENE, O.; DUMITRACHE, I.; MIHU, I. **Medicine expert system dynamic Bayesian Network and ontology based**. Expert Systems with Applications, 2011. v. 38, n. 12, p. 15253–15261. Disponível em: <<http://dx.doi.org/10.1016/j.eswa.2011.05.074>>.
- BEAL, Matthew James. **Variational algorithms for approximate Bayesian inference**. London: University of London, 2003.
- BEZ, Marta Rosecler. **Construção de um Modelo para o Uso de Simuladores na Implementação de Métodos Ativos de Aprendizagem das Escolas de Medicina**. Porto Alegre, 2013. 314 f. Tese (Doutorado em Informática na Educação) – PGIE/UFRGS, Porto Alegre, 2013.
- BEZ, M. R. et al. **HEALTH SIMULATOR: um simulador de casos de estudo para a área da saúde**. Revista Observatório, [S.l.], v. 4, n. 3, p. 283-306, abr. 2018. ISSN 2447-4266. doi: <https://doi.org/10.20873/uft.2447-4266.2018v4n3p283>.
- BIBI, S. et al. **Ontology based Bayesian software process improvenent**. ICISOFT-EA 2014 - Proceedings of the 9th International Conference on Software Engineering and Applications, 2014. p. 568–575. Disponível em: <<http://www.scopus.com/inward/record.url?eid=2-s2.0-84908876208&partnerID=tZOtx3y1>>.
- BIO PORTAL. **Informação do portal**. Available in: <bioportal.bioontology.org>. Accessed in: Fev. 2018.

BLOCH, Francis and Jackson, MATTHEW O. and TEBALDI, Pietro. **Centrality Measures in Networks** (June 1, 2017). Available at SSRN: <https://ssrn.com/abstract=2749124> or <http://dx.doi.org/10.2139/ssrn.2749124>

BUCCI, G.; SANDRUCCI, V.; VICARIO, E. **Ontologies and Bayesian networks in medical diagnosis**. Proceedings of the Annual Hawaii International Conference on System Sciences, 2011. p. 1–8.

CARDOSO, P.C.F., PARDO, T.A.S., NUNES, M.G.V.: Métodos para Sumarização Automática Multidocumento Usando Modelos Semântico-Discursivos. In: Proceedings of the 3rd RST Brazilian Meeting, pp. 59-74. October 26, Cuiabá/MT, Brazil (2013)

CHAN, A.P.C.; WONG, F.K.W.; HON, C.K.H.; CHOI, T.N.Y. **A Bayesian Network Model for Reducing Accident Rates of Electrical and Mechanical (E&M)**. Work. Int. J. Environ. Res. Public Health 2018, 15, 2496.

CHANG, Y. S.; HUNG, W. C.; JUANG, T. Y. **Depression diagnosis based on ontologies and bayesian networks**. Proceedings - 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013, 2013. p. 3452–3457.

CONG, R. H. **About the Studying on the Uncertainty of Ontology on the Bayesian Network**. Applied Mechanics and Materials, 2014. v. 513–517, p. 1717–1721. Disponível em: <<http://www.scientific.net/AMM.513-517.1717>>.

COOPER, D., 2016. What is a mapping study? J. Med. Library Assoc. 104, 76–78.

DING, Z. (2011). **The Development of Ontology Information System Based on Bayesian Network and Learning**. In: Jin D., Lin S. (eds) Advances in Multimedia, Software Engineering and Computing Vol.2. Advances in Intelligent and Soft Computing, vol 129. Springer, Berlin, Heidelberg

FENZ, S. **An ontology-based approach for constructing Bayesian networks**. Data & Knowledge Engineering, 2012. v. 73, p. 73–88. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0169023X11001583>>.

FISCHER, Y.; BEYERER, J. **Ontologies for probabilistic situation assessment in the maritime domain**. 2013 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, CogSIMA 2013, 2013. p. 102–105.

FONSECA, J. M. L. (2013). **Descrição de um simulador baseado em RBs como modelo de avaliação do aprendizado de diretrizes clínicas em ensino a distância para medicina de família e comunidade** (Doctoral dissertation).

FORNAHL, Dirk; HASSINK, Robert; MENZEL, Max-Peter Menzel. (2015) **Broadening Our Knowledge on Cluster Evolution**. European Planning Studies 23:10, pages 1921-1931.

FUJITA, H.; KUREMATSU, M.; HAKURA, J. **Multiviews ontologies based reasoning for medical diagnosis in VDS**. SISY 2011 - 9th International Symposium on Intelligent Systems and Informatics, Proceedings, 2011. p. 397–406.

GRUBER, T. R. **Toward principles for the design of ontologies used for knowledge sharing**. USA, ago 1993. Technical report KSL 93-04.

GRUBER, T. R. **What is an Ontology?** USA, ago 2002. Disponível em: <<http://wwwksl.stanford.edu/kst/what-is-an-ontology.html>>. Acesso em: fev 2018.

GRUBIŠIĆ, A.; STANKOV, S.; PERAIĆ, I. **Ontology based approach to Bayesian student model design.** Expert Systems with Applications, 2013. v. 40, n. 13, p. 5363–5371. Disponível em: <<http://dx.doi.org/10.1016/j.eswa.2013.03.041>>.

HASHIMOTO, S. et al. **A Framework Of Ontology-Based Knowledge Information Processing For Change Detection In Remote Sensing Data.** Graduate School of Information Science and Technology , Hokkaido University Earth Observation Research Center , Japan Aerospace Exploration Agency. 2011. p. 3927–3930.

HELWANGER, A. H; BEZ, M. R; STAHNKE, F. R; BARROS, P. R. M. **Bayes Editor: Desenvolvimento e validação de um editor de RBs para o uso no ensino na saúde.** Informática na Educação: teoria & prática, Porto Alegre, v.20, n.4, dez. 2017.

HELWANGER, F. **Um Editor de RBs com Foco em Usabilidade.** Trabalho de conclusão de curso, Universidade Feevale, 2016.

HERNANDES, Elis et al. **Avaliação da ferramenta StArt utilizando o modelo TAM e o paradigma GQM.** In: Proceedings of 7th Experimental Software Engineering Latin American Workshop (ESELAW 2010). 2010. p. 30.

IEEEXPLORE. **IEEE Xplore digital library.** 2017. Disponível em: <<http://ieeexplore.ieee.org/xpl/aboutUs.jsp>>. Acessado em 16/08/2017.

JULINA, J. K. J.; THENMOZHI, D. **Ontology based EMR for decision making in health care using SNOMED CT.** 2012 International Conference on Recent Trends in Information Technology, 2012. p. 514–519. Disponível em: <<http://uvic.summon.serialsolutions.com/2.0.0/link/0/eLvHCXMwY2BQMDFISrUwSjEzNTI3AGYfYK2QamJpnJZsmQbs4BqB705BbB1DKs3dhBiYUvNEGVTCXEOcPXRBI53EF0DOXIg3ALdmzEyA9YtRomWagYEYAwuwf5wKAGA3Gdk>>.

JUNQUEIRA, K. T. C.; FERNANDES, A. M R. **Análise de Sentimento em Redes Sociais no Idioma Português com Base em Mensagens do Twitter.** Computer on the Beach, Florianópolis: 2018.

KALET, A. M.; DOCTOR, J.N.; GENNARI, J.H.; PHILLIPS, M.H. **Developing Bayesian networks from a dependency-layered ontology: A proof-of-concept in radiation oncology.** Med Phys. 2017 Aug;44(8):4350-4359. doi: 10.1002/mp.12340. Epub 2017 Jun 30.

KITCHENHAM, Elisabete. **Guidelines for performing Systematic Literature Reviews in Software Engineering version 2.** Department of Computer Science University of Durham Durham, UK, 2007.

KOVALENKO, Mykyta; ANTOSHCHUK, Svetlana. **Human action recognition using a semantic-probabilistic network.** In: 2015 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC). IEEE, 2015. p. 67-72.

KÖCHE, José Carlos. Fundamentos de metodologia científica: teoria da ciência e iniciação à pesquisa. 29. ed. [s.l.] : Editora Vozes, 2011.

KOVALENKO, M.; ANTOSHCHUK, S.; HODOVYCHENKO, M. **Event recognition using a semantic-probabilistic network**. 2015 Information Technologies in Innovation Business Conference, ITIB 2015 - Proceedings, 2015. p. 35–38.

LIMA, A. et al. **Projeto para desenvolvimento do Simulador Health Simulator**. Anais do Computer on the Beach, Florianópolis, 2015. 279-288.

LIN, H. et al. **A novel approach to power transformer fault diagnosis based on ontology and Bayesian network**. In: 2014 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC). IEEE, 2014. p. 1-6.

MANNAI, M. **Bayesian information extraction network for Medline abstract**. 2013. p. 1–3.

MEDEIROS, Edna Ramos de. **Revisão Sistemática Sobre os Dispositivos Vestíveis na Área da Saúde**. Trabalho de Conclusão de Curso de Sistemas de Informação. Universidade Feevale, 2016.

MESSAOUD, M. BEN; LERAY, P.; AMOR, N. BEN. **Active learning of causal Bayesian networks using ontologies: A case study**. The 2013 International Joint Conference on Neural Networks (IJCNN), 2013. n. August, p. 1–8. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6706815>>.

MESSAOUD, M. BEN; LERAY, P.; AMOR, N. BEN. **SemCaDo: A serendipitous strategy for causal discovery and ontology evolution**. Knowledge-Based Systems, 2015. v. 76, p. 79–95. Disponível em: <<http://dx.doi.org/10.1016/j.knosys.2014.12.006>>.

MOK, J.; MIN, H. **Ontology-based context-aware model by applying Bayesian network**. 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2015, 2016. p. 2660–2664.

NILSSON, Nils J. **Artificial intelligence: a new synthesis**. San Francisco: Morgan Kaufmann, 1998. 513 p.

NOLLETO, Felipe; BARROS, Paulo Ricardo Muniz; PINHEIRO, Diego. **Desenvolvimento do Módulo de Redes Bayesianas Multi-Seccionadas para o Health Simulator**. Novo Hamburgo: Anais do GamePad, 2017.

OPEN CLINICAL. **Informação do portal**. Available in: <<http://www.openclinical.org/ontologies>>. Accessed in: Fev. 2018.

PAVLIN, J. P. D. V. G.; LASKEY, P. C. K. **A URREF interpretation of Bayesian network information fusion**. Information Fusion (FUSION), 2014 17th International Conference on. Disponível em: <http://ieeexplore.ieee.org/ielx7/6900113/6915967/06916273.pdf?tp=&arnumber=6916273&isnumber=6915967%5Cnhttp://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6916273&tag=1>.

PEARL, Judea. **Probabilistic reasoning in intelligent systems: networks of plausible inference**. Elsevier, 2014.

PILATO, G. et al. **Integration of ontologies and Bayesian networks for maritime situation awareness**. Proceedings - IEEE 6th International Conference on Semantic Computing, ICSC 2012, 2012. p. 170–177.

PINHEIRO, Diego et al. **Exploração do uso de bases de conhecimento e processamento de linguagem natural em um simulador de casos clínicos**. Anais do Computer on the Beach, p. 641-650, 2018.

PINHEIRO, D; MELLO, B; ROCKENBACK, L; BEZ, M; RIGO, S. **Inteligência Artificial Aplicada a um Simulador na Área da Saúde**. XVI Congresso Brasileiro de Informática em Saúde. Fortaleza, 2018. Disponível em: www.sbis.org.br/biblioteca_virtual/cbis/Anais_CBIS_2018_Artigos_Completos.pdf

PINHEIRO, Diego; CERVI, Gustavo; SCHUH, Anderson; BARROS, Paulo R. M.; BEZ, Marta R. **Redes Bayesianas como geração de conhecimento para games**. Novo Hamburgo: Anais do GamePad, 2015.

PRISMA. **Preferred reporting items for Systematic Reviews and Meta-Analyses**. 2015. Disponível em: <<http://prisma-statement.org/PRISMAStatement/Checklist.aspx>> . Acessado em 16/08/2017.

PUBMED. **PubMed**. 2017. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed>>. Acessado em 16/08/2017.

QINGHONG, Y. **A Method of Constructing User Interest Model Based on Semantics**. 2012. n. Iccse.

QUEIROZ-SOUSA, P.O., SALGADO, A.C. AND PIRES, C.E., 2013. **A Method for Building Personalized Ontology Summaries**. Journal of Information and Data Management, 4(3), pp.236–250.

ROBLEDO-GIRALDO, S., DUQUE-MÉNDEZ, N. D., ZULUAGA-GIRALDO, J. I. **Difusión de productos a través de redes sociales: una revisión bibliográfica utilizando la teoría de grafos**. Revista Respuestas, 2013. 18(2), 27–41.

RODRIGUES, F. H.; BEZ, M. R.; FLORES, C. D. **Generating Bayesian networks from medical ontologies**. 2013 8th Computing Colombian Conference, 8CCC 2013, 2013.

ROSENBERG, A. <http://eniac.cs.qc.cuny.edu/andrew/gcml/lecture10.pdf>. **Junction Tree Algorithm CSCI 780 - Machine Learning**. 4 Março 2004. Disponível em: <<http://eniac.cs.qc.cuny.edu/andrew/gcml/lecture10.pdf>>. Acesso em: Novembro 2017.

RUOHONEN, K. **Graph theory**. Tampere University of Technology, Tampere, Finland, J. Tamminen, K.-C. Lee, & R. Pich, eds., 114 p., 2013.

SANTOS, E.; JURMAIN, J. C. **Bayesian knowledge-driven ontologies: Intuitive uncertainty reasoning for semantic networks**. Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics, 2011. p. 856–863.

SAYED, M. S.; LOHSE, N. **Ontology-driven generation of Bayesian diagnostic models for assembly systems**. International Journal of Advanced Manufacturing Technology, 2014. v. 74, n. 5–8, p. 1033–1052.

SETTAS, D.; CERONE, A.; FENZ, S. **Towards automatic generation of ontology-based antipattern Bayesian network models.** Proceedings - 2011 9th International Conference on Software Engineering Research, Management and Applications, SERA 2011, 2011. p. 46–53.

SOUSA, P. O.; PIRES, C. E., SALGADO, A. C. **A Method for Building Personalized Ontology Summaries.** Journal of Information and Data Management (JIDM), v. 4, n. 3, p. 236-250, 2013

SOUZA, A. S., DURAN, A., & VIEIRA, V. (2014). **Uma Ontologia de Domínio para a Metodologia de Aprendizagem Baseada em Problemas.** In Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE) (Vol. 25, No. 1, p. 1253).

STAJDUHAR, Ivan; DALBELO-BASIC, Bojana; BOGUNOVIC, Nikola. **Impact of censoring on learning Bayesian networks in survival modelling.** Artificial Intelligence in Medicine (2011) 47, 199—217.

TROULLINO, Georgia et al. **RDF digest: Efficient summarization of RDF/S KBs.** In: European Semantic Web Conference. Springer, Cham, 2015. p. 119-134.

URBANSKY, S. M. C. **Uma ontologia para representação de conhecimento sobre boas práticas nas infecções relacionadas à assistência à saúde.** Dissertação (Mestrado em Ensino na Saúde) – UFCS, Porto Alegre, 2018.

WAZLAWICK, Raul. **Metodologia de pesquisa para ciência da computação.** Elsevier Brasil, 2017.

WEB OF SCIENCE. **Web Of Science.** 2017. Disponível em: <<http://www.apps-webofknowledge.ez101.periodicos.capes.gov.br>>. Acessado em 16/08/2017.

YAN, L. et al. **Development of a novel asset management system for power transformers based on ontology.** Power and Energy Engineering Conference (APPEEC), 2013 IEEE PES Asia-Pacific, 2013. p. 1–6.

YANG, Y; CALMET, J. **OntoBayes: An Ontology-Driven Uncertainty Model.** International Conference on Computational Intelligence for Modelling, Control and Automation, 2005.

ZACARIAS A. C. I. **Investigação de métodos de sumarização automática multidocumento baseados em hierarquias conceituais.** Dissertação para conclusão de curso em linguística. São Carlos, 2016.

ZACARIAS, A. C. I.; FELIPPO, A. D; **Explorando Hierarquias Conceituais para a Seleção de Conteúdo na Sumarização Automática Multidocumento.** Proceedings of Symposium in Information and Human Language Technology. Natal, RN, Brazil, November 4–7, 2015. Sociedade Brasileira de Computação.

ZANATTA, Eduardo J; RODRIGUES, Fabrício Henrique; CAZELLA, Silvio Cesar; BEZ, Marta Rosecler. **Modelando Ontologias a Partir de Diretrizes Clínicas: Diagnóstico e Tratamento da Cefaléia.** Conference: Proceedings of Joint V Seminar on Ontology Research in Brazil and VII International Workshop on Metamodels, Ontologies and Semantic Technologies, Recife, Brazil, September 19-21, 2013.

ZANETTI, H. A., & BONACIN, R. (2014). **Uma Metodologia Baseada em Semiótica para Elaboração e Análise de Práticas de Ensino de Programação com Robótica Pedagógica.** In Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE) (Vol. 25, No. 1, p. 1233).

APÊNDICES

APÊNDICE A - PROTOCOLO DA REVISÃO SISTEMÁTICA

Título - Revisão sistemática sobre extração de dados de ontologias para conversão em RBs aplicadas à saúde.

Resumo - Esta revisão sistemática utilizará como guia o protocolo criado por Medeiros (2016), tendo como tema desta revisão a extração de dados de ontologias e RBs, buscando também assuntos de interesse referentes às mesmas, como suas ligações com a saúde e técnicas de extração de dados e PLN utilizadas. O protocolo foi desenvolvido com a ajuda de especialistas da área de computação (ligados a saúde e IA) e, caso for necessário, poderá sofrer ajustes posteriormente, sendo todas as modificações justificadas.

Objetivo – PICOC

Para o auxílio na criação da *string* de busca, Kitchenham (2007) recomenda a utilização do PICOC (População, Intervenção, Comparação, Resultados e Contexto). A formulação da pesquisa é apresentada a seguir.

FORMULAÇÃO DE PESQUISA

Foco da questão: Esta revisão sistemática busca encontrar os artigos que trabalham com ontologias e RBs, buscando encontrar técnicas e validações de IA, como PLN e extração de dados, visando principalmente verificar na literatura modos de conversão de ontologias para RBs para serem aplicadas à saúde.

Questões de interesse: Técnicas utilizadas, Abordagens, Estudos já realizados e Aplicações recorrentes.

Palavras-chave: *Bayesian network, Ontology, Health, Natural Language Processing, Data extraction.*

Intervenção: Verificar as tecnologias que são utilizadas nas pesquisas.

Controle: Não será utilizado.

Efeito: Identificar as oportunidades de pesquisa na área de IA, voltando o conhecimento para ontologias e RBs.

Medida do Resultado: Gerar um embasamento teórico para a dissertação, assim como, a publicação de artigos científicos.

População de interesse: Pesquisadores, professores, desenvolvedores e profissionais da computação.

Aplicação: Esta revisão sistemática tem como foco pesquisadores, professores e alunos da área da computação aplicada, pois trará uma visão abrangente de tecnologias e metodologias adotadas até o momento para o desenvolvimento de pesquisas.

Desenho do experimento: Não será desenvolvido.

Financiamento: CAPES.

FORMULAÇÃO DE CRITÉRIOS

Definição dos critérios de seleção das fontes de dados: as fontes de dados foram selecionadas através de trabalhos relacionados e indicações dos orientadores deste projeto. As mesmas possuem artigos relacionados a computação e saúde. Para a área da computação serão utilizadas as bases de dados: Web Of Science, que possui o acesso a periódicos em diversas áreas do conhecimento, tendo conteúdo integral apenas para pesquisadores da CAPES (Web Of Science, 2017); ACM Digital Library, que é uma base de dados bibliográficos com foco na área da computação (ACM, 2017); IEEEExplore, que fornece acesso a algumas das publicações mais citadas no mundo em Ciência da Computação, Eletrônica e Engenharia Elétrica (IEEEEXPLORE, 2017). Para a área da saúde, será utilizada a base de dados PubMed, que conta com publicações biomédicas do MEDLINE (PubMed, 2017);

Idiomas das Fontes de Dados: somente serão consideradas as publicações que estiverem no idioma em inglês.

String de busca: a *string* de busca foi criada conforme as palavras-chave definidas, tendo como elementos obrigatórios da pesquisa o uso de “Ontologias” (Ontology) e “RBs” (bayesian network). Os termos “Saúde” (health), “Processamento de Linguagem Natural” (natural language processing) e “Extração de Dados” (data extraction) serão definidos como elementos auxiliares. A *string* será aplicada nas bases de dados definidas anteriormente.

(“ontology”) AND (“bayesian network”) OR (“health”) OR (“natural language processing”) OR (“NLP”) OR (“data extraction”).

Artigos de controle: optou-se por não utilizar nenhum artigo de controle para esta revisão sistemática.

SELEÇÃO DOS ESTUDOS

Critérios para a inclusão/exclusão dos resultados:

- As publicações selecionadas devem obedecer aos seguintes critérios:
- O ano de publicação do artigo deve estar dentro do período de 2011 e 2017;
- Ser um artigo científico publicado;
- Deve estar escrito em inglês;
- O artigo deve apresentar uma forma de validação;
- A publicação deve estar disponível na íntegra na internet ou disponível através de convênios das instituições de ensino.
- Ser validado pelo software My Tree of Science (ROBLEDO-GIRALDO et al., 2013).

Definição dos tipos de estudo: serão selecionados estudos dos tipos teóricos, qualitativos e quantitativos referentes aos temas.

Procedimentos para seleção dos estudos: na primeira etapa, será executada a *string* de busca nas bases de dados selecionadas. Todas as publicações encontradas serão exportadas no formato BibTex para serem cadastradas na ferramenta [StArt](#). Após o registro, serão validados os critérios de inclusão já definidos. Os artigos da base Web Of Science passarão pela validação do software ToS. Avante, serão realizadas as leituras de todos os títulos, palavras-chave e resumos, passando em seguida para uma validação de introdução e conclusão e, na fase final, a leitura integral dos artigos selecionados.

Fases de seleção de artigos:

Fase 1 - Validar os critérios de inclusão e exclusão;

Fase 2 - Validar as publicações na ferramenta TOS (conforme as bases compatíveis).

Fase 3 - Leitura do título, palavras-chave e resumo;

Fase 4 - Leitura da introdução e conclusão;

Fase 5 - Leitura integral dos artigos e validação das respostas para as perguntas.

Critérios de qualidade das fases da Revisão Sistemática: os critérios de qualidade que serão avaliados nesta revisão sistemática estão descritos a seguir:

- Quais foram as técnicas mais utilizadas?
- Que tipos de metodologias foram usadas?
- Qual a ligação entre os temas propostos?
- O artigo apresenta algum protótipo desenvolvido?
- Que formas de validação foram utilizadas?
- Os artigos foram aplicados à área da saúde?
- Quais as linguagens que foram utilizadas?
- A proposta teve resultados positivos?
- Quais foram as ferramentas utilizadas?

Análises adicionais: não há.

APÊNDICE B – DICIONÁRIO CRIADO PARA ANÁLISE DE CONTEXTO

Português - Sinal, Sinais, Sintoma, Sintomas, Diagnóstico, Diagnósticos, Conduta, Condutas, Tratamento, Tratamentos, Característica Definidora, Características Definidoras, Intervenções, Intervenção, Queixas, Queixa, Características Observadas, Característica Observada, Quadro Observado, Quadros Observados, Bogo, Bogos, Investigação, Investigações, Casos Clínicos, Caso Clínico,

Inglês - *Sign, Signal, Signals, Symptom, Symptoms, Diagnosis, Diagnostics, Conduct, Ducts, Pipelines, Treatment, Treatments Defining Characteristics, Defining Characteristic, Intervention, Interventions, Complaints, Complaint, Characteristics Observed, Characteristic Observed, Observed Frame, Observed Frames, Bogus, Clinical Cases, Clinical Case, Investigation, Investigations.*

Espanhol - *Señal, Señales, Síntoma, Síntomas, Conducta, Tratamiento, Tratamientos, Características Definitórias, Intervención, Intervenciones, Quejas, Queja, Cuadro Observado, Cuadros Observados, Investigación, Investigaciones.*

APÊNDICE C – EXEMPLO DE ONTOLOGIA EXTRAÍDA (NÁUSEA)

```

<?xml version="1.0"?>
<Ontology xmlns="http://www.w3.org/2002/07/owl#"

xml:base="http://www.semanticweb.org/diegopinheiro/ontologies/2018/2/OntologiaNausea"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xml="http://www.w3.org/XML/1998/namespace"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"

ontologyIRI="http://www.semanticweb.org/diegopinheiro/ontologies/2018/2/OntologiaNausea">
  <Prefix name="" IRI="http://www.semanticweb.org/diegopinheiro/ontologies/2018/2/OntologiaNausea#"/>
  <Prefix name="owl" IRI="http://www.w3.org/2002/07/owl#"/>
  <Prefix name="rdf" IRI="http://www.w3.org/1999/02/22-rdf-syntax-ns#"/>
  <Prefix name="xml" IRI="http://www.w3.org/XML/1998/namespace"/>
  <Prefix name="xsd" IRI="http://www.w3.org/2001/XMLSchema#"/>
  <Prefix name="rdfs" IRI="http://www.w3.org/2000/01/rdf-schema#"/>
  <Declaration>
    <Class IRI="#Eliminar_odor_desagradável"/>
  </Declaration>
  <Declaration>
    <Class IRI="#Ensinar_técnicas"/>
  </Declaration>
  <Declaration>
    <Class IRI="#Medicamentos_antieméticos"/>
  </Declaration>
  <Declaration>
    <ObjectProperty IRI="#Dignóstico_de"/>
  </Declaration>
  <Declaration>
    <Class IRI="#Diagnóstico"/>
  </Declaration>
  <Declaration>

```

```
<Class IRI="#Tratamento"/>
</Declaration>
<Declaration>
  <Class IRI="#Vômito"/>
</Declaration>
<Declaration>
  <Class IRI="#Nausea"/>
</Declaration>
<Declaration>
  <Class IRI="#Estômago"/>
</Declaration>
<Declaration>
  <Class IRI="#Quimioterapia"/>
</Declaration>
<Declaration>
  <Class IRI="#Etiologia"/>
</Declaration>
<Declaration>
  <Class IRI="#Sequela_Tratamento"/>
</Declaration>
<Declaration>
  <Class IRI="#Superdosagem_de_drogas"/>
</Declaration>
<Declaration>
  <Class IRI="#Características_Definidoras"/>
</Declaration>
<Declaration>
  <ObjectProperty IRI="#Tratamento_de"/>
</Declaration>
<Declaration>
  <Class IRI="#Infecções"/>
</Declaration>
<Declaration>
  <Class IRI="#Cálculos_renais"/>
</Declaration>
```

```
<Declaration>
  <Class IRI="#Intestino_irritável"/>
</Declaration>
<Declaration>
  <Class IRI="#Pancreatite"/>
</Declaration>
<Declaration>
  <Class IRI="#Gastroenterite_aguda"/>
</Declaration>
<Declaration>
  <ObjectProperty IRI="#Resultado_de"/>
</Declaration>
<Declaration>
  <Class IRI="#Irritação_Gástrica"/>
</Declaration>
<Declaration>
  <Class IRI="#Sintomas"/>
</Declaration>
<Declaration>
  <Class IRI="#Úlcera_péptica"/>
</Declaration>
<Declaration>
  <Class IRI="#Risco_de_Náusea"/>
</Declaration>
<Declaration>
  <Class IRI="#Explicar_causa_e_duração"/>
</Declaration>
<Declaration>
  <Class IRI="#Enjoo"/>
</Declaration>
<Declaration>
  <Class IRI="#Gravidez"/>
</Declaration>
<Declaration>
  <Class IRI="#Enxaquecas"/>
```

```

</Declaration>
<Declaration>
  <Class IRI="#Anestesia"/>
</Declaration>
<Declaration>
  <Class IRI="#Refeições_pequenas"/>
</Declaration>
<Declaration>
  <Class IRI="#Repouso"/>
</Declaration>
<Declaration>
  <Class IRI="#Musculatura_esquelética"/>
</Declaration>
<Declaration>
  <Class IRI="#Outro_Diagnóstico"/>
</Declaration>
<Declaration>
  <Class IRI="#Diarreia"/>
</Declaration>
<EquivalentClasses>
  <Class IRI="#Características_Definidoras"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Diarreia"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Características_Definidoras"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Estômago"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Características_Definidoras"/>

```



```

<ObjectSomeValuesFrom>
  <ObjectProperty IRI="#Resultado_de"/>
  <Class IRI="#Musculatura_esquelética"/>
</ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Características_Definidoras"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Vômito"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Eliminar_odor_desagradável"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Tratamento_de"/>
    <Class IRI="#Nausea"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Eliminar_odor_desagradável"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Tratamento_de"/>
    <Class IRI="#Risco_de_Náusea"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Ensinar_técnicas"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Tratamento_de"/>
    <Class IRI="#Nausea"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Etiologia"/>

```

```

<ObjectSomeValuesFrom>
  <ObjectProperty IRI="#Tratamento_de"/>
  <Class IRI="#Nausea"/>
</ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Explicar_causa_e_duração"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Tratamento_de"/>
    <Class IRI="#Nausea"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Explicar_causa_e_duração"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Tratamento_de"/>
    <Class IRI="#Risco_de_Náusea"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Irritação_Gástrica"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Cálculos_renais"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Irritação_Gástrica"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Enjoo"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Irritação_Gástrica"/>

```

```

<ObjectSomeValuesFrom>
  <ObjectProperty IRI="#Resultado_de"/>
  <Class IRI="#Enxaquecas"/>
</ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Irritação_Gástrica"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Gastroenterite_aguda"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Irritação_Gástrica"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Gravidez"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Irritação_Gástrica"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Infecções"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Irritação_Gástrica"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Intestino_irritável"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Irritação_Gástrica"/>

```

```

<ObjectSomeValuesFrom>
  <ObjectProperty IRI="#Resultado_de"/>
  <Class IRI="#Pancreatite"/>
</ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Irritação_Gástrica"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Superdosagem_de_drogas"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Irritação_Gástrica"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Úlcera_péptica"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Medicamentos_antieméticos"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Tratamento_de"/>
    <Class IRI="#Nausea"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Medicamentos_antieméticos"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Tratamento_de"/>
    <Class IRI="#Risco_de_Náusea"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Nausea"/>

```

```

<ObjectSomeValuesFrom>
  <ObjectProperty IRI="#Dignóstico_de"/>
  <Class IRI="#Características_Definidoras"/>
</ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Nausea"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Dignóstico_de"/>
    <Class IRI="#Irritação_Gástrica"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Nausea"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Dignóstico_de"/>
    <Class IRI="#Sequela_Tratamento"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Outro_Diagnóstico"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Dignóstico_de"/>
    <Class IRI="#Características_Definidoras"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Outro_Diagnóstico"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Dignóstico_de"/>
    <Class IRI="#Irritação_Gástrica"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Outro_Diagnóstico"/>

```

```

<ObjectSomeValuesFrom>
  <ObjectProperty IRI="#Dignóstico_de"/>
  <Class IRI="#Sequela_Tratamento"/>
</ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Refeições_pequenas"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Tratamento_de"/>
    <Class IRI="#Nausea"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Refeições_pequenas"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Tratamento_de"/>
    <Class IRI="#Risco_de_Náusea"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Repouso"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Tratamento_de"/>
    <Class IRI="#Nausea"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Repouso"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Tratamento_de"/>
    <Class IRI="#Risco_de_Náusea"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Risco_de_Náusea"/>

```

```

<ObjectSomeValuesFrom>
  <ObjectProperty IRI="#Dignóstico_de"/>
  <Class IRI="#Características_Definidoras"/>
</ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Risco_de_Náusea"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Dignóstico_de"/>
    <Class IRI="#Irritação_Gástrica"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Risco_de_Náusea"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Dignóstico_de"/>
    <Class IRI="#Sequela_Tratamento"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Sequela_Tratamento"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Anestesia"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<EquivalentClasses>
  <Class IRI="#Sequela_Tratamento"/>
  <ObjectSomeValuesFrom>
    <ObjectProperty IRI="#Resultado_de"/>
    <Class IRI="#Quimioterapia"/>
  </ObjectSomeValuesFrom>
</EquivalentClasses>
<SubClassOf>
  <Class IRI="#Anestesia"/>

```

```

    <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
    <Class IRI="#Características_Definidoras"/>
    <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
    <Class IRI="#Cálculos_renais"/>
    <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
    <Class IRI="#Diarreia"/>
    <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
    <Class IRI="#Eliminar_odor_desagradável"/>
    <Class IRI="#Tratamento"/>
</SubClassOf>
<SubClassOf>
    <Class IRI="#Enjoo"/>
    <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
    <Class IRI="#Ensinar_técnicas"/>
    <Class IRI="#Tratamento"/>
</SubClassOf>
<SubClassOf>
    <Class IRI="#Enxaquecas"/>
    <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
    <Class IRI="#Estômago"/>
    <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>

```



```
<Class IRI="#Etiologia"/>
  <Class IRI="#Tratamento"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Explicar_causa_e_duração"/>
  <Class IRI="#Tratamento"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Gastroenterite_aguda"/>
  <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Gravidez"/>
  <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Infecções"/>
  <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Intestino_irritável"/>
  <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Irritação_Gástrica"/>
  <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Medicamentos_antieméticos"/>
  <Class IRI="#Tratamento"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Musculatura_esquelética"/>
  <Class IRI="#Sintomas"/>
</SubClassOf>
```

```

<SubClassOf>
  <Class IRI="#Nausea"/>
  <Class IRI="#Diagnóstico"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Outro_Diagnóstico"/>
  <Class IRI="#Diagnóstico"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Pancreatite"/>
  <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Quimioterapia"/>
  <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Refeições_pequenas"/>
  <Class IRI="#Tratamento"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Repouso"/>
  <Class IRI="#Tratamento"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Risco_de_Náusea"/>
  <Class IRI="#Diagnóstico"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Sequela_Tratamento"/>
  <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Superdosagem_de_drogas"/>
  <Class IRI="#Sintomas"/>

```

```
</SubClassOf>
<SubClassOf>
  <Class IRI="#Vômito"/>
  <Class IRI="#Sintomas"/>
</SubClassOf>
<SubClassOf>
  <Class IRI="#Úlcera_péptica"/>
  <Class IRI="#Sintomas"/>
</SubClassOf>
<SubObjectPropertyOf>
  <ObjectProperty IRI="#Dignóstico_de"/>
  <ObjectProperty abbreviatedIRI="owl:topObjectProperty"/>
</SubObjectPropertyOf>
<SubObjectPropertyOf>
  <ObjectProperty IRI="#Resultado_de"/>
  <ObjectProperty abbreviatedIRI="owl:topObjectProperty"/>
</SubObjectPropertyOf>
<SubObjectPropertyOf>
  <ObjectProperty IRI="#Tratamento_de"/>
  <ObjectProperty abbreviatedIRI="owl:topObjectProperty"/>
</SubObjectPropertyOf>
</Ontology>
```