



UNIVERSIDADE FEDERAL DE CATALÃO
INSTITUTO DE MATEMÁTICA E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM E OTIMIZAÇÃO



Vinícius Vieira da Silva Dutra

APRENDIZAGEM DE MÁQUINA COM REPRESENTAÇÃO DO
CONHECIMENTO VIA LINGUAGEM OWL NA DETECÇÃO
AUTOMÁTICA DE COMPORTAMENTOS SUSPEITOS EM REDES
SOCIAIS

DISSERTAÇÃO DE MESTRADO

CATALÃO – GO, 2023



UNIVERSIDADE FEDERAL DE GOIÁS
UNIDADE ACADÊMICA ESPECIAL DE MATEMÁTICA E TECNOLOGIA

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

Dissertação Tese Outro*: _____

*No caso de mestrado/doutorado profissional, indique o formato do Trabalho de Conclusão de Curso, permitido no documento de área, correspondente ao programa de pós-graduação, orientado pela legislação vigente da CAPES.

Exemplos: Estudo de caso ou Revisão sistemática ou outros formatos.

2. Nome completo do autor

Vinícius Vieira da Silva Dutra

3. Título do trabalho

Aprendizagem de máquina com representação do conhecimento via linguagem OWL na detecção automática de comportamentos suspeitos em redes sociais

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento SIM NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

- a) consulta ao(à) autor(a) e ao(à) orientador(a);
 - b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.
- O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **VINÍCIUS VIEIRA DA SILVA DUTRA, Discente**, em 19/10/2023, às 15:53, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Vaston Goncalves Da Costa, Professor do Magistério Superior**, em 19/10/2023, às 17:23, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4014284** e o código CRC **F0D6AEF8**.

VINÍCIUS VIEIRA DA SILVA DUTRA

APRENDIZAGEM DE MÁQUINA COM REPRESENTAÇÃO DO
CONHECIMENTO VIA LINGUAGEM OWL NA DETECÇÃO
AUTOMÁTICA DE COMPORTAMENTOS SUSPEITOS EM REDES
SOCIAIS

Dissertação apresentada ao Programa de Pós-Graduação em Modelagem e Otimização, da Unidade do Instituto de Matemática e Tecnologia, da Universidade Federal de Catalão (UFCAT), como requisito para obtenção do título de Mestre em Modelagem e Otimização. Área de concentração: Ciência da Computação. Linha de pesquisa: Inteligência Artificial.

Orientador:
Vaston Gonçalves da Costa

CATALÃO – GO

2023

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFCAT.

Dutra, Vinícius Vieira da Silva

Aprendizagem de máquina com representação do conhecimento via linguagem OWL na detecção automática de comportamentos suspeitos em redes sociais / Vinícius Vieira da Silva Dutra. - 2023. 63, LXIII f.

Orientador: Prof. Vaston Gonçalves da Costa Costa.

Dissertação (Mestrado) - Universidade Federal de Catalão, Instituto de Matemática e Tecnologia, Catalão, Programa de Pós-Graduação em Modelagem e Otimização, Catalão, 2023.

Bibliografia.

Inclui siglas, abreviaturas, tabelas, algoritmos, lista de figuras, lista de tabelas.

1. Inteligência Artificial. 2. Aprendizado de Máquina. 3. Representação do Conhecimento. 4. Detecção de Comportamentos Suspeitos. I. Costa, Vaston Gonçalves da Costa, orient. II. Título.

CDU 004.423.4



UNIVERSIDADE FEDERAL DE GOIÁS

UNIDADE ACADÊMICA ESPECIAL DE MATEMÁTICA E TECNOLOGIA

ATA DE DEFESA DE DISSERTAÇÃO

Ata nº 22 da sessão de Defesa de Dissertação de **Vinícius Vieira da Silva Dutra**, que confere o título de **Mestre(a) em Modelagem e Otimização**, na área de concentração em **Modelagem e Otimização**.

Aos tres dias de outubro de 2023, a partir das **14h**, via Web Conferência (Google Meet - <https://meet.google.com/afd-yjxs-mtf>) e realizada no bloco J - sala 3, reuniram-se os componentes da banca examinadora, professores **Dr. Vaston Gonçalves da Costa (PPGMO / IMTec / UFCAT)**, (**orientador**), **Dr. Marcelo Henrique Stoppa (PPGMO / IMTec / UFCAT)**, **membro titular interno** e **Dr. Júlio César Ferreira (PPGEP IF Goiano)**, **membro titular externo**, para, em sessão pública, procederem a avaliação da Dissertação intitulada “*EMPREGO DE APRENDIZADO DE MÁQUINA E LÓGICA MATEMÁTICA EM REPRESENTAÇÃO DO CONHECIMENTO*”, de autoria de **Vinícius Vieira da Silva Dutra**, discente do Programa de Pós-graduação em Modelagem e Otimização – PPGMO, da Universidade Federal de Catalão - UFCAT. A sessão foi aberta pelo presidente, que fez a apresentação formal dos membros da banca. Em seguida, a palavra foi concedida ao discente que, em 30 min procedeu a apresentação. Terminada a apresentação, cada membro da banca arguiu o examinando. Terminada a fase de arguição, procedeu-se a avaliação da Dissertação, que foi considerado: **(X) Aprovado(a) ou () Reprovado(a)**. Cumpridas as formalidades de pauta, a presidência da mesa encerrou a sessão e, para constar, lavrou-se a presente ata que, depois de lida e aprovada, segue assinada pelos membros da banca examinadora.

TÍTULO SUGERIDO PELA BANCA

Aprendizagem de máquina com representação do conhecimento via linguagem OWL na detecção automática de comportamentos suspeitos em redes sociais.



Documento assinado eletronicamente por **Vaston Goncalves Da Costa, Professor do Magistério Superior**, em 03/10/2023, às 18:53, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Júlio César Ferreira, Usuário Externo**, em 03/10/2023, às 20:00, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Marcelo Henrique Stoppa, Professor do Magistério Superior**, em 03/10/2023, às 23:08, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4014235** e o código CRC **FB31E7A8**.

Dedico esta dissertação ao professor Vaston Gonçalves da Costa pelas valiosas e incontáveis horas dedicadas ao mesmo. Ele me auxiliou na germinação das ideias e durante todo o processo de desenvolvimento desta dissertação.

Agradecimentos

Primeiramente, expresso minha gratidão sincera a Deus, que, com sua luz, guiou meu caminho e me presenteou com a vida. Agradeço por me capacitar a moldar meu próprio destino e por preencher meu horizonte com inúmeras possibilidades inspiradoras.

Minha família, o alicerce da minha existência, merece uma gratidão especial. A minha mãe, Maria Lúcia Vieira da Silva Dutra, meu pai, Margarido Vaz Dutra, e minhas irmãs, Vívian Vieira da Silva Dutra e Amanda Vieira da Silva Dutra, sempre estiveram ao meu lado, oferecendo apoio e serenidade nos momentos de agitação e ansiedade que permearam minha jornada acadêmica. Impossível deixar de mencionar o meu tio, Dorico Vaz Dutra, uma figura inspiradora que sempre me incentivou a prosseguir nos estudos. Ele não foi apenas um profissional brilhante, mas também uma fonte constante de motivação em minha jornada.

Vaston Gonçalves da Costa, meu orientador, teve um papel essencial ao longo da minha jornada acadêmica e na criação deste trabalho. Suas valiosas orientações e revisões foram um verdadeiro impulso para o meu desenvolvimento tanto pessoal quanto profissional. Em momentos cruciais, quando eu precisava de direcionamento e apoio, o Professor Vaston estava sempre presente, oferecendo sua grande experiência e sabedoria. Sua orientação foi importantíssima, ajudando a trilhar o caminho em meio às complexidades da pesquisa e da academia. Quero expressar minha mais sincera gratidão a você.

Gostaria de expressar minha sincera e profunda gratidão à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES pelo apoio financeiro fundamental fornecido. É com grande apreço que reconheço a importância desse suporte na realização do presente trabalho. *Este estudo foi concretizado graças ao apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), cujo Código de Financiamento é 001.* O compromisso da CAPES com o avanço da pesquisa e da educação superior é verdadeiramente notável, e sua contribuição é um alicerce vital para o desenvolvimento acadêmico e científico no Brasil. Agradeço profundamente por tornarem possível este projeto.

Quero expressar minha sincera gratidão a todos os professores que compartilharam

seu conhecimento durante minha jornada acadêmica. Suas valiosas contribuições desempenharam um papel fundamental no meu percurso para a conclusão do curso. Cada um deles, de sua maneira única, ajudou a moldar meu entendimento e meu desenvolvimento acadêmico, e por isso, merecem meu respeito e agradecimento. Suas lições e orientações foram peças essenciais na construção do meu percurso educacional.

Quero estender meu agradecimento aos amigos e colegas de turma que encontrei durante o mestrado, com um destaque especial para Rafael, Jaime e Carlos, com os quais mantive uma estreita relação de ajuda mútua. Sempre houve uma troca valiosa entre nós.

Por fim, agradeço a todos que, de maneira direta ou indireta, contribuíram para minha trajetória. A todos, meu mais sincero agradecimento.

“Um computador merece ser chamado de inteligente se puder enganar um humano fazendo-o acreditar que é humano.”

Alan Turing

RESUMO

O avanço da Inteligência Artificial (IA) teve um impacto transformador no cenário tecnológico e na sociedade em geral. Desde a concepção da Máquina de Turing, que marcou um ponto fundamental nessa trajetória, a compreensão da computação e da capacidade das máquinas em realizar operações lógicas e algorítmicas tem evoluído constantemente. Com o advento do Aprendizado de Máquina, a forma como as máquinas processam informações mudou drasticamente, permitindo a identificação de padrões complexos em conjuntos massivos de dados. Além disso, a Representação do Conhecimento, juntamente com o uso de ontologias, a linguagem OWL e a biblioteca NLTK, oferece a oportunidade de construir sistemas capazes de detectar comportamentos suspeitos em redes sociais. O objetivo deste estudo é aprimorar a análise semântica das interações complexas nas redes sociais, com ênfase na detecção de padrões que indiquem atividades potencialmente prejudiciais ou ameaçadoras. Especificamente, o estudo visa desenvolver uma abordagem que possa contribuir para a identificação de comportamentos suspeitos em plataformas de redes sociais. Para atingir esse objetivo, utilizamos ontologias e a linguagem OWL para modelar o conhecimento e a estrutura das interações sociais. Isso contribui para uma abordagem mais eficaz na identificação e mitigação de comportamentos prejudiciais nas redes sociais, ajudando a criar ambientes mais seguros para os usuários.

Palavras-chaves: Inteligência Artificial, Aprendizado de Máquina, Representação do Conhecimento, Detecção de Comportamentos Suspeitos.

ABSTRACT

The advancement of Artificial Intelligence (AI) has had a transformative impact on the technological landscape and society at large. Since the inception of the Turing Machine, which marked a pivotal point in this trajectory, the understanding of computation and the capacity of machines to perform logical and algorithmic operations have continuously evolved. With the advent of Machine Learning, the way machines process information has undergone a drastic change, enabling the identification of complex patterns within massive datasets. Furthermore, Knowledge Representation, in conjunction with the use of ontologies, the OWL language, and the NLTK library, provides the opportunity to construct systems capable of detecting suspicious behaviors on social networks. The objective of this study is to enhance the semantic analysis of complex interactions on social networks, with a focus on detecting patterns indicative of potentially harmful or threatening activities. Specifically, the study aims to develop an approach that can contribute to the identification of suspicious behaviors on social media platforms. To achieve this objective, we utilize ontologies and the OWL language to model the knowledge and structure of social interactions. This contributes to a more effective approach in identifying and mitigating harmful behaviors on social networks, thereby aiding in the creation of safer environments for users.

Keywords: Artificial Intelligence, Machine Learning, Knowledge Representation, Suspicious Behavior Detection.

LISTA DE FIGURAS

Figura 2.1 – Mensagens postadas pelo indivíduo envolvido na tragédia.	24
Figura 3.1 – ALC sendo utilizada para representar conceitos em DL.	31
Figura 4.1 – DL-learner Protégé plugin.	41
Figura 4.2 – Representação de ontologia simplificada.	45
Figura 4.3 – Pellet no Protégé.	46

LISTA DE TABELAS

Tabela 3.1 – A sentença idêntica representada em distintas lógicas.	30
Tabela 3.2 – Interpretação DL para Teoria dos Conjuntos.	32

LISTA DE QUADROS

Quadro 3.1 – Tipos de Ontologias.	28
Quadro 3.2 – Abordagens para estruturar o conhecimento.	29
Quadro 4.1 – Funcionalidades do NLTK.	46
Quadro 6.1 – Possíveis trabalhos futuros.	57

LISTA DE CÓDIGOS

Código 5.1	Ontologia <i>Tweets</i>	49
Código 5.2	Exemplo de Consulta	50
Código 5.3	Uso da Biblioteca NLTK	51
Código 5.4	Inserção de DL na Ontologia	52
Código 5.5	Restrição de busca	53

LISTA DE ABREVIATURAS E SIGLAS

- ALC — Linguagem de Conceitos Atributivos com Complementos, do inglês *Attributive Concept Language with Complements*
- CNN — Rede Neural Convolutacional, do inglês *Convolutional Neural Network*
- CTL — Lógica de Árvore Computacional do inglês *Computer Tree Logic*
- DECRADI — Delegacia de Crimes Raciais e Delitos de Intolerância
- DL — Lógica Descritiva, do inglês *Description Logic*
- DPCA — Delegacia de Proteção ao Adolescente
- FAIR — Encontreáveis, Acessíveis, Interoperáveis e Reutilizáveis, do inglês *Findable, Accessible, Interoperable and Reusable*
- FCA — Análise de Conceito Formal, do inglês *Formal Concept Analysis*
- ILP — Programação Lógica Indutiva
- ILSVRC — Desafio de Reconhecimento Visual em Grande Escala do ImageNet, do inglês *ImageNet Large Scale Visual Recognition Challenge*
- KL-ONE — Linguagem de Representação de um Conhecimento, do inglês *Knowledge Representation Language One*
- LGPL — Lesser General Public License
- ML — Aprendizado de Máquina, do inglês *Machine Learning*
- NLP — Processamento de Linguagem Natural, do inglês *Natural Language Processing*
- NLTK — Kit de Ferramentas de Linguagem Natural, do inglês *Natural Language Toolkit*
- OWL — Linguagem de Ontologia para Web, do inglês *Ontology Web Language*
- RA — Reação Adversa

RDB — Banco de Dados Relacionais, do inglês *Relational Database*

ReLU — Unidade Linear Retificada, do inglês *Rectified Linear Unit*

SMV — Verificador de Modelo Simbólico, do inglês *Symbolic Model Verifier*

TECMF — Laboratório de Métodos Formais

SUMÁRIO

1	INTRODUÇÃO	17
2	JUSTIFICATIVA	23
3	REFERENCIAL TEÓRICO	27
4	TRABALHOS RELACIONADOS	34
4.1	Análise dos artigos	34
4.2	Análise final dos trabalhos relacionados	39
4.2.1	Sobre o trabalho de Lehmann e Hitzler	39
4.2.2	Protégé	41
4.2.3	O trabalho de Fanizzi	42
4.2.3.1	FOIL	42
4.2.3.2	FOIL e DL	43
4.3	Grupo de pesquisa em Métodos Formais associado à pesquisa	44
4.3.1	Pellet	45
4.3.2	NLTK	45
5	RESULTADOS	48
6	CONCLUSÃO	56
6.1	Trabalhos Futuros	57
	REFERÊNCIAS	59

Capítulo 1

INTRODUÇÃO

O percurso da Inteligência Artificial (IA) é caracterizado por avanços extraordinários que exerceram profunda influência sobre o panorama tecnológico e a sociedade de um modo geral. Ao retornarmos às origens, deparamo-nos com um marco de inicial notável na trajetória da IA: a conferência de Macy. Com o nome de *The Macy Conferences* (Conferência de Macy), aconteceu, entre os anos 1946 e 1953, uma série de dez conferências interdisciplinares que levou à fundação do que hoje conhecemos como cibernética. Essas conferências consolidaram ainda mais as ideias discutidas anteriormente, ampliando os horizontes da IA e estabelecendo conexões com a cibernética, um campo que também se tornou crucial para a compreensão da complexidade dos sistemas inteligentes (BRAGA; SYLLA, 2022).

Um dos trabalhos de destaque durante a conferência foi o artigo de Warren McCulloch, um neurocientista, e Walter Pitts, um matemático, intitulado *A Logical Calculus of Ideas Immanent in Nervous Activity* (Um Cálculo Lógico de Ideias Inerentes à Atividade Nervosa), publicado em (MCCULLOCH; PITTS, 1943). Neste artigo, eles apresentaram um modelo inovador de neurônios artificiais e uma lógica proposicional que poderia representar o funcionamento do sistema nervoso. Esse modelo foi uma peça fundamental para o desenvolvimento das redes neurais artificiais, que agora são uma parte central da IA moderna (BRAGA; SYLLA, 2022).

Além disso, durante a conferência, houve discussões significativas sobre o trabalho revolucionário de Alan Turing, incluindo sua Máquina de Turing. Esta máquina teórica se tornou crucial para a compreensão da computabilidade e serviu como uma base conceitual para a teoria da computação. As discussões sobre a Máquina de Turing foram essenciais para compreender a natureza dos algoritmos e o processamento de informações, aspectos centrais na IA (BRAGA; SYLLA, 2022).

Ainda de acordo com Braga e Sylla (2022), a verdadeira essência da Conferência de Macy residia na interdisciplinaridade e na troca de ideias entre especialistas de diversas áreas. O ambiente propício para essa interação promoveu um diálogo multidisciplinar sobre

tópicos relacionados à mente, ao cérebro e à inteligência. A diversidade de perspectivas e a troca de ideias entre matemáticos, biólogos, engenheiros e filósofos contribuíram para uma visão mais holística e abrangente do estudo da inteligência, estabelecendo as bases para futuras pesquisas em IA.

Um dos principais marcos significativos na história da Inteligência Artificial, especialmente no campo de Visão Computacional, foi a contribuição da equipe liderada por Alex Krizhevsky em (KRIZHEVSKY; SUTSKEVER; HINTON, 2012) no desafio ILSVRC (ImageNet Large Scale Visual Recognition Challenge) de 2012. O ILSVRC é uma competição anual que desafia os pesquisadores a desenvolver algoritmos capazes de reconhecer objetos e categorias em imagens em uma escala muito grande, utilizando um grande conjunto de dados chamado ImageNet. O conjunto de dados do ImageNet contém milhões de imagens rotuladas em milhares de categorias.

A contribuição fundamental de Krizhevsky, conhecida como a arquitetura AlexNet, revolucionou a visão computacional e estabeleceu um novo padrão de desempenho na classificação de imagens. A AlexNet foi a primeira a utilizar uma arquitetura de Rede Neural Convolucional (CNN, do inglês *Convolutional Neural Network*) profundamente empilhada com sucesso em uma tarefa tão complexa como a classificação de imagens em um conjunto de dados tão grande (KRIZHEVSKY; SUTSKEVER; HINTON, 2012).

A arquitetura AlexNet consistia em oito camadas de aprendizado, incluindo cinco camadas convolucionais, seguidas por três camadas totalmente conectadas. Além disso, introduziu técnicas importantes, como a aplicação de funções de Unidade Linear Retificada (ReLU, do inglês *Rectified Linear Unit*) e o uso de técnicas de regularização, como dropout (KRIZHEVSKY; SUTSKEVER; HINTON, 2017).

De acordo com Krizhevsky, Sutskever e Hinton (2017), na competição anual renomada de classificação de imagens, a equipe liderada por A. Krizhevsky obteve uma impressionante taxa de erro de 15,3%. Essa conquista significativa foi marcada pela notável diferença em relação ao segundo colocado, que registrou uma taxa de erro de 26,2%. Esse feito evidenciou a superioridade da arquitetura AlexNet e o poder das redes neurais convolucionais profundas, estabelecendo um novo padrão de desempenho na área de Visão Computacional e impulsionando avanços subsequentes nesse campo dinâmico da Inteligência Artificial.

A partir desse ponto, houve uma explosão de pesquisas e avanços em CNNs e aplicações relacionadas à visão computacional, incluindo detecção de objetos, segmentação semântica, reconhecimento facial e muito mais. A AlexNet e o desempenho excepcional na competição ILSVRC 2012, foi um ponto de inflexão que impulsionou a evolução e o desenvolvimento contínuo da Inteligência Artificial, especialmente na área de Visão Computacional (KRIZHEVSKY; SUTSKEVER; HINTON, 2017).

De acordo com Negnevitsky (2005), esses passos iniciais estabeleceram os alicerces

para a progressão contínua da Inteligência Artificial ao longo das décadas subseqüentes. Entretanto, apenas nas últimas duas décadas, pudemos observar notáveis avanços que catapultaram a Inteligência Artificial a níveis inéditos. No presente momento, a Inteligência Artificial emerge como uma tecnologia fundamental em diversas esferas do cotidiano, manifestando-se em aplicações que abrangem desde assistentes virtuais e veículos autônomos até auxílios computacionais para diagnósticos médicos e previsões climáticas de alta precisão.

A relevância da Inteligência Artificial na contemporaneidade é incontestável. O atual mundo digital e interconectado gera uma profusão incessante de dados a cada instante. A IA se apresenta como uma ferramenta de magnitude singular para extrair percepções de valor a partir desses dados, viabilizando a tomada de decisões mais embasadas e eficazes. Conforme citado por [Negnevitsky \(2005\)](#), o culminar evolutivo desse desenvolvimento manifesta-se no advento do Aprendizado de Máquina (ML, do inglês *Machine Learning* , uma ramificação da IA que revolucionou a forma como máquinas processam e assimilam informações.

O Aprendizado de Máquina possibilitou o desenvolvimento de sistemas automatizados que processam informações em uma escala previamente inimaginável. Através da análise de conjuntos extensos de dados, os algoritmos de Aprendizado de Máquina são capazes de discernir padrões intrincados, identificar tendências e até antecipar resultados futuros com uma precisão notável. Esta aptidão encontra aplicações de amplitude considerável, abrangendo desde a personalização de estratégias de marketing até a detecção de atividades fraudulentas no âmbito financeiro e o aprimoramento das operações em cadeias de suprimentos ([MITCHELL, 1997](#)).

De acordo com [Mitchell \(1997\)](#), uma das facetas mais proeminentes do avanço da IA reside na sua contribuição para a capacidade de raciocinar sobre informações. Sistemas automatizados têm a capacidade de explorar a vastidão da internet, discernindo informações pertinentes dentre volumes colossais de dados. Essa competência se mostra especialmente valiosa em âmbitos como a inteligência de mercado, a segurança cibernética e a pesquisa acadêmica. A IA viabiliza a compilação eficiente de informações cruciais, resultando em uma economia de tempo e recursos inestimáveis.

Essa capacidade intrínseca da Inteligência Artificial para lidar com a imensa quantidade de dados disponíveis é diretamente enriquecida por sua própria evolução. A busca por métodos mais avançados para dotar as máquinas da habilidade de compreender e interpretar o vasto conjunto de informações que caracteriza o mundo contemporâneo tem impulsionado essa evolução. Neste contexto há alinhamento de técnicas de métodos formais e, conforme apontado em ([BRACHMAN; LEVESQUE; REITER, 1992](#)), a Representação do Conhecimento desempenha um papel essencial ao estabelecer estruturas que capacitam as máquinas a capturar, organizar e assimilar informações de maneira mais profunda,

semântica e contextualmente embasada.

Ao combinar essas duas perspectivas, delinea-se um cenário de sinergia entre a capacidade de raciocínio informacional da Inteligência Artificial e a aptidão de compreensão semântica viabilizada pela Representação do Conhecimento. Enquanto a IA viabiliza a análise automatizada de volumes extensos de dados, a Representação do Conhecimento concede a habilidade de apreender as sutilezas e conexões subjacentes presentes nesses dados (CHAPTER *et al.*, 2009).

Nesse sentido, conforme exposto por Chollet (2018), as máquinas não apenas são capazes de raciocinar informações pertinentes de maneira eficaz, mas também possuem a capacidade de interpretá-las e derivar *insights* significativos a partir delas. Esse processo assume particular importância no contexto da detecção de comportamentos suspeitos em plataformas de redes sociais. Ao amalgamar a competência da Inteligência Artificial em explorar e analisar dados com a aptidão de compreender o significado e o contexto por intermédio da Representação do Conhecimento, torna-se viável identificar padrões sutis e comportamentos anômalos que, de outra maneira, poderiam passar despercebidos.

As ontologias, enquanto meios de representação do conhecimento, desempenham um papel fundamental nessa trajetória. Conforme indicado por Levesque (1986), elas oferecem um método sistemático e formal para descrever os conceitos, relações e limitações que delineiam um domínio específico. Essa estruturação do conhecimento é de relevância extrema, pois viabiliza que as máquinas transcendam a compreensão superficial e adentrem um nível mais profundo de análise. Ao estabelecer conexões nítidas entre os elementos inerentes ao domínio, as ontologias capacitam a Inteligência Artificial a apreender as sutilezas e complexidades das interações entre esses elementos, de maneira que, em certo aspecto, emulem o processo de raciocínio humano.

A Linguagem de Ontologia para Web (OWL, do inglês *Ontology Web Language*), se destaca como uma ferramenta de proeminência nesse contexto. Ela provê um conjunto de recursos e sintaxes que viabilizam a formal definição dos elementos de uma ontologia, estabelecendo um fundamento para inferências lógicas. A aptidão de conduzir inferências lógicas assume um papel fundamental na análise e compreensão de informações. Tal habilidade capacita as máquinas a realizar deduções e raciocínios com base nas relações e limitações previamente estabelecidas, enriquecendo substancialmente sua capacidade de extrair percepções significativas dos dados (ANTONIOU; HARMELEN, 2009).

De acordo com Sowa (1999), a relevância da Representação do Conhecimento se amplifica quando voltamos nosso olhar para as aplicações práticas, como a detecção de comportamentos suspeitos em plataformas de redes sociais. As ontologias, ao estabelecerem uma estrutura sólida para a definição de conceitos e suas inter-relações, facultam à Inteligência Artificial compreender não apenas os eventos em curso, mas também os motivos subjacentes e os modos como tais eventos se desenrolam. Essa compreensão contextual as-

sume um caráter crucial na identificação de padrões anômalos ou condutas suspeitas, uma vez que frequentemente tais padrões podem não ser perceptíveis em uma análise superficial.

Dessa maneira, a aplicação de ontologias e da linguagem OWL em sistemas de Aprendizado de Máquina se torna especialmente pertinente ao lidar com tarefas complexas, como a detecção de comportamentos suspeitos em redes sociais. Plataformas de redes sociais compõem ambientes repletos de interações e informações, onde a detecção de padrões sutis e contextuais é de suma importância. Através da Representação do Conhecimento, torna-se viável construir modelos semânticos que delineiam os elementos pertinentes dessas interações, facultando uma análise mais profunda dos comportamentos dos usuários (BRACHMAN; LEVESQUE; REITER, 1992).

Além disso, ainda como uma ferramenta formal, a Lógica Descritiva (DL, do inglês *Description Logic*) desempenha um papel crucial nesse contexto. Ela oferece uma estrutura que amplia a lógica de primeira ordem, permitindo a representação de relações intrincadas entre conceitos e a realização de inferências lógicas acerca dessas relações. Isso enriquece a capacidade dos sistemas de Aprendizado de Máquina de apreender conexões mais profundas e abstratas, facilitando a identificação de padrões não evidentes e comportamentos atípicos (BLACKBURN; BENTHEM; WOLTER, 2006).

Conforme destacado por Blackburn, Benthem e Wolter (2006), a convergência desses conceitos culmina em sistemas de Inteligência Artificial de maior robustez e eficácia na detecção de comportamentos suspeitos. A representação organizada do conhecimento capacita as máquinas a compreenderem o contexto e a semântica subjacentes aos dados, enquanto a lógica descritiva simplifica o processo de raciocínio e inferência a respeito das relações conceituais.

No âmbito de aplicações práticas, como a detecção de comportamentos suspeitos em redes sociais, essa abordagem se traduz em um incremento na precisão e na habilidade de identificação precoce de ameaças. Tal precisão é vital em cenários nos quais a detecção de padrões comportamentais atípicos pode acarretar impactos significativos, como na mitigação de eventos indesejáveis.

Dentre os trabalhos que associam conceitos formais e métodos de Inteligência artificial podemos destacar o de Lehmann e Hitzler (2010) que faz uso de operadores de refinamento e o de Fanizzi (2011) que propõe uma abordagem para indução de conceitos em ontologias de lógica descritiva usando heurísticas baseadas em teoria da informação.

Nesse âmbito, o propósito subjacente a este estudo reside no aprimoramento da avaliação de sistemas de Aprendizado de Máquina, mediante a incorporação dos princípios da Representação do Conhecimento. O objetivo central consiste em habilitar o processo de aprendizado a lidar com informações estruturadas no formato OWL apontando a forma de

representar as informações e como, a partir desta representação extrair o conhecimento. Não se trata, pois, de um trabalho de apresentação de uma aplicação¹ mas, sim, de apresentação de viabilidade de aplicação². Entretanto, mesmo sem o carácter de desenvolvimento de produto final, apresentaremos um exemplo de como se aplicar conceito em uma base de dados reduzida.

O trabalho aborda um tema de preocupação nos momentos atuais de uma maneira diferente da que se conhece até o momento, com está descrito no capítulo 4, o que lhe confere originalidade. Ao associar conceitos de inteligência artificial com métodos formais para a produção de nova abordagens, o trabalho também apresenta inovação.

Para melhor compreensão do trabalho, este se encontra subdividido em capítulos da seguinte forma: no capítulo 2 é apresentada a justificativa para a pesquisa, no capítulo 3 o referencial teórico básico de conceitos e notações, no capítulo 4 os trabalhos relacionados e as pesquisas que mantém relação, mesmo que distantes, do emprego de inteligência artificial e/ou métodos formais para extração de conhecimento; no capítulo 5 apresentamos a abordagem empregando IA, OWL e DL para detecção de comportamento suspeitos em bases de dados, por meio de exemplos, por fim; no capítulo 6 a conclusão.

¹ Criação de app

² Que se pode construir ferramentas teóricas e/ou práticas que englobem o proposto.

Capítulo 2

Justificativa

A chegada da Web 2.0 provocou uma revolução no comportamento dos usuários da Internet, expandindo suas ações para além do consumo passivo, passando a incluir também a criação e o compartilhamento de conteúdo de maneiras sem precedentes. Essa mudança foi impulsionada principalmente pelo surgimento e pela expansão das redes sociais. Essas plataformas online proporcionaram aos usuários a capacidade de estabelecer conexões com pessoas de todo o mundo, compartilhar instantaneamente suas opiniões, vivências, fotografias, vídeos e ideias (SHIRKY, 2008).

Através das redes sociais, os indivíduos não apenas se tornaram meros consumidores de informações, mas também se transformaram em participantes ativos na criação e disseminação de conteúdo. Essa participação ativa reflete o fenômeno da *prosumerização*, uma tendência que combina os papéis de produtor e consumidor. De acordo Zuboff (2019), em sua análise do *capitalismo de vigilância*, destaca que os usuários contemporâneos não apenas consomem o que é oferecido, mas também produzem suas próprias narrativas, compartilham opiniões, criam vídeos, escrevem artigos e muito mais. As redes sociais proporcionam uma plataforma para expressar criatividade, opiniões e perspectivas únicas, permitindo que cada indivíduo tenha uma voz e alcance global. Esse envolvimento ativo promove interações significativas entre os usuários e contribui para a formação de comunidades virtuais em torno de interesses comuns. Além disso, a facilidade de compartilhamento nas redes sociais facilita a rápida disseminação de informações e ideias, amplificando o alcance e o impacto do conteúdo gerado pelos usuários (ZUBOFF, 2019).

Com isso, resultou em um aumento substancial no engajamento online e na formação de comunidades virtuais pautadas por interesses compartilhados. Conforme definido por Kaplan e Haenlein (2010), as mídias sociais são “um grupo de aplicações para internet construídas com base nos fundamentos ideológicos e tecnológicos da Web 2.0, e que permitem a criação e troca de conteúdo gerado pelo utilizador”. Neste contexto, fóruns de discussão, blogs, plataformas para compartilhar fotos e, posteriormente, redes como Facebook, Twitter, Instagram e LinkedIn possibilitaram que as pessoas estabelecessem conexões

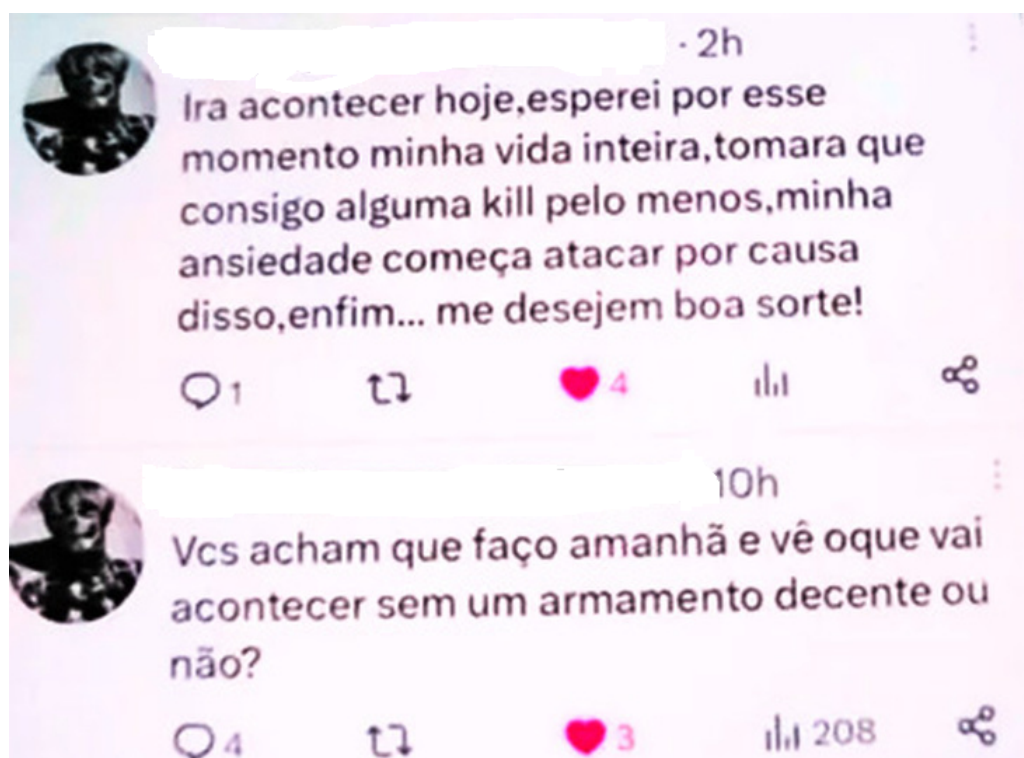
e interagissem de formas previamente inimagináveis (BOYD, 2014).

Essas redes sociais tornaram-se canais de extrema importância para a expressão pessoal, o ativismo, o marketing, a disseminação de informações e até mesmo para a influência política. Elas não apenas oferecem uma plataforma para a manifestação do eu, mas também desempenham um papel crucial na propagação de notícias e informações, influenciando perspectivas e mobilizando movimentos sociais (GLADWELL, 2006).

Entretanto, apesar dos diversos aspectos positivos das redes sociais, também há situações em que seu impacto é profundamente negativo. Um exemplo alarmante disso é o caso de um adolescente de 13 anos que atacou colegas e professores com uma faca em uma escola. Este trágico incidente teve lugar na Escola Estadual Thomazia Montoro e culminou na triste morte de uma docente de 71 anos (LEAL, 2023).

A ironia cruel reside no fato de que o agressor já havia antecipado suas intenções desde o dia anterior ao ataque, através de postagens feitas em um perfil privado em uma rede social, como ilustrado na Figura (2.1). O adolescente manifestou suas intenções trágicas e, ainda mais alarmante, recebeu incentivos de outros usuários com perfis semelhantes. Isso foi corroborado pelas curtidas e comentários presentes nas postagens, evidenciando uma faceta sombria das redes sociais (LEAL, 2023).

Figura 2.1 – Mensagens postadas pelo indivíduo envolvido na tragédia.



Fonte: (LEAL, 2023).

Esse trágico exemplo realça como as redes sociais, apesar de serem poderosas fer-

ramentas positivas, também podem ser exploradas de maneira prejudicial (BIEGING *et al.*, 2013). No entendimento de Coll e Monereo (2010), a tecnologia exerce uma influência crucial no direcionamento do desenvolvimento humano, operando na zona de desenvolvimento proximal do indivíduo. Eles defendem que a utilização das novas tecnologias, especialmente por crianças e adolescentes, pode estimular aspectos fundamentais do desenvolvimento, como criatividade, comunicabilidade e auto-realização.

Essa visão é reforçada por Buckingham (2007), que atribui às novas tecnologias e mídias a capacidade de libertar as crianças, proporcionando-lhes espaço para expressar sua criatividade e se conectar com outros de maneiras inovadoras. Além disso, Castells (2005) argumenta que a interação nas redes, seja por meio de jogos ou conversas online, cria novas formas de socialização, ampliando as possibilidades de interação e aprendizado.

No entanto, é imperativo considerar que, apesar dos benefícios evidentes, a utilização indiscriminada das redes sociais e novas tecnologias pode também apresentar desafios significativos para a sociedade. Conforme destacado por (LEAL, 2023), essas plataformas podem servir como veículos para a disseminação de ideias prejudiciais, o estímulo à violência e a adoção de comportamentos danosos.

A inserção das crianças e adolescentes no universo das mídias e tecnologias é um tema complexo, apresentando uma dualidade marcante: benefícios e desafios decorrentes do uso cotidiano dessas ferramentas. Especialmente nas faixas etárias de sete a onze anos, as crianças estão prontas para iniciar uma aprendizagem mais estruturada, desenvolvendo autonomia e começando a refletir sobre suas ações. No entanto, esse estágio também as torna suscetíveis a manipulações e influências externas (RIBEIRO; BATISTA, 2010).

Por outro lado, como resposta a um alerta emitido pela Interpol, a Polícia Civil do Rio de Janeiro conduziu com sucesso uma operação que resultou na detenção de um menor de 17 anos, suspeito de planejar um ataque a uma escola na região. A ação conjunta envolveu investigadores da Polícia Federal e da Polícia Civil, que empregaram técnicas avançadas de monitoramento e rastreamento de atividades online para identificar indícios preocupantes (NASCIMENTO; BRASIL, 2023).

A investigação revelou que o adolescente vinha exibindo símbolos nazistas em vídeos compartilhados online, e havia manifestado intenções de executar um ataque à escola. O emprego de tecnologias de vigilância digital possibilitou às autoridades detectar esses comportamentos suspeitos, demonstrando a capacidade da implementação de técnicas de Aprendizado de Máquina e análise de dados em larga escala para identificar ameaças em potencial (NASCIMENTO; BRASIL, 2023).

O desenlace deste caso foi alcançado por meio da investigação na chamada *deep web*, uma porção da internet que não é prontamente acessível por meio de mecanismos de busca convencionais. A competência de explorar essa esfera sombria da internet, onde frequente-

mente ocorrem atividades ilícitas, revelou-se crucial para desvendar o plano em questão e prevenir um possível ataque (NASCIMENTO; BRASIL, 2023).

Após a identificação do suspeito, as autoridades agiram de forma ágil e coordenada. O adolescente foi encaminhado à Delegacia de Proteção ao Adolescente (DPCA), onde as investigações prosseguiram. Além disso, a colaboração da Delegacia de Crimes Raciais e Delitos de Intolerância (DECRADI) destaca a importância de abordar aspectos ideológicos que possam estar correlacionados com comportamentos violentos (NASCIMENTO; BRASIL, 2023).

Este episódio contribui de maneira substancial para corroborar a relevância do desenvolvimento de sistemas automatizados capazes de identificar possíveis ameaças de massacres divulgadas em plataformas de redes sociais, com um enfoque especial no Twitter. Tal empreendimento beneficia-se grandemente da aplicação de técnicas avançadas de Aprendizado de Máquina (LAZER *et al.*, 2009).

As capacidades intrínsecas das técnicas de Aprendizado de Máquina oferecem um potencial notável para a identificação automática de termos-chave, *hashtags* ou até mesmo padrões de comportamento específicos que possam indicar atividades perigosas, como ameaças de ataques violentos. Essa análise direcionada pode desempenhar um papel especialmente crucial na prevenção de crimes planejados e orquestrados nas redes sociais (HUBERMAN; ROMERO; WU, 2008).

A perspectiva de empregar tecnologia com o propósito de evitar a ocorrência de crimes amplia ainda mais a relevância do trabalho voltado à detecção automatizada de possíveis planos de massacres, baseado nas informações extraídas das redes sociais. As abordagens de Aprendizado de Máquina se destacam como ferramentas especialmente eficazes para identificar e interromper ações criminosas de natureza violenta que são concebidas e organizadas dentro das plataformas de redes sociais. A exploração dessa tecnologia pode possibilitar a identificação precoce de sinais de alerta, facilitando uma resposta ágil e eficaz que evite a concretização de tragédias (BOYD, 2010).

Capítulo 3

Referencial Teórico

Ontologias, no âmbito da ciência da computação, constituem modelos abstratos e computacionais que têm a finalidade de representar o conhecimento inerente a um domínio específico. Estas estruturas proporcionam uma estratégia fundamental ao disponibilizar um conjunto de conceitos interligados, os quais têm a capacidade de descrever de maneira abrangente os objetos, as ações e as interconexões presentes dentro de um determinado campo de estudo ou problema (GRUBER, 1995).

De forma mais ampla, as ontologias funcionam como um alicerce conceitual que ajuda a criar um entendimento coerente do domínio em questão, fornecendo diretrizes precisas para categorizar e organizar informações. Por meio da criação de relações semânticas entre os conceitos, as ontologias propiciam uma representação semântica rica, permitindo uma análise mais profunda e precisa do conhecimento encapsulado (ANTONIOU; HARMELEN, 2004).

No processo de construção de uma ontologia, um conjunto de termos e relações é definido para retratar a realidade de um domínio específico. Esses termos e relações formam a base da ontologia, estabelecendo uma linguagem comum que auxilia na comunicação entre sistemas de software, humanos e até mesmo máquinas. Com a evolução das tecnologias e a interconexão crescente entre sistemas heterogêneos, as ontologias se tornaram componentes essenciais em campos como a inteligência artificial, a recuperação de informações e a interação homem-máquina (NOY; MCGUINNESS, 2001).

De acordo com Horrocks e Sattler (2001), há variados tipos de ontologias, cada qual adequado para diferentes contextos e finalidades. As ontologias podem ser classificadas quanto à sua complexidade, abrangência e aplicação, conforme mostrado em (3.1). Algumas ontologias são específicas de um domínio altamente especializado, enquanto outras têm um escopo mais amplo, abordando conceitos transversais a várias áreas. Também existem ontologias que focam na representação de conhecimento taxonômico, estabelecendo hierarquias de classes e subclasses, enquanto outras exploram relações mais complexas e

até mesmo axiomas lógicos.

Quadro 3.1 – Tipos de Ontologias.

Ontologias Formais	Estas ontologias usam uma linguagem formal para expressar as relações entre conceitos e também possuem regras para definir o que é permitido ou não. Estas ontologias podem ser usadas para especificar o significado de um domínio de forma precisa e para dar suporte às tarefas de inferência.
Ontologias Semi-formais	Estas ontologias usam uma linguagem semi-formal para representar o conhecimento. Estas ontologias são menos rigorosas que as ontologias formais e permitem que sejam expressas relações complexas entre conceitos.
Ontologias Baseadas em Ontologias	Estas ontologias usam ontologias existentes para descrever relações entre conceitos em um domínio específico. Estas ontologias podem ser usadas para reutilizar informações existentes e para criar novas ontologias mais detalhadas.
Ontologias Baseadas em Descrições	Estas ontologias usam descrições para representar o conhecimento sobre um domínio específico. Estas descrições são escritas em linguagem natural e usadas para fornecer informações sobre um conjunto de conceitos.
Ontologias Baseadas em Texto	Estas ontologias usam texto para representar o conhecimento. Estas ontologias usam palavras-chave e frases para descrever relações entre conceitos. Estes textos são usados para fornecer informações sobre um domínio de forma mais flexível do que as ontologias formais.

Fonte: Baader, Horrocks e Sattler (2008).

Na ampla área da Inteligência Artificial, encontramos diversas subáreas, cada uma desempenhando um papel específico na construção do conhecimento e na resolução de problemas. Uma dessas subáreas fundamentais é a Representação de Conhecimento, que se concentra em desenvolver estruturas formais para expressar e gerenciar o conhecimento em um determinado domínio (BAADER; HORROCKS; SATTLER, 2008).

Nessa busca por representar o conhecimento de entidades inteligentes, que necessitam realizar inferências a partir dessas informações, a escolha das ferramentas e formalismos apropriados é essencial. Uma visão amplamente aceita, é a importância das linguagens formais na modelagem de ontologias e na caracterização de axiomas lógicos (BAADER; HORROCKS; SATTLER, 2008).

Dentre os formalismos amplamente empregados na representação de conhecimento, com ênfase na precisão das terminologias, destacam-se as Lógicas de Descrição. Esta lógica é considerada um subconjunto da Lógica de Primeira Ordem, que, por sua vez, é uma extensão da Lógica Proposicional. A Lógica de Descrição é conhecida por sua expressividade e flexibilidade na representação de conhecimento (BAADER; HORROCKS; SATTLER, 2008).

A Lógica de Primeira Ordem é um sistema lógico que permite a representação e o raciocínio sobre proposições e predicados complexos usando quantificadores (como *para todo* e *existe*) e conectivos lógicos (como *e*, *ou* e *não*). Ela é amplamente usada na matemática, ciência da computação e filosofia para expressar relações e propriedades entre objetos (HARMELEN; LIFSCHITZ; PORTER, 2008).

Por outro lado, a Lógica Descritiva é uma extensão da Lógica de Primeira Ordem que foi desenvolvida para representar e raciocinar sobre conceitos complexos e relações em ontologias, que descrevem o conhecimento sobre um domínio particular. As Lógicas de Descrição apresentam uma abordagem singular para estruturar o conhecimento. Elas adotam três definições para representar o que se pretende (HARMELEN; LIFSCHITZ; PORTER, 2008). Cada uma dessas definições é empregada na elaboração de uma ontologia, conforme apresentado no quadro (3.2). São elas:

Quadro 3.2 – Abordagens para estruturar o conhecimento.

Conceitos	São conhecidos como Classes e representam um grupo de indivíduos. No contexto das ontologias, são equivalentes às Classes.
Papéis	As relações binárias que existem entre os indivíduos podem ser designadas como propriedades. Essencialmente, são as Relações de uma ontologia. Alguns papéis podem desempenhar uma função de caracterização, ou seja, definindo alguns atributos para o conceito. Essa função corresponde a uma Propriedade de uma Classe, em uma ontologia.
Indivíduos	Dentro de um Conceito, eles denotam os indivíduos existentes. Nas ontologias, são referidos como as Instâncias.

Fonte: Harmelen, Lifschitz e Porter (2008).

Diante disso, a Lógica Descritiva é caracterizada por uma tríade (N_C , N_R , N_I), em que N_C representa um conjunto de conceitos fundamentais, N_R é um conjunto de papéis fundamentais, e N_I é um conjunto de nomes de entidades. Com essa diferenciação, o entendimento em uma ontologia é subdividido em duas categorias, utilizando a LD. A primeira delas é a *TBox*, que se refere ao conhecimento terminológico, abrangendo os conceitos, as propriedades e as restrições do domínio de forma intensional. A *TBox* engloba as definições terminológicas. A segunda parte é a *ABox*, que armazena o conhecimento assertivo, ou seja, a informação extensional. Ela contém afirmações sobre as instâncias, descrevendo como elas se relacionam e se enquadram nas definições explícitas presentes na *TBox* (HARMELEN; LIFSCHITZ; PORTER, 2008).

No contexto do desenvolvimento de uma ontologia, foi utilizado as Lógicas de Descrição, a Lógica de Primeira Ordem e a Lógica Proposicional para expressar as relações e características entre entidades. Por exemplo, para afirmar que um *Pássaro* ou um *Mamífero* pertencem à categoria *Animal* na ontologia, foi empregado as sentenças e formalismos apresentados na tabela (3.1), que são ferramentas fundamentais para a representação precisa do

conhecimento no domínio em questão (BAADER; HORROCKS; SATTLER, 2008).

Tabela 3.1 – A sentença idêntica representada em distintas lógicas.

Lógica	Sentença
Lógica Proposital	$p \vee m \rightarrow a$
Lógica de Primeira Ordem	$\forall x(Passaro(x) \vee Mamifero(x) \rightarrow Animal(x))$
Lógica de Descrição	$Passaro \sqcup Mamifero \sqsubseteq Animal$

Fonte: o autor.

Na análise das sentenças acima podemos observar diferentes formas de expressar uma relação lógica. A primeira sentença é formulada na forma proposicional, proporcionando uma visão direta da implicação: *se algo é um pássaro ou um mamífero, então é considerado um animal*. A segunda sentença, na lógica de primeira ordem, vai além, introduzindo a variável x e universalidade. Afirma que *para qualquer entidade x que seja um pássaro ou mamífero, essa entidade também é um animal*. A terceira sentença, em lógica de descrição, expressa a inclusão de conjuntos. Ela indica que *a união dos conjuntos de pássaros e mamíferos está contida no conjunto dos animais*, destacando as relações de conjuntos e subconjuntos.

Percebe-se que a Lógica de Descritiva oferece a possibilidade de interpretar suas sentenças por meio da Teoria dos Conjuntos. A última sentença na tabela sugere que a *União dos conjuntos Pássaro e Mamífero está contida no conjunto Animal*. Contudo, é crucial exercer certa cautela ao compreender isso. Os conceitos são, na realidade, mapeados para conjuntos, os quais, por sua vez, são subconjuntos de um domínio. Esse mapeamento é realizado por uma função de interpretação (BAADER; HORROCKS; SATTLER, 2008).

Os avanços na pesquisa da Representação do Conhecimento levaram ao desenvolvimento das chamadas Linguagens Terminológicas de Representação, como já mencionado. Uma dessas linguagens é a *Knowledge Representation Language One (KL-ONE)*, concebida por Brachman e Schmolze (1989), que se destaca por sua semântica clara e pela distinção entre conhecimento assertivo e terminológico.

Essas pesquisas resultaram no surgimento da *ALC (SCHMIDT-SCHAUSS; SMOLKA, 1991)*, uma linguagem cuja sigla representa *Attributive Concept Language with Complements*, isto é, representa as operações básicas que compõem essa linguagem lógica. De acordo com Schmidt-Schauß e Smolka (1991), o *ALC* é uma linguagem lógica utilizada na Lógica Descritiva, que é usada para representar conceitos e relações em ontologias e sistemas de representação de conhecimento. A lógica *ALC* é altamente utilizada na modelagem de ontologias e descrição de domínios em várias áreas, como inteligência artificial, web semântica, medicina, bioinformática e muitas outras. É conhecida por ser expressiva o suficiente para representar uma ampla gama de conceitos e relações complexas, enquanto ainda mantém a decidibilidade computacional, o que é crucial para aplicações práticas. *ALC* é considerada

uma das Lógicas de Descrição mais elementares, apesar de ser amplamente empregada em atividades que requerem raciocínio lógico (SCHMIDT-SCHAUSS; SMOLKA, 1991).

No contexto mencionado, a organização dos conceitos e propriedades segue as diretrizes da gramática apresentada na figura (3.1). Essa gramática representa a formulação linguística correspondente à linguagem lógica conhecida como *ALC*. Por meio dessa linguagem, pode-se combinar conceitos e papéis utilizando um conjunto específico de operadores para criar novos conceitos. Os operadores utilizados na lógica tradicional *ALC* (SCHMIDT-SCHAUSS; SMOLKA, 1991) contempla os seguintes construtores, para quaisquer conceitos *C* e *D*. A variável *R* na tabela é usada para representar a regra associada à propriedade.

Figura 3.1 – *ALC* sendo utilizada para representar conceitos em DL.

C, D	\rightarrow	A	(conceito atômico)
		\top	(conceito universal)
		\perp	(conceito vazio)
		$\neg C$	(negação)
		$C \sqcap D$	(interseção)
		$C \sqcup D$	(união)
		$\forall R.C$	(restrição de valor)
		$\exists R.C$	(quantificação existencial)

Fonte: Fernando (2010).

Uma das maneiras usuais de interpretar termos de conceitos consiste em traduzi-los para expressões de teoria dos conjuntos.

Neste caso, uma interpretação é definida como um mapeamento *I* de termos de conceitos para conjuntos de indivíduos de um universo Δ e de termos de papéis para relações binárias sobre Δ que satisfaça as seguintes exigências: *I* deve atribuir a cada termo de conceito um conjunto de indivíduos, e a cada termo de papel uma relação binária sobre Δ . A interpretação *I* deve satisfazer as condições da teoria dos conjuntos e as regras de inferência da lógica.

Formalmente, define-se uma interpretação como sendo um mapeamento *I* de termos de conceitos para conjuntos de indivíduos de um universo Δ e de termos de papéis para relações binárias sobre Δ satisfazendo as seguintes condições na tabela (3.2).

Tabela 3.2 – Interpretação DL para Teoria dos Conjuntos.

$I(\top) = \Delta$ $I(\perp) = \emptyset$ $I(\neg C) = \Delta - I(C)$ $I(C \sqcap D) = I(C) \cap I(D)$ $I(C \sqcup D) = I(C) \cup I(D)$ $I(\forall R.C) = \{a \in \Delta \mid \forall b. [(a, b) \in I(R) \Rightarrow b \in I(C)]\}$ $I(\exists R.C) = \{a \in \Delta \mid \exists b. [(a, b) \in I(R) \wedge b \in I(C)]\}$
--

Fonte: [Fernando \(2010\)](#).

Com isso, se um indivíduo a pertencer ao conjunto representado por um termo de conceito C , dizemos que está em C . Uma relação binária R , entre par (a, b) de indivíduos, significa que b preenche a propriedade R de a , ou seja, b é um preenchedor de R . Um termo de conceito da forma $\forall R.C$ representa o conjunto de todos os indivíduos que possuem todos os preenchedores de R , se existirem, em C . Um termo de conceito da forma $\exists R.C$ representa o conjunto de todos os indivíduos que têm pelo menos um preenchedor de R em C .

Essa forma de representação também possibilita a escrita de axiomas terminológicos, tais como os relacionados a fórmula de subsunção $C \sqsubseteq D$, que significa que a classe denotada por C está contida na classe denotada por D . Uma fórmula de equivalência $C \equiv D$ significa que C e D denotam a mesma classe. Uma *TBox* é uma coleção de fórmulas de subsunção e/ou equivalência em DL tomadas como axiomas. Ela é usada para representar a taxinomia de termos em um determinado domínio de conhecimento. Esta taxinomia é definida pela relação de subsunção entre os termos. Usando *OWL* ([BECHHOFFER et al., 2004](#)), estes termos e propriedades são caracterizados por meio de uma fórmula em DL. Assim, uma ontologia é definida como simplesmente uma *DL TBox* ([GRUBER, 1995](#)).

A ontologia pode ser vista como um esquema estrutural, em que relações semânticas entre diversos conceitos são estabelecidas, permitindo que o conteúdo seja formalmente especificado e compartilhado. Assim, ela pode ser comparada a um organograma complexo, aplicável às áreas de medicina, biologia e administração, entre outras. ([SHADBOLT; BERNERS-LEE; HALL, 2006](#)).

As ontologias computacionais são sistemas conceituais declarativos e relacionais que permitem a representação e explicação do conhecimento. Estes sistemas são definidos por um vocabulário, regras e termos, que ao serem combinados criam um domínio. Assim, eles permitem a inferência do conhecimento e a criação de conhecimento a partir de processos computacionais. Essa tecnologia tem sido usada para economizar tempo e dinheiro no processo de análise documental na busca de informações e/ou inconsistências ([BRACHMAN; LEVESQUE; REITER, 1992](#)).

A IA foi criada em 1950, com o conceito de que as máquinas poderiam pensar, agir e realizar ações de forma inteligente. Um sistema inteligente usa raciocínio lógico e a capaci-

dade de aprender. A IA possui várias definições entre os diversos estudiosos da área, pois é uma área ampla, abrangendo várias subáreas (RUSSELL; NORVIG, 2013). A Aprendizagem de Máquina é uma dessas subáreas, desempenhando um papel fundamental no desenvolvimento da habilidade de aprender de um agente inteligente.

No início dos anos 80, os sistemas de IA passaram a ser robustecidos por uma série de regras e regulamentos para a utilização do conhecimento, como por exemplo, os sistemas especialistas. Essa abordagem foi geralmente denominada de Inteligência Artificial Simbólica (CHOLLET, 2018).

Com a ampliação do enredamento nas tarefas como reconhecimento de imagens, classificação de sinais e tradução de idiomas, ficou infactível a aplicação da IA Simbólica, por causa da precisão da identificação de grandiosas e complicadas regras. Assim, tornou-se essencial o avanço de outros métodos de Aprendizado de Máquina.

Com a demanda por tarefas mais complexas, a Inteligência Artificial Simbólica não conseguiu mais atender às necessidades. Assim, o Aprendizado de Máquina passou a ser o método mais adequado para essa finalidade. Turing (1950), o idealizador desse método, questionou se os computadores seriam capazes de pensar e criar regras, apenas olhando para um conjunto de dados. Com isso, foi desenvolvido um novo modelo, em que o antigo conceito de inserir dados, definir regras e obter o resultado foi substituído. O objetivo do ML é fazer com que os computadores aprendam com base na experiência durante a realização de uma tarefa, permitindo que seu comportamento se modifique de acordo com a aquisição de dados (MITCHELL, 1997).

Ainda segundo Mitchell (1997), Aprendizado de Máquina tem como objetivo criar algoritmos capazes de aprender a partir de erros e exemplos. Estes algoritmos são classificados em três categorias: Aprendizado supervisionado, Aprendizado não supervisionado e Aprendizado por reforço.

O conhecimento é um conjunto integrado de dados e relações que, quando interpretados corretamente, geram desempenho eficaz. Esta informação pode ser definida como o conteúdo armazenado ou os modelos que uma pessoa ou máquina usa para interpretar, identificar, prever e responder apropriadamente ao ambiente externo (ARTERO, 2009). Para que esta informação seja salva, recuperada, atualizada, aprimorada, conectada a outros conhecimentos e usada em processos de inferência, é necessária a aplicação de representações adequadas para estas finalidades e para o tipo de processamento do agente inteligente. É importante notar que cada agente, diante da mesma realidade, percebe eventos diferentes. Mesmo que vejam o mesmo evento, interpretam e classificam a experiência de forma distinta.

Um sistema inteligente, portanto, deve representar as informações coletadas do meio externo de uma maneira conveniente para o computador digital, permitindo o desenvolvimento de estratégias de tomadas de decisão e de resolução de problemas.

Capítulo 4

Trabalhos Relacionados

Neste capítulo será apresentado o estado da arte envolvendo o tema da pesquisa. Como se notará nas seções (4.1) e (4.2), há um interesse no emprego de IA para representar conhecimento, o que torna relevante a condução de pesquisas com estes temas. Entretanto, o emprego métodos formais ainda se mostra incipiente. Razão pela qual muito do que está em desenvolvimento se encontra dentro de laboratórios. Nesse sentido, na seção (4.3) será apresentado os trabalhos do grupo de pesquisa TECMF da PUC-Rio, que compartilha linha de pesquisa que relaciona métodos formais para representar e extrair conhecimento de bases de dados, empregando também IA.

4.1 Análise dos artigos

Em uma pesquisa conduzida no Portal de Periódicos Capes, com o intuito de avaliar a disponibilidade de recursos acadêmicos, utilizando o termo de busca “Qualquer campo contém *Ontology* E Qualquer campo contém *machine learning* E Qualquer campo contém *Knowledge representation*”, englobando periódicos revisados por pares, artigos, atas de congressos, livros, conjuntos de dados e capítulos de livro, os resultados retornaram um total de 756 recursos online, 456 periódicos revisados por pares, 513 artigos, 89 atas de congressos, 69 livros, 57 conjuntos de dados e 15 capítulos de livro relacionados aos campos de *Machine Learning*, *Science and Technology*, *Ontology*, *Computer Science* e *Knowledge Representation*. Adicionalmente, foram identificados 224 recursos relacionados ao *Science Citation Index Expanded (Web of Science)*, 194 recursos relacionados ao *Directory of Open Access Journals (DOAJ)*, 162 recursos relacionados ao *Gale Academic OneFile*, 126 recursos relacionados ao *PubMed* e 103 recursos relacionados ao *PubMed Central*. No total, foram encontrados 754 recursos em inglês, 9 recursos em espanhol, 4 recursos em turco e 4 recursos em alemão.

Pode-se aferir do resultado que o tema de pesquisa é atual, pois a restrição a partir do ano 2000 retornou 754 produtos sendo, destes, 513 artigos em 456 periódicos revisados por pares. Ademais, 224 recursos contam com referencia na *Science Citation Index Expanded* o

que mostra o alcance do tema em periódicos de renome.

Ao modificar o termo de busca para “Qualquer campo contém *ontology* E Qualquer campo contém *machine learning* E Qualquer campo contém *Knowledge representation* E Qualquer campo contém *description logic*”, ou seja, incluindo *description logic* à pesquisa anterior, a quantidade de recursos identificados é reduzida para um total de 36 periódicos revisados por pares, 27 livros, 19 artigos, 3 atas de congressos e 2 conjuntos de dados. Considerando que a ferramenta difere Resumos de artigos científicos, razão pela qual há discrepância entre a quantidade de artigos e de periódicos.

Além disso, foram identificados 35 ocorrências do tema *inteligência artificial*, 22 ocorrências de *intellectronics* e 22 ocorrências de *pensamento artificial*. Entre as plataformas, o *Springer Books Online Archive*, *Science Citation Index Expanded (Web of Science)*, *Springer-Link Journals*, *Journals@Ovid* e *DOAJ Directory of Open Access Journals* foram citados 19, 14, 9, 5 e 5 vezes, respectivamente.

Ao focalizar os 19 artigos resultantes da pesquisa, uma análise foi conduzida para verificar se eles possuíam conexão com o escopo da pesquisa proposta.

Em (MEHRI; HAARSLEV; CHINAEI, 2021) é proposta uma nova abordagem de raciocínio OWL que utiliza o ML para implementar estratégias de tomada de decisão ótimas e pragmáticas em cenários com heurísticas incompletas. Disjunções que ocorrem em ontologias são uma fonte de ações não determinísticas nos raciocinadores. Duas abordagens baseadas em ML para reduzir o não determinismo provocado pelo tratamento de disjunções são propostas. A primeira abordagem é limitada à DL proposicional enquanto a segunda é capaz de lidar com DL padrão. Ambas as abordagens podem acelerar o raciocínio baseado em ML em até duas ordens de grandeza em comparação com o raciocínio não ML.

Em (LEHMANN; HITZLER, 2010) os autores apresentam um algoritmo de aprendizado baseado em operadores de refinamento para a Lógica de Descrição *ALCQ*, incluindo suporte para papéis concretos. O algoritmo foi desenvolvido a partir de fundamentos teóricos, identificando possíveis combinações abstratas de propriedades para operadores de refinamento em Lógicas Descritivas. Com essas investigações como base, foi gerado um operador de refinamento completo e útil para a prática. O operador foi então testado no algoritmo de aprendizado DL-Learner, e os resultados da avaliação mostraram que a abordagem é superior a outras abordagens de aprendizado em Lógicas Descritivas e é competitiva com sistemas de Programação Lógica Indutiva (ILP) já existentes.

Em (RETTINGER *et al.*, 2012), os autores investigam abordagens estatísticas para mineração de dados. Discutem métodos como similaridade e distância, máquinas kernel, modelos de previsão multivariados, modelos gráficos relacionais e abordagens de aprendizado probabilístico de primeira ordem, e sua aplicabilidade à representação da Web Semântica. Por fim, apresentam experimentos selecionados realizados em tarefas de mineração da Web

Semântica para alguns dos algoritmos descritos, a fim de demonstrar a amplitude e o potencial desta nova área de pesquisa e aplicação.

[Kuznetsov e Poelmans \(2013\)](#) atestam que no passado a Análise de Conceito Formal (FCA) era frequentemente mal interpretada como um formalismo estático e inadequado para tabelas de dados binárias, no texto pretende-se mostrar que a FCA fornece suporte para o processamento de dados dinâmicos e complexos (incluindo dados incertos) que são aprimorados com conhecimento adicional.

Em [\(AMITH *et al.*, 2022\)](#) tem-se um estudo de interesse para a iniciativa Bridge2AI do Instituto Nacional de Saúde dos Estados Unidos para abordar os desafios FAIR (Encontreáveis, Acessíveis, Interoperáveis e Reutilizáveis, do inglês *Findable, Accessible, Interoperable and Reusable*) com modelos de aprendizagem de máquina baseados em inteligência artificial para pesquisa biomédica. É apresentada uma primeira empreitada em desenvolver uma ontologia para capturar as informações de nível conceitual incorporadas nos relatórios de cartão de modelo.

Em [\(MA; MOLNÁR; BENCZÚR, 2021\)](#), os autores abordam os problemas de colisão semântica e inconsistências entre ontologias e o modelo de dados original ao aprender ontologia a partir de banco de dados relacional (RDB), apresenta-se um método de verificação de consistência semântica semi-automático baseado na representação intermediária de gráfico e verificação de modelo. Inicialmente, o W-Graph, como um modelo intermediário entre bancos de dados e ontologias, foi utilizado para formalizar as correspondências semânticas entre bancos de dados e ontologias, que foram então transformadas na estrutura de Kripke e, eventualmente, codificadas com o programa SMV (Verificador de Modelo Simbólico, do inglês *Symbolic Model Verifier*). Enquanto isso, a DL foi empregada para formalizar as especificações semânticas das ontologias aprendidas, já que o OWL-DL apresentou boa compatibilidade semântica e as DLs apresentaram excelente expressividade. Posteriormente, as especificações foram convertidas em uma fórmula de Lógica de Árvore Computacional (CTL, do inglês *Computer Tree Logic*) para melhorar a leitura da máquina. A tarefa de verificação de consistência semântica pôde ser convertida em um problema de verificação de modelo global que poderia ser resolvido automaticamente pelo verificador de modelo simbólico. Além disso, é dado um exemplo para demonstrar o processo específico de formalização e verificação da consistência semântica entre as ontologias aprendidas e o RDB, e um experimento de verificação foi realizado para verificar a viabilidade do método apresentado. Os resultados mostraram que o método apresentado poderia verificar e identificar corretamente os diferentes tipos de inconsistências entre as ontologias aprendidas e seu modelo de dados original.

Em [\(CARDILLO; STRACCIA, 2022\)](#), dada uma ontologia OWL e uma classe alvo T , os autores abordam o problema de aprender axiomas de inclusão de conceitos *fuzzy* que descrevem condições suficientes para ser uma instância individual de T (e em que grau). Para

isso, apresentam o *Fuzzy OWL-Boost* que se baseia no algoritmo de *boosting Real AdaBoost* adaptado para o caso (*fuzzy*) OWL. Sua eficácia é ilustrada por meio de uma experiência com várias ontologias.

Bresso *et al.* (2021) propõem minerar grafos de conhecimento para identificar características biomoleculares que possam habilitar automaticamente a reprodução de classificações especialistas que diferenciam drogas causadoras ou não de um determinado tipo de Reação Adversa (RA). Por meio de técnicas de classificação simples, como Árvores de Decisão e Regras de Classificação, é extraído um grafo de conhecimento para características; treinados classificadores para distinguir, com base nas características extraídas, drogas associadas ou não a duas reações adversas comumente monitoradas: lesões hepáticas induzidas por drogas e reações adversas cutâneas graves; isoladas características que são eficientes na reprodução de classificações de especialistas e interpretáveis por especialistas (isto é, termos do *Gene Ontology*, alvos de drogas ou nomes de vias); e avaliados manualmente, em um mini-estudo, como elas podem ser explicativas.

Na mesma linha de abordagem de Bresso *et al.* (2021), há o trabalho de Henegar *et al.* (2006), entretanto, ontologia tem um papel periférico no estudo, funcionando como uma base de consulta, com enfoque principal na identificação de classificadores biomoleculares.

Em (DOURLENS; RAMDANE-CHERIF; MONACELLI, 2013) é apresentada uma arquitetura semântica para a resolução de interação multimodal. Tal arquitetura é baseada em sistemas multi-agentes, onde os agentes são puramente semânticos usando ontologias e sistemas de inferência.

Maier, Ma e Hitzler (2013) apresentam Lógicas Descritivas paraconsistentes de quatro valores que podem raciocinar sobre inconsistências com foco em lógicas correspondentes à OWL-DL e seus perfis. O *framework* geral descrito tem a vantagem distinta de permitir que raciocinadores clássicos desenhem conclusões sólidas, mas não triviais, de bases de conhecimento mesmo inconsistentes. Lacunas e excessos de valores de verdade também podem ser eliminados seletivamente dos modelos (por meio da inserção de axiomas adicionais nas bases de conhecimento). Se as lacunas, mas não os excessos, forem eliminados, conclusões clássicas adicionais podem ser desenhadas sem afetar a paraconsistência.

Em (FANIZZI, 2011) é apresentada uma abordagem para a construção de ontologia realizada através da indução de descrições de conceitos expressas em Lógica de Descrição. O autor revisa as bases teóricas das representações padrão para ontologias formais na Web Semântica. Após estabelecer o problema de aprendizagem nesse contexto peculiar, é apresentado um algoritmo semelhante ao FOIL¹ que pode ser aplicado para aprender descrições de conceitos em DL. O algoritmo realiza uma busca por meio de um espaço de definições de conceitos candidatos através de operadores de refinamento. Esse processo é guiado por

¹ FOIL é um algoritmo de Aprendizado de Máquina supervisionado usado para classificar dados.

heurísticas que se baseiam nos exemplos disponíveis. O autor discute aspectos teóricos relacionados à aprendizagem com a incompletude inerente à semântica dessa representação. A avaliação experimental do sistema DL-FOIL, que implementa o algoritmo de aprendizagem, foi realizada em duas séries de sessões em ontologias reais de repositórios padrão para diferentes domínios expressos em diversas Lógicas de Descrição.

[Belazoui, Telli e Arar \(2021\)](#) propõem um sistema de tutor inteligente que pode acessar conteúdos de ensino disponíveis na web automaticamente e oferecê-los aos alunos como fontes de informações adicionais. Ao fazer isso, os autores destacam a abordagem da Lógica de Descrição e sua força de representação de conhecimento que garante a modularização, inferência e consulta sobre uma linguagem de ontologia da web, e reforçam a arquitetura de sistemas de tutoria tradicionais usando ontologias e Lógica de Descrição para permitir que eles acessem várias fontes de dados na web. Por fim, os autores concluem que a combinação de Aprendizado de Máquina com a web semântica proporciona um ambiente de estudo de apoio e melhora as condições de ensino dentro de ensino aberto e à distância.

[Bannour e Hudelot \(2014\)](#) propõe uma metodologia para construir ontologias multimídia fuzzy dedicadas à anotação de imagem.

Em ([MIKROYANNIDI *et al.*, 2012](#)) é demonstrado o uso do RIO; um framework para detectar regularidades sintáticas usando análise de cluster das entidades na assinatura de uma ontologia.

Em ([OZAKI, 2020](#)) é fornecida uma visão geral de abordagens de Aprendizado de Máquina clássico e mineração de dados que foram propostas para (semi-)automatizar a criação de ontologias de DL e como elas foram adaptadas para lidar com ontologias DL. Por fim, são discutido os benefícios e limitações de cada uma delas para aprender ontologias DL.

[Dasiopoulou, Kompatsiaris e Strintzis \(2010\)](#) investigam um quadro de raciocínio baseado em lógica fuzzy, que possibilita a integração de classificações de cenas e objetos em uma interpretação semanticamente consistente, capturando e utilizando as associações semânticas subjacentes. A avaliação com dois conjuntos de classificadores de entrada, configurados de modo a variar em relação à riqueza das interações de conceitos, destaca o potencial da abordagem proposta na presença de associações semanticamente ricas, enquanto delinea as questões e desafios envolvidos.

[Fanizzi, D'Amato e Eposito \(2008\)](#) trata de abordagens não-paramétricas para aprendizado estatístico aplicado a linguagens de representação de conhecimento padrão adotadas no contexto da Web Semântica. São apresentados métodos baseados em inferência epistêmica que são capazes de provocar e explorar a similaridade semântica de indivíduos em bases de conhecimento OWL.

[Atof, Hudelot e Bloch \(2014\)](#) propõe uma maneira original de enriquecer Lógicas de Descrição com serviços de raciocínio abduutivo. Com base nas teorias de conjuntos e de

conjuntos ordenados, são reunidos ingredientes de morfologia matemática, Lógicas de Descrição e análise de conceitos formais para computar as melhores explicações de uma observação através da erosão algébrica sobre o reticulado de conceitos de uma teoria de fundo que é construída de forma eficiente usando ferramentas de análise de conceitos formais.

No estudo de [Monteiro, Silva e Dias \(2017\)](#), são exploradas técnicas de aprendizado de máquina para detectar sinais de bullying em textos do Twitter. O estudo ressalta as diversas consequências negativas do bullying, desde problemas psicológicos até situações extremas como tiroteios em escolas. O objetivo central é guiar pesquisas futuras sobre o bullying em língua portuguesa, especialmente no contexto das redes sociais como o Twitter. O trabalho propõe três principais contribuições: primeiro, a criação de um conjunto de textos em português que contenham indícios de bullying extraídos do Twitter, incluindo a identificação do papel do autor no texto (vítima, praticante ou relator/narrador); segundo, a aplicação de técnicas de aprendizado de máquina para modelar computacionalmente a inferência e detecção de textos com traços de bullying nas redes sociais; terceiro, uma vez que um texto de bullying é identificado, proporcionar a classificação automática do papel do autor como vítima, praticante ou relator/narrador de um incidente de bullying.

4.2 Análise final dos trabalhos relacionados

Dois trabalhos são notáveis por sua relevância à visão de pesquisa, que serão detalhados nesta seção.

4.2.1 Sobre o trabalho de Lehmann e Hitzler

Em ([LEHMANN; HITZLER, 2010](#)) é apresentado um método para aprender novos conceitos a partir de conceitos já existentes, utilizando operadores de refinamento.

Os operadores de refinamento são utilizados para modificar um conceito existente, gerando um novo conceito que é mais específico ou mais geral do que o conceito original. O método proposto pelos autores utiliza uma técnica de busca baseada em heurísticas para explorar o espaço de possíveis refinamentos e encontrar o melhor conceito que se ajusta aos exemplos de treinamento.

[Lehmann e Hitzler \(2010\)](#) apresentam uma avaliação experimental do método proposto em um conjunto de dados sintético e em um conjunto de dados do mundo real, mostrando que o método é capaz de aprender conceitos precisos e eficientes a partir de exemplos de treinamento limitados. Para tal empregam a ferramenta DL-Learner para explorar o espaço de possíveis refinamentos de um conceito existente e encontrar o melhor conceito que se ajusta aos exemplos de treinamento.

DL-Learner é desenvolvida em Java e é fornecida como software livre sob a licença

GNU Lesser General Public License (LGPL). Além disso, DL-Learner tem uma interface gráfica de usuário que permite a interação com a ferramenta sem a necessidade de conhecimento prévio em programação.

DL-Learner tem sido amplamente utilizado em aplicações de Aprendizado de Máquina baseadas em web semântica, como o desenvolvimento de ontologias, a classificação automática de dados, a identificação de relações entre conceitos e a recomendação de informações.

Exemplo 1. Um exemplo de uso de DL-Learner é a tarefa de aprendizado de ontologias a partir de exemplos. Suponha-se que esteja diante de um conjunto de exemplos de indivíduos em um domínio de interesse, juntamente com suas propriedades e valores. Por exemplo, considere o seguinte conjunto de exemplos:

- João é um homem, tem 30 anos, é solteiro e mora em São Paulo.
- Maria é uma mulher, tem 25 anos, é casada e mora em Belo Horizonte.
- Pedro é um homem, tem 40 anos, é casado e mora em Rio de Janeiro.

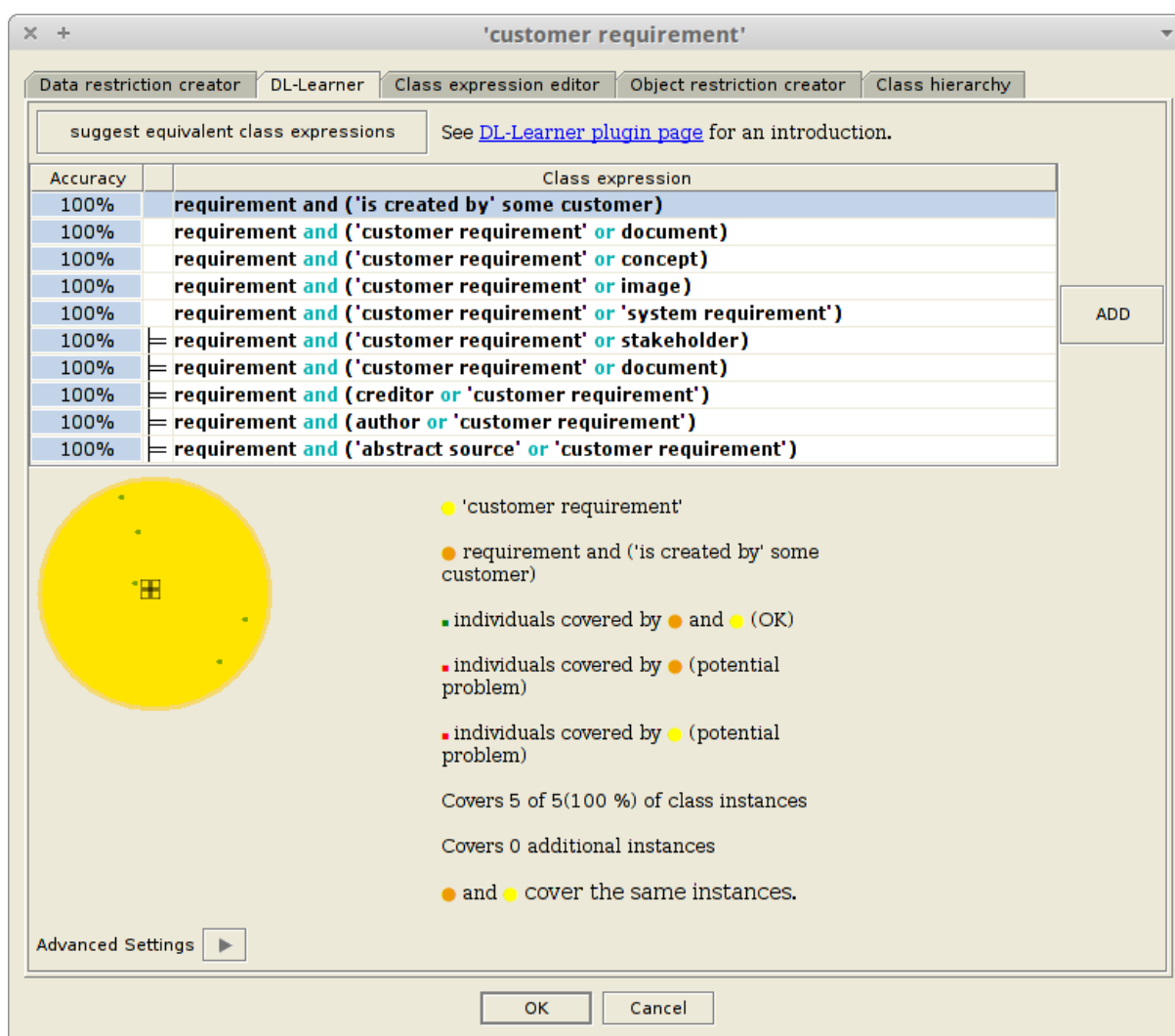
Com base nesses exemplos, podemos usar DL-Learner para aprender uma ontologia que represente o conhecimento sobre o domínio. Para isso, precisamos definir uma linguagem descritiva que permita a representação de conceitos, propriedades e restrições. Uma linguagem comum para isso é a linguagem OWL.

Com DL-Learner, podemos especificar quais propriedades e valores são importantes para a nossa ontologia, e em seguida, executar o algoritmo de aprendizado para gerar um conceito que seja consistente com os exemplos fornecidos. Por exemplo, podemos usar DL-Learner para gerar o seguinte conceito:

Pessoa e (temIdade min 25) e (temIdade max 40) e (temEstadoCivil some Solteiro) e (temEndereco some Brasil)

Este conceito representa uma pessoa com idade entre 25 e 40 anos, solteira e com endereço no Brasil. Esse conceito pode ser refinado ou modificado conforme necessário para representar com mais precisão o conhecimento do domínio. Uma vantagem do DL-learner é sua integração com o Protégé, conforme Figura (4.1).

Figura 4.1 – DL-learner Protégé plugin.



Fonte: Screenshots (2015).

4.2.2 Protégé

O Protégé é uma ferramenta de código aberto utilizada para a construção e manutenção de ontologias. Ele é desenvolvido pela Universidade de Stanford e é amplamente utilizado em pesquisa e aplicações comerciais em todo o mundo (GENNARI *et al.*, 2003).

A ferramenta fornece uma interface gráfica de usuário para criar, editar e visualizar ontologias. Ele suporta vários formatos de ontologia, incluindo OWL, RDE, RDFS e outras linguagens de descrição de conhecimento. O Protégé permite que os usuários criem classes, propriedades, restrições e axiomas em suas ontologias.

Uma das principais características do Protégé é a capacidade de integrar com outras ferramentas de processamento de linguagem natural e de inteligência artificial, como o DL-Learner e o OWL Reasoner. Essas ferramentas permitem a inferência e o aprendizado de novos conhecimentos a partir das ontologias criadas com o Protégé.

O Protégé é utilizado em diversos domínios de aplicação, incluindo biologia, medicina, finanças, indústria automotiva, entre outros. A ferramenta é bastante flexível e personalizável, permitindo que os usuários definam suas próprias visualizações, formas de representação e fluxos de trabalho.

Exemplo 2 (DL-leaener no Protégé). Para utilizar o DL-Learner em conjunto com o Protégé, é necessário primeiro instalar a extensão do DL-Learner no Protégé e, em seguida, exportar a ontologia em um formato que possa ser lido pelo DL-Learner.

Supondo que já temos uma ontologia OWL carregada no Protégé, podemos exportá-la em formato RDF/XML.

Com a ontologia exportada, podemos utilizar o DL-Learner para realizar a tarefa de aprendizado de conceitos. Suponha que temos um conjunto de exemplos que representam a classe *mamífero* e queremos aprender uma ontologia que descreva essa classe, DL-Learner irá executar a tarefa de aprendizado e gerar uma ontologia que representa a classe *mamífero* a partir dos exemplos fornecidos. Com a ontologia aprendida, podemos importá-la de volta para o Protégé e utilizá-la em outras aplicações.

4.2.3 O trabalho de Fanizzi

O artigo de [Fanizzi \(2011\)](#) propõe uma abordagem para indução de conceitos em ontologias de DLs usando heurísticas baseadas em teoria da informação. A ideia é que a teoria da informação possa ser utilizada para determinar a relevância de atributos na descrição de conceitos, o que pode ajudar a simplificar e refinar as ontologias. A abordagem proposta é aplicada em diferentes ontologias de domínio, mostrando resultados positivos em termos de eficiência e qualidade das ontologias induzidas. O autor apresenta um algoritmo de aprendizado semelhante ao FOIL que permite conjugação com DL.

4.2.3.1 FOIL

FOIL (*First-Order Inductive Learner*) é um algoritmo de Aprendizado de Máquina utilizado para indução de regras em Lógica de Primeira Ordem. Ele foi desenvolvido por Ross Quinlan em 1990 e é considerado um dos primeiros algoritmos para aprendizado de regras em Lógica de Primeira Ordem.

FOIL é um algoritmo fundamentado em exemplos, que recebe um conjunto de exemplos de treinamento e uma hipótese inicial como entrada. Ele, então, produz regras que se adequam de forma otimizada aos exemplos de treinamento. As regras geradas pelo FOIL são formuladas em Lógica de Primeira Ordem, podendo ser empregadas para deduções em um determinado domínio.

FOIL funciona de forma incremental, começando com uma hipótese inicial e, em seguida, refinando-a iterativamente. A cada iteração, o algoritmo gera uma regra candidata

que é avaliada com base na sua cobertura e precisão. A cobertura mede a quantidade de exemplos que a regra cobre, enquanto a precisão mede a proporção de exemplos cobertos pela regra que pertencem à classe correta.

As regras geradas pelo FOIL são ordenadas por uma medida de avaliação chamada ganho de informação, que mede a quantidade de informação ganha pela adição da regra à hipótese atual. As regras são geradas de forma recursiva, dividindo-se o espaço de atributos em subespaços mais simples e, em seguida, avaliando-se cada subespaço separadamente.

4.2.3.2 FOIL e DL

A integração entre FOIL e DL envolve a aplicação de conversões de conceitos em DL para a Lógica de Primeira Ordem (FOL, *First Order Logic*). Isso habilita o FOIL a ser empregado diretamente na indução de regras em ontologias em DL. Nesse contexto, é necessário transformar os conceitos em DL em fórmulas de Lógica de Primeira Ordem, tornando possível a utilização do FOIL.

A transformação é realizada através da utilização de um conjunto de regras de conversão, que associam cada conceito em DL com uma fórmula em FOL equivalente. Por exemplo, um conceito atômico em DL como “*Pessoa*” pode ser transformado em uma fórmula em FOL como $\forall x(Pessoa(x) \rightarrow \dots)$. Já conceitos complexos, como $Pai \sqcap \exists temFilho.Pessoa$, podem ser transformados em fórmulas em FOL utilizando regras que envolvem a utilização de quantificadores, conectivos lógicos e funções.

Exemplo 3. Uma definição de conceito na linguagem proposta pode ser: $Pai \equiv \sqcap Homem \sqcap \exists temFilho.\top$ que traduz a frase: “*um pai é um homem que tem alguém como filho.*”². (\top denota o conceito mais geral).

Agora, se definirmos dois novos conceitos:

$$PaiSemFilhosHomens \equiv Homem \sqcap \exists temFilho.\top \sqcap \forall temFilho.(\neg Homem)$$

e

$$Parente \equiv (Homem \sqcup Mulher) \sqcap \exists temFilho.\top$$

Pode-se inferir que: $Pai \sqsupseteq PaiSemFilhos$ e $Parente \sqsupseteq Pai$,

e, também, $Pai \not\sqsupseteq Parente$ e $PaiSemFilhos \not\sqsupseteq Pai$.

As asserções do A-box são fatos concretos como:

$Pai(João)$, $Homem(Pedro)$, $temFilho.Homem(João, Pedro)$, $\geq 1.temFilho(Joao)$, $\exists temFilho.\top(Pedro)$ e assim por diante.

A transformação de conceitos em DL para FOL pode ser realizada de forma automática por ferramentas especializadas, como os sistemas DL2OWL e Onto2FOL, que utiliza um

² no sentido de *child*

conjunto de regras de conversão para realizar a transformação de ontologias em DL em fórmulas em FOL.

Entretanto, as ferramentas são muitas vezes desenvolvidas por laboratórios de pesquisa para emprego em estudos do grupo. Nesse sentido a vinculação e/ou associação com grupos de pesquisa, por se tratar um tema ainda em fases iniciais de desenvolvimento, se faz necessária.

4.3 Grupo de pesquisa em Métodos Formais associado à pesquisa

Como se verifica, o tema de trabalho não possui muitas referências substanciais que permitam produzir resultados mais robustos.

A seleção do tema não foi aleatória, mas sim motivada pela oportunidade de colaborar com grupos de pesquisa no Brasil que atuam nessa área e mantêm intercâmbio com o programa de mestrado, através do professor orientador.

O TECMF, laboratório da PUC-Rio coordenado pelo prof. Edward Hermann Haeusler, desenvolveu uma solução de gestão de bases de conhecimento para a segurança da informação. Para tanto a empresa Módulo *Security* forneceu o problema do ponto de vista da indústria. Neste contexto o projeto tratava do desenvolvimento de uma metodologia de apoio a edição e manutenção de ontologias a partir de texto de padrões de segurança (ditos *Frameworks*) via ferramenta (desenvolvida pela equipe). Além da edição/criação de ontologias a partir de texto, um módulo de análise formal provendo explicações em linguagem natural foi desenvolvido. Este módulo é um provedor de explicações para Teoremas (enunciados na forma de subsunção) provados no sistema lógico de cálculo de seqüentes desenvolvido no TECMF em substituição ao raciocinador clássico em Tableaux normalmente utilizado para Lógica de Descrição (RADEMAKER; HAEUSLER, 2008).

Como continuação do trabalho anterior a equipe do TECMF produziu o artigo (SILVA *et al.*, 2007) em que detalha o processo de extração de explicações em linguagem natural.

Para melhorar o ganho no processamento das informações armazenadas na ontologia desenvolvida, foi criada uma versão simplificada de representação da ontologia o que gerou um ganho de tempo de processamento com o raciocinador Pellet. Um exemplo é apresentado na figura (4.2), em que se representa a ação com a seguinte descrição: Configure o parâmetro *Type* com valor *Nt5DS* no [SO Windows]. O objetivo dessa representação lógica é transformar uma ação complexa e descrita em linguagem natural em uma forma que máquinas possam entender e processar de maneira consistente e automatizada.

Figura 4.2 – Representação de ontologia simplificada.

Action Nested ≡	
1	$\exists hasVerb.(Configure \sqcap$
2	$\exists hasTheme.(TypeParameter \sqcap$
3	$\exists hasValue.(Nt5DS) \sqcap$
4	$\exists hasLocaliton.Windows)$

Fonte: [Silva et al. \(2007\)](#).

4.3.1 Pellet

Pellet é um motor de inferência OWL (Web Ontology Language) que implementa as especificações da linguagem de ontologia OWL 2.0 e é capaz de inferir relações lógicas entre classes, propriedades e indivíduos em ontologias. Ele é capaz de identificar inconsistências nas ontologias, inferir novas relações lógicas a partir das informações existentes e executar consultas sobre as ontologias ([SIRIN et al., 2007](#)).

O Pellet é o primeiro motor de inferência OWL-DL completo e correto, com suporte extensivo para inferência com indivíduos (incluindo suporte nominal e consulta conjuntiva), tipos de dados definidos pelo usuário e suporte para depuração de ontologias. Ele implementa várias extensões ao OWL-DL, incluindo um formalismo de combinação para ontologias OWL-DL, um operador não-monotônico e suporte preliminar para inferência híbrida OWL/Rule. O Pellet é escrito em Java e é de código aberto.

A integração do Pellet com o Protégé permite que os usuários aproveitem o poder de inferência do Pellet em suas ontologias construídas no Protégé. O Protégé fornece uma interface gráfica para carregar ontologias em Pellet e executar inferência, permitindo que os usuários identifiquem inconsistências e descubram novas relações lógicas em suas ontologias. O Pellet é amplamente utilizado como um motor de inferência OWL para ontologias e o Protégé é uma das ferramentas mais populares para a construção de ontologias OWL, tornando a integração entre as duas ferramentas valiosa para os usuários que desejam construir e inferir em ontologias.

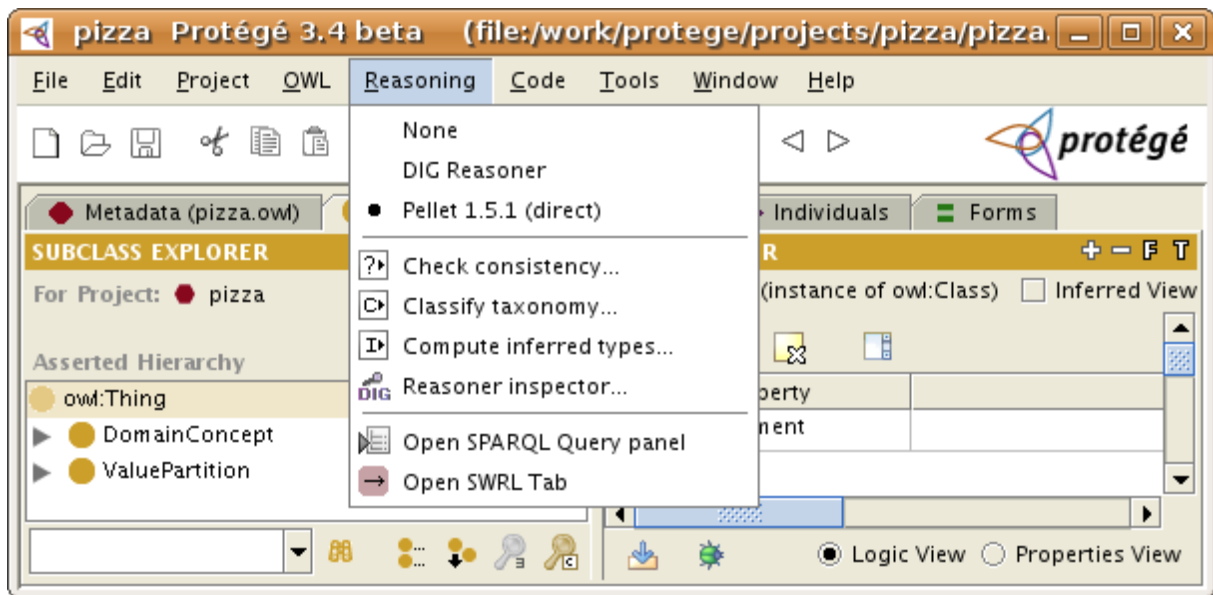
Agora é possível selecionar o Pellet como o motor de raciocínio para sua ontologia. Para fazer isso, vá até o menu *Reasoner* e selecione *Pellet*.

Na Figura (4.3), é apresentado o menu de seleção do raciocinador Pellet no Protégé.

4.3.2 NLTK

A NLTK (Kit de Ferramentas de Linguagem Natural, do inglês *Natural Language Toolkit*) é uma biblioteca de processamento de linguagem natural (NLP, do inglês *Natural Language Processing*) popular para a linguagem de programação Python. Foi desenvolvida para auxiliar os desenvolvedores e pesquisadores na realização de tarefas relacionadas ao proces-

Figura 4.3 – Pellet no Protégé.



Fonte: o autor.

samento de texto e linguagem natural. A NLTK oferece uma variedade de recursos e ferramentas para ajudar na análise e manipulação de textos em linguagem natural.

No quadro (4.1), são apresentados alguns dos recursos e das funcionalidades da NLTK.

Quadro 4.1 – Funcionalidades do NLTK.

Tokenização	A NLTK pode dividir um texto em palavras individuais (tokens), o que é útil para análises posteriores.
Processamento de texto	A biblioteca oferece métodos para transformar texto em letras minúsculas, remover pontuações e realizar outras operações de pré-processamento.
Lemmatização e Stemming	Essas técnicas permitem reduzir palavras a suas formas base, como transformar <i>corrida</i> , <i>correr</i> e <i>corre</i> em <i>correr</i> .
Análise gramatical	A NLTK é capaz de realizar análises gramaticais em textos, identificando partes do discurso, estruturas sintáticas e mais.
Análise de sentimentos	A biblioteca inclui recursos para análise de sentimentos em textos, permitindo determinar se um texto possui um sentimento positivo, negativo ou neutro.

Funcionalidades do NLTK - continuação

Modelos de linguagem probabilística	A NLTK oferece a capacidade de construir e utilizar modelos de linguagem probabilística, que são úteis para prever próximas palavras em uma sequência de texto.
Corpora	A NLTK inclui diversos corpora, que são coleções de textos para treinamento e teste de modelos de processamento de linguagem natural.
Tradução de idiomas	A biblioteca suporta tradução básica entre idiomas usando dicionários e regras gramaticais simples.
Aprendizado de máquina para NLP	A NLTK também oferece funcionalidades para criar modelos de aprendizado de máquina voltados para tarefas de processamento de linguagem natural.
Visualização de dados	A biblioteca inclui recursos para visualização de dados textuais, permitindo criar gráficos e representações visuais de informações extraídas do texto.

Fonte: Bird, Klein e Loper (2009).

A NLTK é amplamente utilizada tanto por iniciantes quanto por especialistas em processamento de linguagem natural e aprendizado de máquina. Ela oferece uma série de ferramentas e recursos prontos para uso.

Capítulo 5

Resultados

Uma das principais vantagens do uso de ontologias é a capacidade de fornecer uma representação semântica precisa do conhecimento. Ao invés de confiar exclusivamente em dados brutos, as ontologias permitem que informações adicionais sobre o domínio sejam incorporadas de forma estruturada. Isso facilita a organização, a categorização e a inferência sobre os dados, melhorando assim a eficiência do aprendizado de máquina. Além disso, a utilização de Lógica Descritiva permite que sejam feitas inferências lógicas sobre as relações entre os elementos do domínio, contribuindo para aprimorar a capacidade preditiva dos modelos.

Outro benefício significativo do uso de ontologias e Lógica Descritiva é a interpretabilidade dos modelos resultantes. Enquanto as abordagens de aprendizado de máquina baseadas apenas em dados podem fornecer resultados precisos, muitas vezes são consideradas *caixas-pretas*, dificultando a compreensão das razões por trás das decisões tomadas pelos modelos. Com a incorporação de ontologias e Lógica Descritiva, é possível obter modelos mais transparentes, nos quais as inferências são baseadas em conhecimento explícito e regras bem definidas. Isso torna o processo de tomada de decisão dos modelos mais compreensível e confiável, permitindo que especialistas do domínio verifiquem e validem as decisões tomadas pelos sistemas de aprendizado de máquina.

A abordagem de criar ontologias sobre domínios específicos permite raciocinar sobre os textos de maneira mais estruturada e semântica. Por exemplo, se construirmos uma ontologia para armazenar conteúdos de *Tweets* conforme o código (5.1), conseguimos organizar as informações de forma hierárquica, definindo as relações entre os elementos presentes nos *Tweets*, facilitando a análise e a interpretação desses dados.

Código 5.1 – Ontología *Tweets*

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
@prefix owl: <http://www.w3.org/2002/07/owl#>.
@prefix tw: <http://example.com/ontologies/twitter#>.

# Ontology declaration
owl:Ontology rdf:about="http://example.com/ontologies/twitter">

# Classes
owl:Class rdf:about="http://example.com/ontologies/twitter#Tweet">
  rdfs:label "Tweet";
  rdfs:comment "A representation of a tweet on Twitter".

# Properties
owl:ObjectProperty rdf:about="http://example.com/ontologies/twitter#
  hasAuthor">
  rdfs:label "has_author";
  rdfs:comment "Indicates the author of a tweet";
  rdfs:range owl:Thing;
  rdfs:domain tw:Tweet.

owl:DatatypeProperty rdf:about="http://example.com/ontologies/twitter#
  hasText">
  rdfs:label "has_text";
  rdfs:comment "Represents the text content of a tweet";
  rdfs:range xsd:string;
  rdfs:domain tw:Tweet.

owl:DatatypeProperty rdf:about="http://example.com/ontologies/twitter#
  hasTimestamp">
  rdfs:label "has_timestamp";
  rdfs:comment "Specifies the timestamp of a tweet";
  rdfs:range xsd:dateTime;
  rdfs:domain tw:Tweet.

owl:ObjectProperty rdf:about="http://example.com/ontologies/twitter#
  hasHashtag">
  rdfs:label "has_hashtag";
  rdfs:comment "Indicates the hashtags associated with a tweet";
  rdfs:range xsd:string;
  rdfs:domain tw:Tweet.
```

Uma consulta para determinar se há *tweet* com a palavra “*Excited*” pode ser feita em Lógica Descritiva a partir da expressão presente no código (5.2).

Código 5.2 – Exemplo de Consulta

```
tw:Tweet and tw:hasText some xsd:string[regex("[Ee]xcited")].
```

Nessa expressão, a combinação da classe *tw:Tweet* ocorre em conjunto com a restrição *tw:hasText some xsd:string[regex("[Ee]xcited")]*. Isso estabelece que o *tweet* precisa apresentar a propriedade *hasText* contendo a palavra “*excited*” (ou suas variações de letras maiúsculas/minúsculas, como “*Excited*” ou “*excited*”).

A combinação de OWL e a consulta em Lógica Descritiva com ML pode ser realizada de várias maneiras para melhorar o aprendizado de máquina em sistemas.

As classes, propriedades, objetos, tipos de dados e suas respectivas relações são definidas na ontologia OWL do *Tweet* exposta no Código (5.2), resultando em uma melhoria na qualidade dos dados de treinamento e fornecendo informações suplementares para os modelos de aprendizado de máquina.

A utilização da Lógica Descritiva para consultas possibilita a realização de análises complexas e refinadas sobre os dados representados na OWL. É viável formular consultas por meio da sintaxe da Lógica Descritiva, visando expressar restrições e padrões específicos que os dados devem cumprir. Essas consultas desempenham o papel de recuperar informações pertinentes de um conjunto de dados ou de efetuar operações de filtragem e seleção dos dados, a fim de prepará-los para treinamento ou inferência de modelos de aprendizado de máquina. No contexto em questão, a busca pela palavra “*excited*” não apresenta conotação ofensiva. Contudo, em outros cenários ou contextos, essa mesma palavra poderia assumir tal conotação. É fundamental considerar o contexto e a forma como certas palavras são utilizadas, além de utilizar ferramentas apropriadas, como expressões regulares, para garantir a precisão na análise dos textos.

A OWL permite a definição de regras lógicas e axiomas que podem ser usados para inferir novas informações a partir dos dados representados. Essas regras podem ser usadas para enriquecer os dados com conhecimento implícito, preencher lacunas nos dados ou deduzir relações ocultas entre as entidades. A inferência lógica pode ajudar a melhorar a qualidade dos dados de treinamento e a capacidade preditiva dos modelos de aprendizado de máquina.

As ontologias em OWL podem ser integradas diretamente aos algoritmos de aprendizado de máquina. Por exemplo, é possível utilizar técnicas de pré-processamento para transformar os dados representados em OWL em formatos adequados para o treinamento de modelos de aprendizado de máquina. Além disso, a estrutura e as relações definidas nas

ontologias podem ser incorporadas nos modelos de aprendizado de máquina, ajudando a melhorar a interpretabilidade, a generalização e o desempenho dos modelos.

Considere no código (5.3) no qual, a biblioteca *NLTK* é utilizada para processamento de texto e o classificador *Naive Bayes* é empregado para efetuar a classificação.

Código 5.3 – Uso da Biblioteca NLTK

```

from owlready2 import *
from nltk.tokenize import word_tokenize
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.naive_bayes import MultinomialNB
onto = get_ontology("path_to_your_ontology.owl").load()
keyword_dict = {
    'positivo': [onto.feliz, onto.alegre, onto.empolgado],
    'negativo': [onto.triste, onto.frustrado, onto.chateado]
}
# Funcao para extrair caracteristicas dos tweets
def extract_features(tweet):
    features = {}
    words = word_tokenize(tweet)
    for keyword, word_list in keyword_dict.items():
        features[keyword] = any(word.lower() in words for word in
            word_list)
    return features
training_data = [
    ('Estou_muito_feliz_hoje!', 'positivo'),
    ('Que_tristeza_perdi_meu_voo.', 'negativo'),
    ('Estou_empolgado_para_o_concerto!', 'positivo'),
    ('Fiquei_chateado_com_a_noticia.', 'negativo'),
]
tweets, labels = zip(*training_data)
vectorizer = TfidfVectorizer(tokenizer=word_tokenize)
X_train = vectorizer.fit_transform(tweets)
classifier = MultinomialNB()
classifier.fit(X_train, labels)
new_tweet = 'Estou_alegre_com_as_boas_noticias!'
new_tweet_features = extract_features(new_tweet)
new_tweet_vector = vectorizer.transform([new_tweet])
predicted_label = classifier.predict(new_tweet_vector)[0]
print('Tweet:', new_tweet)
print('Features:', new_tweet_features)
print('Predicted_Label:', predicted_label)

```

Neste exemplo, as palavras-chave foram substituídas pelo uso de conceitos previamente definidos na ontologia. Isso implica na necessidade de carregar a ontologia adequada

e estabelecer os conceitos pertinentes a cada categoria. No cenário em questão, pressupõe-se que a ontologia já inclui conceitos como *feliz*, *alegre*, *empolgado*, *triste*, *frustrado* e *chateado*, que representam as palavras-chave associadas a cada categoria.

Ao invés de efetuar a representação completa dos *tweets* no formato OWL, a função *extract_features* é empregada para extrair as características dos *tweets*, de maneira semelhante ao exemplo anterior. A distinção reside no fato de que, neste caso, as características são representadas pelos conceitos contidos na ontologia. O código (5.4) apresenta a inserção da regra lógica descritiva na ontologia.

Código 5.4 – Inserção de DL na Ontologia

```
import nltk
from nltk.tokenize import word_tokenize
from owlready2 import *

# Carregar a ontologia
onto = get_ontology("path_to_your_ontology.owl").load()

def extract_features(tweet):
    features = {}
    words = word_tokenize(tweet)
    for word in words:
        # Verificar se o tweet possui alguma propriedade
        # definida na ontologia
        if word in onto.properties():
            features[word] = True
    return features

training_data = [
    ('Estou_muito_feliz_hoje!', 'positivo'),
    ('Que_tristeza_perdi_meu_voo.', 'negativo'),
    ('Estou_empolgado_para_o_concerto!', 'positivo'),
    ('Fiquei_chateado_com_a_noticia.', 'negativo'),
]

nltk.download('punkt')

tweets = [tweet for tweet, _ in training_data]
labels = [label for _, label in training_data]

extracted_features = [extract_features(tweet) for tweet in tweets]
```

 Inserção de DL na Ontologia - continuação

```

with onto:
    class PositiveTweet(Thing):
        equivalent_to = [
            onto.Tweet & onto.hasPositiveProperty.some(onto.boolean),
            onto.Tweet & onto.hasPositiveSentiment.exactly(1)
        ]

    class NegativeTweet(Thing):
        equivalent_to = [
            onto.Tweet & onto.hasNegativeProperty.some(onto.boolean),
            onto.Tweet & onto.hasNegativeSentiment.exactly(1)
        ]

X_train = extracted_features
y_train = labels

# ... restante do código de treinamento e classificação

```

Dentro do código, a função *extract_features* realiza um processo iterativo sobre as palavras contidas em cada *tweet*, verificando se cada palavra corresponde a uma propriedade previamente definida na ontologia carregada. No caso de a palavra coincidir com uma propriedade, essa palavra é incluída como uma característica do *tweet*, sendo armazenada na variável *features*. Portanto, as características extraídas consistem nas palavras que se alinham com as propriedades previamente estabelecidas na ontologia.

Assim, o código (5.5) realiza a extração das palavras que se alinham com propriedades definidas na ontologia, utilizando-as como características dos *tweets*. No entanto, caso seja necessário verificar a presença de uma palavra exclusivamente no conteúdo do *tweet*, é suficiente aplicar uma restrição.

 Código 5.5 – Restrição de busca

```

import nltk
from nltk.tokenize import word_tokenize
from owlready2 import *

# Carregar a ontologia
onto = get_ontology("path_to_your_ontology.owl").load()

```

Restrição de busca - continuação

```
# Funcao para extrair características dos tweets
def extract_features(tweet):
    features = {}
    words = word_tokenize(tweet)
    for word in words:
        # Verificar se o tweet possui a propriedade hasText
        # definida na ontologia
        if word == 'hasText':
            features[word] = True
    return features

# Carregar os dados de treinamento
training_data = [
    ('Estou_muito_feliz_hoje!', 'positivo'),
    ('Que_tristeza_perdi_meu_voo.', 'negativo'),
    ('Estou_empolgado_para_o_concerto!', 'positivo'),
    ('Fiquei_chateado_com_a_noticia.', 'negativo'),
]

# Pre-processamento dos dados
nltk.download('punkt')

tweets = [tweet for tweet, _ in training_data]
labels = [label for _, label in training_data]

# Extrair características dos tweets
extracted_features = [extract_features(tweet) for tweet in tweets]

# Definir a regra de logica descritiva na ontologia
with onto:
    class TextTweet(Thing):
        equivalent_to = [
            onto.Tweet & onto.hasText.some(str)
        ]

# Treinar o classificador
X_train = extracted_features
y_train = labels

# ... restante do código de treinamento e classificacao
```

No contexto do código, a função **extract_features** realiza a verificação da presença da palavra *hasText* no tweet. Se a palavra *hasText* estiver presente, ela é acrescentada como uma característica do tweet à variável **features**. Nesse caso, não se trata da verificação da existência da propriedade *hasText* na ontologia, mas sim da identificação da ocorrência da palavra *hasText* no próprio texto do tweet. Consequentemente, o código extrai a presença da palavra *hasText* como uma característica dos *tweets*.

Capítulo 6

Conclusão

Representação do Conhecimento desempenha um papel fundamental ao estabelecer estruturas que permitem que as máquinas capturem, organizem e assimilem informações de maneira mais profunda, semântica e contextualmente fundamentada.

Ao combinar a competência da Inteligência Artificial em explorar e analisar dados com a habilidade de compreender o significado e o contexto por meio da Representação do Conhecimento, é possível identificar padrões sutis e comportamentos anômalos que, de outra forma, poderiam passar despercebidos.

A Linguagem de Ontologia para a Web destaca-se como uma ferramenta proeminente nesse contexto. Ela fornece um conjunto de recursos e sintaxes que possibilitam a definição formal dos elementos de uma ontologia, estabelecendo um alicerce para inferências lógicas. A capacidade de realizar inferências lógicas é fundamental para a análise e compreensão de informações.

As ontologias, ao estabelecerem uma estrutura sólida para a definição de conceitos e suas inter-relações, capacitam a Inteligência Artificial a compreender não apenas os eventos em andamento, mas também os motivos subjacentes e os modos como esses eventos se desenrolam. Essa compreensão contextual é crucial na identificação de padrões anômalos ou comportamentos suspeitos, muitas vezes não perceptíveis em uma análise superficial.

A pesquisa em questão concentrou-se na aplicação desses conceitos para aprimorar a detecção de atividades suspeitas em plataformas de mídia social, com foco especial na plataforma Twitter. O emprego de ontologias e da linguagem OWL, aliado às capacidades da biblioteca NLTK, se mostrou factível.

Ademais, convergência dessas tecnologias mostrou ser viável para a construção de modelos semânticos, capazes de discernir nuances e padrões em interações sociais complexas, como aquelas presentes nas redes sociais, o que é o objetivo inicial do trabalho.

Este é mais um ganho na condução de trabalhos envolvendo análise de informações, podendo ser aplicado em servidores de dados (Data Center).

6.1 Trabalhos Futuros

O presente estudo oferece uma base para futuras pesquisas que visam aprimorar ainda mais a detecção de comportamentos suspeitos em plataformas de redes sociais, conforme mostrado no quadro (6.1).

Quadro 6.1 – Possíveis trabalhos futuros.

<p>Expansão da Ontologia e Incorporação de Contextos Adicionais</p>	<p>Uma direção interessante é a expansão da ontologia utilizada neste estudo, incorporando elementos adicionais de contexto que possam enriquecer a compreensão dos comportamentos dos usuários. A adição de informações como localização geográfica, relações sociais e tópicos de interesse pode fornecer uma visão mais completa das interações nas redes sociais, contribuindo para a detecção mais precisa de comportamentos suspeitos.</p>
<p>Aplicação do Algoritmo FOIL</p>	<p>O algoritmo FOIL apresenta potencial para complementar a detecção de comportamentos suspeitos. FOIL é um algoritmo de aprendizado de regras de classificação que pode ser aplicado para inferir relações complexas entre os elementos. Sua aplicação poderia resultar na geração automática de regras que identificam padrões específicos associados a atividades suspeitas, melhorando ainda mais a precisão do sistema.</p>
<p>Integração com Bancos de Dados Relacionais</p>	<p>A integração de bancos de dados relacionais pode permitir a coleta e o armazenamento de informações de várias fontes. Isso pode enriquecer a análise, permitindo a correlação de dados de diferentes contextos. Além disso, a aplicação de consultas SQL (Linguagem de Consulta Estruturada, do inglês Structured Query Language) pode ser utilizada para buscar informações relevantes que possam fortalecer a detecção de comportamentos suspeitos.</p>
<p>Utilização da Ferramenta DL-Learner</p>	<p>A ferramenta DL-Learner é uma plataforma amplamente usada para o aprendizado de ontologias a partir de dados. Sua aplicação pode permitir a geração automatizada de axiomas ontológicos a partir das informações coletadas, refinando ainda mais a compreensão semântica dos dados e a capacidade de inferência.</p>
<p>Avaliação de Desempenho e Escalabilidade</p>	<p>Uma área fundamental para futuros estudos é a avaliação de desempenho e escalabilidade da abordagem proposta. Investigar como a metodologia se comporta em cenários de grande escala, com volumes massivos de dados e interações, é essencial para verificar sua viabilidade e aplicabilidade em casos do mundo real.</p>

Fonte: o autor.

Em resumo, os trabalhos futuros podem explorar a expansão e o aprimoramento da abordagem proposta, incorporando técnicas avançadas como o algoritmo FOIL, a integra-

ção com bancos de dados relacionais e a utilização da ferramenta DL-Learner. Essas direções têm o potencial de contribuir significativamente para a detecção de comportamentos suspeitos em plataformas de redes sociais, resultando em sistemas mais eficazes e precisos.

REFERÊNCIAS

AMITH, M. T. *et al.* Toward a standard formal semantic representation of the model card report. *BMC bioinformatics*, BioMed Central Ltd, England, v. 23, n. Suppl 6, 2022. ISSN 1471-2105. Citado na página 36.

ANTONIOU, G.; HARMELEN, F. V. *A semantic web primer*. Cambridge-MA: MIT press, 2004. Citado na página 27.

ANTONIOU, G.; HARMELEN, F. v. Web ontology language: Owl. In: _____. *Handbook on Ontologies*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. ISBN 978-3-540-92673-3. Disponível em: <<https://lnq.com/qk3ye>>. Citado na página 20.

ARTERO, A. O. *Intelifência artificial teórica e pratica*. São Paulo: Editora Livraria da Física, 2009. Citado na página 33.

ATOF, J.; HUDELOT, C.; BLOCH, I. Explanatory reasoning for image understanding using formal concept analysis and description logics : Decision making in human and machine vision. *IEEE transactions on systems, man, and cybernetics. Systems*, Institute of Electrical and Electronics Engineers, New York, NY, v. 44, n. 5, 2014. ISSN 2168-2216. Citado na página 38.

BAADER, E.; HORROCKS, I.; SATTLER, U. Description logics. *Foundations of Artificial Intelligence*, Elsevier, v. 3, 2008. Citado 2 vezes nas páginas 28 e 30.

BANNOUR, H.; HUDELOT, C. Building and using fuzzy multimedia ontologies for semantic image annotation. *Multimedia tools and applications*, Springer US, Boston, v. 72, n. 3, 2014. ISSN 1380-7501. Citado na página 38.

BECHHOFFER, S. *et al.* Owl web ontology language reference. *W3C recommendation*, v. 10, n. 2, 2004. Citado na página 32.

BELAZOUI, A.; TELLI, A.; ARAR, C. Web-based learning under tacit mining of various data sources. *International journal of emerging technologies in learning*, International Association of Online Engineering (IAOE), Vienna, v. 16, n. 16, 2021. ISSN 1863-0383. Citado na página 38.

BIEGING, P. *et al.* *Tecnologia e novas mídias: da educação às práticas culturais e de consumo*. São Paulo: Pimenta Cultural, 2013. Citado na página 25.

BIRD, S.; KLEIN, E.; LOPER, E. *Natural language processing with Python: analyzing text with the natural language toolkit*. Sebastopol, CA: O'Reilly Media, Inc., 2009. Citado na página 47.

BLACKBURN, P.; BENTHEM, J. F. van; WOLTER, F. *Handbook of modal logic*. Netherlands-NL: Elsevier, 2006. Citado na página 21.

BOYD, D. Social network sites as networked publics: Affordances, dynamics, and implications. In: *A networked self*. Oxfordshire-UK: Routledge, 2010. Citado na página 26.

_____. *It's complicated: The social lives of networked teens*. New Haven - CT: Yale University Press, 2014. Citado na página 24.

BRACHMAN, R. J.; LEVESQUE, H. J.; REITER, R. *Knowledge representation*. Cambridge-MA: MIT press, 1992. Citado 3 vezes nas páginas 19, 21 e 32.

BRACHMAN, R. J.; SCHMOLZE, J. G. An overview of the kl-one knowledge representation system. *Readings in artificial intelligence and databases*, Elsevier, 1989. Citado na página 30.

BRAGA, J.; SYLLA, B. *Filosofia da Tecnologia. Introdução ao Pensamento dos Teóricos do Século XX*. [S.l.]: Grácio Editor, 2022. Citado na página 17.

BRESSO, E. *et al.* Investigating adr mechanisms with explainable ai: a feasibility study with knowledge graph mining. *BMC medical informatics and decision making*, BioMed Central, London, v. 21, n. 1, 2021. ISSN 1472-6947. Citado na página 37.

BUCKINGHAM, D. *Crescer na era das mídias eletrônicas*. São Paulo: Edições Loyola, 2007. Citado na página 25.

CARDILLO, F. A.; STRACCIA, U. Fuzzy owl-boost: Learning fuzzy concept inclusions via real-valued boosting. *Fuzzy sets and systems*, Elsevier B.V, Ithaca, v. 438, 2022. ISSN 0165-0114. Citado na página 36.

CASTELLS, M. *A sociedade em rede*. [S.l.]: Paz e terra São Paulo, 2005. v. 1. Citado na página 25.

CHAPTER, I. *et al.* *Artificial Intelligence*. Uttar Pradesh: Tata McGraw Hill, 2009. Citado na página 20.

CHOLLET, F. *Deep Learning mit Python und Keras: Das Praxis-Handbuch vom Entwickler der Keras-Bibliothek*. Cologne: MITP-Verlags GmbH & Co. KG, 2018. Citado 2 vezes nas páginas 20 e 33.

COLL, C.; MONEREO, C. *Psicologia da Educação Virtual: Aprender e ensinar com as tecnologias da informação e da comunicação*. Rio de Janeiro-RJ: Artmed Editora, 2010. Citado na página 25.

DASIOPOULOU, S.; KOMPATSIARIS, I.; STRINTZIS, M. G. Investigating fuzzy dls-based reasoning in semantic image analysis. *Multimedia tools and applications*, Springer US, Boston, v. 49, n. 1, 2010. ISSN 1380-7501. Citado na página 38.

DOURLENS, S.; RAMDANE-CHERIE, A.; MONACELLI, E. Multi levels semantic architecture for multimodal interaction. *Applied intelligence (Dordrecht, Netherlands)*, Springer US, Boston, v. 38, n. 4, 2013. ISSN 0924-669X. Citado na página 37.

- FANIZZI, N. Concept induction in description logics using information-theoretic heuristics. *International journal on semantic web and information systems*, Igi Global, HERSHEY, v. 7, n. 2, 2011. ISSN 1552-6283. Citado 3 vezes nas páginas 21, 37 e 42.
- FANIZZI, N.; D'AMATO, C.; EPOSITO, F. Induction of classifiers through non-parametric methods for approximate classification and retrieval with ontologies. *International journal of semantic computing*, World Scientific Publishing Company, v. 2, n. 3, 2008. ISSN 1793-351X. Citado na página 38.
- FERNANDO, N. Aplicação de lógica descritiva para validação de requisitos em linha de produto de software. *Encontro Anual de Computação*, 2010. Citado 2 vezes nas páginas 31 e 32.
- GENNARI, J. H. *et al.* The evolution of protégé: an environment for knowledge-based systems development. *International Journal of Human-Computer Studies*, Elsevier, v. 58, n. 1, 2003. ISSN 1071-5819. Citado na página 41.
- GLADWELL, M. *The tipping point: How little things can make a big difference*. Lodon-UK: Little, Brown, 2006. Citado na página 24.
- GRUBER, T. R. Toward principles for the design of ontologies used for knowledge sharing? *International journal of human-computer studies*, Elsevier, v. 43, n. 5-6, 1995. Citado 2 vezes nas páginas 27 e 32.
- HARMELEN, F. V.; LIFSCHITZ, V.; PORTER, B. *Handbook of knowledge representation*. Netherlands-NL: Elsevier, 2008. Citado na página 29.
- HENEGAR, C. *et al.* Building an ontology of adverse drug reactions for automated signal generation in pharmacovigilance. *Computers in biology and medicine*, Elsevier Ltd, OXFORD, v. 36, n. 7, 2006. ISSN 0010-4825. Citado na página 37.
- HORROCKS, I.; SATTLER, U. Ontology reasoning in the SHOQ(D) description logic. In: *Proceedings of the 17th International Joint Conference on Artificial Intelligence - Volume 1*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2001. (IJCAI'01), p. 199–204. ISBN 1558608125. Citado na página 27.
- HUBERMAN, B. A.; ROMERO, D. M.; WU, F. Social networks that matter: Twitter under the microscope. *arXiv preprint arXiv:0812.1045*, 2008. Citado na página 26.
- KAPLAN, A. M.; HAENLEIN, M. Users of the world, unite! the challenges and opportunities of social media. *Business horizons*, Elsevier, v. 53, n. 1, 2010. Citado na página 23.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, v. 25, 2012. Citado na página 18.
- _____. _____. *Communications of the ACM*, ACM New York, NY, USA, v. 60, n. 6, 2017. Citado na página 18.
- KUZNETSOV, S. O.; POELMANS, J. Knowledge representation and processing with formal concept analysis. *Wiley interdisciplinary reviews. Data mining and knowledge discovery*, John Wiley & Sons, Inc, Hoboken, USA, v. 3, n. 3, 2013. ISSN 1942-4787. Citado na página 36.
- LAZER, D. *et al.* Computational social science. *Science*, American Association for the Advancement of Science, v. 323, n. 5915, 2009. Citado na página 26.

- LEAL, A. Irá acontecer hoje: aluno anunciou ataque a escola em rede social. *OGlobo*, Mar 2023. Disponível em: <<https://link.dev/CFwig>>. Citado 2 vezes nas páginas 24 e 25.
- LEHMANN, J.; HITZLER, P. Concept learning in description logics using refinement operators. *Machine learning*, Springer US, Boston, v. 78, n. 1-2, 2010. ISSN 0885-6125. Citado 3 vezes nas páginas 21, 35 e 39.
- LEVESQUE, H. J. Knowledge representation and reasoning. *Annual review of computer science*, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 1, n. 1, 1986. Citado na página 20.
- MA, C.; MOLNÁR, B.; BENCZÚR, A. A semi-automatic semantic consistency-checking method for learning ontology from relational database. *Information (Basel)*, MDPI AG, Basel, v. 12, n. 5, 2021. ISSN 2078-2489. Citado na página 36.
- MAIER, F.; MA, Y.; HITZLER, P. Paraconsistent owl and related logics. *Semantic Web*, Ios Press, AMSTERDAM, v. 4, n. 4, 2013. ISSN 1570-0844. Citado na página 37.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, Springer, v. 5, 1943. Citado na página 17.
- MEHRI, R.; HAARSLEV, V.; CHINAEI, H. A machine learning approach for optimizing heuristic decision-making in web ontology language reasoners. *Computational intelligence*, Wiley Subscription Services, Inc, Hoboken, v. 37, n. 1, 2021. ISSN 0824-7935. Citado na página 35.
- MIKROYANNIDI, E. *et al.* Analysing syntactic regularities and irregularities in snomed-ct. *Journal of biomedical semantics*, Springer Nature, LONDON, v. 3, n. 1, 2012. ISSN 2041-1480. Citado na página 38.
- MITCHELL, T. *Machine learning*. McGraw hill Burr Ridge, 1997. Citado 2 vezes nas páginas 19 e 33.
- MONTEIRO, G.; SILVA, N. F. F.; DIAS, M. de S. Aplicando técnicas de aprendizado de máquina para detecção automática de bullying no twitter. 2017. Citado na página 39.
- NASCIMENTO, R.; BRASIL, M. Após alerta da interpol, menor suspeito de planejar ataque a escola no rio é apreendido. *g1 Rio e TV Globo*, Mar 2023. Disponível em: <<https://link.dev/YgjPz>>. Citado 2 vezes nas páginas 25 e 26.
- NEGNEVITSKY, M. *Artificial Intelligence: A Guide to Intelligent Systems*. Upper Saddle River, NJ: Pearson, 2005. ISBN ISBN-13: 978-0321204660. Citado 2 vezes nas páginas 18 e 19.
- NOY, N.; MCGUINNESS, D. Ontology development 101: A guide to creating your first ontology. *Knowledge Systems Laboratory*, v. 32, 01 2001. Citado na página 27.
- OZAKI, A. Learning description logic ontologies: Five approaches. where do they stand? *KI. Künstliche Intelligenz (Oldenbourg)*, Springer Berlin Heidelberg, Berlin/Heidelberg, v. 34, n. 3, 2020. ISSN 0933-1875. Citado na página 38.
- RADEMAKER, A.; HAEUSLER, E. H. Toward short and structural \mathcal{ALC} -reasoning explanations: A sequent calculus approach. In: ZAVERUCHA, G.; COSTA, A. L. da (Ed.). *Advances in Artificial Intelligence - SBIA 2008*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 167–176. ISBN 978-3-540-88190-2. Citado na página 44.

- RETTINGER, A. *et al.* Mining the semantic web: Statistical learning for next generation knowledge bases. *Data mining and knowledge discovery*, Springer US, Boston, v. 24, n. 3, 2012. ISSN 1384-5810. Citado na página 35.
- RIBEIRO, A. C.; BATISTA, A. J. A influência da mídia na criança/pré-adolescente e a educação como mediadora desse contato. *Alcar. Associação Brasileira de Pesquisadores de História da Mídia. I Encontro de História da Mídia da Região Norte. Universidade Federal do Tocantins*, 2010. Citado na página 25.
- RUSSELL, S. J.; NORVIG, P. *Inteligência artificial*. Netherlands-NL: Elsevier, 2013. Citado na página 33.
- SCHMIDT-SCHAUSS, M.; SMOLKA, G. Attributive concept descriptions with complements. *Artificial intelligence*, Elsevier, v. 48, n. 1, 1991. Citado 2 vezes nas páginas 30 e 31.
- SCREENSHOTS. *DL-Learner*, 2015. Disponível em: <<https://dl-learner.org/about/screenshots/>>. Citado na página 41.
- SHADBOLT, N.; BERNERS-LEE, T.; HALL, W. The semantic web revisited. *IEEE intelligent systems*, IEEE, v. 21, n. 3, 2006. Citado na página 32.
- SHIRKY, C. *Here comes everybody: The power of organizing without organizations*. London-UK: Penguin Group, 2008. Citado na página 23.
- SILVA, G. M. H. da *et al.* Dealing with the formal analysis of information security policies through ontologies: A case study. In: *Proceedings of the Third Australasian Workshop on Advances in Ontologies - Volume 85*. AUS: Australian Computer Society, Inc., 2007. (AOW '07), p. 55–60. ISBN 9781920682668. Citado 2 vezes nas páginas 44 e 45.
- SIRIN, E. *et al.* Pellet: A practical owl-dl reasoner. *Journal of Web Semantics*, v. 5, n. 2, 2007. ISSN 1570-8268. Software Engineering and the Semantic Web. Disponível em: <<https://ury1.com/17KQh>>. Citado na página 45.
- SOWA, J. F. *Knowledge representation: logical, philosophical and computational foundations*. California-US: Brooks/Cole Publishing Co., 1999. Citado na página 20.
- TURING, A. Computing machinery and intelligence. *Mind*, v. 59, n. 236, 1950. Citado na página 33.
- ZUBOFF, S. *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. New York - NY: PublicAffairs, 2019. Citado na página 23.