

Ontologies in Hearing Impairment: Structural Analysis and Perspectives on Semantic Interoperability

Ferreira Ana Paula Lopes de Abreu , Rigo Sandro José , and Maldaner Luis Felipe 

Abstract—Hearing impairment requires solutions that promote semantic interoperability and clinical data integration. In this context, ontologies play a fundamental role in the formal representation and standardization of biomedical knowledge, supporting data sharing across heterogeneous health information systems. This study maps, characterizes, compares, and analyzes ontologies related to hearing impairment, aiming to identify conceptual overlaps and potential points of interoperability. An exploratory-descriptive approach was adopted, following PRISMA guidelines to identify and evaluate ontologies associated with hearing health. Based on thematic relevance and conceptual scope, the Hearing Impairment Ontology (HIO) and the Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT) were selected for comparative analysis. Ontology alignment was performed using AgreementMakerLight (AML) with a similarity threshold of 0.5. The comparison identified 39 valid correspondences among 797 classes in SNOMED CT and 495 classes in HIO, indicating a low direct semantic overlap between the ontologies. Correspondences were concentrated in established clinical concepts related to hearing loss, while differences were observed in genetic, phenotypic, and therapeutic domains. The results suggest complementarity between the ontologies and reinforce the importance of semantic alignment strategies to support interoperability among clinical and biomedical data in hearing healthcare and computational audiology.

Link to graphical and video abstracts, and to code: <https://latam.ieceer9.org/index.php/transactions/article/view/10652>

Index Terms — Hearing Health; Interoperability; Ontology.

I. INTRODUÇÃO

A deficiência auditiva constitui um dos desafios mais relevantes da saúde no Brasil, com prevalência crescente à medida que aumenta a expectativa de vida. Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE) [1], cerca de 31 mil crianças entre dois e nove anos apresentam perda auditiva diagnosticada, enquanto, entre indivíduos com 60 anos ou mais, o número ultrapassa 1,5 milhão de casos, o que corresponde a 4,3% dessa população. Esses

The associate editor coordinating the review of this manuscript and approving it for publication was Luis Camarinha-Matos (*Corresponding author: Ana Paula Lopes De Abreu Ferreira*).

Ana Paula Lopes De Abreu Ferreira, S. J. Rigo, and L. F. Maldaner are with the Graduate Program in Applied Computing, Universidade do Vale do Rio dos Sinos, São Leopoldo, RS, Brazil (anaabreu@edu.unisinos.br, rigo@unisinos.br, and fmaldaner@unisinos.br).

indicadores reforçam não apenas a importância, mas também a urgência de estratégias que ampliem o acesso ao diagnóstico, aprimorem a precisão clínica e otimizem os fluxos operacionais dos centros auditivos brasileiros.

Apesar dessa prevalência, os dados audiológicos permanecem dispersos entre sistemas heterogêneos, o que limita o compartilhamento e a rastreabilidade das informações clínicas. Portanto, tal cenário evidencia a necessidade de estratégias que promovam a interoperabilidade semântica e a integração efetiva dos dados em ecossistemas digitais de saúde auditiva.

A transformação digital na área da saúde, impulsionada pelos avanços em inteligência artificial e interoperabilidade semântica [2], [3], delineia um caminho promissor para enfrentar os desafios relacionados à integração e ao uso inteligente de dados clínicos. Entre as abordagens emergentes, destacam-se os modelos de Geração Aumentada por Recuperação (*Retrieval-Augmented Generation* – RAG), que viabilizam a interação entre sistemas de processamento de linguagem natural e bases de conhecimento estruturadas, como ontologias biomédicas, diretrizes clínicas e registros eletrônicos de saúde [4]. Ao combinar a capacidade inferencial dos modelos de linguagem com a precisão semântica das ontologias, o RAG amplia o potencial dos sistemas de apoio à decisão clínica, permitindo consultas automatizadas a fontes externas, interpretação contextual das informações e recomendações terapêuticas fundamentadas [5].

Nesse contexto, as ontologias desempenham papel fundamental na consolidação da interoperabilidade semântica em saúde, ao possibilitar a representação formal e compartilhável do conhecimento biomédico, tornando-o inteligível tanto para humanos quanto para sistemas computacionais [6]. Sua utilidade reside na capacidade de padronizar conceitos, eliminar ambiguidades terminológicas e promover o reuso e a integração de dados entre diferentes plataformas e contextos assistenciais [3]. Assim, as ontologias constituem a infraestrutura semântica essencial que sustenta o avanço de soluções baseadas em inteligência artificial, garantindo consistência e confiabilidade às inferências clínicas automatizadas.

No domínio da saúde auditiva, a integração de dados clínicos provenientes de prontuários eletrônicos ainda constitui uma lacuna crítica para a consolidação de ecossistemas digitais interoperáveis. A ausência de modelos ontológicos padronizados e de infraestrutura semântica unificada limita o compartilhamento e a reprodutibilidade dos dados,

comprometendo a rastreabilidade diagnóstica. Estudos recentes destacam que o avanço da audiologia computacional tem potencial para transformar esse cenário, ao promover o uso de dados interoperáveis e inteligência artificial em processos de diagnóstico e acompanhamento clínico. A adoção de fluxos digitais integrados, associados à análise automatizada, pode ampliar a precisão e a eficiência desses serviços, sobretudo quando apoiada por estruturas de dados interoperáveis [2]. A interoperabilidade entre dados conectados, registros eletrônicos de saúde e plataformas de telessaúde representa um requisito central para a modernização da prática clínica [4]. Sendo assim, a padronização semântica e o uso de ontologias biomédicas interoperáveis constituem elementos estratégicos para viabilizar uma audiologia baseada em evidências e sustentada por dados [3],[7]. Dessa forma, observa-se uma convergência crescente entre as áreas de informática biomédica e saúde auditiva, indicando que a adoção de modelos ontológicos interoperáveis poderá eliminar barreiras semânticas, potencializando tanto a integração de dados clínicos quanto a personalização terapêutica.

Nesse cenário, a representação ontológica da deficiência auditiva torna-se um elemento fundamental para a interoperabilidade e o reuso de dados em escala nacional e internacional para o segmento da saúde auditiva. A criação e o uso de ontologias específicas, como a *Hearing Impairment Ontology* (HIO), e de terminologias clínicas amplas, como a *Systematized Nomenclature of Medicine – Clinical Terms* (SNOMED CT), viabilizam a padronização semântica necessária ao intercâmbio de informações entre os sistemas. No entanto, ainda são escassos os estudos que analisam as relações, sobreposições e lacunas conceituais entre essas ontologias.

Sendo assim, o presente estudo insere-se nesse contexto e tem como propósito mapear, caracterizar, comparar e analisar ontologias relacionadas à deficiência auditiva, com ênfase na análise de suas estruturas conceituais e potenciais pontos de interoperabilidade entre dados clínicos.

O estudo buscou levantar, de forma sistemática, as ontologias existentes que tratam de deficiência auditiva, bem como caracterizá-las e categorizá-las quanto ao escopo, à estrutura conceitual, à granularidade e aos padrões de representação adotados. Na sequência dos fatos, comparar e analisar criticamente as ontologias quanto a sobreposições, lacunas e potenciais pontos de interoperabilidade, destacando oportunidades para a futura integração semântica entre os dados. Por fim, mapear e analisar as ontologias relacionadas à deficiência auditiva.

A. Estrutura do Texto

Após a introdução, o trabalho estrutura-se de forma a conduzir o leitor da fundamentação metodológica à análise comparativa entre as ontologias estudadas. A Metodologia de Pesquisa descreve o delineamento exploratório-descritivo e o protocolo PRISMA adotado, detalhando as estratégias de busca, critérios de inclusão e exclusão e etapas de sistematização dos dados. Em seguida, a seção de Resultados da Pesquisa Exploratória apresenta o mapeamento das ontologias identificadas no domínio da saúde auditiva, justifica a seleção da *Hearing Impairment Ontology* (HIO) e da SNOMED CT para a etapa

experimental e registra o achado relevante da *Ontology of Deafness and Related Auditory Disorders* (ODRAH) e suas limitações de reuso. O texto prossegue com a descrição do Experimento de Comparação, em que se aplica a ferramenta *AgreementMakerLight* para alinhar classes ontológicas e avaliar correspondências segundo critérios de similaridade lexical. Na sequência, os Resultados e Discussão examinam o baixo percentual de sobreposição entre as ontologias e interpretam as divergências à luz dos princípios do realismo ontológico, do modelo SNAP-SPAN e das diretrizes OBO e FHIR-RDF, ressaltando a complementaridade entre abordagens terminológicas e biomédicas. Por fim, as Considerações Finais sintetizam as contribuições do estudo para a interoperabilidade semântica em saúde auditiva, reconhecendo limitações metodológicas e apontando direções para trabalhos futuros.

II. METODOLOGIA

Esta pesquisa adotou uma abordagem exploratória-descritiva, adequada ao objetivo de mapear e analisar ontologias relacionadas à deficiência auditiva e aos aparelhos auditivos. Esse delineamento metodológico permite ampliar o conhecimento sobre o estado da arte nesse domínio e descrever, de forma sistemática, as estruturas conceituais identificadas, bem como suas relações, sobreposições e lacunas.

Para assegurar rigor científico e reprodutibilidade, foi utilizado um protocolo de levantamento sistemático, contemplando definição de fontes e estratégias de busca, critérios de inclusão e exclusão. Portanto, a condução do estudo seguiu as diretrizes do protocolo PRISMA – *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* [8]. A utilização desse protocolo justifica-se por sua ampla aceitação internacional como referencial metodológico para revisões sistemáticas, o que busca assegurar transparência, reprodutibilidade e rigor científico em todas as etapas do processo investigativo.

A. Estratégia de Busca

A busca inicial foi conduzida na plataforma SciSpace, selecionada por sua curadoria de periódicos indexados e revisados por pares, o que assegura qualidade e relevância científica às fontes. Além disso, a plataforma disponibiliza mecanismos de ranqueamento bibliométrico e filtros temáticos que favorecem a recuperação de estudos relevantes em domínios interdisciplinares, como ontologias biomédicas e saúde digital. A escolha da plataforma também considerou sua capacidade de agregação semântica de conteúdos científicos provenientes de diferentes bases indexadoras, ampliando a cobertura temática da pesquisa.

Foram utilizados os descritores combinados ‘*ontology for health and hear loss*’, definidos de forma intencional para recuperar estudos relacionados simultaneamente ao uso de ontologias em saúde e ao domínio específico da deficiência auditiva, delimitando o escopo temático da investigação. Adicionalmente, aplicaram-se os seguintes filtros: ordenamento por relevância bibliométrica, período de publicação entre 2015 e 2025 e restrição a estudos classificados pela plataforma como de alta qualidade (*High*

Quality).

B. Critérios de Inclusão e Exclusão

Inicialmente, foram considerados para triagem os 70 estudos mais relevantes retornados pela busca, ordenados por relevância bibliométrica, priorizando-os com maior fator de impacto e reconhecimento científico. A aplicação dos critérios de exclusão visou assegurar a relevância temática e a qualidade metodológica dos registros analisados. Foram excluídos: (i) estudos duplicados; (ii) trabalhos sem relação direta com a aplicação de ontologias no domínio da saúde; e (iii) ontologias com escopo específico não relacionado à saúde auditiva, perda auditiva ou deficiência auditiva. Essa triagem inicial buscou garantir a representatividade e a consistência do conjunto de estudos elegíveis para as etapas subsequentes de análise.

A lista completa dos 70 registros inicialmente triados, incluindo título, periódico, ano, métricas bibliométricas e status de inclusão/exclusão, encontra-se depositada em repositório institucional, disponível mediante solicitação aos autores.

C. Procedimentos de Seleção dos Estudos

Essa etapa consistiu em uma conferência adicional para validar as ontologias encontradas, com o objetivo de verificar sua aplicabilidade no contexto da saúde auditiva.

Portanto, permitiu a identificação das ontologias com aderência temática ao estudo e o refinamento para a etapa de comparação entre elas.

TABELA I

DESCRIPTIVO METODOLÓGICO DE SELEÇÃO DOS ESTUDOS

Item	Descritivo
Pesquisa	Utilizaram-se os descritores ' <i>ontology for health and hear loss</i> ' como palavras-chave para a filtrar os estudos na plataforma SciSpace. Além disso, aplicou-se o filtro de ordenação relevância bibliométrica e ativou-se a opção <i>High Quality</i> .
Dados	Exportação dos dados da pesquisa em arquivo tipo xls.
Classificação	Após a ordenação inicial dos estudos com base na relevância bibliométrica, foram definidos critérios para inclusão e exclusão, os quais permitiram a categorização dos registros em dois status: incluído e excluído.
Pré-selecionados	Os estudos que não atendiam aos critérios de exclusão foram classificados como incluído, conforme os critérios de inclusão previamente estabelecidos.
Análise de conteúdo	Nesta etapa, todos os estudos classificados como incluído na etapa de pré-selecionados foram avaliados com base nos campos extraídos da exportação de dados da plataforma SciSpace, tais como <i>Insights</i> , <i>Abstract</i> e <i>Referenced Ontology</i> . A análise desses campos permitiu identificar os estudos que abordavam, de forma explícita, o conceito ligado diretamente ao termo de pesquisa, <i>ontology for health and hear loss</i> , o que possibilitou confirmar sua pertinência temática e a adequação à categoria de pré-selecionados.
Selecionados	Concluída a análise de conteúdo, os estudos que atenderam aos critérios temáticos e metodológicos foram formalmente definidos como estudos incluídos.

D. Estratégia de Análise e Síntese dos Dados

Após a seleção dos estudos elegíveis, procedeu-se à

sistematização dos dados com base em uma abordagem descritiva, conforme recomendam as diretrizes do protocolo PRISMA [8]. Com os dados organizados em planilha eletrônica, contendo as variáveis bibliométricas conforme disponibilizadas pela plataforma SciSpace.

A síntese descritiva consistiu na exposição estruturada das informações obtidas nos estudos incluídos, sem a realização de análises inferenciais ou de categorização interpretativa. Foram observadas, de forma sistemática, a presença das ontologias explicitamente ligadas ao termo de pesquisa utilizado '*ontology for health and hear loss*', bem como os tipos específicos de ontologias utilizadas.

E. Resultado da Pesquisa Exploratória

A análise sistemática das informações organizadas na tabela eletrônica possibilitou identificar e classificar ontologias relevantes ao domínio da saúde auditiva, conforme disponibilizado na Quadro II. Com base nesse mapeamento, foram selecionadas, para a etapa experimental, as ontologias que apresentaram maior alinhamento ao objetivo da pesquisa, considerando o escopo conceitual, a disponibilidade pública e a relevância temática. Assim, foram eleitas a *Hearing Impairment Ontology* (HIO) e a *Systematized Nomenclature of Medicine – Clinical Terms* (SNOMED CT), por abordarem, sob perspectivas complementares, o domínio da deficiência auditiva, a primeira sob enfoque biomédico e a segunda sob enfoque clínico e terminológico.

Durante a etapa de pesquisa exploratória, também foi possível identificar a *Ontology of Deafness and Related Auditory Disorders* (ODRAH) como um achado relevante. A ODRAH constitui um modelo ontológico voltado à integração do conhecimento clínico e genético sobre a surdez e distúrbios auditivos relacionados, propondo uma estrutura abrangente capaz de representar desde mecanismos moleculares e modos de herança até fenótipos auditivos e estratégias terapêuticas [9]. Sua arquitetura baseia-se em princípios herdados da *Sickle Cell Disease Ontology* (SCDO), adotando classes superiores como *Diagnostics*, *Mode of Inheritance*, *Phenotype* e *Therapeutics*, o que demonstra a continuidade metodológica da equipe de Mazandu e Wonkam na construção de ontologias biomédicas [9].

Diante da análise comparativa das ontologias biomédicas de referência HIO e ODRAH, evidenciam-se distintos níveis de granularidade, cobertura e maturidade ontológica, refletindo abordagens complementares de modelagem do conhecimento em saúde. A maturidade de uma ontologia biomédica depende de sua capacidade de representar com precisão entidades do domínio, aderir a princípios de interoperabilidade e integrar-se a *frameworks*, como o *OBO Foundry* [10], [11]. A HIO apresenta granularidade média-alta, orientada à representação biomédica e clínica da deficiência auditiva, porém, a ODRAH demonstra cobertura ampliada em mecanismos genéticos e moleculares, contudo, menor maturidade computacional devido à limitação de sua disponibilização formal em OWL. Portanto, essa comparação reforça que a integração entre ontologias biomédicas exige a articulação de modelos de diferentes níveis para garantir coerência semântica e interoperabilidade integral em ecossistemas de saúde digital.

TABELA II
NÚMERO DE OCORRÊNCIAS DA ONTOLOGIA COMO PRINCIPAL

Ontologia	N de vezes ^a	Domínio principal
SNOMED CT	4	Clínica geral e terminologia médica padronizada
MHMO	2	Saúde mental e comportamento
SMASH	2	Internet das Coisas (IoT) em saúde
AHSO	1	Monitoramento de estado de saúde
<i>Building an Ontology-Based Electronic Health Record System</i>	1	Registro eletrônico de saúde (EHR)
<i>Case Profile Ontology, mHealth Ontology</i>	1	Gestão de casos clínicos
<i>CHEAR Ontology</i>	1	Exposição ambiental e saúde infantil
ChEBI	1	Bioquímica e substâncias químicas
FOAF	1	Identidade e redes sociais semânticas
FOAR	1	Nutrição e agricultura
<i>Granular Ontology Model</i>	1	Modelagem ontológica hierárquica
HCMM	1	Monitoramento clínico e sensores
HDDO	1	Dicionários de dados em saúde
HIO	1	Deficiência auditiva
HPO	1	Fenótipos humanos
HSM-Onto	1	Monitoramento de sistemas de saúde
ICF	1	Funcionalidade e incapacidade humana
MIMO	1	Imagem médica e radiologia
OBI	1	Experimentos e protocolos biomédicos
ODRAH	1	Saúde auditiva e surdez
OFFF	1	Aspectos sociais em saúde
OGMS	1	Estrutura conceitual médica
OMRSE	1	Entidades sociais e legais
OntoDrug	1	Farmacologia
OntoGSM	1	Gestão de serviços médicos
ONTOPROG	1	Processos ontológicos genéricos
OSH-Core	1	Saúde ocupacional
PACO	1	Cuidados de pacientes
PhenX	1	Fenótipos e exposições clínicas
PHM	1	Autogerenciamento da saúde
POF	1	Processos clínicos e administrativos
<i>Public Health Ontology</i>	1	Saúde pública
SAREF4health	1	Internet das Coisas em saúde
SHM-Ontology	1	Smart Health Monitoring
SKOS	1	Organização de vocabulários
SNIK	1	Gestão hospitalar e sistemas de informação
TEO	1	Temporalidade em processos clínicos
UTPO	1	Telemedicina
Wuxing	1	Saúde tradicional chinesa
Não evidenciou uma ontologia principal	35	-

(a) Número de vezes que apareceu como principal na pesquisa.

Embora apresente escopo temático mais amplo e detalhamento genético superior ao da HIO, a ODRAH ainda se encontra em fase de consolidação e maturação, com documentação parcial e disponibilidade limitada em formato OWL para reuso computacional. Tais restrições metodológicas inviabilizam sua aplicação direta em experimentos de alinhamento automático e dificultam a reprodutibilidade de testes de interoperabilidade

semântica. Além disso, o enfoque da ODRAH é predominantemente biomédico e molecular, o que a distância do objetivo desta pesquisa, centrado na integração de dados clínicos e técnicos (provenientes de sistemas audiológicos e dispositivos assistivos).

Dessa forma, a escolha da HIO como ontologia representativa para comparação com a SNOMED CT fundamenta-se em critérios de maturidade ontológica, de acessibilidade e de aderência ao domínio aplicado da saúde auditiva. A HIO apresenta documentação consolidada, compatibilidade com o padrão OBO *Foundry*, interoperabilidade comprovada com ontologias biomédicas amplamente utilizadas, como por exemplo, *Orphanet* e *Disease Ontology* (HPO) e ampla disponibilidade pública no NCBO BioPortal e em repositórios GitHub [12]. Assim, a HIO foi considerada a ontologia mais adequada para a análise comparativa com a SNOMED CT, assegurando rigor metodológico, reprodutibilidade experimental e coerência conceitual, com o propósito de investigar a interoperabilidade semântica entre dados clínicos e técnicos na saúde auditiva.

III. CARACTERIZAÇÃO DAS ONTOLOGIAS SELECIONADAS

A *Hearing Impairment Ontology* (HIO) constitui um modelo ontológico abrangente e formalmente estruturado dedicado à representação do conhecimento sobre deficiência auditiva, integrando dimensões genéticas, clínicas e terapêuticas do fenômeno [12]. Desenvolvida por um consórcio internacional liderado pela Universidade da Cidade do Cabo (UCT), a HIO foi derivada metodologicamente da *Sickle Cell Disease Ontology* (SCDO), preservando sua arquitetura lógica, mas expandindo o escopo para contemplar a complexidade etiológica e fenotípica da perda auditiva [12].

Referente ao escopo da HIO, ele abrange fatores genéticos, ambientais e sociais que influenciam o desenvolvimento e a progressão da deficiência auditiva, descrevendo entidades biomédicas, mecanismos fisiopatológicos, atributos pessoais e abordagens terapêuticas. Sua estrutura conceitual organiza-se em torno do conceito central de *Hearing Impairment*, desdobrando-se em dez classes superiores: *Diagnostics*, *Gene Product*, *Genetic Phenomena*, *Mode of Inheritance*, *Personal Attribute*, *Phenotype*, *Therapeutics*, *Disease Attribute*, *Modifier* e *Environmental Factor* [12]. Essas classes interconectam-se por relações hierárquicas e semânticas do tipo ‘*is_a* e *part_of*’, totalizando 495 termos e 543 relações formais.

Quanto à granularidade, a HIO permite descrever desde entidades abstratas, como ‘modo de herança’, até condições clínicas altamente específicas, como *Cisplatin-Induced Hearing Impairment* e *Prelingual Hearing Impairment*, oferecendo múltiplos níveis de detalhamento conceitual e temporal. No que se refere aos padrões de representação, a ontologia foi desenvolvida em OWL (*Ontology Web Language*) e disponibilizada em formato OBO (*Open Biomedical Ontologies*), assegurando compatibilidade com ferramentas de edição como Protégé e WebProtégé, além de integração com o NCBO BioPortal e repositórios do GitHub [12].

A adoção da HIO favorece a interoperabilidade semântica e a padronização terminológica entre bancos de dados biomédicos e sistemas de apoio à decisão, além de possibilitar o mapeamento cruzado com outras ontologias de referência, como HPO (*Human Phenotype Ontology*), *Orphanet*, *Disease Ontology* (DO) e MeSH, consolidando-se como um recurso estratégico para a integração de dados clínicos e genômicos na saúde auditiva [12].

A *Systematized Nomenclature of Medicine – Clinical Terms* (SNOMED CT) constitui uma terminologia clínica estruturada e formalmente organizada, destinada à representação abrangente e computacional do conhecimento médico. Reconhecida como a ontologia biomédica de maior cobertura semântica na área da saúde, ela foi desenvolvida originalmente pelo *College of American Pathologists* (CAP) e, atualmente, é mantida pela *SNOMED International*, resultante da fusão entre as nomenclaturas SNOMED RT e *Clinical Terms Version 3* (CTV3). Seu objetivo é padronizar a representação de informações clínicas e viabilizar a interoperabilidade semântica entre sistemas heterogêneos de saúde, consolidando-se como uma referência global nesse domínio [13],[14].

O escopo da SNOMED CT abrange, de forma transversal, todo o domínio biomédico, permitindo representar entidades anatômicas, processos fisiopatológicos, achados clínicos, procedimentos terapêuticos, produtos farmacológicos, substâncias químicas, organismos e fatores ambientais. Sua estrutura conceitual é hierárquica e poliaxial, organizada em dezenove hierarquias principais (*top-level hierarchies*), incluindo *Clinical finding*, *Procedure*, *Body structure*, *Observable entity*, *Pharmaceutical/biologic product* e *Substance*, o que favorece a descrição semântica detalhada de eventos clínicos e relações ontológicas complexas [15].

A versão internacional de setembro de 2024 compreende aproximadamente 365 mil conceitos únicos, 1,6 milhão de relações ontológicas e mais de 1,1 milhão de descritores linguísticos, incluindo termos preferidos e sinônimos [16]. Cada conceito é identificado por um código numérico estável e participa de múltiplas relações semânticas, como *is-a*, *part-of*, *associated with* e *finding site*, permitindo inferência automática e verificação de consistência lógica. A granularidade conceitual da SNOMED CT é ampla, possibilitando a descrição de entidades abstratas, como *Hearing disorder (Disorder)* (15188001), e condições altamente específicas, como *Noise-induced hearing loss (Disorder)* (1091891000119106) e *Sensorineural hearing loss of left ear (Disorder)* (47338001), o que assegura múltiplos níveis de detalhamento semântico e anatômico relevantes para domínios especializados, como a saúde auditiva.

No que se refere aos padrões de representação, a SNOMED CT é formalizada em OWL 2 (*Description Logic EL++*), o que garante consistência semântica e interoperabilidade computacional com outros modelos formais. Embora alguns estudos proponham o alinhamento conceitual da SNOMED CT a ontologias de topo, como a *Basic Formal Ontology* (BFO) e a *Ontology for General Medical Science* (OGMS), a literatura especializada indica que esse alinhamento não é formal nem integral. Pesquisas conduzidas demonstram que apenas parte da hierarquia superior pode ser mapeada aos tipos ontológicos da BFO, devido a diferenças estruturais e semânticas, não havendo integração oficial consolidada [10], [17]. Assim, a SNOMED

CT é caracterizada como uma ontologia terminológica independente, de base clínica, que utiliza lógica descritiva para estruturar conceitos e relações, sem depender de *frameworks* superiores.

A ampla adoção da SNOMED CT em sistemas de registro eletrônico de saúde, projetos de interoperabilidade e plataformas de apoio à decisão evidencia sua relevância prática e científica. Revisões sistemáticas recentes, confirmam a maturidade crescente do modelo e sua aplicação em mineração de dados clínicos, aprendizado de máquina e integração semântica de registros, consolidando-o como um núcleo semântico de padronização terminológica amplamente utilizado em mais de oitenta países [15].

IV. DESCRIÇÃO DO EXPERIMENTO DE COMPARAÇÃO

A etapa experimental deste estudo teve como propósito principal comparar e analisar, semanticamente, as ontologias escolhidas na etapa anterior, com foco em identificar convergências conceituais, sobreposições e lacunas estruturais entre modelos ontológicos distintos.

Logo, a escolha dessas ontologias foi fundamentada em critérios de representatividade e escopo: a HIO apresenta granularidade elevada e especialização conceitual na descrição de entidades, sintomas e classificações associadas à deficiência auditiva; por sua vez, a SNOMED CT, ainda que de natureza mais ampla, constitui o vocabulário clínico mais difundido internacionalmente, abrangendo terminologias interoperáveis utilizadas em sistemas de registros eletrônicos de saúde em mais de 80 países.

A comparação entre ambas foi realizada por meio do *AgreementMakerLight* (AML), versão 3.2, uma ferramenta amplamente reconhecida na literatura científica como uma das mais eficientes para alinhamento semântico automatizado (*ontology matching*).

O *AgreementMakerLight* (AML) é uma ferramenta de *ontology matching* desenvolvida para promover interoperabilidade semântica entre ontologias de grande porte, com destaque para o domínio biomédico. O sistema combina eficiência computacional, precisão nos alinhamentos e mecanismos de automação que reduzem a intervenção humana no processo de correspondência semântica [18].

Sua arquitetura é modular e escalável, composta por estruturas específicas, como o *RelationshipMap*, que armazena relações semânticas; o *Lexicon*, que organiza rótulos e sinônimos ponderados pela confiabilidade; e o *Alignment*, que consolida os mapeamentos resultantes. Essas camadas permitem que o AML mantenha um desempenho elevado mesmo em ontologias com milhões de axiomas [18].

Um dos diferenciais do AML é a seleção automática de conhecimento de fundo (*background knowledge*), fundamentada no indicador *mapping gain*, que avalia o ganho obtido ao integrar ontologias auxiliares, como MeSH e DOID. Essa técnica demonstrou aumento significativo na precisão dos alinhamentos e na métrica F-measure em contextos biomédicos [19]. Além disso, o sistema incorpora métodos de reparo de alinhamentos incoerentes, baseados em modularização e heurísticas de confiança, garantindo coerência lógica sem perda de cobertura [20]. Durante a

execução, o AML gerou relatórios automáticos de alinhamento, contendo o número total de correspondências (*mappings*) identificadas, os valores de confiança (*confidence values*) atribuídos a cada par de classes correspondentes e a matriz de similaridade associada, conforme mostram as figuras 1 e 2.

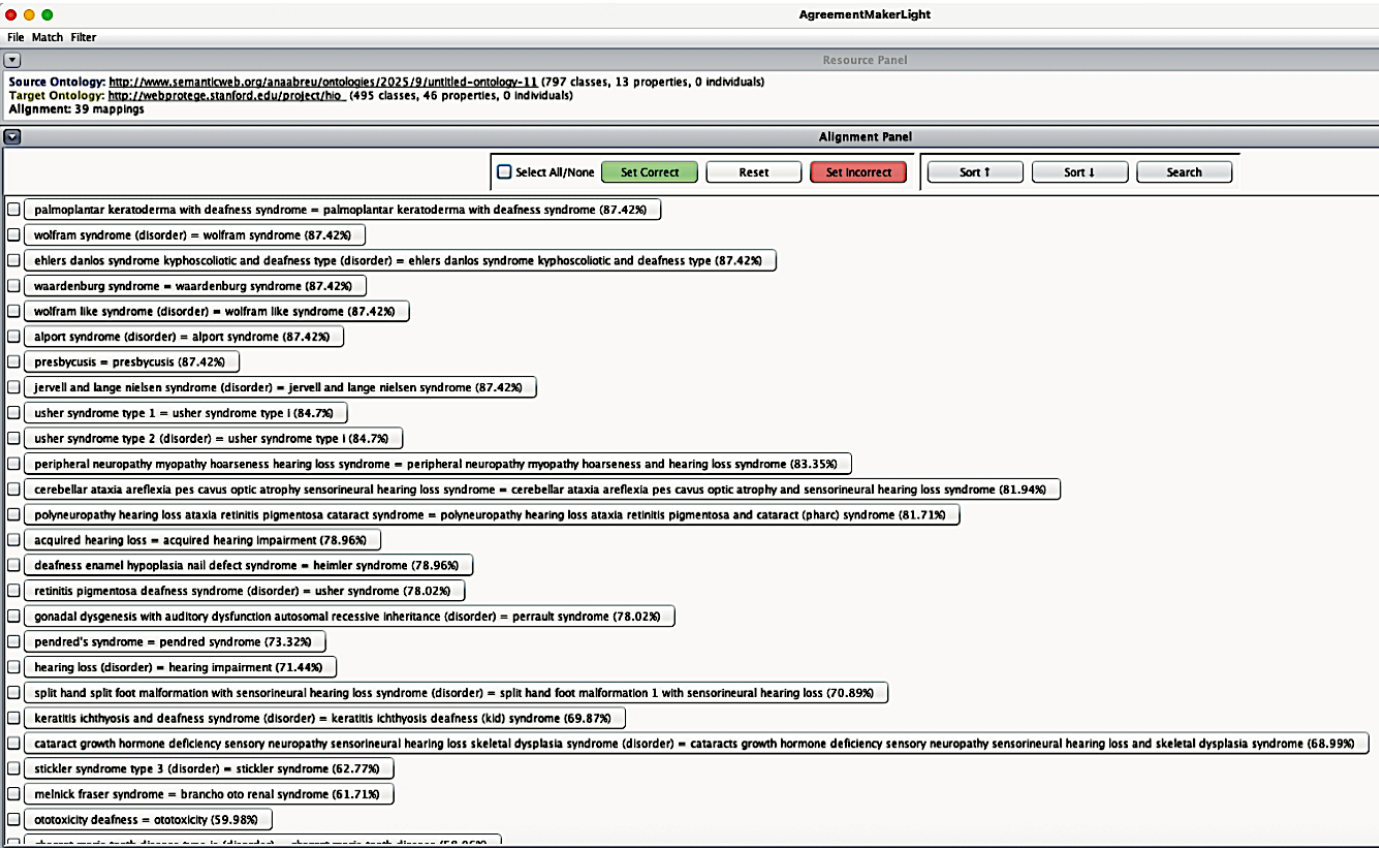


Fig. 1. Interface de comparação entre ontologias utilizando AML.

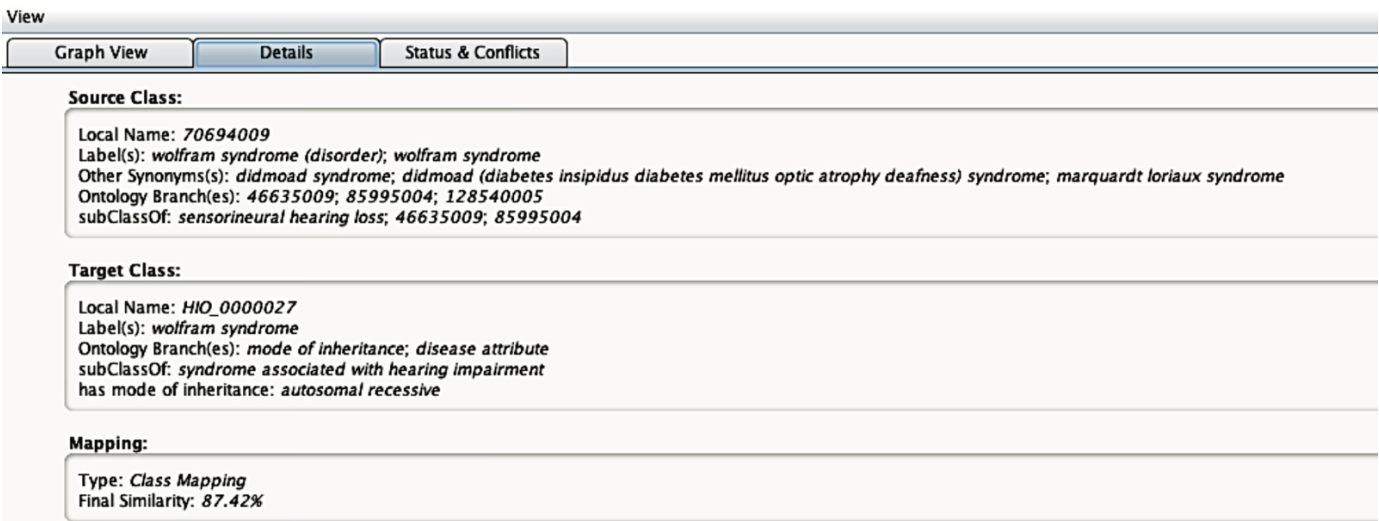


Fig. 2. Interface de detalhamento da comparação entre ontologias utilizando AML.

Avaliações realizadas com o AML indicam desempenho consistente em diversos cenários de alinhamento, especialmente nas trilhas biomédicas da *Ontology Alignment Evaluation Initiative* (OAEI). Estudos apontam que o uso de estratégias lexicais ponderadas e mediação semântica aprimora a robustez dos resultados [21]. Dessa forma, o AML se consolida como uma das ferramentas mais maduras e

rigorosamente validadas para integração semântica no campo da saúde digital.

V. RESULTADOS

O experimento realizado com o *AgreementMakerLight* (AML) identificou 39 correspondências válidas entre as 797

classes da SNOMED CT e as 495 classes da HIO, com um limiar de 0,5. Essa quantidade representa 4,9% do total de classes analisadas na SNOMED CT e 7,9% do conjunto da HIO, o que sugere uma baixa intersecção semântica direta entre as ontologias. As correspondências concentram-se em conceitos fundamentais de perda auditiva, como *Hearing loss*, *Sensorineural hearing loss*, *Conductive hearing loss* e *Mixed hearing loss*, indicando alinhamentos em categorias clínicas consolidadas, mas divergências em níveis genéticos, fenotípicos e terapêuticos.

TABELA III
SÍNTESE DOS RESULTADOS

Ontologia	Classes	Threshold	Matches	% De classes mapeadas	Domínio predominante
SNOMED CT	797	0,5	39	4,9 %	Diagnóstico clínico e procedimentos
HIO	495	0,5	39	7,9 %	Fenótipos, genótipos e terapias auditivas

Contudo, essa proporção reduzida de correspondências não deve ser interpretada como ausência de compatibilidade, mas sim como evidência da divergência epistemológica e ontológica entre os modelos comparados. A SNOMED CT estrutura-se sob um paradigma terminológico e clínico, voltado à codificação e interoperabilidade de registros de saúde em larga escala, enquanto a HIO adota uma abordagem ontológica e biomédica, direcionada à representação formal dos mecanismos genéticos, fenotípicos e terapêuticos da deficiência auditiva.

Portanto, o alinhamento identificado entre a HIO e a SNOMED CT, predominantemente, lexical e estrutural, baseado em similaridades terminológicas e relações hierárquicas entre classes correspondentes [22]. As correspondências mais fortes ocorreram em conceitos clínicos consolidados, sugerindo convergência semântica em categorias diagnósticas de alto nível. A análise foi conduzida com o *AgreementMakerLight* (AML), cujas métricas de similaridade lexical e o módulo de distância de Levenshtein permitem identificar correspondências mesmo em casos de variações morfológicas [21].

Por outro lado, os domínios genético e terapêutico da HIO não encontram equivalentes diretos na SNOMED CT, reforçando sua complementaridade epistemológica e a necessidade de mapeamentos híbridos, lexical e conceitual, para alcançar interoperabilidade plena.

VI. DISCUSSÃO

A análise comparativa entre a *Hearing Impairment Ontology* (HIO) e a *Systematized Nomenclature of Medicine – Clinical Terms* (SNOMED CT) revelou uma baixa sobreposição semântica direta, com 4,9% das classes da SNOMED CT e 7,9% das classes da HIO apresentando correspondências válidas. Essa discrepância não deve ser interpretada como limitação do método, mas como evidência de uma divergência

epistemológica entre modelos ontológicos fundamentados em paradigmas distintos.

A SNOMED CT é essencialmente terminológica e operacional, concebida para padronizar registros clínicos e viabilizar interoperabilidade entre sistemas de saúde heterogêneos. Sua estrutura é predominantemente extensional, voltada à descrição e codificação de entidades clínicas e procedimentos [13], [15]. Em contraste, a HIO adota uma perspectiva intencional e biomédica, voltada à representação formal das causas genéticas, fenótipos e terapias auditivas, fundamentando-se nos princípios da ontologia biomédica [11]. Assim, a baixa sobreposição semântica reflete diferenças de granularidade, propósito e nível ontológico, e não ausência de compatibilidade conceitual.

Sob a ótica epistemológica, essas diferenças sugerem que ontologias de distintos níveis não competem, mas se complementam na estruturação do conhecimento biomédico [23]. A SNOMED CT opera em níveis clínico e administrativo, enquanto a HIO descreve dimensões biomédicas e moleculares. Essa assimetria reforça a necessidade de mecanismos de mediação semântica, capazes de articular conceitos oriundos de camadas ontológicas heterogêneas.

A aplicação do *AgreementMakerLight* (AML) confirmou essa distinção ao identificar alinhamentos robustos apenas entre conceitos clínicos consolidados, como *hearing loss*, *sensorineural hearing loss* e *conductive hearing loss*, com menor desempenho em domínios genéticos e terapêuticos. Tal comportamento é coerente com as limitações conhecidas de métodos puramente lexicais de *ontology matching* em cenários biomédicos [20].

Diante disso, o avanço da interoperabilidade entre modelos biomédicos e clínicos requer a adoção de *frameworks* ontológicos superiores, como o OBO *Foundry*, que estabelece princípios de interoperabilidade, modularidade e reuso [24]. Complementarmente, o uso de FHIR RDF *mappings*, que integram o padrão HL7 FHIR ao *Resource Description Framework* (RDF), mostra-se promissor para a representação semântica unificada de dados clínicos [25].

Em uma perspectiva epistemológica, esta pesquisa ancora-se nos princípios do realismo ontológico, que parte da premissa de que as entidades representadas nas ontologias biomédicas correspondem a aspectos efetivamente existentes na realidade, e não apenas a construções linguísticas ou abstrações computacionais [11]. Nesse sentido, a HIO adota uma postura realista ao buscar representar os fenômenos biomédicos associados à deficiência auditiva, descrevendo entidades como fenótipos, mecanismos genéticos e processos terapêuticos enquanto instâncias de classes ontológicas fundamentadas na realidade biológica. Essa aderência epistemológica pode ser compreendida à luz do modelo SNAP-SPAN (continuant-ocorrentes) que distingue entidades persistentes no tempo (SNAP, ou continuantes) de entidades processuais ou que se desenvolvem no tempo [22]. Na HIO, essa distinção manifesta-se na representação simultânea de elementos estáticos, como estruturas anatômicas e predisposições genéticas, e de processos dinâmicos, como a evolução clínica da perda auditiva e as respostas terapêuticas. Assim, o modelo SNAP-SPAN fornece a estrutura conceitual subjacente que confere coerência ao tratamento das entidades na HIO, garantindo compatibilidade com o realismo ontológico defendido pelo

OBO *Foundry* e assegurando consistência semântica entre as dimensões biomédicas e clínicas representadas na ontologia.

VII. CONSIDERAÇÕES FINAIS

A análise comparativa entre a *Hearing Impairment Ontology* (HIO) e a SNOMED CT revela que a interoperabilidade semântica entre ontologias biomédicas e terminológicas permanece um desafio relevante no domínio da saúde auditiva. Os resultados mostraram baixa sobreposição direta, mas revelaram complementaridade estrutural e semântica que pode ser explorada em futuras iniciativas de alinhamento conceitual.

Portanto, compreender as diferenças entre a HIO e a SNOMED CT é fundamental para identificar os limites e as potencialidades da integração semântica entre níveis distintos de representação do conhecimento. Enquanto a SNOMED CT organiza o domínio clínico em uma estrutura terminológica extensiva voltada ao registro e codificação de achados [2], [10], a HIO estrutura o domínio biomédico da deficiência auditiva de forma intencional, modelando relações causais, genéticas e terapêuticas [12]. Reconhecer essa distinção permite compreender que as ontologias não são concorrentes, mas complementares: a SNOMED CT fornece a linguagem clínica interoperável necessária aos sistemas de registro e prescrição, enquanto a HIO oferece a base conceitual que sustenta o raciocínio biomédico e a inferência computacional.

Então, conhecer suas diferenças é o ponto de partida para delinear estratégias de alinhamento ontológico capazes de articular o conhecimento clínico e biomédico em ecossistemas de saúde digital.

Nesse contexto, o estudo contribui para a compreensão das lacunas ontológicas e reforça a importância de estratégias de harmonização semântica baseadas em *frameworks* de referência, como o OBO *Foundry* e o FHIR RDF, que promovem padrões abertos e interoperáveis de representação do conhecimento biomédico.

A investigação realizada permitiu responder à questão central desta pesquisa ao demonstrar que a integração semântica entre ontologias de saúde contribui para a melhoria da interoperabilidade dos dados clínicos, possibilitando o alinhamento conceitual entre modelos terminológicos e biomédicos. A análise comparativa entre a HIO e a SNOMED CT mostrou que, embora a sobreposição conceitual direta seja limitada, há complementaridade suficiente para sustentar mapeamentos semânticos consistentes. Essa integração favorece o desenvolvimento de sistemas de suporte à decisão clínica mais robustos e a intercambialidade segura de dados entre instituições, representando um avanço teórico e metodológico para a consolidação da audiologia computacional orientada por ontologias.

Entre as limitações desta pesquisa, destacam-se o foco em apenas duas ontologias e a adoção de um processo automatizado de correspondência lexical, o que restringe a generalização dos resultados e não abrange relações semânticas de natureza ontológica mais profunda. Pesquisas futuras poderão ampliar o escopo empírico por meio da inclusão de outras ontologias do domínio biomédico, da utilização de abordagens híbridas de correspondência (lexical e estrutural) e

da incorporação de curadoria manual como etapa de validação semântica.

As contribuições deste estudo manifestam-se de forma articulada ao longo de suas etapas. No levantamento sistemático, a pesquisa exploratória proporcionou um mapeamento das ontologias aplicadas à deficiência auditiva, consolidando evidências sobre o estado da arte e identificando, além da *Hearing Impairment Ontology* (HIO) e da *Systematized Nomenclature of Medicine – Clinical Terms* (SNOMED CT), o achado relevante da *Ontology of Deafness and Related Auditory Disorders* (ODRAH), que evidenciou lacunas de interoperabilidade e oportunidades de reuso conceitual [8]. Na etapa experimental, a aplicação do *AgreementMakerLight* permitiu a comparação objetiva entre as ontologias selecionadas, resultando na identificação de 39 correspondências válidas entre 797 classes da SNOMED CT e 495 da HIO, o que demonstrou uma complementaridade semântica superior à sobreposição direta [20]. Essa constatação permitiu reinterpretar as diferenças ontológicas não como limitações, mas como reflexos de propósitos distintos, terminológicos e clínicos, no caso da SNOMED CT [9], [14], e biomédicos e intencionais, na HIO [2], [12].

A análise, conduzida sob a perspectiva do realismo ontológico [11] e do modelo SNAP-SPAN [23], contribuiu para compreender a coexistência de diferentes níveis de granularidade e intencionalidade, reforçando que as ontologias comparadas são complementares, e não concorrentes. Em síntese, o estudo contribui com o estado da arte ao demonstrar empiricamente a viabilidade de integração entre modelos biomédicos e clínicos, oferecendo evidências e diretrizes conceituais para futuras estratégias de alinhamento baseadas em *frameworks* de referência, como o *Open Biological and Biomedical Ontology Foundry* (OBO Foundry) e o FHIR-RDF, e fortalecendo as bases para a construção de ecossistemas interoperáveis de saúde auditiva digital.

Dessa forma, conclui-se que a integração semântica entre ontologias da saúde auditiva e sistemas clínicos representa uma via promissora para fortalecer a interoperabilidade de dados e sustentar o desenvolvimento de sistemas inteligentes de apoio à decisão. Ao evidenciar empiricamente lacunas e convergências entre a HIO e a SNOMED CT, este trabalho oferece subsídios teóricos e metodológicos para a evolução de práticas e tecnologias na audiologia computacional e para o avanço da interoperabilidade semântica em saúde.

REFERENCES

- [1] Instituto Brasileiro de Geografia e Estatística (IBGE), “Pessoas com deficiência auditiva, por sexo e situação do domicílio,” 2019. [Online]. Available: <https://sidra.ibge.gov.br/tabela/8217>
- [2] J. W. A. Wasmann, C. P. Lanting, W. J. Huinck, E. A. M. Mylanus, J. W. M. van der Laak, P. J. Govaerts, D. W. Swanepoel, D. R. Moore and D. L. Barbour, “Computational Audiology: New Approaches to Advance Hearing Health Care in the Digital Age,” *Ear Hear.*, vol. 42, no. 6, pp. 1499–1507, 2021, doi: 10.1097/AUD.0000000000001041.
- [3] C. Bouton et al., “Ontologies in biomedical sciences and healthcare: a review,” *Yearb. Med. Inform.*, vol. 30, no. 1, pp. 237–247, 2021, doi: 10.1055/s-0041-1726481.

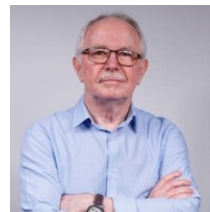
- [4] D. W. Swanepoel and J. L. Clark, "Hearing healthcare in remote or resource-constrained environments," *J. Am. Acad. Audiol.*, vol. 30, no. 7, pp. 578–584, 2019, doi: 10.3766/jaaa.17064.
- [5] M. Lewis, W.-T. Yih, T. Rocktäschel, S. Riedel, and D. Kiela, "Retrieval-augmented generation for knowledge-intensive NLP tasks," in *Proc. 34th Conf. Neural Inf. Process. Syst. (NeurIPS)*, Vancouver, Canada, 2020. [Online]. Available: <https://proceedings.neurips.cc/paper/2020/file/6b493230205f780e1bc26945df7481e5-Paper.pdf>
- [6] W. Ceusters, A. Michelotti, K. G. Raphael, J. Durham, and R. Ohrbach, "Perspectives on next steps in classification of orofacial pain – part 1: role of ontology," *J. Oral Rehabil.*, vol. 42, no. 12, pp. 926–940, Dec. 2015, doi: 10.1111/joor.12340.
- [7] D. R. Moore, "Computational audiology: Big data and machine learning in hearing health care," *Front. Digit. Health*, vol. 4, Art. no. 841735, 2022, doi: 10.3389/fdgth.2022.841735.
- [8] D. Moher et al., "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement," *PLoS Med.*, vol. 6, no. 7, Art. no. e1000097, 2009, doi: 10.1371/journal.pmed.1000097.
- [9] G. K. Mazandu, N. Manyisa, S. M. Adadey, and A. Wonkam, "Ontology of Deafness and Related Auditory Disorders (ODRAH): Integrating clinical and genetic knowledge of hearing loss," *Genes*, vol. 11, Art. no. 1456, 2020, doi: 10.3390/genes11121456.
- [10] W. Ceusters, "Axiomatizing SNOMED CT disorders: Should there be a 'Common Disease Ontology'?" *Appl. Ontol.*, vol. 18, no. 2, pp. 157–174, 2023, doi: 10.3233/AO-230018.
- [11] B. Smith and W. Ceusters, "Ontology as the core discipline of biomedical informatics," *J. Biomed. Inform.*, vol. 43, no. 5, pp. 786–791, 2010, doi: 10.1016/j.jbi.2010.07.003.
- [12] J. Hotchkiss et al., "The Hearing Impairment Ontology: A tool for unifying hearing impairment knowledge to enhance collaborative research," *Genes*, vol. 10, Art. no. 960, 2019, doi: 10.3390/genes10120960.
- [13] R. Cornet and N. de Keizer, "Forty years of SNOMED: A literature review," *BMC Med. Inform. Decis. Mak.*, vol. 8, suppl. 1, Art. no. S2, 2008, doi: 10.1186/1472-6947-8-S1-S2.
- [14] D. Lee, R. Cornet, and N. de Keizer, "Literature review of SNOMED CT use," *J. Am. Med. Inform. Assoc.*, vol. 21, suppl. 1, pp. e11–e19, 2014, doi: 10.1136/amiajnl-2013-001636.
- [15] E. Chang and J. Mostafa, "The use of SNOMED CT, 2013–2020: A literature review," *J. Am. Med. Inform. Assoc.*, vol. 28, no. 9, pp. 2017–2027, 2021, doi: 10.1093/jamia/ocab084.
- [16] SNOMED International, SNOMED CT International Release 2024-09, London, U.K.: SNOMED International, 2024. [Online]. Available: <https://www.snomed.org/snomed-ct/releases/international-release>.
- [17] W. R. Hogan, "Aligning the top level of SNOMED CT with Basic Formal Ontology," *Nature Precedings*, 2008, doi: 10.1038/npre.2008.2373.1. [Online]. Available: <https://www.nature.com/articles/npre.2008.2373.1>
- [18] D. Faria et al., "AgreementMakerLight," *Semantic Web, IOS Press*, 2025, doi: 10.3233/SW-233304.
- [19] D. Faria et al., "Automatic background knowledge selection for matching biomedical ontologies," *PLoS One*, vol. 9, no. 11, Art. no. e111226, 2014, doi: 10.1371/journal.pone.0111226.
- [20] E. Santos et al., "Ontology alignment repair through modularization and confidence-based heuristics," *PLoS One*, vol. 10, no. 12, Art. no. e0144807, 2015, doi: 10.1371/journal.pone.0144807.
- [21] D. Faria et al., "Tackling the challenges of matching biomedical ontologies," *J. Biomed. Semantics*, vol. 9, Art. no. 1, 2018, doi: 10.1186/s13326-017-0170-9.
- [22] P. Shvaiko and J. Euzenat, "Ontology matching: State of the art and future challenges," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 1, pp. 158–176, 2013, doi: 10.1109/TKDE.2011.253.
- [23] P. Grenon and B. Smith, "SNAP and SPAN: Towards dynamic spatial ontology," *Spatial Cogn. Comput.*, vol. 4, no. 1, pp. 69–104, 2004, doi: 10.1207/s15427633scc0401_5.
- [24] B. Smith et al., "The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration," *Nat. Biotechnol.*, vol. 25, no. 11, pp. 1251–1255, Nov. 2007, doi: 10.1038/nbt1346.
- [25] C. Bouton et al., "FHIR RDF: Semantic integration of health data using linked data standards," *J. Biomed. Inform.*, vol. 128, Art. no. 104072, 2022, doi: 10.1016/j.jbi.2022.104072.



Ana Paula Lopes de Abreu Ferreira received her M.Sc. degree in Management and Business from the Universidade do Vale do Rio dos Sinos (Unisinos) and the Université de Poitiers, France, in 2024. She is currently a Ph.D. candidate in Applied Computing at the Universidade do Vale do Rio dos Sinos, Brazil. Her research interests include semantic interoperability in healthcare, biomedical ontologies, and intelligent systems applied to hearing healthcare.



Sandro José Rigo received his B.Sc. degree in Computer Science from the Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Brazil, in 1990, the M.Sc. degree in Computer Science from the Universidade Federal do Rio Grande do Sul (UFRGS), Brazil, in 1993, and the Ph.D. degree in Computer Science from the Universidade Federal do Rio Grande do Sul (UFRGS), Brazil, in 2008. He completed a postdoctoral fellowship at the Friedrich-Alexander Universität Erlangen–Nürnberg, Germany, in 2018.



Luis Felipe Maldaner received his B.A. degree in Business Administration from Universidade Feevale, Brazil, the M.Sc. degree in Business Administration from the Universidade do Vale do Rio dos Sinos (Unisinos), Brazil, and the Ph.D. degree in Latin American Studies from Hankuk University of Foreign Studies, Seoul, South Korea.