





Uso de IA para criação de Ontologias: uma proposição para o domínio “Saúde Indígena”

Use of AI for Ontology creation: a proposition for the "Indigenous Health" domain



Luiz Rogério Lopes Silva¹

 <http://lattes.cnpq.br/6377367856927679>
 <https://orcid.org/0000-0002-7457-9778>

Marcelo Schiessl²

 <https://lattes.cnpq.br/4402964117780625>
 <https://orcid.org/0000-0002-3895-5844>

Milton Shintaku³

 <http://lattes.cnpq.br/8605833104600600>
 <https://orcid.org/0000-0002-6476-4953>

Resumo

Introdução: Propõe o uso de Inteligência Artificial (IA), especificamente a LLM ChatGPT-4, para a criação de Ontologia de domínio em “saúde indígena” de forma semi-automatizada, a fim de auxiliar o desenvolvimento de um sistema de informação para a Secretaria de Saúde Indígena (SESAI). Metodologia: Adotou-se o framework LLMs4OL, adaptando-o para o contexto específico. Inicialmente, foi definido o escopo do domínio e coletado um corpus controlado de documentos referentes à SESAI. Em seguida, foram elaborados *prompts* para o ChatGPT-4, com o objetivo de extrair conceitos e relações do corpus. As respostas do modelo foram avaliadas e os *prompts* foram refinados iterativamente para melhorar a precisão e a completude da ontologia. Resultados: Uma série de *prompts* são disponibilizados para criação de ontologias em outros domínios. A ontologia criada é composta por 180 axiomas, dos quais 54 são axiomas de declaração e 58 são axiomas lógicos. Há 44 classes distintas, 11 propriedades de objeto e 2 propriedades de anotação, além de 14 indivíduos identificados. Conclusão: O uso da IA demonstrou eficácia na extração de termos, relações e na criação de uma estrutura ontológica inicial, mostrando um potencial promissor para criação de ontologias. Entretanto, a supervisão humana garantiu a adequação cultural e a precisão dos dados, provando que a combinação de IA e expertise humana é uma alternativa para organização do conhecimento.

Palavras-chave: inteligência artificial; ontologia; chatGPT-4; saúde indígena; aprendizagem de ontologia.

¹ Doutor em Gestão da Informação, Universidade Federal do Paraná (UFPR), Curitiba, PR, Brasil. Professo colaborador do Programa de Pós-graduação em Gestão da Informação da Universidade Federal do Paraná (UFPR), Curitiba, PR, Brasil. Pesquisador (Bolsista), Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Brasília, DF, Brasil. luiz.lopes@ufpr.br.

² Doutor em Ciência da Informação, Universidade de Brasília (UnB), Brasília, DF, Brasil. Pesquisador (Bolsista), Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Brasília, DF, Brasil. joseschiessl@ibict.br.

³ Doutor em Ciência da Informação, Universidade de Brasília (UnB), Brasília, DF, Brasil. Coordenador de Tecnologias para Informação, Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Brasília, DF, Brasil. shintaku@ibict.br.

Abstract

Introduction: This paper proposes the use of Artificial Intelligence (AI), specifically the LLM ChatGPT-4, for the semi-automated creation of a domain ontology in "indigenous health," in order to assist in the development of an information system for the Special Secretariat of Indigenous Health (SESAI). **Methodology:** The LLMs4OL framework was adopted, adapting it to the specific context. Initially, the domain scope was defined, and a controlled corpus of documents related to SESAI was collected. Then, prompts were designed for ChatGPT-4 to extract concepts and relations from the corpus. The model's responses were evaluated, and the prompts were iteratively refined to improve the accuracy and completeness of the ontology. **Results:** A series of prompts is provided for creating ontologies in other domains. The created ontology consists of 180 axioms, 54 of which are declaration axioms and 58 are logical axioms. There are 44 distinct classes, 11 object properties, and 2 annotation properties, as well as 14 identified individuals. **Conclusion:** The use of AI proved effective in extracting terms, relations, and creating an initial ontological structure, demonstrating promising potential for ontology creation. However, human supervision ensured cultural appropriateness and data accuracy, proving that the combination of AI and human expertise is a viable alternative for knowledge organization.

Keywords: artificial intelligence; ontology; chatGPT-4; indigenous health; ontology learning

1 INTRODUÇÃO

A Secretaria Especial de Saúde Indígena (SESAI) é responsável por coordenar e executar a Política Nacional de Atenção à Saúde dos Povos Indígenas no Brasil, o que inclui a gestão de obras e edificações destinadas ao atendimento médico nas aldeias. Nesse sentido, Benevides, Portillho e Nascimento (2014) afirmam que a SESAI é responsável por planejar, executar e avaliar ações relacionadas ao saneamento e edificações nos Distritos Sanitários Especiais Indígenas (DSEI).

Araújo e Felipe (2021) apresentam, mesmo que de forma discreta, a relação entre as obras de saneamento e melhoria da saúde indígena nas aldeias, assim como as diferenças entre os distritos indígenas, distribuídos por todo o território nacional. Com isso, ressaltam a necessidade de um sistema de informação que gerencie as atividades relacionadas às edificações, que podem estar relacionadas ao saneamento ou à saúde.

Diante deste problema, a SESAI firmou projeto de pesquisa com o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) para propor um ecossistema de informação que possa melhorar a eficiência das suas ações relacionadas a obras de engenharia nas aldeias. Uma das etapas do projeto é criar uma ontologia de domínio da saúde indígena, processo visto como facilitador para estruturar um entendimento formal e compartilhado do campo.

Todavia, criar uma ontologia exige esforço em analisar e extrair conhecimento especializado do domínio. A SESAI, suas coordenações e os Distritos Sanitários Especiais Indígenas (DSEIs) produzem volume extenso de documentos e informações relacionados

à gestão, assistência e monitoramento da saúde indígena, incluindo os processos de obras de engenharia, o que por sua vez, intensifica a complexidade em organizar o conhecimento.

No contexto deste artigo, os autores encontram nos estudos de *Ontology Learning*, ou aprendizado de ontologias, uma alternativa para reduzir a dependência de trabalho manual intensivo e de expertise específica para desenvolver uma ontologia do domínio "Saúde Indígena". A literatura aponta que uma das maneiras de automatizar a extração e estruturação do conhecimento contido em textos variados é o uso de Modelos de Linguagem de Grande Escala (LLMs). Tais modelos são capazes de identificar padrões e extrair informações relevantes mesmo de textos não estruturados, organizando conceitos, relações e regras que do modo tradicional demandaria mais tempo e dedicação.

Assim, o objetivo deste trabalho é apresentar os procedimentos usados na criação de uma ontologia de domínio da saúde indígena que utilizou de LLM para automatizar a extração e estruturação de conhecimento. O trabalho se justifica tendo em vista a necessidade de equilibrar a precisão e a riqueza cultural exigidas pela saúde indígena com a eficiência operacional necessária para a gestão em larga escala. A aplicação de inteligência artificial em contextos socioculturais complexos, como é o caso da SESAI, promove uma inovação que pode servir de modelo para outras iniciativas da organização da informação e do conhecimento.

2 ONTOLOGY LEARNING

Aprendizado de Ontologias (AO) é uma disciplina que combina Inteligência Artificial e engenharia de conhecimento para construir automaticamente uma ontologia a partir de fontes textuais estruturadas ou semiestruturadas (Giglou; D'Souza; Auer, 2023). As tarefas do AO incluem identificar conceitos, significados e relações em textos, além de extrair axiomas, o que pode reduzir a necessidade de construção manual (Giglou; D'Souza; Auer, 2023).

Na literatura, modelos distintos têm sido propostos para auxiliar na engenharia de ontologias. *Ontology Learning Layer Cake* é um modelo introduzido por Philipp Cimiano, em 2006, que estrutura o processo de AO em camadas sequenciais. Começa com

a extração de termos lexicais e prossegue com a identificação de sinônimos, construção de taxonomias, extração de relações não hierárquicas e, finalmente, formulação de axiomas e regras. Cada camada do modelo concentra-se em um aspecto diferente da ontologia, desde a identificação básica de termos até a definição de relações complexas e restrições. Esse modelo permite uma análise detalhada e modular, sendo útil para compreender as etapas individuais do processo de criação de ontologias.

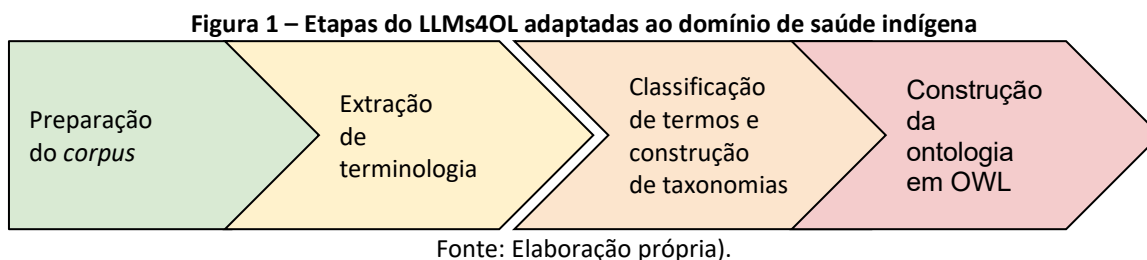
O modelo *End-to-End Ontology Learning com Large Language Models (LLMs)*, proposto por Lo *et al.* (2024), busca automatizar todas as etapas do AO de forma contínua. Utiliza técnicas de Processamento de Linguagem Natural (PLN) para extrair e formalizar conhecimento diretamente das fontes de dados, sem a necessidade de intervenção manual em cada etapa. Este modelo se diferencia por tentar capturar todas as nuances necessárias para a construção de ontologias em uma única fase, o que pode acelerar significativamente o processo de desenvolvimento.

O paradigma *Large Language Models for Ontology Learning (LLMs4OL)*, desenvolvido por Giglou, D'Souza e Auer, em 2023, utiliza LLMs para realizar tarefas específicas de AO. Isso inclui preparação de corpus, extração de terminologia, classificação de termos e construção de taxonomias. O LLMs4OL reconhece que, embora os LLMs possam realizar essas tarefas, a validação por especialistas é essencial para garantir a relevância e a precisão dos axiomas gerados.

Para o desenvolvimento da ontologia de gestão de obras e edificações em comunidades indígenas, o paradigma LLMs4OL foi escolhido como o modelo a ser seguido. Essa escolha se justifica pela capacidade dos LLMs de processar dados textuais e pela importância da validação especializada, oferecendo um equilíbrio entre automação e expertise humana.

3 PROCEDIMENTOS METODOLÓGICOS

Conforme o objetivo proposto, este trabalho é uma pesquisa aplicada adaptando modelos da IA para a construção de uma ontologia específica. Assim, o estudo segue o fluxo apresentado na Figura 1, com base no paradigma LLMs4OL.



A primeira etapa, a preparação do corpus, envolveu a coleta de documentações oficiais da SESAI relacionadas a edificações, que abrangem informações sobre Estrutura e Organização Institucional, Fluxos Processuais, Políticas Públicas e Sistemas de Informação. Esses documentos foram integrados em um único conjunto de dados por meio de um *script* em *Python*, permitindo que o método fosse replicável em diferentes contextos.

Com o conjunto de dados preparado, foram realizados testes utilizando quatro modelos de LLMs: *ChatGPT-4*, *Maritalk*, *QWEN* e *Deepseek*. Assim, para avaliar o desempenho desses modelos no processo de criação de ontologias, adotamos um conjunto de métricas que envolvem tanto aspectos de qualidade semântica quanto de eficiência computacional. Entre as principais métricas utilizadas, destacam-se:

- Cobertura conceitual: verifica se os termos, classes e relações extraídos cobrem os conceitos principais de “saúde indígena”;
- coerência e consistência dos axiomas: verifica se há conflitos ou duplicações nos axiomas para garantir a viabilidade da ontologia;
- eficiência de processamento: inclui tanto o tempo de resposta dos modelos na extração de conceitos e relações quanto a quantidade de passos de refinamento de prompts necessários para atingir um resultado satisfatório.

Nas análises empíricas, o ChatGPT-4 superou os demais em termos de eficácia – melhor equilíbrio entre precisão e cobertura conceitual – e eficiência – menor tempo médio de iterações até a geração de um resultado coerente. Além disso, foi o modelo que apresentou menor incidência de relações redundantes ou inconsistentes, reduzindo o esforço de validação manual.

A definição das categorias conceituais relevantes para o domínio de “saúde indígena” foi guiada por um processo semiestruturado que combinou:

- a) Análise prévia de documentos e políticas da SESAI: estrutura organizacional, fluxos processuais e termos técnicos recorrentes;
- b) metodologias formais de *Ontology Learning*, em especial o paradigma LLMs4OL, que orienta a elaboração de ontologias a partir de corpus;
- c) validação por especialistas: discussões com pesquisadores do tema "saúde indígena" para revisar e confirmar a pertinência das categorias.

A técnica utilizada para a criação dos *prompts* uniu referências formais, como as etapas sugeridas pelo LLMs4OL, com ajustes ad hoc baseados na resposta inicial do modelo. Em outras palavras, partimos de uma estrutura metodológica clara – organizando a coleta de dados, extração de termos e modelagem – e, a cada iteração, refinamos os comandos de entrada – *prompts* – conforme analisávamos a qualidade das saídas geradas. Esse ciclo de realimentação contínua buscou melhorar a aderência dos resultados à realidade do domínio.

Portanto, na etapa de extração de terminologia, os termos relevantes foram identificados a partir do corpus preparado. O *ChatGPT-4* foi empregado para auxiliar na extração desses termos, aproveitando sua capacidade de processar e entender a linguagem natural. Em seguida, na classificação de termos, os conceitos extraídos foram agrupados em categorias conceituais relevantes, identificando classes ou tipos a que os termos pertencem.

A construção de taxonomias foi realizada ao estabelecer relações hierárquicas "is-a" entre os conceitos identificados. Além disso, foram extraídas relações semânticas não-taxonômicas, capturando relações complexas entre os conceitos.

As relações semânticas não taxonômicas foram identificadas durante o processo de extração e classificação de termos, tendo como base:

- a) Verbos e expressões verbais como executar, gerir, fiscalizar que indicam ações ou vínculos funcionais entre os conceitos. Por exemplo: "Distrito Sanitário Especial Indígena (DSEI)" –[executa]→ "Obras de Saneamento". Isso significa que o DSEI realiza, na prática, a execução ou coordenação de obras de saneamento no território indígena;
- b) expressões funcionais ou de dependência como por exemplo, depende de, pertence a, está contido em ou oferece suporte a. Desse modo, a relação

“Equipe de Saúde” – [pertence a] → “DSEI” pode ser inferida da frase: “a equipe de saúde está formalmente vinculada ao DSEI”.

Quando identificadas relações mais complexas (p. ex., que envolvem três ou mais conceitos), utilizamos regras como “Se (X fazParteDe Y) e (Y supervisionadoPor Z) então (X supervisionadoPor Z)”, formalizando-as em formato de *Semantic Web Rule Language* (SWRL) ou em axiomas OWL adicionais. Tal regra representa a herança da relação de supervisão, típica do contexto de unidades de saúde subordinadas a uma secretaria maior.

Por fim, foram criados axiomas e regras para definir restrições que regem a ontologia, garantindo sua consistência e validade. O conhecimento foi representado formalmente em triplas *Resource Description Framework* (RDF) e *Web Ontology Language* (OWL), permitindo uma estrutura que pode ser facilmente compartilhada e reutilizada. O detalhamento de cada uma das etapas é apresentado a seguir.

4 PROPOSIÇÃO DE ONTOLOGIA COM ChatGPT-4

O presente estudo é uma proposição e, por isso, a formação do *corpus* deu-se com um pequeno conjunto de documentos, de forma a facilitar os processos seguintes, na medida em que há conhecimento sobre eles. Como primeiro resultado, o Quadro 1 apresenta os *prompts* utilizados em cada etapa e os resultados alcançados.

Quadro 1 – Etapas de desenvolvimento

Etapa	Prompts	Resultado Alcançado	Exemplo
Extração de Termos	Análise o texto fornecido e extraia a maior quantidade possível de entidades nomeadas (NER), incluindo pessoas, localidades, organizações, funções e atividades. Liste os principais substantivos, verbos (convertidos para o infinitivo) e siglas e acrônimos (proponha a forma expandida, se disponível ou inferível). Organize cada lista em ordem alfabética, numerando os itens e evitando duplicatas. Adicione a quantidade total de termos do resultado.	Extraíram-se 282 termos dos dados, categorizados em 95 substantivos, 70 verbos e 117 siglas.	O termo “DSEI Yanomami” foi classificado como organização/localidade; verbo “monitorar” indica ação típica das equipes de saúde; e sigla “SIS (Sistema de Informação em Saúde)”, interpretada como sistema computacional.

<p>Geração de Taxonomia</p>	<p>Com base nos termos extraídos no Passo 1, classifique cada termo em tipos conceituais apropriados (ex: 'instituição', 'procedimento', 'equipamento', 'pessoa', 'organização', 'conceito abstrato'). Crie uma hierarquia taxonômica, definindo relações de superclasse-subclasse (ex: 'organização' > 'empresa' > 'startup'). Se um termo for polissêmico, crie ramificações adicionais. Estruture o resultado em tabela com:</p> <ol style="list-style-type: none"> 1. Termo original. 2. Tipo conceitual associado. 3. Relação hierárquica (supertipo e subtipo). <p>Evite repetições, exceto para termos com múltiplos significados. Adicione a quantidade total de termos classificados do resultado.</p>	<p>Criação de uma estrutura taxonômica de 31 termos.</p>	<p>Foi criada a relação "Organização > Secretaria > SESAI", pois a SESAI foi reconhecida como um tipo específico de Secretaria na hierarquia.</p>
<p>Extração de Relações Não Taxonômicas</p>	<p>Analise todos os termos e a taxonomia gerada nos Passos 1-2. Identifique relações semânticas não taxonômicas entre os tipos conceituais. Organize as relações em uma tabela no formato: #(número da relação)– Tipo A – [relação] – Tipo B (com breve justificativa ou evidência textual). Evite duplicações e inclua apenas relações semanticamente válidas. Adicione a quantidade total de relações encontradas.</p>	<p>Identificação de 54 propriedades relacionais.</p>	<p>Identificou-se, por exemplo, 'DSEI' –[implementa]→ 'Programa de Saneamento', mostrando uma vinculação operacional; e 'Obra de Edificação' –[exige]→ 'Licença Ambiental', indicando a necessidade legal.</p>
<p>Extração de Axiomas</p>	<p>Com base nos passos anteriores, defina todos os axiomas possíveis. Crie regras (ex: SWRL) para comportamentos complexos, como:</p> <p><i>Se X trabalha em Y, e Y pertence a Z, então X está relacionado a Z</i></p> <p>Organize as definições numeradas em tabelas:</p> <ol style="list-style-type: none"> 1. Axiomas de Classe. 2. Axiomas de Propriedade. 3. Regras Unárias/Binárias. 4. Regras N-árias. <p>Garanta consistência lógica e evite contradições. Adicione a quantidade total de axiomas encontrados.</p>	<p>Identificação de 13 axiomas.</p>	<p>Axioma "(DSEI SubClassOf Organização)" formaliza que cada DSEI é uma instância de "Organização".</p>
<p>Criação de triplas RDF</p>	<p>Com base em todos os termos, relações e axiomas dos Passos 1-4,</p>	<p>Identificação de 36 triplas.</p>	<p><DSEI_Yanomami> <rdfs:subClassOf> <DSEI> -</p>

	gere triplas RDF no formato <sujeito> <predicado> <objeto> Evite duplicações e organize as triplas em seções ou listas, agrupando por tipo de relação ou termo. Se houver ambiguidade, indique-a claramente. Adicione a quantidade total de triplas geradas.		<SESAI> <supervisiona> <DSEI> - <ObraEdificacao> <exige> <LicencaAmbiental> Cada tripla espelha uma relação fundamental no domínio, seja ela hierárquica (subclasse) ou funcional (coordena, supervisiona).
Criação da ontologia em owl	Com base nas triplas RDF do Passo 5, organize as classes em uma estrutura hierárquica (superclasses e subclasses). Defina atributos (propriedades) relevantes para cada classe e crie axiomas para restringir e descrever melhor as classes e propriedades. Formule regras para comportamentos complexos ou inferências que não possam ser expressas apenas com axiomas OWL. Garanta que as definições sejam coerentes e não gerem contradições lógicas e gere a ontologia OWL mais completa possível, com base nos passos até aqui. Além disso, estenda a ontologia com instâncias (indivíduos) adicionais e inclua as triplas RDF listadas anteriormente.	Ontologia gerada em OWL.	Na ontologia final, por exemplo, "SESAI" e "DSEI" são classes com restrições de domínio e alcance, além de incluir indivíduos como <DSEI_Xavante> e <DSEI_Yanomami>. As propriedades de objeto "coordena", "supervisiona" e "exige" possuem domínios e alcances compatíveis, prevenindo inconsistências lógicas.
Exportação e visualização da ontologia	Considere a ontologia construída nos Passos 1-6, contendo classes, propriedades, axiomas e regras. Salve a ontologia em um arquivo OWL (ex: 'minha_ontologia.owl'). Gere, também, um código de visualização para Graphviz, criando um diagrama simplificado com: <ol style="list-style-type: none"> 1. Nós representando classes. 2. Arestas representando relações (taxonômicas e não taxonômicas). 3. Cores e layouts intuitivos para facilitar a interpretação. 	Arquivo OWL gerado e código para GraphViz fornecido.	O código para Graphviz cria nós para cada classe principal (por exemplo, "SESAI", "DSEI", "ObraSaneamento", "LicencaAmbiental") e arestas rotuladas ("coordena", "supervisiona"), evidenciando a hierarquia de classes e as conexões centrais do domínio de "saúde indígena".

Fonte: Elaboração própria.

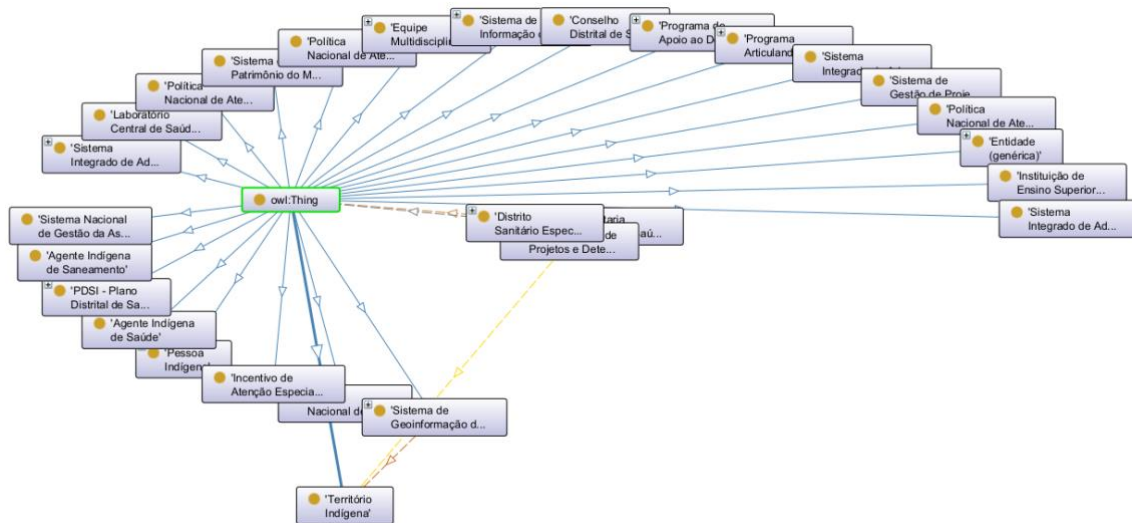
Inicialmente, a capacidade de compreensão e representação lógico-semântica do ChatGPT-4 evidenciou um resultado parcial e satisfatório do domínio. A ferramenta demonstrou eficácia na definição de termos e na criação da ontologia, analisando

documentos rapidamente e superando métodos manuais. Sua escalabilidade permitiu a extração de informações, identificando padrões e relações complexas.

Em um segundo momento, o achado de pesquisa esteve no contexto da engenharia de *prompts* e na necessidade de refinar continuamente as instruções fornecidas ao ChatGPT-4. Os autores avaliaram que a simples aplicação dos princípios LLMs4OL foi insuficiente para o domínio "Saúde Indígena", exigindo que fosse otimizada a interação para melhorar a qualidade das respostas.

No que compete à ontologia gerada em OWL, Figura 2, o artefato é composto por 180 axiomas, dos quais 54 são axiomas de declaração e 58 são axiomas lógicos. Há 44 classes distintas, 11 propriedades de objeto e 2 propriedades de anotação, além de 14 indivíduos identificados. A ontologia também detalha os tipos de axiomas presentes, como *SubClassOf* e *ClassAssertion*. No entanto, foi identificada uma diferença entre o número de termos, relações e axiomas gerados anteriormente pelo ChatGPT-4 e o que a ferramenta exportou em OWL. Neste momento, os pesquisadores não investigaram o motivo da discrepância, mas consideram realizar outros testes na sequência para tentar identificar a lógica utilizada.

Figura 2 – Representação da ontologia



Fonte: Elaboração própria.

No campo ético da pesquisa, questões de transparência e soberania do conhecimento merecem atenção. A natureza proprietária dos algoritmos e a falta de divulgação dos dados de treinamento da LLM criam uma "caixa preta", dificultando a

análise crítica de vieses e pressupostos na ontologia gerada. Isso compromete a confiança na representatividade fiel do conhecimento do domínio. Para tanto, o próprio modelo LLMs4OL exige a validação de especialistas, o que aconteceu graças ao conhecimento dos próprios pesquisadores sobre o domínio. Vale reforçar que a equipe de desenvolvedores está envolvida com a SESAI desde o ano de 2024, participando de eventos, analisando documentos e visitas *in loco* aos distritos indígenas.

5 CONSIDERAÇÕES FINAIS

A criação de uma ontologia de saúde indígena para a SESAI, utilizando o ChatGPT, identificou avanços e preocupações. A capacidade da LLM de processar volumes de dados, identificar relações e gerar representações de conhecimento em tempo reduzido é um avanço. A ferramenta demonstrou eficácia na extração de termos, relações e criação de uma estrutura ontológica inicial. Por outro lado, a questão da explicabilidade é um dilema: a opacidade no processo de geração das respostas, principalmente na criação do arquivo em OWL, provocou inquietações e receios.

Inicialmente, observou-se uma discrepância entre a quantidade de classes, axiomas e relações gerados pelos *prompts* nas fases anteriores e os gerados no OWL. Tal situação exige investigações futuras para evitar que a criação de ontologias por meio de LLMs seja uma "caixa preta". Dada a novidade do tema, ainda há uma lacuna na realização de testes que observem o desempenho das ferramentas na proposição de terminologias confiáveis para domínios específicos. Um gerador pode criar um texto que parece plausível, mas isso não garante que ele esteja utilizando conhecimento sobre o domínio, o que pode ser resultado do pouco rigor na seleção e confiabilidade das fontes mapeadas. Não basta reduzir o tempo e o esforço humano envolvidos no processo, é necessário um processo de geração de conhecimento que seja suficientemente transparente e confiável.

Após a criação da ontologia, a avaliação focou em critérios como precisão, pertinência, abrangência, correção e a estruturação da informação. A versão inicial do artefato demonstrou ser apropriada para representar o domínio em questão. Contudo, os testes prosseguem com o objetivo de aprimorar essa primeira versão e assegurar que as representações sejam confiáveis e verdadeiras para o domínio da saúde indígena. A

equipe responsável valoriza e busca manter um equilíbrio entre a automação proporcionada pela IA e a supervisão e co-criação humanas.

REFERÊNCIAS

ARAÚJO, A. N. S.M.; FELIPE, E. C. Análise qualitativa das ações de saneamento e saúde em áreas indígenas. **Brazilian Journal of Development**, Curitiba, v. 7, n. 5, p. 53203-53213, 2021. DOI: <http://doi.dx/10.34117/bjdv.v7i5.30478>. Disponível em: <https://ojs.brazilianjournals.com.br/ojs/index.php/BRJD/article/view/30478>. Acesso em: 10 mar. 2025.

BENEVIDES, L.; PORTILLO, J. A. C.; NASCIMENTO, W. F. A atenção à saúde dos povos indígenas do Brasil: das missões ao subsistema. **Tempus Actas de Saúde Coletiva**, Brasília, v. 8, n. 1, p. 29-39, mar. 2014. DOI: <http://doi.dx/10.18569/tempus.v8i1.1450>. Disponível em: <http://www.tempusactas.unb.br/index.php/tempus/article/view/1450>. Acesso em: 10 mar. 2025.

CIMIANO, P. **Ontology learning and population from text: algorithms, evaluation and applications**. New York: Springer, 2006. DOI: <http://doi.dx/10.1007/978-0-387-39252-3>. Disponível em: <http://link.springer.com/10.1007/978-0-387-39252-3>. Acesso em: 10 mar. 2025.

GIGLOU, H. B.; D'SOUZA, J.; AUER, S. LLMs4OL: large language models for ontology learning. *In*: PAYNE, T. R.; PRESUTTI, V.; QI, G.; VILLALÓN-POVEDA, M.; STOILLOS, G.; HOLLINK L.; KAOUDI, Z.; CHENG, G.; LI, J. (ed). **The semantic web: ISWC 2023**. Athens: Springer, 2023. v. 14265, p. 408-427. DOI: http://doi.dx/10.1007/978-3-031-47240-4_22. Disponível em: https://link.springer.com/10.1007/978-3-031-47240-4_22. Acesso em: 10 mar. 2025.

LO, A.; JIANG, A. Q.; LI, W.; JAMNIK, M. End-to-end ontology learning with large language models. *In*: CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 38., 2024, Vancouver. **Conference [...]**. Vancouver: NeurIPS, 2024. DOI: <http://doi.dx/10.48550/arXiv.2410.23584>. Disponível em: <http://arxiv.org/abs/2410.23584>. Acesso em: 10 mar. 2025.



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

Como citar este trabalho:

SILVA, Luiz Rogério Lopes; SCHIESSL, Marcelo; SHINTAKU, Milton. *Uso de IA para criação de Ontologias: uma proposição para o domínio "Saúde Indígena"*. *In*: WORKSHOP DE INFORMAÇÃO DADOS E TECNOLOGIA, 8., 2025, Marília, SP. **Anais [...]**. Marília, SP: Universidade de Marília, 2025. DOI: <https://doi.org/10.22477/viii.widat.260>.