

Gerando uma Ontologia usando *ChatGPT*: Um Relato de Experiência

Daniel Lichtnow¹, Ana Marilza Pernas², Leonardo Vianna do Nascimento³,
Guilherme Medeiros Machado⁴, José Palazzo Moreira de Oliveira⁵

¹Universidade Federal de Santa Maria (UFSM)
Santa Maria/RS, Brazil

²Universidade Federal de Pelotas (UFPel)
Pelotas/RS, Brazil

³Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS)
Alvorada/RS, Brazil

⁴LyRIDS lab, ECE Engineering School
Paris, France

⁵Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Porto Alegre – RS – Brazil

daniel.lichtnow@ufsm.br, marilza@inf.ufpel.edu.br,
leonardo.nascimento@alvorada.ifrs.edu.br, gmedeirosmachado@ece.fr,
palazzo@inf.ufrgs.br

Abstract. *This paper presents a report on experiences using ChatGPT in the development of an ontology, made through the application of a well-known methodology: Ontology Development 101. This methodology defines steps to be followed to generate an ontology. For each of these steps are experimental, their results are commented and the adjustments made. A summarized version of the account of this experience was presented in an article published at a conference in which, due to available space, many details, considerations and analyzes about the ontology generation process were omitted.*

Resumo. *Este artigo apresenta um relato de experiências do uso do ChatGPT no desenvolvimento uma ontologia, feito a partir da aplicação de uma metodologia bastante conhecida: a Ontology Development 101. Esta metodologia define etapas a serem seguidas para geração de uma ontologia. Para cada uma destas etapas são apresentadas tentativas, comentados seus resultados e os ajustes realizados. Uma versão resumida do relato desta experiência foi apresentada em um artigo publicado em uma conferência no qual, em função do espaço disponível, muitos detalhes, considerações e análises sobre o processo de geração da ontologia foram omitidos.*

1. Introdução

No contexto da Computação, ontologias estão fortemente associadas a Web Semântica para a qual os cenários descritos há alguns anos (ver [Berners-Lee et al. 2001]) não che-

garam a se concretizar como esperado, sendo esta questão alvo de discussões em trabalhos como [Neto 2024].

A não concretização dos cenários previstos é fruto em parte do fato de que a construção de ontologias com recursos da Web Semântica é um trabalho que consome tempo, exige conhecimento do domínio sobre o qual está se definindo a ontologia, bem como conhecimento dos padrões da Web Semântica e a busca por vocabulários e ontologias já existentes relacionados a ontologia em desenvolvimento. Ter a disposição ferramentas que favoreçam e apoiem este processo é, portanto, algo extremamente desejável. Dentre estas ferramentas estão as *Generative Artificial Intelligence* como *ChatGPT*¹ e *Gemini*².

Quando foram realizados os experimentos descritos neste artigo, constatou-se que não haviam muitas iniciativas de utilização de *LLMs* na geração de ontologias. Isto, aliado ao fato de que os autores estavam envolvidos na busca por ontologias para *Smart Campus*, motivou a tentativa de gerar uma ontologia para este domínio com o apoio de *LLMs*.

Uma questão que surgiu no início do trabalho foi como fornecer ao agente os requisitos para construção de uma ontologia. Inicialmente, foram feitos testes elaborando textos com descrição dos requisitos que foram submetidos ao *ChatGPT*, mas os resultados obtidos indicaram a necessidade de uma definição sobre as etapas de construção da ontologia que reduzisse a necessidade de fornecer para o *ChatGPT* informações relacionadas ao domínio da ontologia. Em função disto, aplicou-se como estratégia a adoção de uma metodologia bem conhecida para desenvolvimento de ontologias, a *Ontology Development 101* [Noy and McGuinness 2001].

Este artigo objetiva portanto fazer um relato desta experiência. Cabe ressaltar parte deste relato foi apresentado em [Lichtnow et al. 2025] no qual em função do espaço disponível, muitos detalhes, considerações e análises sobre o processo de geração da ontologia foram omitidos, sendo aqui detalhados.

Este artigo está organizado da seguinte forma. A Seção 2 apresenta conceitos sobre ontologias e Web Semântica, além de trabalhos que fazem uso de *LLMs* para geração de modelos conceituais, modelos de bancos de dados e ontologias. A Seção 3 apresenta a metodologia usada, a *Ontology Development 101* [Noy and McGuinness 2001], sendo após descrita sua aplicação. A Seção 4 resume as lições aprendidas e os trabalhos futuros.

2. Conceitos e Trabalhos Relacionados

Uma ontologia é caracterizada como “uma especificação de um vocabulário de representação para um domínio compartilhado de discurso”[Gruber 1993]. No contexto da Ciência da Computação e da Ciência da Informação, uma ontologia é tipicamente estruturada como um conjunto de conceitos dentro de um domínio e as relações entre eles. Assim uma ontologia é usada para definir vocabulários compartilhados e fornecer um entendimento comum de um domínio específico, facilitando a comunicação e a interoperabilidade entre diferentes sistemas e aplicações.

A concepção de ontologias está no cerne da arquitetura da Web Semântica proposta por Tim Berners-Lee, sendo que neste contexto, ontologias são frequentemente

¹<https://openai.com/ChatGPT/>

²<https://gemini.google.com/>

representadas usando *RDF - Resource Description Framework* e *OWL - Web Ontology Language*, que procuram fornecer a estrutura e a padronização necessárias para determinar dados conectados de forma aberta e consistente.

Um dos principais objetivos da Web Semântica é fornecer soluções para integração e intercomunicação de dados. Para atingir esse objetivo, foi introduzido o conceito de Dados Abertos Conectados (*Linked Open Data - LOD*), que pode ser definido como um conjunto de melhores práticas para publicação e conexão de conjuntos de dados estruturados na web, para criar uma “Web de Dados” [Isotani and Bittencourt 2015]. A definição de Dados Abertos segue os princípios propostos por Tim Berners-Lee, conhecidos como Sistema 5 Estrelas, que classifica o nível de abertura dos dados usando estrelas. Um sistema 5 estrelas está disponível na internet com uma licença aberta, estruturado, em um formato não proprietário, dentro dos padrões definidos pelo *W3C (World Wide Web Consortium)* e conectado a outros de forma contextualizada. Existem alguns exemplos de aplicações em *Linked Open Data* disponíveis na Web. Um dos mais conhecidos é o *DBPedia*, localizado no centro da Nuvem *Linked Open Data - LOD*³), que consiste em um esforço comunitário de *crowdsourcing* para extrair conteúdo estruturado das informações criadas em vários projetos da *Wikimedia*. Ele está disponível na Web como um gráfico de conhecimento, acessível por meio de navegadores da Web, rastreadores ou por meio de consultas *SPARQL* por meio de *endpoints* [DBPedia 2024].

Gerar modelos conceituais para bancos de dados e ontologias para Web Semântica vem sendo um desafio. E ao longo do tempo metodologias e ferramentas vem sendo aplicadas nesta tarefa. Uma possibilidade é o uso de Processamento de Linguagem Natural (PLN) e Inteligência Artificial (IA) para apoiar a modelagem conceitual, algo que não é novidade. Desde a década de 1980 existem trabalhos que utilizam sistemas baseados em PLN e IA para apoiar o projeto de bancos de dados.

A novidade nos últimos anos foi o surgimento/popularização de *LLMs* como *BERT* [Devlin et al. 2018], *RoBERTa (Robustly Optimized BERT Approach)* [Liu et al. 2019] ou como o *ChatGPT* que consistem em modelos de aprendizado de máquina que memorizam fatos e conhecimentos contidos no corpus de treinamento [Petroni et al. 2019]. *LLMs* como *ChatGPT* e *Gemini* podem ser vistos como uma enorme base de conhecimento que apresentam uma interface baseada em PLN para interagir com seus usuários. Cabe ressaltar que estas ferramentas apresentam problemas e limitações, como o fato de que nem sempre possuem todo conhecimento necessário ou mesmo este conhecimento pode ter sido obtido de fontes não confiáveis. Neste sentido, estratégias como o uso da chamada (*Retrieval-Augmented Generation - RAG*) surgiram como uma forma de incorporar conhecimento de bancos de dados externos [Gao et al. 2023].

Considerando este cenário, o advento das *LLMs* e o fato das *LLMs* terem sido, até o momento da execução dos experimentos descritos neste artigo, pouco usadas na geração de ontologias para Web Semântica, nossa motivação para a realização dos experimentos relatados aqui foi avaliar o quanto *LLMs* podem apoiar a construção de modelos conceituais/ontologias.

³<https://lod-cloud.net/dataset/dbpedia>

2.1. Exemplos de Uso de LLMs para Geração de Ontologias

A seguir, alguns exemplos de trabalhos que utilizam LLMs para apoiar diferentes tarefas relacionadas à modelagem conceitual, seja para modelagem de banco de dados ou para desenvolvimento de ontologias.

O trabalho de [Rodrigues et al. 2023] investiga o uso do *ChatGPT* para classificar termos de domínio de acordo com categorias de duas ontologias. O trabalho apresenta uma análise preliminar, sendo necessários mais experimentos para desenvolver abordagens específicas relacionadas a essa tarefa.

Uma abordagem de verificação de fatos para o *ChatGPT* é descrita em [Mountantonakis and Tzitzikas 2023]. A abordagem usa o *ChatGPT* para gerar triplas *RDF* que descrevem as informações digitadas pelo usuário em um *prompt*. Essas triplas são validadas usando um grafo de conhecimento como DBPedia.

O trabalho de [Helskyaho 2023] analisou o projeto automático de banco de dados pelo *ChatGPT* usando o *Data Vault 2.0*. Os autores concluem que os LLMs parecem ser uma solução possível para automatizar o projeto de banco de dados bruto. No entanto, o *ChatGPT* cometeu erros em todas as rodadas de testes. Uma possível explicação para isso é que os dados da internet usados para treinar o modelo LLM possuem dados corretos suficientes sobre Linguagens de Descrição de Dados (DDLs), mas não possuem dados suficientes sobre a metodologia e/ou apresentam muitas práticas incorretas. Para melhorar a precisão do modelo, os autores sugerem o uso de ajuste fino ou *Retrieval-Augmented Generation - RAG*.

Pode-se considerar o trabalho de [Babaei Giglou et al. 2023] que propõe a abordagem *LLMs4OL*, como mais próximo do proposto no presente artigo, uma vez que nele são utilizados LLMs para aprendizagem de ontologias a partir de textos em linguagem natural. Os resultados obtidos mostram que os LLMs, no momento da realização dos experimentos pelos autores, não são suficientemente adequados para a construção de ontologias que exigem um alto grau de habilidades de raciocínio e expertise no domínio, mas que podem funcionar como assistentes.

Por fim, cabe ressaltar que no primeiro semestre de 2024, durante a realização dos experimentos descritos neste artigo, foi observado que geralmente os trabalhos cobriam apenas uma parte do desenvolvimento de ontologias e não ontologias completas. Também não foram identificados trabalhos que fizessem uso de metodologias para construção de ontologias.

Porém, mais recentemente e após a realização dos experimentos descritos no presente artigo, foi publicado um trabalho no qual é realizado um mapeamento sistemático com o objetivo de sintetizar pesquisas existentes sobre a aplicação de técnicas de *Deep Learning* na construção de ontologias, abrangendo 2.765 artigos publicados entre 2015 e setembro de 2024, sendo selecionados 47 artigos para revisão e análise. Destes 47 artigos 31,3% envolvem construção de ontologias [Amalki et al. 2025]. Este trabalho revela um crescente interesse pela aplicação de LLMs no auxílio a construção de ontologias.

3. Usando LLMs para o desenvolvimento de uma ontologia

Objetivando avaliar o quanto LLMs podem auxiliar na construção de ontologias, nesta seção é feita a descrição do uso do *ChatGPT* para a geração de uma ontologia para *Smart*

Campus. Inicialmente é feita a caracterização do domínio da ontologia, sendo após feito um relato dos experimentos iniciais que não envolveram o uso da metodologia *Ontology Development 101* para então descrever o desenvolvimento da ontologia aplicando a metodologia.

3.1. Caracterização do domínio da ontologia: *Smart Campus*

A crescente procura por aprendizagem personalizada e adaptativa impulsionou a utilização de Tecnologias de Informação e Comunicação (TIC) no sector educativo. Cada vez mais instituições de ensino superior utilizam estas tecnologias para reformar as experiências de ensino, aprendizagem e investigação dos estudantes e funcionários universitários, incluindo modos alternativos de frequência e colaboração.

Um *Campus Inteligente* ou, como mais comumente referenciado, *Smart Campus* é um Ambiente Inteligente. Um Ambiente Inteligente se caracteriza por possuir sensores e dispositivos conectados tendo, a partir desta estrutura, a possibilidade de identificar características do ambiente e dos usuários para selecionar ações que venham a auxiliar seus usuários [Cook et al. 2009].

Em termos de abrangência geográfica/espacial, um Ambiente Inteligente pode abranger uma sala (*Smart Room*), um prédio (*Smart Building*) ou mesmo uma cidade (*Smart City*). Em termos de abrangência geográfica/espacial uma *Smart City* possui forte relação com o conceito de *Smart Campus*, sendo que a distinção entre *Smart City* e *Smart Campus* consiste especialmente no fato de que em um *Smart Campus* o objetivo final é favorecer o aprendizado, ainda que este favorecimento possa ser feito indiretamente (e.g. conforto ambiental, proporciona ambientes mais adequados para o aprendizado e isto pode ser controlado em um Ambiente Inteligente).

Smart Campus é um termo que não possui uma definição amplamente aceita, sendo adotada neste artigo a definição considerada em [Maran et al. 2023], que é apresentada em [Muhamad et al. 2017a]: “a ideia básica do *Smart Campus* consiste de um esforço para integrar um conjunto de tecnologias avançadas inteligentes para universidades a fim de aprimorar a performance, a qualidade dos graduados [...], não apenas para atividades de aprendizado mas cobrindo um amplo espectro que inclui: interações sociais, ambiente, gestão de escritório, economia de energia, etc.”.

Os dados de interesse para um *Smart Campus* não estão limitados aos dados obtidos pelos sensores presentes no campus. Dados externos podem ser úteis, como os de estações climáticas que podem servir para definir atividades no campus e também dados dos sistemas da instituição (dados de alunos, servidores, matrículas, etc.) [Maran et al. 2023]. Assim, o compartilhamento de dados entre sistemas/agentes distintos é uma questão importante quando se trata de aplicações em *Smart Campus*. Para facilitar este compartilhamento, os sistemas/agentes devem utilizar um padrão de representação compreensível para todos eles, sendo esta também uma motivação para a utilização de ontologias que podem servir para integrar estes dados em um modelo conceitual compartilhado.

Outra motivação para a criação de uma ontologia para *Smart Campus* está relacionada ao fato de que, conforme pode ser visto em [Maran et al. 2023] onde foi feita uma revisão sistemática da literatura buscando identificar ontologias para *Smart Campus*, a maioria dos trabalhos que discutem ontologias para *Smart Campus* apresentam soluções

relacionadas à representação de dispositivos como sensores e dados por eles produzidos nos ambientes da universidade, sem enfatizar questões de aprendizagem. Outro ponto destacado em [Maran et al. 2023] é o de que foi constatado também que poucos trabalhos abordam como representar objetos físicos existentes em um campus que possam ser usados como recursos de aprendizagem. Assim, pela análise feita, constata-se que na grande maioria dos trabalhos um *Smart Campus* pouco difere de uma *Smart City* e há falta de uma ontologia amplamente aceita para *Smart Campuses*.

3.2. Experimentos iniciais sem uso da *Ontology Development 101*

As primeiras tentativas de gerar uma ontologia para *Smart Campus* usando *ChatGPT*, feitas por alguns dos autores, consistiram em enviar uma descrição, produzida por um interessado, solicitando a geração de uma ontologia. Estas tentativas foram testes visando uma avaliação inicial da possibilidade de usar um *LLM* para esta tarefa.

Um exemplo de requisição enviada para o *ChatGPT* foi: *Please, create an ontology using OWL. Use the following information. The campus has places. These places can refer to two kinds of places: outdoor places and indoor places. Indoor places are basically buildings and rooms. It is important to note that a room belongs to a building. A building can have one or more rooms. A room can be a classroom, a laboratory, a snack bar, a library, a office, a conference room, auditorium, indoor gardens, indoor parking, sports courts, squares. Outdoor places can be gardens, parking, sports courts, squares. An outdoor place can have buildings inside. It is important to keep geographic data about places. It is necessary to have properties represent latitude and longitude. Besides it is necessary to have the possibility of maintaining geographic data related to a polygon that represents a place..*

Uma dificuldade encontrada nesta requisição refere-se ao fato de que a descrição deve ser precisa e produzida por alguém que tenha conhecimento do domínio da ontologia. Além disto, esta descrição produziu uma ontologia onde “Places” acabou ficando como subclasse de “Campus”, conforme mostrado na Figura 1. Ainda cabe mencionar que a descrição apresentada não é a inicialmente elaborada, é fruto de um refinamento feito sobre descrições anteriormente fornecidas. Houve dificuldade em expressar todos os requisitos da ontologia em uma descrição, algo esperado, já que uma descrição precisa é algo que mesmo especialistas no domínio teriam dificuldade de elaborar.

3.3. Experimentos com o uso da *Ontology Development 101*

Em função das dificuldades constadas nos experimentos iniciais, relacionadas especialmente a necessidade de fornecer uma descrição precisa dos requisitos, foi feita a opção pelo uso de uma metodologia para geração da ontologia. A seguir, a metodologia é sucintamente descrita e após é feito o relato de sua aplicação com o *ChatGPT*.

3.3.1. Breve caracterização da *Ontology Development 101*

Os resultados dos experimentos iniciais indicaram a necessidade de seguir alguma metodologia para conduzir o processo e diminuir a dependência de requisitos precisos a serem fornecidos por especialistas na forma de textos. A opção foi utilizar a *Ontology Development 101* [Noy and McGuinness 2001] que é bastante conhecida, sendo direcionada para aqueles que estão desenvolvendo suas primeiras ontologias.

```

# Classes
campus: a owl:Class ;
    rdfs:label "Campus"@en .

building: a owl:Class ;
    rdfs:label "Building"@en ;
    rdfs:subClassOf campus: .

room: a owl:Class ;
    rdfs:label "Room"@en ;
    rdfs:subClassOf campus: ;
    rdfs:subClassOf [
        a owl:Restriction ;
        owl:onProperty campus:belongsTo ;
        owl:someValuesFrom building:
    ] .

place: a owl:Class ;
    rdfs:label "Place"@en ;
    rdfs:subClassOf campus: .

```

Figura 1. Resultado de um dos primeiros experimentos

Embora voltada para iniciantes, o uso da *Ontology Development 101* exige a participação pessoas com conhecimento sobre o domínio relacionado a ontologia. Mas ao empregar *LLMs*, espera-se auxiliar os especialistas no domínio ou mesmo pessoas que não possam ser caracterizadas como especialistas.

O processo definido na *Ontology Development 101* consiste em 7 etapas, indicadas para aquele(s) que pretendem gerar uma ontologia:

1. Determine o domínio e o escopo da ontologia;
2. Considere reutilizar ontologias existentes e reutilizar termos existentes, identificando os termos já existentes em ontologias ou vocabulários existentes que podem ser reutilizados;
3. Enumere os termos importantes na ontologia: liste os termos mais importantes no domínio da ontologia;
4. Defina as classes e a hierarquia de classes;
5. Defina as propriedades das classes;
6. Defina as facetas dos slots (valores permitidos, tipo de dados, cardinalidade);
7. Verifique a definição da ontologia através da criação de instâncias.

3.3.2. Experimentos de criação da ontologia para *Smart Campus* usando *Ontology Development 101*

O processo aqui descrito compreende as etapas 1 a 6, excluindo a sétima etapa, que se refere à criação das instâncias. Em cada etapa, são apresentados resultados, sendo explicado o que foi feito para aprimorar as respostas inicialmente obtidas.

Etapa 1. Determinar o domínio e o escopo da ontologia

Para executar esta etapa, na *Ontology Development 101* são sugeridas algumas questões para serem respondidas:

1. Qual é o domínio que a ontologia deve cobrir?
2. Para que será usada a ontologia?
3. Para quais tipos de perguntas as informações da ontologia devem fornecer respostas?
4. Quem usará e irá manter a ontologia?

Já neste ponto, foi utilizado o *ChatGPT*, buscando responder estas questões. Ao contrário do que foi feito no primeiro experimento, busca-se não contar com conhecimento prévio das pessoas que estejam construindo a ontologia, conhecimento este expresso na forma de textos enviados para o *ChatGPT*.

Para responder a questão 1 proposta na metodologia, foi formulada a seguinte questão: O que é um *Smart Campus*? Cabe destacar que toda interação com o *ChatGPT* foi feita em inglês, portanto a questão enviada foi *What is a Smart Campus?*

O objetivo ao submeter esta questão está em identificar as informações sobre o *Smart Campus* coletadas e produzidas pelo *ChatGPT* visando obter conhecimento sobre o domínio, com menor dependência de especialistas. A resposta produzida pelo *ChatGPT* foi: “*A smart campus is an educational institution that utilizes various technologies and digital infrastructure to enhance the learning experience, improve operational efficiency, and foster a more connected and sustainable environment. These technologies can include Internet of Things (IoT) devices, sensors, data analytics, and communication networks.*”

Seguindo as questões sugeridas na metodologia, foi submetida a seguinte questão: *For what we are going to use a Smart Campus ontology?* A resposta obtida foi: “*An ontology for a Smart Campus would serve as a structured representation of all the entities, concepts, and relationships relevant to the campus environment, particularly in the context of IoT (Internet of Things) and smart technologies.*”

Um aspecto nesta resposta que chamou a atenção, refere-se ao fato de aspectos da aprendizagem não serem explicitados, sendo explicitados aspectos relacionados a *IoT (Internet of Things)* que estão relacionados a qualquer ambiente inteligente. Nesse sentido, o trabalho de [Maran et al. 2023] observa que muitos trabalhos sobre *Smart Campus* não consideram (ou enfatizam) aspectos de aprendizagem, relacionando e quase que limitando um *Smart Campus* a ideia de uma *Smart City*. A resposta do *ChatGPT* parece reproduzir esse cenário, algo que enfatiza a necessidade de conhecimento do domínio e do fornecimento de informação adicionais (e.g. *RAG*).

A requisição foi reformulada, suprimindo a palavra *smart*, pois nossa percepção foi de que a palavra *smart* contribuiu para que a resposta fosse direcionada para aspectos o uso de tecnologia. Assim a nova requisição feita foi: *For what are we going to use a Campus ontology?*. A resposta foi semelhante à obtida na questão anterior, mas agora é possível identificar referências à aprendizagem: “*A Campus ontology is typically used to model the various entities, relationships, and attributes within a university or college campus environment in a structured and organized manner. This ontology can be utilized for a variety of purposes, including [...] Enabling semantic web applications that utilize campus-related data in innovative ways, such as campus navigation systems, personalized learning environments, or recommendation systems for courses and activities. [...]*”

Foi submetida ainda uma questão ao *ChatGPT* baseada na terceira questão sugerida na metodologia: *For what types of questions the information in the smart campus ontology should provide answers?*. Sendo a resposta retornada: “[...] *In summary, the smart campus ontology should provide answers to questions related to navigation, resource availability, events, safety, facilities management, environmental conditions, security, administrative matters, campus amenities, academic support, and sustainability initiatives*”. Novamente aprendido não aparece como algo destacado.

Etapas 2. Considerar o reuso de ontologias existentes

Para identificar ontologias que pudessem ser reutilizadas, a seguinte questão foi submetida: *Are there ontologies for Smart Campus?*. A resposta do *ChatGPT* foi: “*Yes, there are ontologies designed specifically for Smart Campuses. These ontologies serve to structure and organize the vast amount of data generated and utilized within a smart campus environment.[...]*”. Neste sentido, a resposta do *ChatGPT* chega a indicar que “*Several organizations and research institutions are actively working on developing ontologies tailored to the needs of smart campuses. These ontologies help streamline data integration, interoperability, and decision-making processes, ultimately contributing to the efficient management and operation of campus facilities and services.*”. Entretanto, em [Maran et al. 2023] os autores destacam que não foi encontrada uma ontologia específica pertencente ao domínio de *Smart Campus*.

A resposta do *ChatGPT*, baseada em nosso conhecimento prévio, levou à formulação de uma nova questão: *Give me examples of these ontologies.* A resposta do *ChatGPT* foi “*Certainly! Here are a few examples of ontologies that have been developed for smart campus environments: Smart-M3 Ontology: Smart-M3 is an open-source platform for creating smart environments, including smart campuses. Its ontology provides a framework for representing and managing various entities and interactions within a smart environment, such as devices, services, users, and events. Linked Sensor Middleware Ontology (LSMO): LSMO is an ontology developed to support the integration of heterogeneous sensor data in smart environments. It provides a standardized representation for sensor data and metadata, enabling interoperability and semantic integration across different sensor networks. Semantic Sensor Network Ontology (SSN): SSN is an ontology developed by the W3C Semantic Sensor Network Incubator Group. It provides a standard vocabulary for describing sensors, observations, and sensor-related phenomena, which can be used to model sensor data in smart campus applications. CityPulse Ontology: CityPulse is a project focused on developing smart city solutions, which can be adapted for smart campus environments. The CityPulse ontology provides a comprehensive framework for modeling various aspects of urban environments, including transportation, energy, environment, and public safety. Smart Campus Framework Ontology (SCFO): SCFO is an ontology developed specifically for smart campus applications. It provides a modular framework for representing different aspects of campus infrastructure, services, and resources, allowing for flexible and extensible modeling of smart campus environments. These are just a few examples, and there are many other ontologies and frameworks available that can be adapted or extended for use in smart campus applications. The choice of ontology depends on the specific requirements and use cases of the smart campus project.*”.

Os exemplos dados pelo *ChatGPT* consistem em um conjunto de ontologias que

podem ser usadas não apenas em um *Smart Campus*, com exceção da última:

1. *Smart-M3 Ontology*;
2. *Linked Sensor Middleware Ontology (LSMO)*;
3. *Semantic Sensor Network Ontology (SSN)*;
4. *CityPulse Ontology*;
5. *Smart Campus Framework Ontology (SCFO)*

Para validar as respostas, essas ontologias foram buscadas utilizando o *Google*. Cabe salientar que, no momento em que os experimentos foram feitos, recursos para realizar pesquisas na Web dentro do *ChatGPT* não estavam disponíveis (ver *ChatGPT Search*⁴). Algumas das ontologias indicadas na resposta foram encontradas, mas não a *Smart Campus Framework Ontology (SCFO)*, que pareceu ser a mais fortemente relacionada a *Smart Campus*.

A partir disto, foi feita a seguinte pergunta para o *ChatGPT*: *Where can I obtain the Smart Campus Framework Ontology (SCFO)?* A resposta foi: *“As of my last update, there isn’t a widely recognized or standardized ontology called ”Smart Campus Framework Ontology (SCFO)” in the academic or research community. However, such an ontology may have been developed by a specific research group, organization, or company for their smart campus projects.”*

Neste ponto, dada a incapacidade do *ChatGPT* de verificar a existência de ontologias específicas para *Smart Campus*, optamos por primeiro enumerar os termos importantes da ontologia (próxima etapa da metodologia) e posteriormente tentar associar estes termos à ontologias previamente existentes. Mas esta etapa forneceu indicativos de limitação do *ChatGPT*, algumas conhecidas, para a construção de ontologias.

Etapa 3. Enumerar termos importantes na ontologia

Nesta etapa considera-se criar uma lista de termos que representem classes, propriedades e relações na ontologia. Evidentemente esta identificação exige conhecimento de diversos especialistas sobre o domínio da ontologia. Considerando que o *ChatGPT* reúne informações de muitas fontes produzida por muitas pessoas, soliciou-se esta listagem de termos ao *ChatGPT*, assumindo a presença do conhecimento do domínio por especialistas e buscando verificar se isto diminuiria a necessidade de especialistas.

A solicitação dos termos importantes para a ontologia foi feita fornecendo a seguinte requisição para o *ChatGPT*: *Enumerate important terms that I can use for defining a smart campus ontology.* O *ChatGPT* retornou 20 termos: *Smart Campus, Internet of Things (IoT), Sensors, Data Analytics, Big Data, Artificial Intelligence, Machine Learning, Edge Computing, Cloud Computing, Wireless Networks, Smart Buildings, Energy Management, Environmental Monitoring, Smart Transportation, Digital Twin, Augmented Reality (AR), Cybersecurity, Privacy Protection, User Experience (UX), and Interoperability.*

Ao observar os termos, não foram retornados termos que enfatizam aspectos relacionados ao aprendizado. Algo similar ao que ocorreu na execução da Etapa 1. Assim, decidiu-se reformular a questão utilizando uma estratégia simplificada de *Retrieval-augmented Generation - RAG* que consistiu em fornecer à requisição informações extraídas

⁴<https://help.openai.com/en/articles/9237897-ChatGPTsearch>

de outras fontes. Neste caso, foi utilizado o conceito de *Smart Campus* apresentado em [Muhamad et al. 2017b].

Dessa forma, a requisição enviada para o *ChatGPT* foi: *The basic idea of a smart campus is an effort to integrate a set of advanced intelligence technologies by the university to improve performance, the quality of graduates, and the ease of life through the provision of information technology services that are valuable, dynamic, and user-oriented to support automation and real-time reporting. This encompasses not only learning activities but also broader aspects such as social interaction, environment, office management, energy saving, etc. Based on this text, enumerate important terms that I can use for defining a smart campus ontology.*

O retorno consistiu de 15 termos, mas agora aspectos de aprendizado estão presentes: *Smart Campus, Advanced Intelligence Technology, Performance Improvement, Quality of Graduates, Information Technology Services, Automation, Real-Time Reporting, Learning Activities, Social Interaction, Environmental Management, Office Management, Energy Saving, User-Oriented Services, Dynamic Services, Ontology.*”.

Ainda foi feita uma nova requisição usando a definição de *Smart Campus* fornecida pelo *ChatGPT* na Etapa 1. A requisição feita foi: *A smart campus is an educational institution that utilizes various technologies and digital infrastructure to enhance the learning experience, improve operational efficiency, and foster a more connected and sustainable environment. Based on this text, enumerate important terms that I can use for defining a smart campus ontology.* O retorno produzido pelo *ChatGPT* teve 15 termos, sendo um deles *Learning Experience* e outro *Remote Learning*. A presença destes termos indica maior ênfase em aprendizagem.

Considerando que o número de termos retornados era insuficiente (apenas 15), foi enviada uma nova requisição sendo explicitado na requisição um número de termos a serem gerados: *A smart campus is an educational institution that utilizes various technologies and digital infrastructure to enhance the learning experience, improve operational efficiency, and foster a more connected and sustainable environment. Based on this text, enumerate the 100 most important terms I can use to define a smart campus ontology.*

A resposta retornou 100 termos e a partir da análise dos termos retornados foi tomada a decisão de aprimorar a requisição de forma a obter termos relacionados a aspectos distintos de um *Smart Campus*, não ficando com termos relacionados quase que tão somente a um aspecto, como estrutura física (*Places*), por exemplo. Assim, nova requisição foi formulada explicitando que os termos deveriam estar relacionados a *learning places, people* e *resources*. A requisição enviada foi a seguinte: *A smart campus is an educational institution that utilizes various technologies and digital infrastructure to enhance the learning experience, improve operational efficiency, and foster a more connected and sustainable environment. Based on this text, enumerate the 100 most important terms I can use to define a smart campus ontology. Take into account places, people, and resources related to learning.*

O retorno obtido para esta nova requisição foi: “[...] *This list covers a wide range of terms relevant to defining the ontology of a smart campus, encompassing technology, infrastructure, learning resources, people, and places*”. Conforme pode ser visto na resposta, o *ChatGPT* acabou gerando a lista de termos para **cinco** aspectos relacionados

a *Smart Campus* e não apenas aos **três** aspectos solicitados na requisição. A Tabela 1 apresenta todos termos retornados na resposta⁵.

Tabela 1. Termos gerados para aspectos distintos de um *Smart Campus*.

Aspecto	Termo
Technology	Internet of Things (IoT) Wireless Connectivity Sensors Data Analytics Artificial Intelligence (AI) Machine Learning Automation Cloud Computing Big Data Virtual Reality (VR) Augmented Reality (AR) Mixed Reality (MR) Wearable Technology Smart Devices
Infrastructure	Smart Buildings Energy Management Environmental Monitoring Renewable Energy Green Spaces Mobility Solutions Smart Transportation Parking Management Traffic Flow Optimization Electric Vehicles (EVs) Sustainable Infrastructure Smart Grid Energy Efficiency Waste Management Recycling Programs Water Conservation Smart Irrigation Indoor Air Quality Smart Lighting Digital Signage Interactive Displays
Learning Resources	Smart Classrooms E-learning Online Learning Platforms Virtual Learning Environments Blended Learning

⁵<https://chat.openai.com/share/8c219c59-6070-410d-b818-163990d62ae8>

Tabela 1. Termos gerados para aspectos distintos de um *Smart Campus*.

Aspecto	Termo
	Flipped Classroom Learning Management System (LMS) Student Information System (SIS) Academic Calendar Course Catalog Faculty Directory Student Services Academic Advising Library Services Digital Libraries Open Educational Resources (OER) Research Repositories Laboratory Facilities
People	Student Organizations Faculty Administrators Staff Researchers Visitors Alumni Industry Partners Interns Counselors Fitness Trainers Coaches Campus Safety Personnel Health and Wellness Professionals Career Advisors International Students Diversity and Inclusion Advocates
Places	Campus Events Smart Campus Educational Institution Maker Spaces Innovation Labs Entrepreneurship Centers Campus Facilities Recreational Areas Community Engagement Spaces Student Housing Dining Facilities Retail Services Financial Services Administrative Offices

Tabela 1. Termos gerados para aspectos distintos de um Smart Campus.

Aspecto	Termo
	Academic Departments
	Libraries
	Laboratories
	Collaborative Spaces
	Co-working Spaces
	Fitness Centers
	Sports Facilities
	Health Centers
	Counseling Centers
	Campus Grounds
	Lecture Halls
	Study Areas
	Dormitories
	Cafeterias
	Offices
	Classrooms

Etapa 4. Definir classes e hierarquia de classes

Para definir as classes, considerando a última resposta obtida na Etapa 3, foi enviada a seguinte requisição: “*Please redo and classify these terms into technology, infrastructure, learning resources, people, and places*”. A seguir, solicitamos ao *ChatGPT* para “*organize the terms into a class hierarchy to compose the ontology of the smart campus*”. Parte do que foi gerado pelo *ChatGPT* é apresentado na Figura 2.

Foram constatados alguns problemas. Por exemplo, na categorização mostrada na Figura 2 não faz sentido uma classe *Smart Campus*. Foi feita então uma nova requisição: *Please, remove category smart campus and redo*. Outro refinamento esteve relacionado as classes *Smart Buildings* e *Smart Classroom*. No esquema gerado, a classe *Smart Buildings* foi colocada como uma subclasse de *Infraestrutura* e *Smart Classroom* considerada uma subclasse de *Learning Resources*. Optamos por classificá-los como subclasses de *Places* e solicitamos isto ao *ChatGPT* fazendo a seguinte requisição: *I think Smart Buildings and Smart Classroom are places. Please, redo*.

O número de classes dificultou a análise. Por causa disto, foi feita a opção de gerar separadamente a hierarquia de classes para cada um dos 5 aspectos (*Technology, Infrastructure, Learning Resources, People, and Places*). A opção feita foi iniciar por *Places*.

A partir desta decisão, foi feita a seguinte a seguinte requisição: *Please redo the class hierarquy just for places*. Na Figura 3 é mostrado o resultado produzido pelo *ChatGPT*. Importante notar que a hierarquia foi detalhada em relação a inicialmente produzida, sendo geradas oito subclasses: *Smart Buildings, Campus Facilities, Recreational Areas, Innovation and Learning Spaces, Administrative Offices, Event Spaces, Student Housing e Campus Grounds*.

Portanto, focar apenas em um dos aspectos do *Smart Campus*, representado pela

1. Smart Campus

- Technology
 - Internet of Things (IoT)
 - Wireless Connectivity
 - Sensors
 - Data Analytics
 - Artificial Intelligence (AI)
 - Machine Learning
 - Automation
 - Cloud Computing
 - Big Data
 - Virtual Reality (VR)
 - Augmented Reality (AR)
 - Mixed Reality (MR)
 - Wearable Technology
 - Smart Devices
- Infrastructure
 - Smart Buildings
 - Energy Management
 - Environmental Monitoring



Figura 2. Primeira Tentativa de Criar uma Hierarquia

1. Places

- Smart Buildings
 - Smart Classrooms
 - Laboratories
 - Lecture Halls
 - Study Areas
 - Dormitories
 - Offices
 - Classrooms
- Campus Facilities
 - Libraries
 - Collaborative Spaces
 - Co-working Spaces
 - Fitness Centers
 - Sports Facilities
 - Health Centers
 - Counseling Centers
 - Dining Facilities
 - Cafeterias
 - Retail Services
 - Financial Services
- Recreational Areas
 - Green Spaces
 - Community Engagement Spaces
- Innovation and Learning Spaces
 - Maker Spaces
 - Innovation Labs
 - Entrepreneurship Centers
- Administrative Offices



Figura 3. Hierarquia de classes - apenas classe Places

classe *Places*, tornou mais fácil a realização de correções e refinamentos e o próprio retorno do *ChatGPT* foi melhor. Para classe *Places*, os refinamentos/correções foram feitos realizando sucessivas requisições para o *ChatGPT*:

- *There are some mistakes... Smart Classrooms and Classrooms are not buildings...*
- *Laboratories, Offices and Administrative Offices are not smart buildings*
- *Please remove Smart Buildings and include Buildings*
- *Laboratories, Offices, Administrative Offices are not buildings. Laboratories, Offices Administrative Offices are inside a building*
- *Laboratories, Offices, and Administrative Offices are not subclasses of Building. Please redo*
- *Please change Parking Management by Parking.*

Feitas estas requisições, a hierarquia final obtida através do *ChatGPT* para *Places* é apresentada na Figura 4.

Foi feita então uma tentativa de retornar a Etapa 2 buscando ontologias existentes para as classes identificadas. A requisição feita foi: *Are there ontologies for places?*. O retorno mostrou um conjunto de ontologias como *GeoSPARQL*, *DBPedia Ontology*, *LinkedGeoData*, *WGS84 Geo Positioning*. Foi solicitado que o *ChatGPT* associasse as classes de *palces* geradas para a ontologia do *Smart Campus* para as existentes nestas ontologias e vocabulários. O resultado não foi bom, pois foram identificadas classes na *DBPedia*, por exemplo, que não existiam de fato - algo similar ao que aconteceu na Etapa 2 quando uma ontologia não existente foi indicada pelo *ChatGPT*.



Figura 4. Hierarquia final gerada para Places

Etapa 5. Definir as propriedades

Na metodologia é sugerido que alguns dos termos identificados na Etapa 3 representem propriedades. Mas no caso de nosso experimento, quase todos termos da Etapa 3 foram usados na definição das classes da ontologia. Assim, juntamente com a hierarquia final de classes foi enviada a seguinte requisição para o *ChatGPT*: “*For each class of*

For each class of this hierarchy, please generate properties.

1. **Places**

- **Buildings**
- **Classrooms**
 - Smart Classrooms
 - Lecture Halls
- **Academic Facilities**
 - Laboratories
 - Libraries
 - Study Areas
 - Collaborative Spaces
 - Co-working Spaces
 - Innovation Labs
 - Maker Spaces
 - Entrepreneurship Centers
- **Student Housing**
 - Dormitories
- **Health and Wellness Facilities**
 - Fitness Centers
 - Sports Facilities
 - Health Centers
 - Counseling Centers
- **Dining and Retail**
 - Dining Facilities
 - Cafeterias
 - Retail Services
- **Administrative Spaces**
 - Offices

Figura 5. Geração de propriedades - Requisição inicial

this hierarchy, please generate properties”. Na Figura 5 é mostrado parte da requisição enviada.

Nesta primeira tentativa, alguns tipos de propriedades das classes não foram geradas. Um exemplo são os relacionamentos entre as classes. A Figura 6 mostra parte do que foi gerado. Em função deste resultado, foi feita nova geração das propriedades, mas agora passando como argumento definições sobre os tipos de propriedades que estão presentes na descrição da metodologia.

Assim a requisição feita, usando a descrição das propriedades presente em [Noy and McGuinness 2001] foi: *In general, several types of object properties can become slots in an ontology: “intrinsic properties such as the flavor of a wine; extrinsic properties such as a wine’s name, and area it comes from; parts, if the object is structured; these can be both physical and abstract “parts” (e.g., the courses of a meal) relationships to other individuals; these are the relationships between individual members of the class and other items (e.g., the maker of a wine, representing a relationship between a wine and a winery, and the grape the wine is made from.)” These classes do not have relationship properties (e.g. Classroom belongs to Building). Please redo and for each class try to include relationship properties.*

O resultado é mostrado na Figura 7 e demonstra claramente o quanto as definições para propriedades presentes no artigo que descreve a metodologia auxiliaram. Novamente, conhecimento externo serviu para aprimorar o resultado.

Etapa 6 - Defina as facetas dos slots (valores permitidos, tipo de dados, cardinalidade)

Sure, here are some properties for each class:

1. Buildings:

- Number of floors
- Total square footage
- Construction material
- Year built
- Architectural style

2. Classrooms:

- Seating capacity
- AV equipment availability
- Whiteboard/Chalkboard availability
- Technology integration level
- Accessibility features

3. Smart Classrooms:

- Interactive displays
- Video conferencing capabilities
- Integrated sound systems
- Automated lighting controls
- Smartboards or touch screens



Figura 6. Geração de propriedades - Resultado da Tentativa 1

1. **Buildings**:

- Intrinsic properties: Architectural style, construction material
- Extrinsic properties: Name, address
- Parts: Floors, rooms
- Relationships: Owned by institution, managed by facilities department

2. **Classrooms**:

- Intrinsic properties: seating capacity, technology integration level
- Extrinsic properties: Room number, building location
- Relationships: Belongs to a building, assigned to specific courses

3. **Smart Classrooms**:

- Intrinsic properties: Interactive displays, video conferencing capabilities
- Extrinsic properties: Room number, building location
- Relationships: Belongs to a building, equipped with technology by IT department

Figura 7. Geração de propriedades - Resultado da Tentativa 2

Nesta etapa, a tentativa foi produzir uma versão da ontologia utilizando os Recursos da Web Semântica. Isto foi feito separadamente para cada uma das classes principais (*Places, People, Technology, Infrastructure, Learning Resources*).

A requisição passada para o *ChatGPT* foi: “*Now, using each class and properties, generate an ontology using OWL and RDF*”. O *ChatGPT* gerou apenas algumas classes. Em função disto, requisitamos que a ontologia fosse gerada usando todas as classes e propriedades. No entanto, obter toda a ontologia em uma única resposta é impossível porque, de acordo com o próprio *ChatGPT*, “O limite de caracteres para uma única resposta é de aproximadamente 4.096 caracteres.”. Se uma resposta exceder esse limite, ela poderá ser dividida em várias partes para garantir que todas as informações sejam fornecidas.

A alternativa para contornar esta limitação foi a cada resposta gerada incompleta, em função do número máximo de caracteres resposta por partes, solicitar que o *ChatGPT* desse prosseguimento em outro retorno que correspondia a uma segmento da resposta completa. Ao final o conjunto de segmentos foram colocados em um único arquivo.

Alguns problemas surgiram em função da nossa decisão de iniciar um novo chat com o *ChatGPT* para geração do arquivo *OWL/RDF* e do fato do *ChatGPT* ser *stateless*. Um exemplo foi o de que *Smart Room* voltou a ser uma subclasse de *Building*. Resolvemos esse problema editando a ontologia usando o *Protégé*, alternativa não automatizada mas a mais ágil no momento. Logicamente, está longe de ser a ideal.

Ainda cabe destacar que geramos versões em *OWL/RDF* para para cada uma das classes (*Places, People, Technology, Infrastructure, Learning Resources*) separadamente, colocando cada uma em um arquivo e após reunimos todas em um único arquivo. Porém, quando tentamos gerar uma nova ontologia a partir desta, solicitando que fosse levado em conta as relações entre as classes (*Places* e *People*, por exemplo), isto não não foi possível via *prompt* em função do tamanho do arquivo da ontologia.

Uma parte da do código *OWL/RDF* da ontologia é mostrado na Figura 8. O arquivo completo com toda ontologia gerada está disponível⁶.

4. Considerações Finais

Este artigo descreveu o uso de *LLMs* para a criação de uma ontologia seguindo uma metodologia. Cabe ressaltar que o objetivo não foi produzir a versão final de uma ontologia, mas avaliar como um *LLM* como o *ChatGPT* poderia auxiliar no processo. É possível fazer algumas observações sobre o experimento realizado.

Fornecer conceitos, sejam de fontes externas ou gerados pelo próprio *ChatGPT*, mostrou-se útil em nosso caso (Etapa 3 e Etapa 5). No entanto, para conceitos fruto de respostas do *ChatGPT* o cuidado deve ser maior. Assim, é preciso ressaltar que *LLMs* não podem ser a única fonte de conhecimento, sendo necessário uso de recursos externos — via *RAG* ou via acréscimo de definições de um artigo relevante e reconhecido, como demonstrado em diversos momentos do experimento.

É preciso verificar se o *ChatGPT* não está indicando algo inexistente. Neste sentido, como relatado na Etapa 2, as respostas iniciais do *ChatGPT* incluíam uma ontologia inexistente. Uma possibilidade que pode ser utilizada em trabalhos futuros é o uso de re-

⁶https://drive.google.com/file/d/1NmJY07kWcXAmlLr4-ehc6EcdK7A5L2S/view?usp=drive_link

```

▼<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:owl="http://www.w
xmlns:xsd="http://www.w3.org/2001/XMLSchema#" xmlns:ex="http://example.org/">
  <!-- Define the ontology -->
  <owl:Ontology rdf:about="http://example.org/PlacesOntology"/>
  <!-- Classes -->
  <owl:Class rdf:about="http://example.org/Places"/>
  <owl:Class rdf:about="http://example.org/Buildings">
    <rdfs:subClassOf rdf:resource="http://example.org/Places"/>
  </owl:Class>
  <owl:Class rdf:about="http://example.org/Classrooms">
    <rdfs:subClassOf rdf:resource="http://example.org/Buildings"/>
  </owl:Class>
  <owl:Class rdf:about="http://example.org/SmartClassrooms">
    <rdfs:subClassOf rdf:resource="http://example.org/Classrooms"/>
  </owl:Class>
  <owl:Class rdf:about="http://example.org/LectureHalls">
    <rdfs:subClassOf rdf:resource="http://example.org/Classrooms"/>
  </owl:Class>
  <owl:Class rdf:about="http://example.org/AcademicFacilities">
    <rdfs:subClassOf rdf:resource="http://example.org/Buildings"/>
  </owl:Class>
  <owl:Class rdf:about="http://example.org/Laboratories">
    <rdfs:subClassOf rdf:resource="http://example.org/AcademicFacilities"/>
  </owl:Class>
  <owl:Class rdf:about="http://example.org/Libraries">
    <rdfs:subClassOf rdf:resource="http://example.org/AcademicFacilities"/>
  </owl:Class>

```

Figura 8. Class with properties - relationships

curso como o *ChatGPT Search*, que permite ao *ChatGPT*⁷ acessar conteúdo atualizado da internet e utilizar o método *Chain-of-Verification - CoVe* [Dhuliawala et al. 2023].

É por vezes preciso indicar ao *ChatGPT* o tamanho da resposta desejada. Na Etapa 3, onde os termos relevantes para a ontologia são identificados, solicitamos arbitrariamente 100 termos, caso isso não fosse feito o número de termos gerados seria insuficiente. Como dimensionar quando uma resposta tem o tamanho adequado é um trabalho futuro e talvez dependente de especialistas.

No que diz respeito a construção de ontologias, na Etapa 4, solicitar a categorização dos termos, seguida pela criação de hierarquias e seu refinamento incremental, como feito com a classe *Places* gerou resultados interessantes. Em trabalhos futuros, na Etapa 3, cabe avaliar se pode ser benéfico solicitar termos relacionados às classes da ontologia, fornecendo previamente definições para as classes superiores da hierarquia.

O *ChatGPT* não mantém memória ou um estado persistente entre interações individuais. Isso é um problema para a construção de ontologias (certamente uma tarefa que leva muitos dias). Uma possibilidade é usar o Letta⁸.

Na Etapa 3, após enumerar os termos para a ontologia na Etapa 3, tentamos vinculá-los a vocabulários conhecidos, como *DBPedia* ou a ontologia *SSN*, e o *ChatGPT* estabeleceu, em alguns casos, relações corretas. No entanto, alcançar a interoperabilidade perfeita entre sistemas diversos geralmente requer um esforço considerável no alinhamento e mapeamento de ontologias, abordando discrepâncias entre diferentes padrões e garantindo uma interpretação consistente dos vocabulários definidos [Amini et al. 2024].

Considerando que a construção de uma ontologia é normalmente uma tarefa conjunta, seria importante poder compartilhar o chat entre aqueles que estão definindo a ontologia. No entanto, isso não é possível no *ChatGPT*, onde apenas uma pessoa pode interagir.

⁷<https://help.openai.com/en/articles/9237897-ChatGPT-search>

⁸<https://www.letta.com/>

Em trabalhos futuros, pretendemos:

- Buscar formas de estabelecer relacionamentos mais eficazes entre a ontologia em desenvolvimento e as existentes;
- Explorar formas para mitigar o fato do *ChatGPT* ser *stateless*;
- Desenvolver estratégias para criar relacionamentos entre classes de uma ontologia com classes e vocabulários pré-existentes usando *LLMs*;
- Desenvolver estratégias para criar relacionamentos entre classes da própria ontologia, pois no trabalho, para cada uma das classes superiores da hierarquia da ontologia para *Smart Campus*, foi gerado um arquivo *RDF/OWL* separado, sendo esses arquivos reunidos em um único. Isto prejudicou a definição de relacionamentos entre classes e deve ser algo explorado em trabalhos futuros;
- Gerar ontologias menores e menos complexas do que a gerada no trabalho atual e realizar o processo de validação de ontologias, solicitando a avaliação de especialistas.

Por fim, de forma similar ao concluído por [Babaei Giglou et al. 2023], consideramos que atualmente *LLMs* podem fornecer assistência a algumas das etapas da construção de ontologias. No caso da metodologia usada neste artigo a Etapa 1 e a Etapa 3 foram as mais beneficiadas, mas mesmo assim necessitaram de conhecimento externo como definição de *Smart Campus*. Existem portanto oportunidades a serem exploradas no sentido de estabelecer procedimentos que possam melhorar a assistência prestada pelas *LLMs* na construção de ontologias.

Referências

- Amalki, A., Tatane, K., and Bouzit, A. (2025). Deep learning-driven ontology learning: A systematic mapping study. *Engineering, Technology & Applied Science Research*, 15(1):20085–20094.
- Amini, R., Norouzi, S. S., Hitzler, P., and Amini, R. (2024). Towards complex ontology alignment using large language models. In *International Knowledge Graph and Semantic Web Conference*, pages 17–31. Springer.
- Babaei Giglou, H., D’Souza, J., and Auer, S. (2023). Llms4ol: Large language models for ontology learning. In *International Semantic Web Conference*, pages 408–427. Springer.
- Berners-Lee, T., Hendler, J., and Lassila, O. (2001). *Scientific american*. 284(5):34–43.
- Cook, D. J., Augusto, J. C., and Jakkula, V. R. (2009). Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing*, 5(4):277–298.
- DBPedia, A. (2024). About dbpedia. Accessed on May 10, 2024.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Dhuliawala, S., Komeili, M., Xu, J., Raileanu, R., Li, X., Celikyilmaz, A., and Weston, J. (2023). Chain-of-verification reduces hallucination in large language models. *arXiv preprint arXiv:2309.11495*.

-
- Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., and Wang, H. (2023). Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997*.
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220.
- Helskyaho, H. (2023). Towards automating database designing. In *2023 34th Conference of Open Innovations Association (FRUCT)*, pages 41–48. IEEE.
- Isotani, S. and Bittencourt, I. (2015). *Dados Abertos Conectados*. São Paulo: Novatec, 1 edition.
- Lichtnow, D., Fleischmann, A. M. P., do Nascimento, L. V., Machado, G. M., and de Oliveira, J. P. M. (2025). Pipeline for ontology construction using a large language model: A smart campus use case. In Filipe, J., Smialek, M., Brodsky, A., and Hammoudi, S., editors, *Proceedings of the 27th International Conference on Enterprise Information Systems, ICEIS 2025, Porto, Portugal, April 4-6, 2025, Volume 2*, pages 97–104. SCITEPRESS. DOI: <https://doi.org/10.5220/0013096000003929>.
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.
- Maran, V., Lunardi, G., Machado, G., Nascimento, L., Lichtnow, D., Cezar, N., Gasparini, I., Oliveira, E., and Oliveira, J. (2023). Smart campuses e ontologias: Um mapeamento sistemático da literatura e caminhos para a pesquisa. In *Anais do XXXIV Simpósio Brasileiro de Informática na Educação*, pages 1180–1190, Porto Alegre, RS, Brasil. SBC.
- Mountantonakis, M. and Tzitzikas, Y. (2023). Real-time validation of chatgpt facts using rdf knowledge graphs. *ISWC Demo Paper*.
- Muhamad, W., Kurniawan, N. B., Suhardi, and Yazid, S. (2017a). Smart campus features, technologies, and applications: A systematic literature review. In *International Conference on Information Technology Systems and Innovation (ICITSI)*, pages 384–391.
- Muhamad, W., Kurniawan, N. B., Yazid, S., et al. (2017b). Smart campus features, technologies, and applications: A systematic literature review. In *2017 International conference on information technology systems and innovation (ICITSI)*, pages 384–391. IEEE.
- Neto, W. C. B. (2024). Web semântica e educação distância: Euforia, frustração e recomeço. *Revista Brasileira de Informática na Educação*, 32:75–97.
- Noy, N. F. and McGuinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology. Technical report.
- Petroni, F., Rocktäschel, T., Lewis, P., Bakhtin, A., Wu, Y., Miller, A. H., and Riedel, S. (2019). Language models as knowledge bases? *arXiv preprint arXiv:1909.01066*.
- Rodrigues, F. H., Lopes, A. G., dos Santos, N. O., Garcia, L. F., Carbonera, J. L., and Abel, M. (2023). On the use of chatgpt for classifying domain terms according to upper ontologies. In *International Conference on Conceptual Modeling*, pages 249–258. Springer.