

Ontologia Pinakes: uma análise com Modelos de Linguagem de Grande Escala

Bruno Carlos da Cunha Costa¹, Greicy Kely C. dos Santos¹,
Ana Carolina Novaes de Mendonça¹, Gabriel Moraes de Oliveira¹,
Dayane Onaga Ferreira Machado^{1,2} and Ana Carolina Simionato Arakaki^{2,3}

¹Instituto Brasileiro de Informação em Ciência e Tecnologia, SAUS, Quadra 5, Lote 6, Bloco H, Brasília, DF, 70.070-912, Brasil

²Universidade Federal de São Carlos, Rod. Washington Luiz, km. 235, São Carlos, SP, 13565-905, Brasil

³Universidade de Brasília, Asa Norte, Brasília, DF, 70910-900, Brasil

Resumo

A construção de ontologias é essencial para a representação formal do conhecimento, mas permanece um processo complexo e dependente de especialistas. Os Modelos de Linguagem de Grande Escala (LLMs) surgem como alternativas promissoras para apoiar essa tarefa. Este estudo exploratório avaliou o uso dos LLMs GPT-4o (OpenAI), Claude 4 Sonnet (Anthropic) e Gemini (Google) na geração automática de ontologias a partir do Documento de Especificação de Requisitos (ORSD) da Ontologia Pinakes, desenvolvida manualmente no Protégé para representar os serviços bibliográficos do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), com foco no Catálogo Coletivo Nacional de Publicações Seriadadas (CCN). As ontologias geradas em RDF (*Resource Description Framework*) foram comparadas à Pinakes, tomada como referência. Os resultados indicam ganhos em prototipagem e cobertura conceitual, mas revelam limitações em coerência lógica, tipagem de dados e aderência a padrões bibliográficos.

Abstract

The Ontology construction is essential for the formal representation of knowledge but remains a complex process that relies heavily on experts. Large Language Models (LLMs) have emerged as promising alternatives to support this task. This exploratory study evaluated the use of GPT-4o (OpenAI), Claude 4 Sonnet (Anthropic), and Gemini (Google) for the automatic generation of ontologies based on the Ontology Requirements Specification Document (ORSD) of the Pinakes Ontology. The Pinakes was manually developed in Protégé to represent the bibliographic services of the Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), with a focus on the Catálogo Coletivo Nacional de Publicações Seriadadas (CCN). The ontologies generated in RDF (*Resource Description Framework*) were compared against Pinakes, which served as the reference. The results indicate advantages in prototyping and conceptual coverage, but also reveal limitations in logical consistency, data typing, and compliance with bibliographic standards.

Keywords

Large Language Models, Ontology Engineering, Pinakes Ontology, Competency Questions,

1. Introdução

A engenharia de ontologias demanda tempo e conhecimento, desde a elicitación de requisitos até a modelagem formal e validação. Documentos de requisitos, como o *Ontology Requirements Specification Document* (ORSD), e conjuntos de Questões de Competência (QCs) são instrumentos consagrados para orientar esse processo [1, 2]. Com a popularização dos *Large Language Models* (LLMs), surge a possibilidade de empregá-los para apoiar a engenharia de ontologias, acelerando etapas iniciais [3],

Proceedings of the 18th Seminar on Ontology Research in Brazil (ONTOBRAS 2025) and 9th Doctoral and Masters Consortium on Ontologies (WTDO 2025), São José dos Campos (SP), Brazil, September 29 – October 02, 2025.

✉ brunocosta@ibict.br (B. C. d. C. Costa); greicysantos@ibict.br (G. K. C. d. Santos); anamendonca@ibict.br (A. C. N. d. Mendonça); gabrieloliveira@ibict.br (G. M. d. Oliveira); dayanemachado@ibict.br (D. O. F. Machado); ana.arakaki@unb.br (A. C. S. Arakaki)

ORCID 0000-0001-9655-8698 (B. C. d. C. Costa); 0009-0008-3604-4298 (G. K. C. d. Santos); 0009-0004-0285-9932 (A. C. N. d. Mendonça); 0009-0005-4454-2340 (G. M. d. Oliveira); 0000-0001-5772-418X (D. O. F. Machado); 0000-0002-0140-9110 (A. C. S. Arakaki)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

como a transformação de documentos de requisitos e QCs em rascunhos ontológicos, desde que a curadoria humana assegure qualidade e aderência semântica [4, 5].

Apesar do potencial, ainda há pouca evidência empírica sobre a eficácia dos LLMs em domínios especializados, como os serviços bibliográficos, nos quais predominam abordagens manuais. Nesse contexto, este estudo explora o uso de LLMs na geração de ontologias a partir de engenharia de *prompts* e do ORSD da Ontologia Pinakes, comparando as ontologias produzidas pelos modelos com a própria Pinakes, construída manualmente no *Protégé* [6], um software *open source* para desenvolvimento e validação de ontologias.

A Ontologia Pinakes é fruto de um projeto de pesquisa desenvolvido na Coordenação de Serviços Bibliográficos (COBIB) do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), com o objetivo de promover a integração semântica dos dados e a interoperabilidade dos serviços tradicionais do Instituto. Sua construção baseou-se no modelo de domínio do Catálogo Coletivo Nacional de Publicação Seriada (CCN), alinhado a modelos conceituais de referência internacional, o que permite sua expansão para representar o conhecimento de outros serviços bibliográficos. Este estudo não teve como propósito validar a Pinakes, mas sim verificar a capacidade dos LLMs de interpretar requisitos formais e gerar estruturas ontológicas consistentes. A análise busca contribuir para o avanço da pesquisa em Ciência da Computação e Ciência da Informação, especialmente na geração automatizada de ontologias, com potencial de apoiar futuramente a modernização dos serviços bibliográficos do Ibict e o fortalecimento da infraestrutura semântica nacional.

2. Trabalhos Relacionados

A engenharia de ontologias é reconhecidamente complexa e demanda tempo, desde a eliciação de requisitos à formalização em OWL (*Web Ontology Language*). Nesse contexto, os LLMs têm sido investigados como alternativas para reduzir os esforços iniciais de modelagem.

Michael Funk, Jan Hosemann, Arne Jung e Carsten Lutz exploraram o uso do *GPT-3.5* para gerar hierarquias conceituais a partir de *seed concepts*, restringindo-se à estrutura taxonômica sem apoio de documentos de requisitos [3]. Anna Sofia Lippolis, Mohammad Javad Saeedizade, Robin Keskisärkkä, Aldo Gangemi, Eva Blomqvist e Andrea Giovanni Nuzzolese avançaram ao avaliar *GPT-4-1106*, *OpenAI o1-preview* e *LLaMA-3.1-405B-Instruct-16b* em um *benchmark* com dez ontologias, cem questões de competência e *user stories*, comparando as estratégias *Memoryless CQ-by-CQ* e *Ontogenia* [4]. Em trabalho posterior [7], os mesmos autores ampliaram os testes com LLMs de raciocínio, como o DeepSeek, incluindo formulação lógica além da geração de classes e propriedades. Mohammad Javad Saeedizade, Eva Blomqvist e Karl Hammar testaram técnicas de engenharia de *prompt* (*Zero-shot*, *Few-shot*, *Chain of Thought*, *Graph of Thought*, *Decomposed* e *CQ-by-CQ*) com *GPT-3.5*, *GPT-4*, *Bard* e *LLaMA-2-70B*, usando o *SPARQL Protocol and RDF Query Language (SPARQL)* como parte da avaliação da qualidade das ontologias geradas pelos modelos [5]. Em escala mais ampla, Jiayi Li, Daniel Garijo e María Poveda-Villalón realizaram uma revisão sistemática, destacando o uso recorrente da série *GPT*, *Claude*, *LLaMA* e *Mistral* em domínios diversos, como saúde, patrimônio cultural, finanças, gestão de emergências e organização do conhecimento acadêmico [8].

Apesar dos avanços, não foram encontrados estudos que empreguem LLMs na construção de ontologias voltadas à Biblioteconomia e aos serviços bibliográficos, com base em um Documento de Requisitos de Ontologia (ORSD) formalmente estruturado. Nesses domínios, predominam abordagens conceituais manuais, como o *IFLA Library Reference Model* [9] e o *PRESSoo* [10], aplicados inclusive na Ontologia Pinakes [6]. Projetos institucionais, como *BIBFRAME* [11] e a *Wikidata Bibliographic Task Force* [12], avançaram em direção ao *Linked Data*, mas com alinhamentos majoritariamente manuais. Catálogos internacionais, como o *British National Bibliography* [13] e o *BnF Data* [14], oferecem dados em RDF, porém sem geração automática via LLMs. Este estudo, portanto, contribui de forma inédita ao aplicar LLMs à geração de ontologias no domínio de serviços bibliográficos, tomando o CCN como caso.

3. Modelos Utilizados e Critérios de Escolha

A escolha dos modelos considerou critérios identificados na literatura e relevantes ao contexto institucional. Destacaram-se: (i) geração de saídas RDF/OWL válidas, demonstrada para GPT-3.5 e GPT-4 [5]; (ii) estabilidade e reprodutibilidade, frente a relatos de inconsistências em execuções [3, 5]; e (iii) diversidade de fornecedores e arquiteturas, fundamental para reduzir vieses e comparar abordagens distintas [4, 5, 7].

Também foram incluídos critérios pouco explorados, mas essenciais: (iv) custo, pela viabilidade institucional, e (v) suporte multilíngue, já que o domínio envolve recursos bibliográficos brasileiros alinhados a modelos internacionais. A relevância é reforçada por Li et al. [8], que evidenciaram a superioridade do GPT-4 em múltiplos domínios, por Lippolis et al. [4], que atestaram sua consistência estrutural, e por Luis Miguel Vieira da Silva, Aljosha Köcher, Felix Gehlhoff e Alexander Fay [15], que mostraram a capacidade do *Claude 3* e *Gemini Pro* de gerar axiomas OWL a partir de descrições em linguagem natural.

Para este estudo, foram selecionados o *GPT-4o* (OpenAI) e o *Claude 4 Sonnet* (Anthropic), pela precisão, raciocínio avançado e robustez conceitual; e o *Gemini 2.5 Flash* (Google) gratuito, por viabilidade de custo. Todos se enquadram nas versões já avaliadas na literatura, o que reforça sua relevância para fins comparativos. A intenção é analisar como cada um responde aos desafios da engenharia de ontologias, sem estabelecer hierarquia entre eles.

4. Metodologia

Este estudo exploratório analisa como diferentes modelos de inteligência artificial generativa interpretam e implementam um conjunto comum de requisitos ontológicos descritos no ORSD, tendo a Ontologia Pinakes como referência. Sem hipóteses rígidas, buscou-se identificar padrões e limitações ao longo de um processo dividido em quatro etapas.

A primeira etapa consistiu na definição do insumo de entrada para os LLMs, composto pelo ORSD da Ontologia Pinakes, cuja amostra está na Tabela 1, e por um prompt fornecendo contexto e orientações para geração da ontologia. Para os LLMs o ORSD incluiu: (i) especificações iniciais como objetivo, escopo e usabilidade; (ii) 59 Questões de Competência (QCs) com exemplos de respostas extraídas de registros reais do CCN; (iii) lista de termos derivados das QCs para enriquecer a terminologia; e (iv) quadro de alinhamento entre entidades dos modelos LRM, PRESSoo e CCN, produzido na fase conceitual da ontologia manual e incluído apenas como orientação estrutural para guiar os LLMs na geração de hierarquias compatíveis com padrões bibliográficos.

Na segunda etapa, de geração automática, o ORSD foi submetido ao *Claude Sonnet 4*, *Gemini 2.5 Flash* e *GPT-4o*, acompanhado do mesmo *prompt*, estruturado para gerar saídas em RDF/XML. Para fins descritivos, o conteúdo do prompt foi organizado em três seções: instruções iniciais, instruções complementares, com exemplos do domínio e orientações finais. Essa estrutura está resumida na Tabela 2 e aproxima-se da técnica *few-shot prompting* descrita por Lippolis et al. [7], diferenciando-se da abordagem tradicional ao utilizar um conjunto consolidado de QCs no ORSD e instruções finais para geração completa da ontologia. O *prompt* completo, o ORSD e todas as saídas RDF/XML estão disponíveis em repositório público¹.

Na terceira etapa, as ontologias geradas pelos LLMs foram carregadas no *Protégé* e comparadas à Pinakes. A análise comparativa considerou a cobertura conceitual, a organização hierárquica, o uso de metadados textuais e o esforço de edição. Para verificação preliminar de cobertura das QCs, estas foram convertidas em consultas *SPARQL*, um dos meios de validação de requisitos ontológicos, e executadas no *GraphDB*, um *triplestore* para dados em RDF. Os resultados dessa verificação foram tratados como indicadores, uma vez que o avaliador semiautomático, calibrado para a ontologia manual, pode ter descartado respostas corretas com variações de grafia ou estrutura. Paralelamente, na última etapa, conduziu-se uma análise qualitativa para identificar pontos fortes e limitações de cada LLM.

¹<https://github.com/cobib-ibict/OntologiaPinakescomLLMs>

Tabela 1

Amostra de elementos incluídos no ORSD da Ontologia Pinakes [6], empregados como insumos semânticos para os LLMs

Componentes	Exemplos (amostra)
(i) Especificações iniciais	Objetivo: Desenvolver um modelo de dados em OWL flexível que represente o domínio dos serviços Bibliográficos do Ibict Escopo: O nível de granularidade está diretamente relacionado às questões de competência e aos termos identificados.
(ii) Questões de Competência (QCs)	QC1: "Qual é o título de uma determinada obra?" Resposta esperada: <i>Revista Ciência da Informação</i> .
(iii) Lista de termos	Classes: <i>PublicacaoSeriada</i> , <i>Biblioteca</i> ; Propriedades de Dados: <i>tituloProprio</i> , <i>nome</i> ; Propriedades de Objeto: <i>temISSN</i> (<i>hasISSN</i>), <i>temEditora</i> (<i>hasPublisher</i>); Instâncias: <i>Ibict</i> , <i>Ciência da Informação</i> .
(iv) Alinhamento LRM-PRESSoo-CCN	LRM: <i>Obra</i> ↔ <i>PRESSoo</i> : <i>Serial Work</i> ↔ <i>CCN</i> : <i>Publicação Seriada</i> .

Tabela 2

Estrutura resumida do *prompt* fornecido aos LLMs

Seção	Elemento	Conteúdo resumido
Instruções Iniciais	Objetivo e escopo	Gerar ontologia em OWL com base nos elementos do documento de especificação de requisitos, representando o domínio do CCN do Ibict, alinhada conceitualmente a LRM e PRESSoo; idioma pt-BR; saída válida para uso direto no <i>Protégé</i>
Instruções Complementares	Estrutura Taxonômica	Construir hierarquia com base no Quadro de alinhamento conceitual LRM-PRESSoo-CCN do ORSD.
	Exemplos estruturais	Exemplo de cadeia de especialização: <i>Biblioteca</i> ⊆ <i>AgenteColetivo</i> ⊆ <i>Agente</i> ⊆ <i>Res</i> , onde ⊆ indica relação de tipo (<i>subClassOf</i>).
	Diretrizes de modelagem	Criar todas as <i>classes</i> listadas no ORSD; criar <i>propriedades de dados e objetos</i> e <i>instâncias</i> listadas e extrair as não listadas a partir das QCs e de suas respostas; criar inversas quando aplicável; adicionar restrições de domínio e alcance e relações complexas quando inferíveis das QCs.
	Lista de termos	Exemplos de termos extraídos das QCs (e.g., <i>Classes</i> : <i>ISSN</i> , <i>Idioma</i> , <i>propriedades de dados</i> : <i>nome</i> , <i>tipologiaDocumental</i> , <i>propriedades de objeto</i> : <i>hasISSN</i> , <i>hasPublisher</i> , <i>instâncias</i> : <i>Ciência da Informação</i> , <i>Ibict</i>).
Orientações Finais	Regras de nomenclatura	Classes em <i>camelCase</i> (<i>PublicacaoSeriada</i>); propriedades em <i>camelCase</i> (<i>tituloProprio</i>); rótulos e comentários em pt-BR; evitar ambiguidade.
	Instruções de Saída	Modelar apenas conceitos do ORSD, descrever anotações usando bases semânticas de recursos bibliográficos da web. RDF/XML: <i>OntologiaPinakes_[NomeLLM].owl</i> , pronto para cópia ou download.

5. Resultados e Discussões

A seguir, apresentam-se os principais resultados da comparação entre as ontologias geradas e a Pinakes construída manualmente, com destaque para estrutura, cobertura e desempenho dos LLMs.

5.1. Cobertura conceitual e estrutural

A Tabela 3 mostra que *Claude* e *Gemini* foram os modelos que mais se aproximaram quantitativamente da Pinakes manual, sobretudo no número de propriedades e instâncias. Já o *GPT-4o* apresentou cobertura limitada, mesmo após instruções reiteradas para maximizar o escopo.

Tabela 3

Comparativo entre os modelos gerados por LLMs e a Ontologia Pinakes manual

Modelo	Classes	Propriedades de Objeto	Propriedades de Dados	Instâncias
Claude Sonnet 4	36	52	45	63
Gemini 2.5 Flash	36	56	45	27
GPT-4o	36	40	34	24
Ontologia Pinakes	38	113	57	92

Apesar de gerar 56 propriedades de objeto, o *Gemini* deixou metade delas sem domínio e alcance declarados, o que comprometeu a precisão semântica, embora o raciocinador tenha inferido parte dessas restrições. O *Claude*, ao criar 63 instâncias, incorreu em excesso ao interpretar 18 valores literais como entidades independentes, inflando a contagem. Já o *GPT-4o* foi mais conservador, mas sacrificou cobertura: produziu menos de 25 instâncias e menos da metade das propriedades da ontologia manual.

Na análise preliminar de cobertura das QCs via consultas *SPARQL*, o *Gemini* apresentou a melhor correspondência ($\approx 58\%$), seguido do *Claude* ($\approx 51\%$), enquanto o *GPT-4o* ficou restrito ($\approx 27\%$). Embora esses percentuais ofereçam uma visão inicial do desempenho dos modelos, variações de grafia das respostas esperadas podem ter influenciado os resultados, especialmente por se tratar de consultas calibradas para a ontologia de referência.

5.2. Organização hierárquica e metadados

A comparação das hierarquias (Figura 1) mostrou diferenças relevantes. *Claude* representou múltiplos níveis, distinguindo *Obra*, *Manifestação* e *Item*, além de subdividir *Agente* e *Lugar*, aproximando-se do LRM. O *GPT-4o*, embora com menos elementos, preservou a estrutura central a partir da superclasse *Res* do LRM. Já o *Gemini* priorizou quantidade, mas com organização mais superficial. Nos metadados textuais, *Claude* e *Gemini* forneceram *labels* e *comments* semanticamente adequados, ainda que breves. O *GPT-4o*, em contraste, limitou-se a repetições pouco informativas, reduzindo a expressividade conceitual.

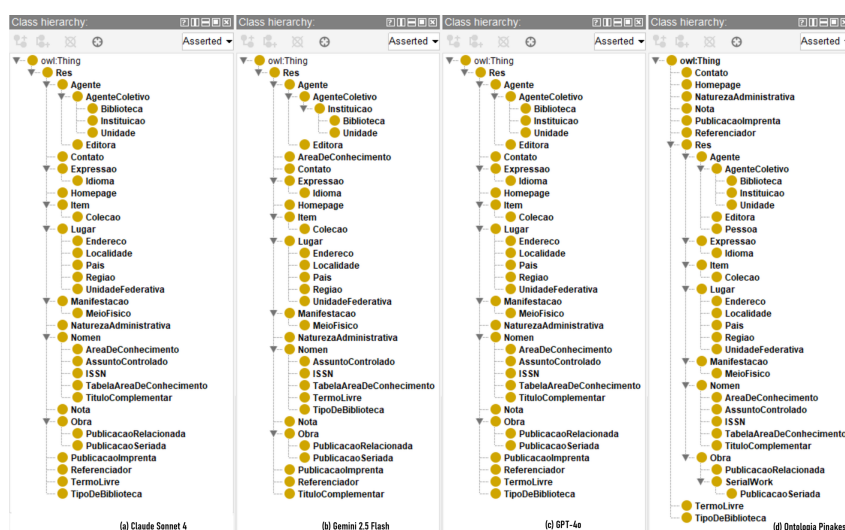


Figura 1: Comparação visual das hierarquias geradas pelos LLMs: (a) Claude, (b) Gemini e (c) GPT-4o, em relação ao modelo manual da (d) Ontologia Pinakes. Figura elaborada pelos autores, a partir da captura de tela do software *Protégé*.

Quanto às restrições, apenas o *Gemini* inferiu cardinalidades, embora algumas aplicadas incorretamente a literais. O *Claude* criou algumas inversas incompletas, corrigidas pelo raciocinador. Nenhum modelo gerou a classe *SerialWork* (PRESSoo), presente no ORSD, já a ausência da classe Pessoa (LRM) é compreensível, pois embora prevista para futura expansão da Pinakes, não foi incorporada ao ORSD por não estar alinhada às classes do CCN, o que limita sua inferência automática.

5.3. Aspectos práticos e limitações dos modelos

O *Claude* demonstrou boa capacidade de expansão conceitual, mas limitado pelo tamanho máximo das respostas, o que exigiu divisão do RDF em blocos. O *Gemini*, mesmo em versão gratuita, produziu uma ontologia relativamente robusta, mas sem exportação direta em RDF/XML. O *GPT-4o* foi o mais prático, permitindo importação imediata no *Protégé*, mas entregou menor cobertura conceitual.

De forma geral, todos os modelos respeitaram o escopo do ORSD e demonstraram ganhos expressivos na prototipagem, embora careçam de tipagem rigorosa e restrições adequadas, confirmando a necessidade de curadoria humana.

5.4. Observações sobre o esforço de modelagem

A construção manual da Ontologia Pinakes demandou dias de trabalho intensivo para modelar mais de 200 entidades, além da inserção manual de instâncias necessárias para validar as QCs, atividade trabalhosa e custosa de revisar. Em contraste, os LLMs geraram ontologias completas em poucas horas, dentro de um único dia de execução. Embora tenham exigido ajustes, especialmente em instâncias com atributos mal tipados e propriedades sem domínio ou alcance, o esforço global foi consideravelmente menor, evidenciando o potencial dos LLMs para acelerar fases iniciais da engenharia de ontologias.

6. Considerações Finais

A automação da construção de ontologias com LLMs mostrou-se promissora ao reduzir tempo e esforço nas etapas iniciais de modelagem. Este estudo evidenciou que LLMs geram classes, propriedades, instâncias e anotações válidas em RDF/XML a partir de um documento de requisitos (ORSO). Nenhum modelo, porém, alcançou integralmente a aderência hierárquica e conceitual nem a completude da Ontologia Pinakes. Seu uso ainda exige revisão e complementação especializada.

O trabalho contribui para o debate sobre a integração da Inteligência Artificial (IA) na engenharia de ontologias, destacando avanços e limitações. Como próximos passos, pretende-se explorar abordagens híbridas que combinem LLMs com extração automática de termos e relações, integradas a scripts em Python, além de aplicar outras técnicas de *prompt engineering* como *CQ-by-CQ* e *Chain of Thought* (CoT). Também será necessário avançar para aplicação em sistemas reais de gestão bibliográfica, avaliando o impacto na interoperabilidade e recuperação da informação, com aprimoramento do ORSD e do *prompt*, tornando-os mais claros, padronizados e alinhados às regras semânticas do domínio.

Declaração sobre IA Generativa

Os autores não empregaram ferramentas de IA Generativa; apenas aquelas inerentes ao objeto de análise do estudo.

Referências

- [1] M. Grüninger, M. S. Fox, Methodology for the design and evaluation of ontologies, in: Workshop on basic ontological issues in knowledge sharing, Montreal, 1995. URL: https://www.researchgate.net/publication/2288533_Methodology_for_the_Design_and_Evaluation_of_Ontologies.

- [2] M. C. Suárez-Figueroa, A. Gómez-Pérez, B. Villazón-Terrazas, How to write and use the ontology requirements specification document, in: *Lecture notes in computer science*, volume 5871, Springer, Berlin, Heidelberg, 2009, pp. 966–982. URL: https://link.springer.com/chapter/10.1007/978-3-642-05151-7_16.
- [3] M. Funk, J. Hosemann, A. Jung, C. Lutz, Towards ontology construction with language models, in: *Joint proc. of the 1st workshop on knowledge base construction from pre-trained language models (KBC-LM) and the 2nd challenge on language models for knowledge base construction (LM-KBC)*, co-located with ISWC 2023, volume 3577, CEUR-WS, Athens, Greece, 2023. URL: <https://ceur-ws.org/Vol-3577>.
- [4] A. S. Lippolis, M. J. Saeedizade, R. Keskiärrkkä, A. Gangemi, E. Blomqvist, A. G. Nuzzolese, Ontology generation using large language models, *arXiv preprint arXiv:2503.05388* (2025). URL: <https://arxiv.org/abs/2503.05388>.
- [5] M. J. Saeedizade, E. Blomqvist, K. Hammar, Navigating ontology development with large language models, in: *European Semantic Web Conference (ESWC 2024)*, Springer, Cham, 2024, pp. 19–36. URL: https://link.springer.com/chapter/10.1007/978-3-031-60635-9_2.
- [6] A. C. S. Arakaki, et al., Ontologia pinakes: formalização semântica para o catálogo coletivo nacional de publicações seriadas, in: *Anais do VIII ISKO Brasil – 1º ISKO Brasil*, 2025.
- [7] A. S. Lippolis, M. J. Saeedizade, R. Keskiärrkkä, A. Gangemi, E. Blomqvist, A. G. Nuzzolese, On the use of large language models to generate capability ontologies, in: *Proceedings of the extended semantic web conference (ESWC 2025)*, 2025. URL: <https://arxiv.org/abs/2503.08055>.
- [8] J. Li, D. Garijo, M. Poveda-Villalón, Large language models for ontology engineering: A systematic literature review, *Semantic Web Journal* (2025). URL: <https://www.semantic-web-journal.net/system/files/swj3864.pdf>.
- [9] P. Riva, P. Le Boeuf, M. Žumer, IFLA library reference model: A conceptual model for bibliographic information, Ifla, Netherlands, 2017. URL: https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017_rev201712.pdf.
- [10] PRESSOO Review Group, Definition of PRESSoo: A conceptual model for bibliographic information pertaining to serials and other continuing resources, version 1.3 ed., Ifla, Netherlands, 2017. URL: <https://repository.ifla.org/items/4fcb0b59-1872-41ff-a9f9-1168bfb17d1a>.
- [11] K. Tharani, Linked data in libraries: A case study of harvesting and sharing bibliographic metadata with bibframe, *Information Technology and Libraries* 34 (2015) 5–19. doi:10.6017/ital.v34i1.5664.
- [12] Wikidata, Wikidata:wikiproject bibliographies, 2025. URL: https://www.wikidata.org/wiki/Wikidata:WikiProject_Bibliographies.
- [13] C. Deliot, Publishing the british national bibliography as linked open data, *Catalogue & Index* (2014) 13–18. URL: https://microblogging.infodocs.eu/wp-content/uploads/2014/10/publishing_bnb_as_lod.pdf, [Online].
- [14] E. Grimaldi, The evolution of data.bnf.fr: Past, present and future of the bnf linked open data project, *JLIS.it* 15 (2024) 119–133. doi:10.36253/jlis.it-588.
- [15] L. M. V. da Silva, A. Köcher, F. Gehlhoff, A. Fay, Assessing the capability of large language models for domain-specific ontology generation, in: *Proceedings of the international conference on industrial ontologies (IO 2024)*, Springer, 2024. doi:10.1109/ETFA61755.2024.10710775.