

# Universidade de Pernambuco

## Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

### Proposta de Dissertação de Mestrado

**Área: Inteligência Computacional**

**Título: Avaliação Comparativa de Embeddings Baseados em Grafos e Ontologias no Domínio Jurídico: Um Estudo com Node2Vec e OWL2Vec\*\***

**Orientador: Cleyton Mário de Oliveira Rodrigues ([cleyton.rodrigues@upe.br](mailto:cleyton.rodrigues@upe.br))**

A representação computacional de conhecimento jurídico permanece um desafio central para a área de Inteligência Artificial, considerando a complexidade lógica e a ambiguidade das fontes normativas. As ontologias jurídicas se destacam como ferramentas estruturantes de conceitos, axiomas e relações legais, mas sua aplicação em métodos vetoriais ainda é limitada. Avanços recentes em técnicas de graph embeddings como o Node2Vec e em embeddings híbridos como o OWL2Vec\* abrem caminho para explorar a representação semântica de conceitos jurídicos de forma integrada, com potencial para aprimorar tarefas como inferência ontológica, organização de conceitos e análise de coerência semântica.

Este projeto propõe comparar sistematicamente técnicas de embeddings puramente estruturais e híbridos em uma ontologia jurídica de porte intermediário, avaliando sua efetividade em tarefas de classificação de relações, preenchimento ontológico e clustering semântico. Entre os objetivos, destacam-se: (i) revisar a literatura sobre embeddings em ontologias jurídicas; (ii) selecionar e preparar ontologias de referência; (iii) gerar embeddings com Node2Vec, OWL2Vec\* e, complementarmente, RDF2Vec e Onto2Vec; (iv) realizar experimentos com classificadores supervisionados (Random Forests, MLPs) e técnicas de agrupamento (K-Means, DBSCAN); (v) analisar estatisticamente os resultados com métricas como MRR, Silhouette Score e AUC-ROC, complementados por visualizações via t-SNE e UMAP.

A pesquisa visa contribuir com evidências empíricas para orientar o uso de embeddings em tarefas de Ontology Completion, destacando limites, potencialidades e trade-offs entre abordagens. Como produto, será disponibilizada uma pipeline reprodutível, documentação técnica e código aberto, alinhados às diretrizes de ciência aberta. Além de avançar no estado da arte em IA Jurídica, os resultados visam apoiar a construção de sistemas de decisão mais interpretáveis e semanticamente robustos, com aplicação prática para instituições jurídicas e pesquisadores do domínio.

### Referências Bibliográficas:

- BREUKER, J.; VALENTE, A. Knowledge Engineering and Legal Theory. In: *International Conference on Artificial Intelligence and Law*. ACM, 1994.
- CHEN, J. et al. OWL2Vec\*: Embedding of OWL Ontologies. *Machine Learning*, v. 110, n. 7, p. 1813-1845, 2021.
- GANGEMI, Aldo et al. Sweetening ontologies with DOLCE. In: *International conference on knowledge engineering and knowledge management*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. p. 166-181.
- GAO, Shang et al. How legal knowledge graph can help predict charges for legal text. In: *International Conference on Neural Information Processing*. Singapore: Springer Nature Singapore, 2023. p. 408-420.
- GROVER, Aditya; LESKOVEC, Jure. node2vec: Scalable feature learning for networks. In: *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. 2016. p. 855-864.
- PALMIRANI, Monica; CONTISSA, Giuseppe; RUBINO, Rossella. Fill the gap in the legal knowledge modelling. In: *Rule Interchange and Applications: International Symposium, RuleML 2009, Las Vegas, Nevada, USA, November 5-7, 2009. Proceedings 3*. Springer Berlin Heidelberg, 2009. p. 305-314.
- RODRIGUES, C. M. O. et al. Legal Ontologies Over Time: Legal ontologies over time: A systematic mapping study. *Expert Systems with Applications*, v. 130, p. 12-30, 2019.
- SARTOR, Giovanni et al. Argumentation in AI and law. In: *Research Handbook on Legal Argumentation*. Edward Elgar Publishing Ltd., 2025. p. 1-21.
- VAN ENGERS, Tom et al. Ontologies in the legal domain. *Digital Government: E-Government Research, Case Studies, and Implementation*, p. 233-261, 2008.