

Universidade de Pernambuco

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

Proposta de Dissertação de Mestrado

Área: Modelagem Computacional

Título: Engenharia de Requisitos Orientada a Dados para Sistemas de IA

Orientadora: Prof. Dra. Maria Lencastre (mlpmcomp.poli.br)

Este projeto tem como objetivo propor e avaliar um processo de Engenharia de Requisitos centrado nos dados que alimentam sistemas de Inteligência Artificial (IA). A ideia é tratar os conjuntos de dados como artefatos de software com ciclo de vida próprio — incluindo definição, projeto, testes, manutenção e documentação — e integrá-los explicitamente ao processo de requisitos. A motivação decorre do fato de que a qualidade, a transparência e a conformidade dos dados são determinantes para o desempenho, a explicabilidade e a não discriminação dos modelos de IA. Assim, os requisitos de dados devem ser considerados desde o início do projeto, em conjunto com os requisitos funcionais e não funcionais tradicionais de software. As perguntas de pesquisa que orientam o trabalho são: RQ1: Quais tipos de requisitos de dados são atualmente considerados em projetos de IA e como são documentados? RQ2: Como integrar requisitos de dados — como qualidade, representatividade, conformidade legal e plano de monitoramento e manutenção — ao processo de elicitação de requisitos de negócio e de sistema? RQ3: Quais benefícios e custos essa abordagem traz em termos de clareza dos requisitos, qualidade dos conjuntos de treino e percepção de transparência por parte dos envolvidos? A metodologia será desenvolvida em quatro etapas. Primeiro, será realizado um levantamento da literatura sobre requisitos de dados em projetos de IA, com foco em práticas de documentação e gestão. Em seguida, um estudo empírico inicial, por meio de entrevistas ou estudo de caso em uma ou duas organizações, permitirá compreender como os requisitos de dados são tratados hoje. A terceira etapa consiste na proposição da abordagem, com definição de um processo de engenharia de requisitos voltado a dados, incluindo passos de elicitação específicos e modelos de documentos (*templates*) para requisitos de dados, plano de manutenção de conjuntos de treino, política de explicabilidade e requisitos de não discriminação. Por fim, a quarta etapa será a avaliação inicial, realizada por meio da aplicação piloto da abordagem em um projeto acadêmico ou em parceria com uma única organização, até a fase de protótipo de modelo. Essa avaliação envolverá entrevistas e questionários com os participantes para verificar utilidade, clareza de papéis e esforço adicional, além da análise da qualidade dos requisitos de dados em termos de completude, rastreabilidade e antecipação de problemas como vieses ou ausência de métricas claras. As contribuições esperadas incluem a definição de um processo de engenharia de requisitos centrado em dados para sistemas de IA, um conjunto de *templates* para especificação, teste e manutenção de conjuntos de treino, e evidências empíricas iniciais sobre vantagens e limitações dessa integração, incluindo aspectos éticos e legais.

Referências Bibliográficas:

- [1] Kaveen Hiniduma, Suren Byna, and Jean Luca Bez. 2025. Data Readiness for AI: A 360-Degree Survey. *ACM Comput. Surv.* 57, 9, Article 219 (2025), 39 pages. <https://doi.org/10.1145/3722214>
- [2] Habiba, Ue., Haug, M., Bogner, J. *et al.* How mature is requirements engineering for AI-based systems? A systematic mapping study on practices, challenges, and future research directions. *Requirements Eng* 29, 567–600 (2024). <https://doi.org/10.1007/s00766-024-00432->
- [3] Pei, Z., Liu, L., Wang, C., and Wang, J. Requirements Engineering for Machine Learning: A Review and Reflection (2022). DOI:10.48550/arXiv.2210.00859 License CC BY 4.0
- [4] Hutchinson, B., Smart, A., Hanna, A., Denton, R., Greer, C., Kjartansson, O. and Mitchell, M. (2021). Towards Accountability for Machine Learning Datasets: Practices from Software Engineering and Infrastructure. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). ACM, NY, USA, 560–575. <https://doi.org/10.1145/3442188.3445918>