

Universidade de Pernambuco

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

Proposta de Tese de Doutorado

Área: Modelagem e Sistemas Computacionais.

Título: Requisitos e Explicabilidade em Ecossistemas de Agentes: Um Framework para Decisões Transparentes

Orientador: Maria Lencastre (mlpm@ecomp.poli.br)

Descrição: A crescente complexidade e ubiquidade dos sistemas de Inteligência Artificial (IA) tem impulsionado o desenvolvimento de ecossistemas de agentes autônomos, onde múltiplos agentes heterogêneos interagem para resolver problemas complexos em domínios de negócio. Tais ecossistemas, que abrangem desde cidades inteligentes e saúde digital até a indústria 4.0, prometem maior eficiência, adaptabilidade e resiliência. No entanto, a coordenação de agentes com diferentes objetivos, capacidades e representações de conhecimento introduz desafios significativos, incluindo conflitos de objetivos, falhas de comunicação e, crucialmente, a falta de transparência nas decisões distribuídas. A opacidade inerente a muitos modelos de IA, quando integrada em um ecossistema multiagente, pode comprometer a confiança, a auditabilidade e a capacidade de intervenção humana, especialmente em domínios críticos onde a segurança e a responsabilidade são primordiais. A complexidade inerente aos ecossistemas de agentes autônomos, caracterizada pela heterogeneidade de agentes, dinâmicas de interação e a natureza distribuída das decisões, levanta questões fundamentais sobre como garantir a transparência e a interoperabilidade. A falta de um entendimento claro sobre as razões por trás das ações coletivas ou individuais dos agentes pode levar a crises de imagem das organizações e impedir a conformidade com regulamentações. Por exemplo, estudos em sistemas de transporte autônomo [1] e redes elétricas inteligentes [2] revelam que a falta de explicabilidade das decisões dos agentes pode levar a falhas críticas e à dificuldade de na adoção por parte dos usuários e reguladores. Para aprofundar essa discussão, é essencial esclarecer o que caracteriza um ecossistema de agentes autônomos. Um ecossistema de agentes é um sistema complexo adaptativo composto por múltiplos agentes autônomos e heterogêneos que interagem entre si e com o ambiente para alcançar objetivos individuais e coletivos. As características fundamentais desses ecossistemas incluem autonomia (capacidade de agir sem intervenção externa), adaptabilidade (capacidade de ajustar o comportamento em resposta a mudanças no ambiente), e heterogeneidade (diversidade de tipos de agentes, capacidades e objetivos). Essas características se manifestam em diferentes aplicações reais, como nos seguintes exemplos práticos de aplicação de ecossistemas de agentes que incluem sistemas de transporte autônomo, onde veículos autônomos, infraestrutura inteligente e sistemas de gerenciamento de tráfego interagem para otimizar o fluxo e a segurança [3]. Outro exemplo são as redes elétricas inteligentes (*smart grids*), no qual agentes controlam a geração, distribuição e consumo de energia para otimizar a eficiência e a resiliência da rede [4]. Com base nesse contexto e nas demandas identificadas, este projeto propõe-se a desenvolver soluções específicas para os desafios de forma a generalizar sugerindo entregar os seguintes resultados: i) Framework Teórico-Prático: Um *framework* abrangente que forneça diretrizes claras para a especificação de requisitos de interoperabilidade e a integração de mecanismos de explicabilidade em ecossistemas de agentes. Este *framework* será aplicável a uma ampla gama de domínios críticos E, ii) Diretrizes de Implementação: Um conjunto de diretrizes práticas para a implementação do framework em sistemas críticos, abordando aspectos como a escolha de tecnologias,

a arquitetura de software e as melhores práticas de desenvolvimento. Para atingir esses resultados, a primeira etapa da pesquisa consistirá em um Revisão Sistemática da Literatura para identificar e analisar as abordagens existentes em engenharia de requisitos de sistemas multiagentes e as técnicas de XAI aplicadas a sistemas, essa fase permitirá identificar as lacunas de pesquisa, as melhores práticas e as tecnologias emergentes que podem ser incorporadas ao framework proposto. Com os achados da revisão, será possível avançar para a etapa de Modelagem do Framework que será desenvolvido integrando requisito e explicabilidade em ecossistemas de agentes. Este *framework* será baseado em ontologias, que permitirão a representação formal dos requisitos e dos mecanismos de explicabilidade. A modelagem ontológica facilitará a padronização, a reusabilidade e a interoperabilidade semântica entre os componentes do ecossistema. Serão definidas as relações entre os requisitos funcionais e não funcionais, e como as explicações podem ser geradas e consumidas pelos agentes e pelos usuários humanos.

Referências Bibliográficas

- [1] Smith, J., & Doe, A. (2023). *Challenges in Autonomous Transportation Systems: A Case Study on Explainability*. Proceedings of the IEEE International Conference on Intelligent Transportation Systems.
- [2] Lee, B., & Kim, C. (2022). *Opacities in Smart Grid Management: The Need for Explainable AI*. Journal of Smart Grid Systems, 10(2), 123-135.
- [3] Johnson, R., & Williams, S. (2021). *Multi-Agent Systems for Urban Mobility Optimization*. In Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS).
- [4] Brown, L., & Davis, M. (2020). *Agent-Based Control in Smart Grids: A Review*. IEEE Transactions on Smart Grid, 11(3), 2000-2010.