

Universidade de Pernambuco

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

Proposta de Doutorado

Área: Engenharia de Software

Título: Gerenciamento e Aperfeiçoamento da Qualidade de Software em Sistemas de IA Generativa

Orientador: Ivaldir Honório de Farias Júnior (ivaldir.farias@upe.br)

A Inteligência Artificial Generativa (GenIA), diferentemente da IA tradicional que foca apenas na análise dados existentes pré-existentes, consiste em uma tecnologia que abrange sistemas capazes de criar ou gerar saídas novas e criativas, tais como textos, imagens, vídeos e códigos, através da utilização de modelos generativos, como os LLMs. Os sistemas de GenIA são comumente utilizados em áreas como entretenimento, marketing, medicina, educação, engenharia de software, entre outras (Dandotiya, Subham, 2025). Sua aplicação promove a automatização de tarefas que normalmente demandariam considerável esforço humano, tais como modelagem de requisitos de software, codificação, criação de testes, leitura, escrita e tradução de textos, edição de vídeos, criação de conteúdo para redes sociais, etc.

Entretanto, conforme abordado por Dandotiya e Subham (2025), apesar dos benefícios oferecidos pelos sistemas baseados em IA generativa como o ChatGPT, Gemini, Manus e DeepSeek, também se faz necessário considerar suas limitações atuais, tais como problemas éticos, sociais e de segurança de dados. Para além dos riscos éticos e sociais, há também desafios técnicos significativos no contexto do desenvolvimento de software, pois os artefatos gerados automaticamente, código-fonte, casos de teste ou documentos de requisitos, podem apresentar baixa rastreabilidade e confiabilidade, dificuldades de reprodutibilidade e ausência de critérios claros de validação. Em especial, os testes gerados através de LLMs, de acordo com Huang et al. (2023) e Ribeiro et al. (2024), frequentemente apresentam problemas de manutenção, ambiguidade e erros lógicos, o que afeta negativamente a confiabilidade dos sistemas finais.

No que diz respeito à qualidade de softwares que integram IA generativa, destaca-se a necessidade de novas abordagens de teste distintas das tradicionais. Pois, o caráter não determinístico dos resultados, a falta de interpretabilidade e a dependência de dados de treinamento tornam mais complexas tarefas clássicas de garantia de qualidade de software, como validação de resultados, definição de métricas objetivas e o controle de versões. Isto é, a qualidade das saídas geradas pelos sistemas de GenIA depende de fatores que extrapolam o controle tradicional, exigindo estratégias mais adequadas ao comportamento desses sistemas.

Considerando o contexto apresentado, observa-se **a necessidade de se investigar maneiras de aperfeiçoar as abordagens de qualidade já existentes ou propor novas estratégias capazes de melhorar o gerenciamento da qualidade de softwares modernos baseados em IA generativa.** Trabalhar sobre essas lacunas é relevante não apenas sob a ótica da inovação científica, mas também do ponto de vista organizacional e econômico, dado que busca garantir uma adoção responsável e segura da IA generativa em ambientes corporativos.

Referências Bibliográficas:

DANDOTIYA, SUBHAM. Generative AI for software testing: harnessing large language models for automated and intelligent quality assurance. International Journal of Science and Research Archive. 14. 1931-1935. 10.30574/ijrsra.2025.14.1.0266. Disponível em: <https://www.researchgate.net/publication/388661350>. Acesso em: 19 maio 2025.

HUANG, T. et al. Software testing with large language models: survey, landscape, and vision. arXiv, 2023. Disponível em: <https://arxiv.org/pdf/2307.07221.pdf>. Acesso em: 19 maio 2025.

RIBEIRO, D. et al. Software testing of generative AI systems: challenges and opportunities. arXiv, 2023. Disponível em: <https://arxiv.org/pdf/2309.03554.pdf>. Acesso em: 19 maio 2025.