

Uso da inteligência artificial generativa em testes de software

Josiane Pietrobon
Lucas Moreira



Software Testing With Large Language Models: Survey, Landscape, and Vision

Junjie Wang¹, Member, IEEE, Yuchao Huang², Chunyang Chen³, Zhe Liu²,
Song Wang², Member, IEEE, and Qing Wang², Member, IEEE

Abstract—Pre-trained large language models (LLMs) have recently emerged as a breakthrough technology in natural language processing and artificial intelligence, with the ability to handle large-scale datasets and exhibit remarkable performance across a wide range of tasks. Meanwhile, software testing is a crucial undertaking that serves as a cornerstone for ensuring the quality and reliability of software products. As the scope and complexity of software systems continue to grow, the need for more effective software testing techniques becomes increasingly urgent, making it an area ripe for innovative approaches such as the use of LLMs. This paper provides a comprehensive review of the utilization of LLMs in software

I. INTRODUCTION

SOFTWARE testing is a crucial undertaking that serves as a cornerstone for ensuring the quality and reliability of software products. Without the rigorous process of software testing, software enterprises would be reluctant to release their products into the market, knowing the potential consequences of delivering flawed software to end-users. By conducting thorough and meticulous testing procedures, software enterprises can minimize the occurrence of critical software failures, usability issues, or security breaches that could potentially lead to

Compreender como a inteligência artificial generativa está sendo utilizada na área de testes de software.

1. Definição das Perguntas de Pesquisa:

RQ1. Como a inteligência artificial generativa está sendo aplicada nos testes de software?

RQ2. Quais os desafios encontrados na adoção da inteligência artificial generativa nos testes de software?

RQ3: Quais os benefícios que a inteligência artificial generativa trouxe para os testes de software?

2. Estratégia de Busca:

Realizar buscas nas bases de dados

IEEE Xplore, ScienceDirect, Scopus, ACM, DBLP e arXiv.

String de busca: ("software test" OR "software testing") AND ("Large Language Model" OR "Generative artificial intelligence" OR "Generative AI")

Filtro de data: nov/2023 a ago/2024.

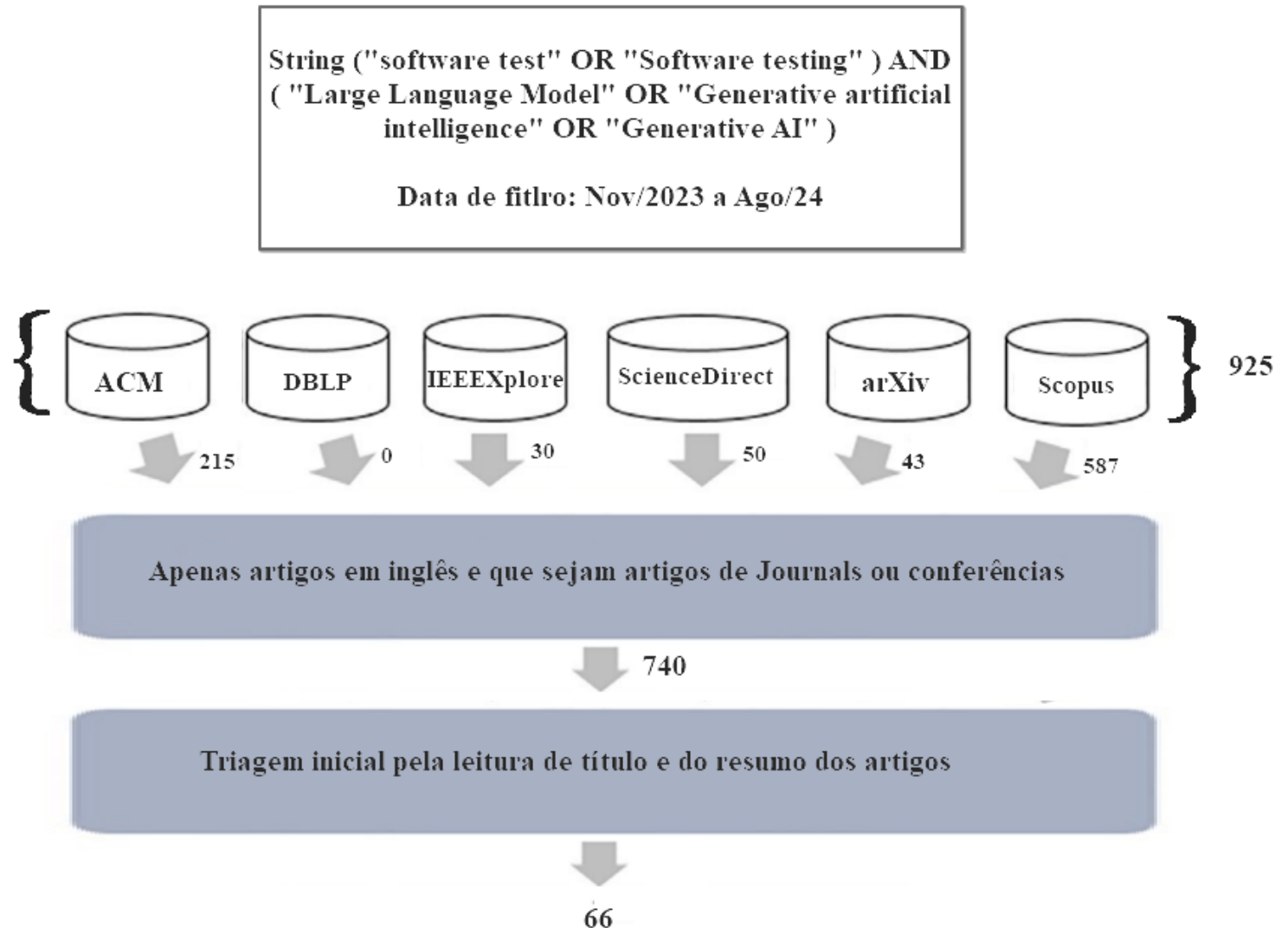
3. Critérios de Seleção de Busca:

Incluir apenas artigos em inglês, que foram publicados em revistas ou conferências.

Realizar a triagem inicial pela leitura do título e do resumo dos artigos.

Método

Detalhamento do processo de filtragem dos artigos



5. Extração e Análise de Dados:

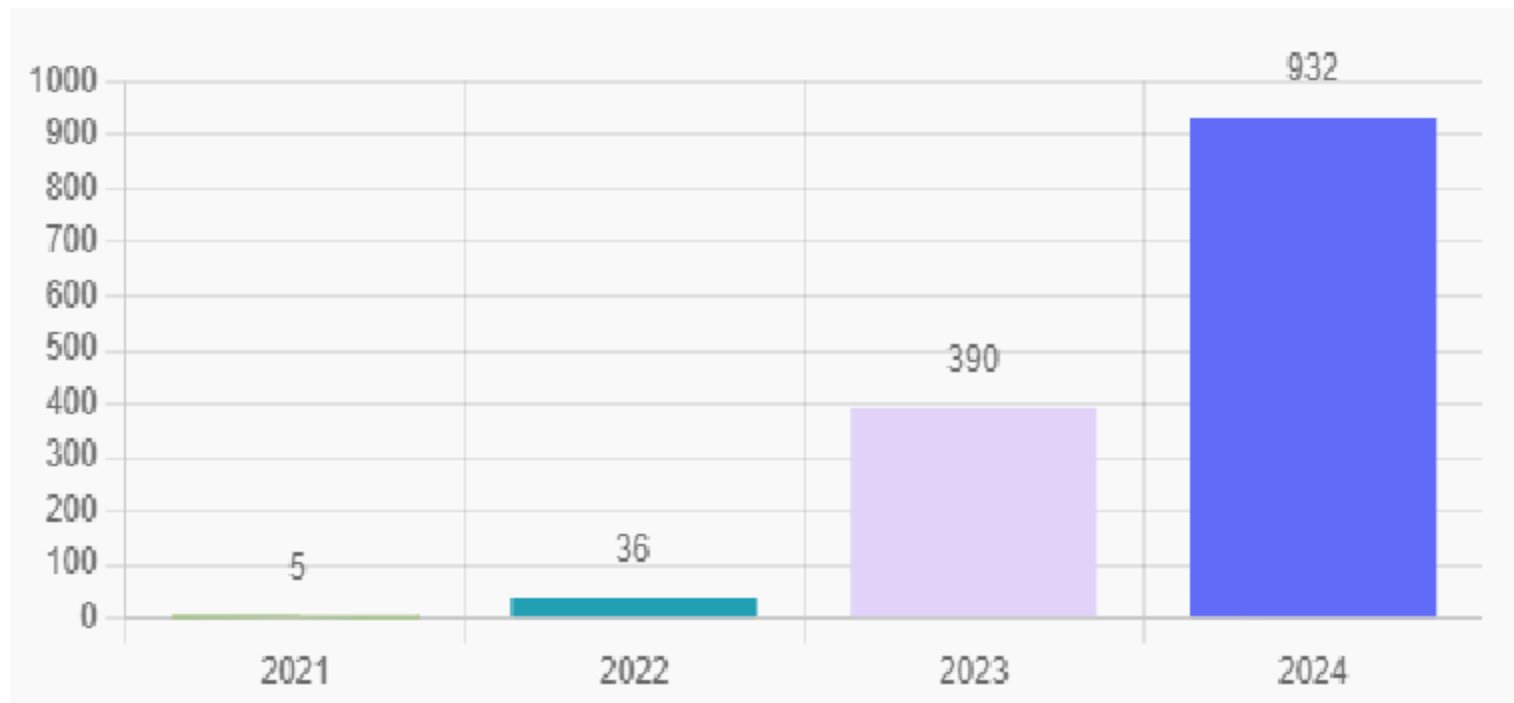
- Extrair dados relevantes de utilização de LLMs em testes de software.
- Organizar e analisar os dados para responder às perguntas de pesquisa.

6. Síntese dos Resultados:

- Sintetizar os resultados para identificar as tendências de utilização de LLM na área de testes de software.

Resultados

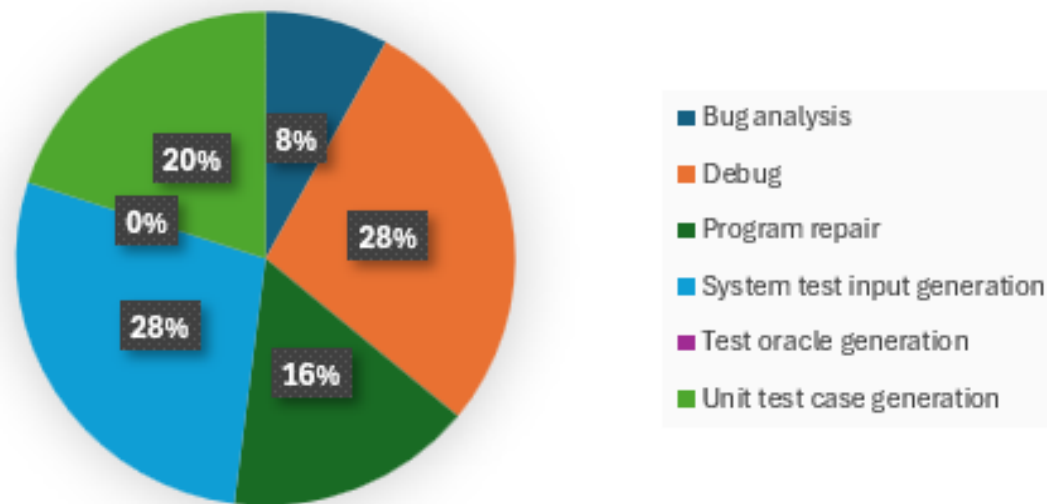
Evolução das pesquisas sobre LLMs em testes de software



Resultados

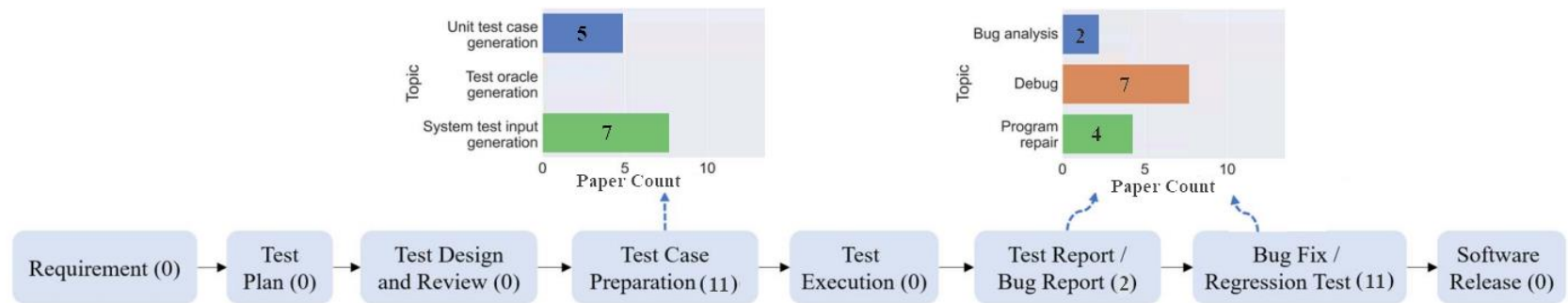
1. Como a inteligência artificial generativa está sendo aplicada nos testes de software?

Distribuição de LLMs por tópicos



Resultados

Distribuição de tarefas de testes com LLMs



Distribuição de tarefas de teste com LLMs (alinhadas com o ciclo de vida de testes de software, o número entre colchetes indica o número de estudos coletados por tarefa, e um artigo pode envolver múltiplas tarefas)

2. Quais os desafios encontrados na adoção da inteligência artificial generativa nos testes de software?

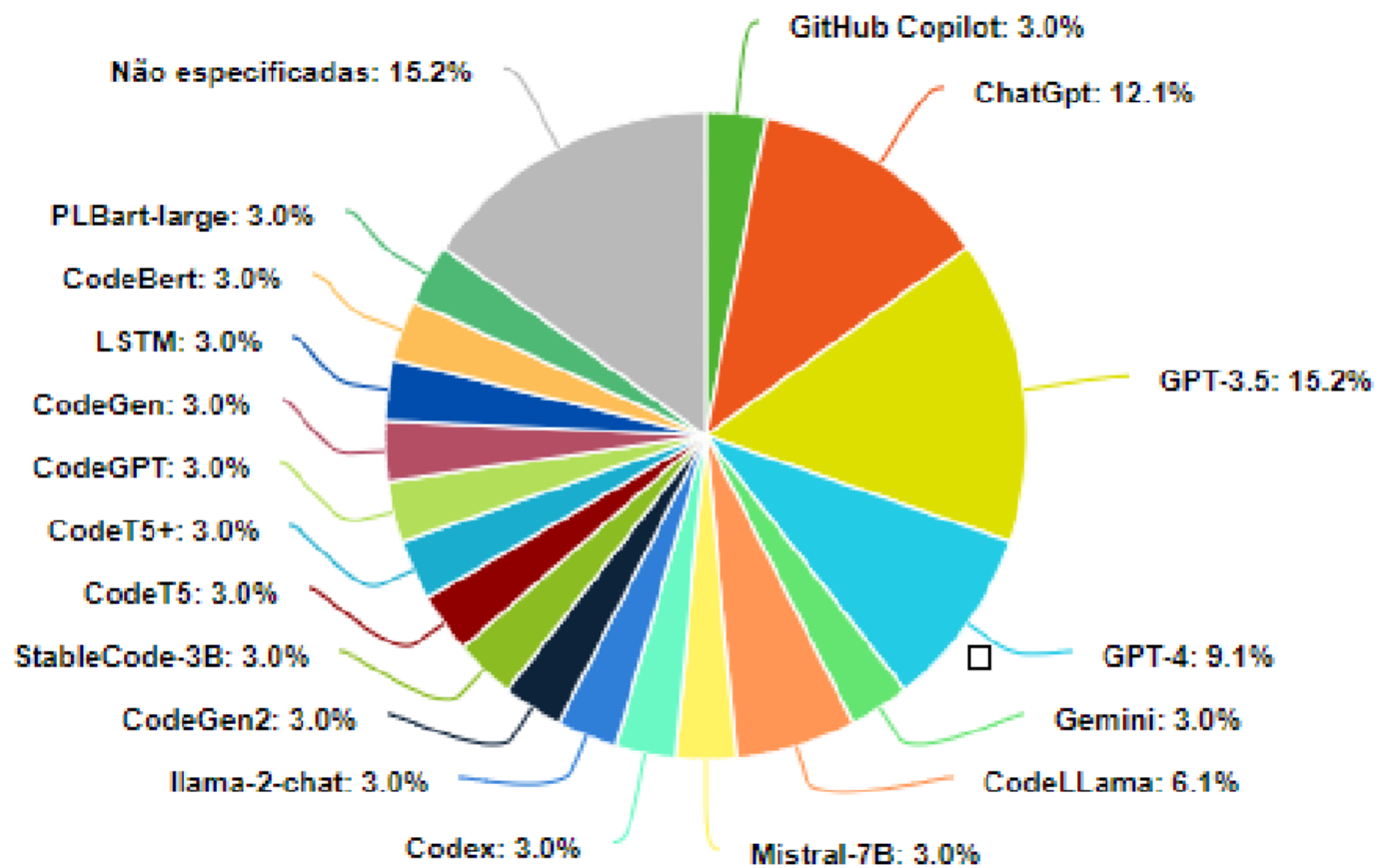
- Alucinações
- Escolha da LLM a ser utilizada
- Lógicas complexas

3. Quais os benefícios que a inteligência artificial generativa trouxe para os testes de software?

- Aumento na cobertura dos casos de teste.
- Descoberta mais rápida de bugs.
- Redução do tempo de desenvolvimento de software

Discussão dos Resultados

LLMs Utilizadas





Conclusão

- Na análise da quantidade de artigos publicados ano a ano, é possível verificar um aumento de 139% em comparação com o ano anterior (2023).
- O foco principal das pesquisas se restringe a algumas etapas do ciclo de testes de software, como a preparação de casos de teste e o debug.
- O uso de LLMs em testes de software traz, como principal benefício, a redução no tempo de desenvolvimento do produto.



Conclusão - Ameaça a validade

- Esta pesquisa foi realizada em um curto período de tempo de tempo, e, por isso, foram considerados apenas os 20 artigos mais recentes nos resultados apresentados.
- A base de dados da ACM foi utilizada para a construção desse ; entretanto, apenas os artigos de acesso público foram considerados, devido à indisponibilidade temporária do acesso à ACM no Brasil.
- Não foi feito a revisão por pares para essa apresentação.



Conclusão - Trabalhos futuros

- Pretendemos expandir a pesquisa, incluindo a análise dos demais artigos, com o objetivo de realizar uma publicação em conferência.
- Nos artigos lidos percebemos que ainda há espaço para pesquisa sobre como a LLM está impactando em etapas do ciclo de teste que ainda não foram exploradas.