

# LLMs, Engenharia de Requisitos e Simulação: Desafios e Oportunidades de Pesquisa

Kauan Oliveira Perdigão Lopes, Lucas Moura, Emanuel Ferreira Coutinho

Universidade Federal do Ceará (UFC) – Quixadá – CE – Brasil

kauanopl@alu.ufc.br, lucasmoura07@alu.ufc.br, emanuel.coutinho@ufc.br

**Abstract.** *Requirements Engineering (RE) is an essential discipline in the software development process to gather stakeholders' needs. Large-Scale Language Models (LLMs) have been used to automate and optimize this process, reducing ambiguities and inconsistencies. This research explores the applications of LLMs in RE, such as automatic elicitation, classification of requirements, generation of specifications and impact analysis of changes. It also discusses challenges such as response accuracy, data security and integration with external tools, as well as possible solutions with specialized models.*

**Resumo.** *A Engenharia de Requisitos (ER) é uma disciplina essencial no processo de desenvolvimento de software para coletar as necessidades dos stakeholders. Os Modelos de Linguagem de Larga Escala (LLMs) vêm sendo usados para automatizar e otimizar esse processo, reduzindo ambiguidades e inconsistências. Esta pesquisa explora as aplicações dos LLMs na ER, como a elicitação automática, classificação dos requisitos, geração de especificações e análise do impacto de mudanças. Também discute desafios como a precisão das respostas, a segurança dos dados e a integração com ferramentas externas, além de possíveis soluções com modelos especializados.*

## 1. Introdução

A Engenharia de Requisitos (ER) é uma das etapas mais importantes do desenvolvimento de software, definindo funcionalidades e restrições que um sistema deve atender. Essa fase abrange a elicitação, análise, especificação, validação e gestão de requisitos, atividades fundamentais para garantir que as necessidades dos *stakeholders* sejam corretamente compreendidas em um produto de qualidade [Kotonya and Sommerville 1998]. No entanto, a ER enfrenta desafios, como ambiguidade, inconsistência, incompletude e dificuldades na comunicação entre desenvolvedores e usuários [Mohan and Chenoweth 2011].

Com a crescente complexidade dos sistemas e a necessidade de coletar grandes volumes de dados para embasar decisões de requisitos, tem se tornado essencial a aplicação de tecnologias inteligentes para apoiar esse processo. A ER depende de processos manuais e da experiência dos engenheiros de requisitos para identificar e documentar corretamente as necessidades dos usuários. No entanto, erros nessa fase podem ter grandes impactos no desenvolvimento, resultando em retrabalho, aumento de custos e falhas nos sistemas desenvolvidos [Ferreira et al. 2022]. E nesse cenário, utilizar simulações podem melhorar muito o desempenho das atividades de ER.

O avanço dos *Large Language Models* (LLMs), como o GPT-4, trouxe novas possibilidades para a automação e otimização da Engenharia de Software (ES), incluindo a

ER. Esses modelos são treinados em grandes quantidades de dados textuais, capazes de compreender, gerar e refinar textos com um bom nível de coerência e contexto [Ferrari et al. 2017]. Com capacidade de processar linguagem natural, os LLMs apresentam um grande potencial para auxiliar na elicitação, análise e validação de requisitos, reduzindo ambiguidades e inconsistências que frequentemente surgem em documentos de requisitos. Entre as possíveis aplicações dos LLMs na ER, destacam-se a extração automática de requisitos a partir de documentos e transcrições de reuniões, detecção de ambiguidade e inconsistências nos requisitos especificados e geração automática da documentação [Rocha et al. 2023]. Ademais, os LLMs podem atuar como assistentes inteligentes para analistas de requisitos, sugerindo reformulações e verificando se os requisitos estão de acordo com normas e padrões estabelecidos [Menegazzi and Silva 2023]. Tais aplicações podem trazer pontos positivos para o desenvolvimento de software, tornando o processo mais eficiente, preciso e menos propenso a falhas humanas.

Porém, o uso dos LLMs na ER ainda apresenta desafios importantes. Questões como tendências ruins dos modelos, segurança, privacidade dos dados e a integração desses modelos com ferramentas de Engenharia de Software precisam ser discutidas para garantir que a aplicação seja confiável e eficaz [Ferreira et al. 2022]. Ademais, é necessário que esses modelos sejam adaptados ao contexto específico da ER, considerando conceitos técnicos e as necessidades de cada tipo de aplicação [Ferrari et al. 2017]. Desse modo, esse trabalho tem como objetivo explorar as oportunidades de pesquisa relacionadas ao uso de LLMs na ER, analisando como essas tecnologias podem otimizar as principais atividades desse processo e discutir os desafios que ainda precisam ser superados.

## 2. Aplicação de LLMs na ER e Possíveis Impactos

LLMs vêm sendo aplicados na ER para melhorar processos e estruturar informações em linguagem natural, otimizando atividades como a redução de inconsistências, melhorando a clareza dos requisitos [Mohan and Chenoweth 2011]. Alguns desses processos são exemplificados a seguir.

**Elicitação e Refinamento de Requisitos:** A elicitação dos requisitos de software envolve comunicação entre *stakeholders* e analistas, e como consequência dessa comunicação, requisitos ambíguos ou incompletos podem vir à tona devido a dificuldades na formulação dos requisitos do sistema e das necessidades dos usuários [Ferrari et al. 2017]. Os LLMs podem ser utilizados para automatizar a extração de requisitos a partir de documentos técnicos, atas de reuniões e pela comunicação entre as equipes. Além disso, podem auxiliar na reformulação de requisitos redundantes, gerando descrições mais claras e objetivas.

**Geração e Classificação Automática de Requisitos:** A divisão dos requisitos entre funcionais (RFs) e não funcionais (RNFs) é essencial para construir a documentação do sistema e definir as características [Zhou et al. 2022]. Os LLMs podem automatizar essa divisão, garantindo mais estabilidade e confiança. Além disso, os modelos podem converter as descrições textuais em documentos mais estruturados, permitindo a geração desses documentos técnicos e também a conversão dos requisitos presentes nesses documentos em diagramas UML ou BPMN para melhorar a visualização da arquitetura do sistema.

**Geração e Refinamento de User Stories:** *User Stories* representam requisitos na visão do usuário, mas a criação delas pode variar com a experiência da equipe, comprometendo critérios como clareza e coerência [Arora et al. 2015]. LLMs podem auxiliar na geração

clara e coerente das *User Stories*, para seguirem um padrão claro e coerente incluindo informações como descrição do objetivo do usuário e critérios de aceitação. Isso facilita a comunicação entre equipes de desenvolvimento, melhorando e otimizando a qualidade e confiança dos requisitos e do *backlog* do produto ao longo do projeto.

**Validação e Análise do Impacto de Mudanças:** Incluir novos requisitos ou modificar existentes pode gerar impactos em outras partes do sistema. Os LLMs podem ser utilizados para identificar dependências entre os requisitos, prever os impactos das mudanças e sugerir reformulações para minimizar os problemas. Ademais, esses modelos podem auxiliar na geração automática de testes para novos requisitos serem implementados sem comprometer as funcionalidades existentes. Portanto, podem contribuir para reduzir o re-trabalho e aumentar a confiabilidade do sistema durante o desenvolvimento do software.

Tais aplicações podem se beneficiar de simulação. A simulação na ER pode auxiliar na geração, refinamento e validação. Usar LLMs pode reduzir custos e acelerar o processo de ER, sendo alinhado com vantagens de usar simulação. A avaliação de alternativas e o impacto em ER também pode ser bem trabalhada com simulação.

### 3. Desafios e Limitações

Apesar das vantagens dos LLMs na ER, a implementação ainda enfrenta desafios técnicos e conceituais que precisam ser superados para garantir sua confiabilidade e aplicabilidade.

**Dispersão e Qualidade das Respostas:** LLMs são treinados em grandes bases de dados textuais, o que pode resultar em respostas dispersas ou imprecisas [Ferrari et al. 2017]. Em cenários críticos, como a definição de requisitos técnicos, esse problema pode comprometer a qualidade e a precisão da documentação. O treinamento de modelos próprios para ER com bases de requisitos reais é uma forma de otimização e mitigação do problema para esse cenário específico.

**Segurança e Privacidade dos Dados:** A criação de requisitos pode envolver informações sensíveis e confidenciais, como dados estratégicos de empresas ou regulamentações do setor. O uso dos LLMs hospedados em nuvens públicas possui riscos de vazamento e pode não estar em conformidade com legislações, como a LGPD. A adoção dos LLMs locais pode ser uma solução para garantir que os dados sejam processados internamente, sem exposição a sistemas e pessoas externas, reduzindo os riscos de vazamento.

**Integração com Ferramentas da Engenharia de Requisitos:** Para que os LLMs sejam aplicáveis à ER, precisam ser integrados a ferramentas específicas, como JIRA e IBM DOORS. Porém, a maioria dos modelos disponíveis operam de forma independente, sem compatibilidade. A criação de APIs específicas para essa integração seria uma alternativa para incorporar os LLMs ao fluxo de trabalho das equipes, permitindo que os modelos atuem diretamente nas plataformas utilizadas pelas organizações, sendo confiáveis e mais consolidadas pelos usuários internos.

**Dependência da Qualidade dos Dados de Treinamento:** A eficácia dos LLMs depende da qualidade e diversidade dos dados de treino. Treinamentos com dados genéricos podem não capturar especificidades da ER por serem dispersos e não focados na ER, levando a sugestões pouco precisas ou fora de contexto [Mohan and Chenoweth 2011]. O uso de datasets focados para ER com requisitos reais de diferentes domínios, pode aumentar a precisão e a confiabilidade das respostas fornecidas.

## 4. Possíveis Soluções

A adoção dos LLMs na ER ainda enfrenta desafios, mas alguns avanços podem tornar sua aplicação mais eficiente e confiável.

**Modelos Especializados para ER:** O uso de LLMs genéricos, como o GPT-4, é poderoso, mas eles não foram treinados para uso específico na ER. O treinamento de LLMs com datasets focados em ER pode aumentar a precisão das respostas, garantindo que os modelos compreendam padrões, termos técnicos e boas práticas da ER. Essas práticas podem melhorar a precisão das respostas, reduzir a geração de informações irrelevantes ou genéricas e facilitar a identificação da ambiguidade, inconsistência e incompletude dos requisitos [Zhao et al. 2023].

**Integração com Ferramentas de ER:** Implementar APIs compatíveis com ferramentas de gestão pode permitir a automação da documentação e geração de requisitos a partir de linguagem natural, além de sugestões de melhoria nas plataformas utilizadas por equipes de desenvolvimento, permitindo que LLMs sejam diretamente parte do trabalho ao invés de serem utilizados apenas como ferramentas externas [Gu et al. 2023].

**Segurança e Conformidade com Regulamentações:** A adoção de LLMs locais é uma alternativa mais segura para lidar com documentos de requisitos com políticas de segurança robustas, pois os documentos podem conter informações sensíveis, viabilizando a conformidade com as regulamentações como a LGPD, reduzindo os riscos de exposição dessas informações sensíveis.

## 5. Engenharia de Requisitos com LLMs: Simulações de Cenários de Uso

A aplicação de Inteligência Artificial (IA) na ER permite automatizar atividades como elicitacão, validacão e análise de impacto [Zhou et al. 2022]. Esses avanços abrem espaço para simulações que representam interações, mudanças e comportamentos do sistema, favorecendo uma compreensão mais precisa desde as etapas iniciais. Nesse cenário, simulações com apoio de LLMs ampliam os benefícios da IA na ER, permitindo validar antecipadamente a clareza, completude e impacto dos requisitos. A seguir, são apresentadas sete categorias de simulações aplicáveis à ER com o uso de LLMs.

**1. Simulação de Diálogos:** Simular interações entre usuários e sistemas, sendo útil nas fases iniciais de elicitacão. LLMs geram conversas plausíveis com base em personas, antecipando dúvidas e revelando necessidades implícitas. **2. Simulação de Requisitos de Software:** Simula a coleta, análise e documentacão de requisitos, com foco em ambiguidade, redundância e inconsistências. LLMs comparam versões, sugerem reformulações e verificam conformidade com padrões. **3. Simulação de Casos de Uso:** Reproduz cenários funcionais do sistema, testando o comportamento previsto. Permite avaliar a completude e coerência dos requisitos. LLMs auxiliam ao gerar fluxos e convertê-los em representações formais. **4. Simulação de Testes de Usabilidade:** Avalia como diferentes perfis interagem com a interface. Simulações antecipam problemas de acessibilidade e navegacão. LLMs geram scripts representando diversos níveis de familiaridade do usuário. **5. Simulação de Modelagem de Sistemas:** Representa o comportamento técnico do sistema, como desempenho, escalabilidade e confiabilidade. LLMs sugerem arquiteturas e respondem a questionamentos sobre requisitos não funcionais. **6. Simulação de Análise de Impacto:** Avalia efeitos de mudanças em requisitos sobre o

restante do sistema. LLMs mapeiam dependências e sugerem ajustes para preservar a integridade e reduzir retrabalho. **7. Simulação de Geração de Documentação:** Automatiza a criação de artefatos como requisitos formais, especificações técnicas e user stories. LLMs transformam descrições naturais em documentos padronizados e consistentes.

## 6. Considerações Finais

O objetivo dessa pesquisa é destacar o potencial de automatizar e aprimorar a ER com o uso de LLMs, tornando a elicitação, documentação e validação mais eficientes, possivelmente com o apoio de simulações. Desafios como a precisão nas respostas geradas e a validação, a segurança das informações sensíveis e a falta de integração ainda precisam ser superados. Modelos especializados, integração com ferramentas da ER e mais transparência nas respostas são avanços que podem tornar viável um uso mais seguro e confiável. Contudo, os LLMs podem ser essenciais para otimizar a ER e o desenvolvimento de software, e com simulações seu desempenho e qualidade podem ser melhores.

## Referências

- Arora, C., Sabetzadeh, M., et al. (2015). Automated extraction and formalization of legal norms in requirements documents. *IEEE Transactions on Software Engineering*, 41(12):1103–1127.
- Ferrari, A., Gnesi, S., and Spagnolo, G. O. (2017). Detecting ambiguity in requirements engineering using nlp techniques. *Requirements Engineering*, 22(3).
- Ferreira, L., Okano, M. T., Aguiar, F., De Castro Lobo dos Santos, H., and Ursini, E. L. (2022). A panorama of the implementation of the general law for the protection of personal data (lgpd) in brazil: An exploratory survey. In *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)*.
- Gu, Y., Shu, Y., Yu, H., Liu, X., Dong, Y., Tang, J., Srinivasa, J., Latapie, H., and Su, Y. (2023). Middleware for llms: Tools are instrumental for language agents in complex environments. *arXiv preprint arXiv:2402.14672*.
- Kotonya, G. and Sommerville, I. (1998). *Requirements Engineering: Processes and Techniques*. Wiley Publishing.
- Menegazzi, D. and Silva, C. (2023). Conformidade com a lgpd por meio de requisitos de negócio e requisitos de solução. In *WER23 - Workshop em Engenharia de Requisitos*.
- Mohan, S. and Chenoweth, S. (2011). Teaching requirements engineering to undergraduate students. In *Proceedings of the 42nd ACM Technical Symposium on Computer Science Education, SIGCSE '11*.
- Rocha, L. D., Silva, G. R. S., and Canedo, E. (2023). Privacy compliance in software development: A guide to implementing the lgpd principles. In *Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing, SAC '23*.
- Zhao, W. X., Zhu, J., Wang, D., Hou, Y., Chang, B., Li, J., Wang, X., Zhou, K., Zhou, Z., and Wen, J.-R. (2023). A survey of large language models. *arXiv preprint arXiv:2303.18223*.
- Zhou, J., Wang, Y., et al. (2022). Requirements engineering in the era of ai: Opportunities and challenges. *ACM Transactions on Software Engineering and Methodology*.