

Uso do ChatGPT na Priorização de Requisitos: Uma Experiência Educacional em Engenharia de Software

Francisca Karolina Queiroz¹, Márcia Sampaio Lima¹

¹Universidade do Estado do Amazonas (UEA/EST)
Av. Darcy Vargas, 1.200– 69050-020 – Manaus – AM – Brasil

fkdnq.lic21@uea.edu.br, msllima@uea.edu.br

Abstract. *The teaching of Software Engineering (SE) has been transformed by Generative Artificial Intelligence (AI), expanding possibilities in the teaching-learning process of computing. In SE, AI supports various software development tasks. This study reports an experience with 16 students using ChatGPT to prioritize software requirements for the MVP (Minimum Viable Product). Initially, the students elicited and prioritized the software requirements for MVP implementation. Subsequently, they used ChatGPT to review the prioritization and propose new requirements. ChatGPT corroborated 86.02% of the students' choices. Despite acknowledging the agility provided, the students reported needing more confidence in the AI's results.*

Resumo. *O ensino de Engenharia de Software (ES) tem sido transformado pela Inteligência Artificial Generativa (IA-Generativa), ampliando possibilidades no processo ensino-aprendizagem de computação. Na ES, as IAs apoiam várias tarefas de desenvolvimento de software. Este trabalho relata uma experiência com 16 alunos utilizando o ChatGPT para priorizar requisitos de software, visando o MVP (Produto Mínimo Viável). Primeiramente, os alunos elicitaram e priorizaram os requisitos de software implementados no MVP. Em seguida, usaram o ChatGPT para revisar a priorização e propor novos requisitos. O ChatGPT corroborou 86,02% das escolhas dos alunos. Apesar de reconhecerem agilidade proporcionada, os alunos relataram falta de confiança nos resultados da IA.*

1. INTRODUÇÃO

O ensino de Engenharia de Software (ES) tem enfrentado uma transformação significativa com o advento da Inteligência Artificial Generativa (IA-Generativa), representada por ferramentas como o ChatGPT¹ e o Gemini² [Petrovska et al. 2024]. Essa evolução tecnológica abre novas possibilidades para apoiar o aprendizado de ES, oferecendo vantagens que complementam métodos tradicionais [Mittal et al. 2024]. Modelos de Linguagem em Larga Escala (LLM), uma categoria da IA-Generativa, estão sendo rapidamente adotados por desenvolvedores de software e aplicados para gerar código [Sun et al. 2022] e outros artefatos relacionados à ES [Beganovic et al. 2023]. Tais modelos são treinados com grandes quantidades de dados, permitindo a geração de textos e códigos. Exemplos populares de ferramentas baseadas em LLM incluem ChatGPT, o Gemini e GitHub Copilot³. Essas ferramentas de inteligência artificial (IA) têm sido úteis em uma série de tarefas comuns de desenvolvimento e ES [White et al. 2024].

¹<https://chatgpt.com/>

²<https://gemini.google.com/>

³<https://copilot.github.com>

Uma das atividades cruciais no processo de desenvolvimento de software se destaca a concepção do *Minimum Viable Product* (MVP). O MVP é definido como uma versão inicial de um produto que permite a coleta do máximo possível de aprendizado validado sobre os clientes com o menor esforço necessário [Ries 2011]. Esse conceito, originado no contexto do *Lean Startup*, tem como objetivo principal viabilizar experimentação e aprendizado rápido, reduzindo riscos e otimizando recursos durante o desenvolvimento do produto. O MVP é projetado não apenas para entregar valor ao cliente, mas também para testar hipóteses de negócio e gerar *feedback* que guie as próximas etapas de desenvolvimento [Melegati et al. 2020, Alonso et al. 2023].

No entanto, um desafio específico dentro do contexto da ES reside no processo de ensino-aprendizagem de identificação dos requisitos prioritários a serem implementados no MVP devido à complexidade intrínseca da atividade, que exige habilidades analíticas e estratégicas, muitas vezes não plenamente desenvolvidas em alunos. Sendo assim, ensinar a priorização de requisitos visando o desenvolvimento do MVP é fundamental, pois a priorização de requisitos permite que equipes foquem no que é mais relevante para os objetivos do negócio e para o valor percebido pelos usuários, garantindo o uso eficiente de recursos e tempo. Dessa forma, ensinar a priorização de requisitos não apenas aprimora a formação técnica, mas também capacita os futuros engenheiros de software a lidar com a complexidade e as demandas do mercado atual.

Sendo assim, o objetivo deste trabalho é relatar uma experiência conduzida em sala de aula para investigar como o ChatGPT pode ser utilizado como ferramenta de suporte na definição dos requisitos prioritários no desenvolvimento do MVP.

Para tanto, uma experiência, com 16 alunos de graduação matriculados na disciplina Engenharia de Software, foi executada em duas etapas distintas. Na primeira, os alunos foram divididos em equipes e realizaram a elicitação de requisitos e a priorização de requisitos de sistemas de software cujos contextos foram escolhidos pelas próprias equipes. Na segunda etapa, as mesmas equipes utilizaram o ChatGPT para validar os requisitos anteriormente selecionados e para gerar novos requisitos prioritários, baseando-se em *prompts* previamente estabelecidos. Por fim, cada estudante registrou suas percepções sobre o uso do ChatGPT no processo de elicitação de requisitos, destacando vantagens e desvantagens do uso da IA na priorização de requisitos de software visando o MVP.

Os resultados deste estudo demonstraram que o ChatGPT endossou, em média, 86,02% dos requisitos priorizados pelas equipes para o MVP, indicando concordância com as escolhas realizadas pelos alunos. Além disso, o ChatGPT gerou novos requisitos, dos quais 72,5% foram considerados inéditos pelas equipes. Apesar das contribuições, os participantes apontaram a necessidade de supervisão humana e de maior confiança nos resultados gerados pela IA.

2. TRABALHOS RELACIONADOS

No contexto da ES, a área de Engenharia de Requisitos (ER) tem explorado as aplicações de IA, como o ChatGPT, para auxiliar em diversas tarefas [Mendonça et al. 2021]. Marques *et al.* conduziram uma revisão da literatura sobre o uso do ChatGPT em ER, destacando seus benefícios na elicitação de requisitos, como a melhoria da comunicação entre os *stakeholders* e a precisão na captura das necessidades dos usuários [Marques et al. 2024]. Sampaio *et al.* realizaram uma experiência com estudantes uti-

lizando o ChatGPT 3.5 para apoiar a eliciação de requisitos. Os resultados mostraram que a ferramenta gerou requisitos inovadores e relevantes, demonstrando potencial do ChatGPT para auxiliar nesse processo [Sampaio et al. 2024a].

A aplicação de IA na educação em ES também tem sido investigada. Daun e Brings discutem o potencial do ChatGPT para transformar a educação em ES, destacando seus benefícios, como a possibilidade de fornecer *feedback* personalizado aos alunos e a necessidade de adaptar os currículos para refletir essas novas ferramentas [Daun and Brings 2023]. É importante considerar as limitações e desafios do uso de LLM em ER. Arora *et al.* alertam para a necessidade de *prompts* bem definidos para obter resultados desejados do ChatGPT e apontam que a ferramenta pode gerar requisitos amplos e, por vezes, confundir requisitos funcionais e não funcionais [Arora et al. 2024]. Sampaio *et al.* (2024) também destacam a necessidade de supervisão humana para garantir a precisão e especificidade dos requisitos gerados por LLM.

Alinhado com a visão reportada por Melegati *et al.* que ressalta a importância de ensinar a priorização de requisitos como uma habilidade crítica para validar hipóteses de mercado e construir produtos efetivos, este trabalho se concentra no uso do ChatGPT para a priorização de requisitos de software, visando o MVP. O relato de experiência aqui reportado investiga e evidencia o potencial educacional do uso do ChatGPT em um ambiente de aprendizado de ES.

3. CONDUÇÃO DA EXPERIÊNCIA

Esta seção descreve os procedimentos realizados no ensino da disciplina de ES, na qual o ChatGPT foi utilizado como ferramenta de apoio para a priorização de requisitos de MVP de software. Os procedimentos adotados basearam-se nos adotados por Sampaio *et al.* em [Sampaio et al. 2024b]. A disciplina, com carga horária de 60 horas, tem como objetivo capacitar os alunos em tópicos principais de ES. A experiência com o ChatGPT foi direcionada à geração e validação de requisitos prioritários de software para o MVP. Participaram 16 alunos, organizados em quatro equipes.

A experiência foi organizada em duas etapas. Na primeira, os alunos foram divididos em equipes para realizar a eliciação e a priorização de requisitos de sistemas de software, com contextos definidos pelas próprias equipes (Seção 3.1). Na segunda etapa, as equipes utilizaram o ChatGPT para validar os requisitos selecionados anteriormente e gerar novos requisitos prioritários, utilizando *prompts* pré-estabelecidos. Além disso, registraram suas percepções sobre o uso do ChatGPT, incluindo vantagens e desvantagens da aplicação da IA na priorização de requisitos de software para o MVP (Seção 3.2). Os artefatos produzidos nessa experiência estão disponíveis para acesso público ⁴.

3.1. Primeira Etapa

Nesta etapa os alunos foram divididos em equipes para realizarem um trabalho prático que simulava uma situação real de engenharia de requisitos. No total, foram criadas quatro equipes: *Auxilium*, *MentorMath*, *Banco Amigão* e *Gestão de Assinaturas*. Cada equipe recebeu um prazo de 20 dias para entregar um relatório técnico com os seguintes elementos: **1) Escopo inicial:** descrição do problema a ser abordado e motivação para o

⁴<https://figshare.com/s/9f63000d4af8cc88f4be>

sistema proposto, incluindo informações sobre dois perfis iniciais de usuários, bem como os benefícios esperados; **2) Requisitos Funcionais e Não Funcionais:** identificação e detalhamento dos RFs e RNFs, com atenção especial a requisitos inovadores que destacassem o sistema frente a alternativas existentes; **3) Regras de Negócio:** descrição das regras associadas ao funcionamento do sistema; e **4) MVP (Produto Mínimo Viável):** seleção de requisitos prioritários que seriam entregues na primeira versão do sistema e justificativa para a seleção de tais requisitos.

Para cumprir essas entregas, as equipes foram instruídas a utilizar, no mínimo, três técnicas diferentes de elicitação. O relatório também deveria apresentar a rastreabilidade entre as técnicas aplicadas e os requisitos gerados, detalhando a origem de cada requisito identificado. Após concluir o trabalho, os resultados foram apresentados por cada equipe em uma sessão de 20 minutos.

3.2. Segunda Etapa

A segunda etapa ocorreu presencialmente, com duração de 100 minutos. O objetivo foi usar o ChatGPT para validar os requisitos prioritários definidos na primeira etapa e gerar novos requisitos prioritários não identificados pelas equipes. Destaca-se que, nesta etapa, as equipes foram divididas em duplas. O objetivo dessa divisão foi permitir uma análise comparativa de concordância e discordância das respostas das duplas sobre os novos requisitos prioritários gerados pelo ChatGPT.

Para o andamento da atividade, as equipes interagiram com a IA seguindo um roteiro previamente definido. Primeiramente, as equipes receberam orientações sobre a realização da experiência e necessidade de preenchimento do TCLE (Termo de Consentimento Livre e Esclarecido). Em seguida, foram apresentados os instrumentos usados na experiência: (1) um documento contendo o roteiro de atividades a serem executadas (chamado de *Script* de Atividades), e (2) o relatório de elicitação de requisitos gerado na primeira etapa da experiência (cada dupla utilizou o relatório de sua própria autoria). Os instrumentos utilizados estão disponíveis para acesso público.⁵

A interação das equipes com o ChatGPT foi guiada pelo documento ***Script de Atividades***, composto por seis atividades:

Atividade 1: registro do horário de início da atividade.

Atividade 2: caracterização da equipe e assinatura do TCLE.

Atividade 3: contextualização da tarefa no ChatGPT — as duplas forneceram ao ChatGPT um contexto detalhado sobre o sistema, incluindo escopo inicial, perfis de usuários, benefícios esperados e requisitos inovadores, extraídos do relatório da primeira etapa. Em seguida, o ChatGPT foi orientado a agir como um engenheiro de software, com foco na priorização dos requisitos para o MVP (Figura 1). Por fim, o ChatGPT recebeu a definição do contexto do sistema a ser desenvolvido (Figura 2).

Atividade 4: Validação e Geração de requisitos prioritários do MVP — As equipes apresentaram ao ChatGPT uma lista inicial de requisitos prioritários para o MVP (obtida no relatório de elicitação de requisitos gerado na primeira etapa da experiência), solicitando a validação e justificativa para cada um (Figura 3, Comando 1). Em seguida,

⁵<https://figshare.com/s/9f63000d4af8cc88f4be>

pediram à IA que gerasse 10 novos requisitos para o MVP, também acompanhados de justificativas (Figura 3, Comando 2). Todos os novos requisitos sugeridos pelo ChatGPT foram analisados pelas equipes, que decidiram pela inclusão ou exclusão no MVP, justificando suas decisões.

Atividade 5: avaliação de autopercepção dos alunos sobre o uso do ChatGPT, refletindo sobre as vantagens, dificuldades e a aplicabilidade do ChatGPT na validação e geração de requisitos prioritários.

Atividade 6: registro do horário de término da atividade.

Olá, ChatGPT! Estou trabalhando atualmente no desenvolvimento de um sistema de software. Eu sou um analista de requisitos e gostaria do seu apoio durante a atividade de priorização de requisitos do sistema a ser desenvolvido visando o desenvolvimento do MVP (Minimum Viable Product, ou Produto Mínimo Viável). O MVP é a versão inicial de um produto que possui as funcionalidades essenciais, visando testar a proposta de valor do software e validar a aceitação do sistema pelo público-alvo. O processo de desenvolvimento do MVP envolve definir o problema a ser resolvido, identificar as funcionalidades essenciais para resolver esse problema, priorizar os requisitos que compõem essas funcionalidades e desenvolver uma primeira versão do sistema. Diante deste contexto, aja como um engenheiro de software ou um profissional da área de desenvolvimento de software que esteja interessado em gerar e validar os requisitos prioritários de um MVP de software.

Considere o seguinte escopo inicial do software a ser desenvolvido <<escopo inicial>>. Este sistema foi solicitado/idealizado por << informe aqui a fonte do sistema>>. A motivação para a construção do software é << informe aqui a motivação >>. Inicialmente, foram identificados 2 perfis de usuários do sistema: << perfil 1>> e << perfil 2>>. Os benefícios esperados com o desenvolvimento do sistema são << informe aqui>>. Inicialmente, como inovação para o sistema visualiza-se <<Requisitos inovadores>>. Os requisitos funcionais elicitados para este software foram: << Informe todos RFs >>. Os requisitos não-funcionais elicitados para este software foram: << Informe todos RNFs>>. A lista abaixo descreve os requisitos prioritários, elegidos pela equipe de desenvolvimento, que devem fazer parte do MVP do sistema: <<lista requisitos prioritários>>.

Figura 1. Prompt de interação - Contexto atividade.

Figura 2. Prompt de interação - Contexto sistema.

Comando 1: Você considera que o requisito << informar os requisitos prioritários >> é adequado para estar no MVP do sistema (sim ou não)? Justifique sua resposta.

Comando 2: Com base no contexto acima descrito, gere 10 novos requisitos que poderiam ser inseridos no MVP do sistema. Para cada novo requisito gerado, justifique o motivo da geração do requisito.

Figura 3. Prompt de interação - Comandos.

Por fim, foram conduzidas análises quantitativas e qualitativas com base nos dados coletados durante a experiência. A análise quantitativa, baseada nos resultados da Atividade 4, focou na identificação da quantidade de (1) requisitos prioritários, determinados na primeira etapa, que foram endossados pela avaliação do ChatGPT e (2) novos requisitos prioritários gerados pelo ChatGPT e o julgamento das duplas quanto à sua inclusão ou exclusão no MVP, acompanhada de justificativas (Seção 4). Já a análise qualitativa, fundamentada nos registros individuais da Atividade 5, organizou os comentários dos participantes em categorias que refletiam suas percepções sobre o uso do ChatGPT, destacando vantagens e desvantagens do uso da IA (Seção 5).

4. RESULTADOS QUANTITATIVOS

Esta seção apresenta os resultados da análise quantitativa dos dados coletados durante a experiência com o ChatGPT, focando na avaliação (Seção 4.1) e geração de novos requisitos prioritários pelo ChatGPT (Seção 4.2).

4.1. Avaliação do ChatGPT dos requisitos prioritários selecionados pelas Equipes

Nesta subseção, é apresentada a análise da concordância geral do ChatGPT com as equipes na priorização dos requisitos para o MVP. A Tabela 1 mostra a validação do ChatGPT em relação aos requisitos prioritários estabelecidos pelas equipes na primeira etapa da experiência, sendo apresentada a validação por duplas de cada equipe (D1 e D2).

Tabela 1. Avaliação da concordância do ChatGPT em relação aos Requisitos Prioritários definidos na primeira etapa da experiência

Dupla	Auxilium	MentorMath	Banco Amigão	Gestão de Assinaturas	Total de concordância (%)
D1 - Sim	7/7 (100%)	7/7 (100%)	7/8 (87,5%)	10/15 (66,7%)	31/37 (83,8%)
D2 - Sim	6/7 (85,7%)	7/7 (100%)	6/8 (75%)	11/15 (73,3%)	30/37 (81,1%)
Média total de concordância	-	-	-	-	86,02%

O ChatGPT demonstrou alta concordância com as escolhas das equipes em relação aos requisitos prioritários definidos por elas, corroborando em média com 86,02% (32 de 37) dos requisitos prioritários definidos na primeira etapa da experiência. Contudo, observa-se que apenas uma equipe (MentorMath) obteve 100% de concordância na avaliação do ChatGPT. A equipe Gestão de Assinaturas foi a que obteve a menor taxa de aceite na avaliação do ChatGPT, obtendo taxa média de aceite de 70%. Uma análise nas justificativas geradas pelo ChatGPT por não ter concordado com a seleção do requisito prioritário mostra que a IA discorda com a seleção de requisitos que (1) não são essenciais para a validação do MVP, (2) representam funcionalidade avançada ou complexa e (3) podem ser implementados em versões futuras. A leitura das justificativas geradas pela IA permite com que as duplas reflitam sobre a priorização feita, contribuindo para uma reflexão crítica das escolhas feitas pela equipe.

A Tabela 2 mostra o indeterminismo do ChatGPT na avaliação dos requisitos prioritários. A coluna “Concordâncias (Iguais)” apresenta o número de requisitos em que o ChatGPT concordou com a inclusão no MVP para ambas as duplas da mesma equipe. A coluna “Discordâncias (Iguais)” mostra os casos em que o ChatGPT discordou sobre a inclusão de um requisito no MVP entre as duplas da mesma equipe. Por fim, a coluna “Inconsistências Identificadas” reflete o número de requisitos iguais nos quais o ChatGPT forneceu respostas contraditórias, ou seja, aprovando a inclusão para uma dupla e rejeitando para a outra.

Tabela 2. Inconsistências observadas na avaliação do ChatGPT em requisitos iguais por equipe

Equipe	Total de Requisitos Avaliados	Concordâncias (Iguais)	Discordâncias (Iguais)	Inconsistências Identificadas
Auxilium	7	6	0	1
MentorMath	7	7	0	0
Banco Amigão	8	6	1	1
Gestão de Assinaturas	15	9	3	3
Total	37	28	4	5

Embora a avaliação do ChatGPT sobre a priorização dos requisitos pelas equipes tenha sido consistente, com 86,48% dos casos (32 de 37) (Tabela 1), análises mostraram discrepâncias nas respostas fornecidas pela IA ao comparar os resultados entre as duplas de cada equipe, representando 13,5% (5 de 37), conforme Tabela 2.

Um exemplo de indeterminismo foi identificado na equipe Auxilium, no requisito prioritário “RF03 - O sistema deve permitir que o tutor e o tutorado troquem mensagens de texto, áudio e vídeo”. Para a Dupla 1, o ChatGPT considerou o requisito adequado para o MVP, justificando que a funcionalidade é essencial para garantir uma comunicação eficaz e validar a proposta de valor do sistema. Em contrapartida, para a Dupla 2, o ChatGPT

não recomendou a inclusão do requisito no MVP, argumentando que a complexidade de implementar suporte a multimídia (áudio e vídeo) na primeira versão tornava mensagens de texto suficientes para o objetivo inicial. Essa discrepância pode estar associada à característica não determinística da LLM. Padrões semelhantes foram observados em outras equipes, identificadas na Tabela 2.

4.2. Novos Requisitos Gerados pelo ChatGPT

Na Atividade 4 do *Script* de Atividade (Seção 3.2), o ChatGPT gerou 10 novos requisitos para cada dupla, as duplas avaliaram se já haviam elicitado o requisito e optado por não incluí-lo no MVP e se concordavam com sua inclusão após a sugestão do ChatGPT. A Tabela 3 sumariza o quantitativo dos novos requisitos prioritários gerados pelo ChatGPT. A coluna “Total de Requisitos Gerados” mostra a quantidade total de requisitos gerados pelo ChatGPT para cada equipe (considerando as duas duplas), sendo 20 no total. A coluna “Total de Requisitos Novos” indica a quantidade de requisitos considerados novos pelas duplas das equipes, ou seja, que não foram identificados na etapa inicial. Por fim, a coluna “Requisitos Semelhantes” indica a quantidade de requisitos considerados similares/semelhantes, identificados ao comparar a lista de novos requisitos gerados pelo ChatGPT para as duplas de cada equipe.

Tabela 3. Quantitativo dos novos requisitos prioritários gerados pelo ChatGPT

Equipe	Total de Requisitos Gerados	Total de Requisitos Novos	Requisitos Semelhantes
Auxilium	20	16	3
MentorMath	20	13	5
Banco Amigão	20	12	4
Gestão de Assinaturas	20	17	3
Total	80	58	15

Uma média de 72,5% (58 de 80) dos requisitos gerados pelo ChatGPT foram considerados novos pelas equipes, indicando a capacidade da ferramenta de propor funcionalidades prioritárias não consideradas inicialmente pelas equipes. Foi identificada a ocorrência de 15 requisitos semelhantes entre as duplas, o que representa a média de 18,75% do total de requisitos gerados. Essa semelhança evidencia a capacidade do ChatGPT de identificar necessidades recorrentes a partir da especificação de um mesmo contexto de software, porém em interações diferentes, mesmo considerando sua natureza não determinística. O quantitativo apresentado na coluna “Total de Requisitos Novos” da Tabela 3 sugere que a IA é eficiente na identificação de elementos essenciais do sistema, independentemente das variações nas interações realizadas pelos usuários.

Tabela 4. Análise dos requisitos semelhantes gerados pelo ChatGPT

Equipe	Total de Requisitos Semelhantes	Novos para Ambas as Duplas	Discordância entre Duplas	Inclusão no MVP
Auxilium	3	3	0	1
MentorMath	5	3	1	2
Banco Amigão	4	2	1	3
Gestão de Assinaturas	3	2	1	2
Total	15	10	3	8

A Tabela 4 aprofunda a análise dos 15 requisitos identificados como semelhantes entre as duplas de uma mesma equipe. Cada coluna da tabela fornece informações específicas sobre esses requisitos. A coluna 2 mostra o total de requisitos semelhantes entre as duplas de cada equipe. A coluna 3, indica quantos dos requisitos semelhantes foram considerados novos pelas duas duplas da equipe. A coluna 4, quantifica os requisitos semelhantes em que houve discordância entre as duplas da equipe em relação à sua novidade. Por fim, a coluna 5 indica a quantidade de requisitos semelhantes que as equipes consideraram adequados para inclusão no MVP. A análise da Tabela 4 revela que 66,67% (10 de 15) dos requisitos semelhantes foram considerados novos para ambas as duplas, o que reforça a capacidade do ChatGPT em gerar ideias inovadoras. A discordância entre as duplas em relação à novidade dos requisitos (20%, 3 de 15) evidencia a necessidade de uma análise crítica humana para validar as sugestões do ChatGPT. A decisão final sobre quais requisitos serão incluídos no MVP deve ser tomada pela equipe, considerando o contexto específico do projeto. A inclusão de 53,33% (8 de 15) dos requisitos semelhantes no MVP pelas equipes demonstra o potencial do ChatGPT em auxiliar na priorização de funcionalidades. Os dados utilizados na geração dos valores apresentados na Tabela 4 estão disponíveis para consulta ⁶.

5. RESULTADOS QUALITATIVOS

Com base nos comentários dos 16 alunos coletados na Atividade 5 (Seção 3.2), foram identificadas percepções positivas e negativas sobre o uso do ChatGPT na priorização de requisitos para MVPs de software.

Vantagens: os alunos destacaram a **validação útil** do ChatGPT em relação aos requisitos avaliados na primeira etapa e os gerados na segunda etapa, conforme evidenciado pelos participantes: P10 - *"O ChatGpt auxiliou no processo criativo de novos requisitos, além de ter validado os existentes"*. A IA também demonstrou capacidade de gerar **soluções novas** que não haviam sido considerados pelas equipes durante a elicitação tradicional, conforme destaca P3 - *"Conseguiu gerar soluções que não haviam sido pensadas ou que acabaram passando despercebidas no MVP"*. A **economia de tempo** foi outra vantagem destacada: P5 - *"É rápido e prático"*. A ferramenta também auxiliou na **condensação de informações** relevantes ao sistema, P8 - *"Ele consegue condensar grandes volumes de texto ..., ajudando o analista de requisitos"*.

Desvantagens: a **confiabilidade dos resultados** gerados pelo ChatGPT foi questionada: P3 - *"O modelo também sugeriu soluções genéricas que não necessariamente se enquadram numa versão mínima do sistema"* e P6 - *"Apesar das muitas recomendações do ChatGPT muitas não se enquadraram para estar no MVP, pois não eram essenciais para o funcionamento do produto de acordo com o que foi proposto"*. Os participantes enfatizaram a necessidade de **prompts bem definidos** para interação com a LLM: P1 - *"A principal dificuldade é que o texto deve ser enviado em uma estrutura cuidadosa, ou o ChatGPT pode confundir as informações enviadas"* e P12 - *"Se a interação com o chat for feita erroneamente, a resposta fornecida é descartável"*. A **necessidade de intervenção humana** para avaliar as repostas do ChatGPT foi outro ponto destacado: P15 - *"Desvantagem seria a necessidade humana de verificar suas respostas e algumas discordâncias."* e P3 - *"Porém a curadoria humana sem dúvidas ainda é necessária"*.

⁶<https://figshare.com/s/9f63000d4af8cc88f4be>

6. LIÇÕES APRENDIDAS

Os resultados desta experiência destacam o ChatGPT como ferramenta de apoio no processo ensino-aprendizagem de priorização de requisitos de software visando o MVP. As lições aqui apresentadas podem ajudar professores a planejar o uso do ChatGPT em atividades práticas, a serem executadas com alunos, objetivando a criação de novas experiências de aprendizado. São elas:

Preparação de *prompts*: a qualidade das respostas geradas pelo ChatGPT está diretamente relacionada à clareza e especificidade dos *prompts* fornecidos. Professores devem planejar e elaborar *prompts* bem estruturados para maximizar a precisão das respostas da IA. Como destaca P12 - *"Se a interação com o chat for feita erroneamente, a resposta fornecida é descartável"*.

Necessidade de supervisão humana: embora o ChatGPT seja útil para avaliar e gerar novos requisitos prioritários, a supervisão humana é indispensável para interpretar criticamente as sugestões da ferramenta e garantir que estejam alinhadas ao contexto do projeto. Como pode-se concluir a partir da citação de P6 - *"Apesar das muitas recomendações do ChatGPT muitas não se enquadraram para estar no MVP, pois não eram essenciais para o funcionamento do produto de acordo com o que foi proposto"*.

O ChatGPT promove a reflexão crítica dos alunos: a IA pode promover uma reflexão crítica ao apresentar justificativas e alternativas que ajudam os alunos a reconsiderarem suas decisões, incentivando o aprendizado ativo. Como evidenciado em P6 - *"O ChatGPT foi muito efetivo para criar requisitos que não havíamos pensado antes"*.

Oportunidade para desmistificação da IA: professores podem executar experiências semelhantes com o objetivo de promover discussões sobre as limitações da IA, destacando que o ChatGPT pode apresentar respostas inconsistentes ou gerar informações genéricas, reforçando a importância de não depender exclusivamente da ferramenta.

7. LIMITAÇÕES

Este trabalho apresenta limitações que devem ser consideradas ao interpretar os resultados reportados. As limitações estão relacionadas ao número de participantes na experiência, ao contexto acadêmico, à metodologia adotada e às características da IA utilizada.

O estudo foi conduzido com apenas 16 alunos, o que limita a generalização dos resultados. A experiência foi realizada em um contexto acadêmico, onde os alunos possivelmente não enfrentam os cenários complexos de equipes de desenvolvimento de software da indústria. Além disso, a interação com o ChatGPT foi guiada por *prompts* estabelecidos previamente, o que pode introduzir vieses nos resultados. O ChatGPT apresentou respostas inconsistentes entre interações, o que pode impactar a confiabilidade das análises realizadas e a replicabilidade dos resultados. Por fim, os resultados reportados basearam-se nas perspectivas dos alunos.

Contudo, destaca-se que o objetivo desta experiência é promover a reflexão crítica dos alunos sobre as vantagens e desvantagens do uso do ChatGPT no apoio à tarefa de priorização de requisitos de software, visando o MVP do sistema.

8. CONCLUSÕES

O objetivo da experiência relatada neste artigo foi investigar o uso do ChatGPT como uma ferramenta de apoio na priorização de requisitos para MVPs (Minimum Viable Product) de software. A experiência envolveu 16 estudantes de graduação e foi conduzida em duas etapas. Na primeira etapa, os participantes, organizados em equipes, realizaram a elicitação e priorização de requisitos de sistema sem o uso da IA. Na segunda etapa, as equipes foram divididas em duplas e utilizaram o ChatGPT para avaliar os requisitos prioritários definidos anteriormente e gerar novos requisitos, os quais foram analisados quanto à originalidade e aceitação no MVP do sistema. Por fim, os participantes registraram suas percepções sobre o uso do ChatGPT na avaliação e geração de requisitos prioritários, destacando vantagens e limitações da ferramenta. A avaliação da experiência conduzida foi realizada por meio de análises quantitativas e qualitativas dos dados coletados.

Os resultados obtidos demonstraram que o ChatGPT corroborou 86,02% das escolhas feitas pelos estudantes, validando a maioria dos requisitos priorizados pelas equipes na primeira etapa. Cerca de 72,5% dos novos requisitos gerados pelo ChatGPT foram considerados inéditos pelas equipes, evidenciando a capacidade da LLM propor funcionalidades relevantes que não haviam sido identificadas na primeira etapa. Na análise qualitativa, os participantes destacaram a capacidade do ChatGPT em sugerir novos requisitos que não haviam sido identificados na primeira etapa. A capacidade de condensar informações foi vista como uma eficiência do ChatGPT, além de sua agilidade em tarefas relacionadas à avaliação e geração de requisitos, otimizando o tempo das atividades. Por outro lado, reconheceram que, apesar de útil, a IA frequentemente apresentou requisitos com descrições genéricas, evidenciando a necessidade de uma análise humana para validar as sugestões da ferramenta. A dependência de *prompts* bem definidos também foi destacada como uma limitação. Respostas inconsistentes ou irrelevantes ao escopo do sistema tornaram evidente a necessidade de supervisão por parte das equipes.

Concluimos que o ChatGPT pode ser uma ferramenta valiosa no contexto educacional, ao complementar atividades de priorização de requisitos e fomentar a inovação. No entanto, seu uso deve ser aliado a uma análise crítica e supervisão humana para garantir a qualidade e relevância dos resultados através do controle e validação das decisões tomadas durante o processo. Além disso, é fundamental considerar os aspectos éticos envolvidos no uso de LLM. Trabalhos futuros podem focar na minimização das limitações aqui identificadas e, ainda, considerar a perspectiva dos professores na avaliação do uso da IA como ferramenta complementar no processo de ensino-aprendizagem de ES.

AGRADECIMENTOS

Agradecemos a todos os participantes do estudo empírico. Gostaríamos de agradecer o suporte financeiro concedido pelo CNPq 445029/2024-2 e pela Universidade do Estado do Amazonas, através do Programa de Produtividade Acadêmica 01.02.011304.026472/2023-87.

Referências

Alonso, S., Kalinowski, M., Ferreira, B., Barbosa, S. D., and Lopes, H. (2023). A systematic mapping study and practitioner insights on the use of software engineering practices to develop mvps. *Information and Software Technology*, 156:107144.

- Arora, C., Grundy, J., and Abdelrazek, M. (2024). Advancing requirements engineering through generative ai: Assessing the role of llms. In *Generative AI for Effective Software Development*, pages 129–148. Springer.
- Beganovic, A., Jaber, M. A., and Abd Almisreb, A. (2023). Methods and applications of chatgpt in software development: a literature review. *Southeast Europe Journal of Soft Computing*, 12(1):08–12.
- Daun, M. and Brings, J. (2023). How chatgpt will change software engineering education. In *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*, pages 110–116.
- Marques, N., Silva, R. R., and Bernardino, J. (2024). Using chatgpt in software requirements engineering: A comprehensive review. *Future Internet*, 16(6):180.
- Melegati, J., Chanin, R., Sales, A., Prikladnicki, R., and Wang, X. (2020). Mvp and experimentation in software startups: a qualitative survey. In *2020 46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, pages 322–325. IEEE.
- Mendonça, G. D., de Souza Filho, I. P., and Guedes, G. T. A. (2021). A systematic review about requirements engineering processes for multi-agent systems. *ICAART (1)*, pages 69–79.
- Mittal, U., Sai, S., Chamola, V., et al. (2024). A comprehensive review on generative ai for education. *IEEE Access*.
- Petrovska, O., Clift, L., Moller, F., and Pearsall, R. (2024). Incorporating generative ai into software development education. In *Proceedings of the 8th Conference on Computing Education Practice*, pages 37–40.
- Ries, E. (2011). *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. Crown Currency.
- Sampaio, S., Lima, M., Rodrigues, E., Meireles, M. A., Pessoa, M., and Conte, T. (2024a). Exploring the use of large language models in requirements engineering education: An experience report with chatgpt 3.5. *SBQS 2024*.
- Sampaio, S., Lima, M., Rodrigues, E., Meireles, M. A., Pessoa, M., and Conte, T. (2024b). Exploring the use of large language models in requirements engineering education: An experience report with chatgpt 3.5.
- Sun, J., Liao, Q. V., Muller, M., Agarwal, M., Houde, S., Talamadupula, K., and Weisz, J. D. (2022). Investigating explainability of generative ai for code through scenario-based design. In *Proceedings of the 27th International Conference on Intelligent User Interfaces*, pages 212–228.
- White, J., Hays, S., Fu, Q., Spencer-Smith, J., and Schmidt, D. C. (2024). Chatgpt prompt patterns for improving code quality, refactoring, requirements elicitation, and software design. In *Generative AI for Effective Software Development*, pages 71–108. Springer.