

Fostering Creativity with ChatGPT in Requirements Engineering Education: An Academic Experience

Marcos Costa

Instituto de Computação

Universidade Federal do Amazonas

Manaus, Amazonas, BR

marcos.araujo@icomp.ufam.edu.br

Maria Meireles

Instituto de Computação

Universidade Federal do Amazonas

Manaus, Amazonas, BR

maria.meireles@icomp.ufam.edu.br

Márcia Lima

Universidade do Estado do Amazonas (UEA)

Manaus, Amazonas, BR

msllima@uea.edu.br

Tayana Conte

Instituto de Computação

Universidade Federal do Amazonas

Manaus, Amazonas, BR

tayana@icomp.ufam.edu.br

ABSTRACT

In the context of Software Engineering education, the Requirements Engineering (RE) phase plays a key role, as it directly influences the quality of the final product. Among the key competencies to be developed in this process is creativity, which is essential for developing innovative solutions that meet user needs. Since the emergence of Large Language Models (LLMs), their use has become more widely adopted in various stages of software engineering (SE), including RE. Thus, stimulating creative thinking through the use of LLMs in the classroom is a promising strategy for preparing future professionals capable of facing complex challenges and proposing original solutions from the early stages of software development. Based on this, the present study investigated the relationship between the use of LLMs, specifically ChatGPT-3.5, and the stimulation of creativity during software RE. For this purpose, we conducted a study with 42 students enrolled in an introductory SE course. Participants did a creative requirements generation activity using the Brainwriting technique, conducted in two stages: the first without the support of tools, and the second with the use of ChatGPT. The requirements generated in both stages were evaluated based on four dimensions from the Torrance Test of Creative Thinking (TTCT): fluency, flexibility, originality, and elaboration. The results showed that using ChatGPT helped increase the number of ideas, the variety of categories covered, the level of detail, and, in many cases, the originality of the requirements. Despite this, it also highlighted limitations, such as the need for intensive filtering of the generated requirements and the superficiality of some proposals. These findings reinforce the potential of LLMs, when critically mediated, as a pedagogical tool to foster creative and reflective skills in the training of future software engineers.

KEYWORDS

LLM, ChatGPT, Creativity, Requirements Engineering, Brainwriting, TTCT, Software Engineering

1 Introdução

A Engenharia de Requisitos (ER) é considerada uma etapa fundamental no desenvolvimento de sistemas e, por isso, ocupa um papel essencial na formação de engenheiros de software, pois, quando

bem conduzida, contribui para o desenvolvimento de soluções de melhor qualidade e reduz consideravelmente a possibilidade de falha [4, 12, 16]. O estímulo à criatividade durante os processos de ER, em especial o de elicitação de requisitos, tem despertado interesse crescente, especialmente com o avanço de ferramentas baseadas em inteligência artificial, como os Modelos de Linguagem de Grande Escala (LLMs)[8, 12].

Entre essas ferramentas, destaca-se o *ChatGPT-3.5*, desenvolvido pela OpenAI, amplamente utilizado por seu potencial em apoiar atividades da engenharia de software [12]. Sua capacidade de adaptação contínua e compreensão de linguagem natural permite melhorar a colaboração com *stakeholders*, acelerar o processo de elicitação de requisitos e contribuir para resultados mais eficientes e bem-sucedidos no desenvolvimento de software [9, 12].

Dada essa relevância, este artigo investiga a relação entre o uso de LLMs, especificamente o *ChatGPT-3.5*, e o estímulo à criatividade durante a elicitação de requisitos de software. Para isso, foi conduzido um estudo experimental com 42 estudantes de Computação matriculados na disciplina de Introdução à Engenharia de Software (IES) da Universidade Federal do Amazonas. Os participantes foram organizados em equipes de 3 a 5 pessoas formadas por conveniência e realizaram uma atividade de elicitação de requisitos criativos em duas etapas: na primeira, o uso do *ChatGPT-3.5* não era permitido; na segunda, o uso da ferramenta foi incentivado. Ao final de cada etapa, as equipes consolidaram uma lista de requisitos. Esses requisitos foram avaliados com base em critérios adaptados do Teste Torrance de Pensamento Criativo (TTCT), contemplando as dimensões de fluência, flexibilidade, originalidade e elaboração. Além disso, aplicou-se um questionário para coletar percepções subjetivas dos participantes quanto ao uso do *ChatGPT-3.5* e à experiência de execução da atividade.

Os resultados obtidos permitem refletir sobre o impacto do uso de LLMs no processo de elicitação de requisitos, tanto sob a perspectiva quantitativa das dimensões de criatividade quanto sob o ponto de vista qualitativo das experiências relatadas pelos participantes. Os resultados indicam que o *ChatGPT-3.5* pode atuar como um recurso complementar à criatividade humana, ampliando o repertório de ideias e melhorando a estruturação dos requisitos, sobretudo nas dimensões de fluência e elaboração. Em relação à flexibilidade, observou-se uma maior variedade de categorias e enfoques nos

requisitos gerados com o apoio da ferramenta, sugerindo que o modelo contribuiu para expandir as possibilidades de abordagem do problema. Por outro lado, observou-se que muitos estudantes se perceberam mais originais ao trabalhar sem o auxílio da ferramenta, o que revela a importância de manter o protagonismo humano no processo criativo.

Este artigo está estruturado da seguinte forma: a Seção 2 apresenta os principais fundamentos teóricos que basearam o estudo. A Seção 3 descreve o planejamento metodológico do estudo, incluindo os participantes, a atividade realizada e os critérios adotados para avaliação. Na Seção 4, são discutidos os resultados obtidos, com base tanto nos dados quantitativos quanto nas percepções dos estudantes. A Seção 5 reúne as principais lições aprendidas com a experiência educacional. Por fim, a Seção 6 traz as considerações finais, refletindo sobre as limitações do estudo e sugerindo caminhos para pesquisas futuras.

2 Fundamentação Teórica

2.1 Criatividade em Engenharia de Software

A criatividade é uma habilidade essencial para engenheiros de software e, embora historicamente pouco explorada no campo da Engenharia de Software (ES), seu estudo vem ganhando destaque nos últimos anos [6, 8]. Criatividade é um conceito amplo, discutido em diversas áreas do conhecimento, pois compreender os elementos que compõem o pensamento criativo e como estimulá-lo é fundamental para a construção de soluções verdadeiramente inovadoras [6–8].

Raramente a criatividade ocorre em completo isolamento, sendo frequentemente potencializada pelo trabalho colaborativo em equipes de desenvolvimento de software diversificadas, compostas por pessoas com diferentes formações, especializações e abordagens de resolução de problemas, o que pode favorecer significativamente a geração de ideias criativas [5, 7]. O poder criativo individual é limitado, então as organizações devem incentivar ambientes de colaboração, nos quais ideias possam circular e se conectar de maneira dinâmica [5]. Contudo, o ambiente social dentro das equipes pode também prejudicar a criatividade, especialmente quando surgem problemas de convivência e comunicação [5].

A criatividade é um conceito multidimensional, fundamental para que engenheiros possam lidar com problemas complexos, embora não seja simples de definir e apresente diferentes interpretações em áreas distintas [1, 6]. Neste artigo, adotamos a definição de criatividade como a habilidade de gerar ideias, requisitos ou artefatos que sejam novos, surpreendentes ou valiosos [7, 8]. Essa abordagem contempla tanto o ato criativo quanto o valor do produto gerado, uma perspectiva especialmente relevante no contexto do desenvolvimento de software, onde a inovação precisa se materializar em soluções concretas, úteis e com qualidade [7, 8].

2.2 Teste Torrance de Pensamento Criativo

O Teste Torrance de Pensamento Criativo (*Torrance Test of Creative Thinking – TTCT*) foi desenvolvido por Ellis Paul Torrance em 1966, e até hoje é bastante recomendado, tanto no âmbito educacional quanto no corporativo, sendo o mais utilizado e referenciado teste de criatividade [3, 10, 14].

O teste objetiva avaliar o pensamento criativo utilizando dimensões, destacam-se: fluência, flexibilidade, originalidade e elaboração, que podem ser definidas da seguinte forma:

- **Fluência:** Refere-se ao número de ideias relevantes [2, 10], ou seja, à capacidade de gerar um grande volume de ideias ou soluções para um estímulo ou problema. Mede a habilidade que a pessoa ou grupo possui de produzir muitas ideias [3].
- **Flexibilidade:** Diz respeito ao número de categorias diferentes ou à mudança de perspectivas [2]. Mede a diversidade das ideias geradas e reflete a habilidade de abordar um problema a partir de diferentes contextos ou ângulos [3].
- **Originalidade:** Representa a habilidade de gerar respostas incomuns, únicas, imaginativas e inventivas, que se distanciam do senso comum [2, 3, 10].
- **Elaboração:** Corresponde à adição de detalhes além do mínimo necessário para compor a ideia [2]. Avalia a capacidade de expandir ideias, acrescentando profundidade, complexidade e sofisticação [3].

2.3 Ensino de Elicitação de Requisitos

A elicitação de requisitos é a fase da ER responsável por identificar os usuários ou *stakeholders* e compreender suas necessidades, expectativas e objetivos em relação ao sistema a ser desenvolvido [12, 16]. Essa etapa concentra grande parte do esforço do processo de ER, pois envolve descobrir, extraer e tornar explícitos desejos e expectativas das partes potencialmente interessadas e a qualidade com que o processo é conduzido influencia diretamente a qualidade dos requisitos gerados e posteriormente da solução [4, 16].

Uma elicitação eficaz é determinante para o sucesso ou fracasso de um projeto de software [16]. Por conta disso, tem ganhado cada vez mais atenção tanto na prática profissional quanto no meio acadêmico [4]. No contexto educacional, o ensino dessa etapa representa um desafio particular, pois requer um equilíbrio entre a abordagem teórica, que tende a ser técnica, normativa e densa, e a realização de atividades práticas que permitam aos estudantes experimentar, na prática, a coleta de requisitos e desenvolver sensibilidade quanto à aplicação adequada das técnicas disponíveis [4].

Para esse fim, diversas técnicas podem ser empregadas. No desenvolvimento deste estudo, utilizou-se a técnica de *Brainwriting*, que será explicada a seguir.

2.3.1 Brainwriting. É um método alternativo ao amplamente conhecido *brainstorming*. É particularmente recomendado em contextos em que o grupo conta com pessoas mais reservadas ou que, em uma sessão de grupo aberta, como as do *brainstorming*, teriam menos probabilidade de compartilhar suas ideias espontaneamente [11, 13].

Em uma sessão de *brainwriting* bem-sucedida, os participantes constroem coletivamente sobre as ideias uns dos outros, contribuindo com ideias e sugestões para gerar soluções variadas [13]. No entanto, essa técnica também apresenta limitações recorrentes, como o medo de julgamento, a acomodação social, o bloqueio de produção e a influência de personalidades dominantes, que podem inibir a participação equitativa e comprometer a diversidade de ideias geradas [11, 13].

Nesse sentido, o *brainwriting* busca mitigar algumas limitações frequentemente associadas ao *brainstorming* ao propor um processo

paralelo e simultâneo, em que os participantes escrevem suas ideias ou soluções em resposta a um estímulo ou problema apresentado, individualmente e ao mesmo tempo, antes de compartilhá-las com os demais [13]. Somente após todos concluírem essa etapa inicial, as ideias são compartilhadas com o grupo, permitindo que os participantes contribuam uns com os outros, melhorando e expandindo as propostas [13]. Constatou-se que a quantidade de ideias de alta qualidade em uma sessão de *brainwriting* frequentemente supera a de um *brainstorming* tradicional [11, 13].

2.4 Uso de Large Languages Models para ER

Os avanços da Inteligência Artificial Generativa, especialmente no Processamento de Linguagem Natural (NLP), deram origem aos Modelos de Linguagem de Grande Escala (LLMs), os quais têm provocado impactos significativos na sociedade contemporânea, incluindo transformações no processo criativo [3, 8, 15]. Embora esses modelos demonstrem notável capacidade para gerar textos coerentes e adequados ao contexto, LLMs também podem produzir informações imprecisas, ambíguas ou enviesadas, o que representa um risco à qualidade e confiabilidade dos requisitos de software gerados [9, 12].

A interação com LLMs ocorre, em geral, por meio de comandos textuais denominados *prompts*, que alimentam o modelo com informações e direcionam a geração de respostas desejadas [9, 12]. Entretanto, como as respostas dos LLMs nem sempre correspondem às expectativas do usuário, surgiu o campo da Engenharia de *Prompts* (*Prompt Engineering*), que consiste no processo de desenvolver e refinar *prompts* com o objetivo de obter respostas mais precisas, relevantes e de alta qualidade [3, 12].

Entre os LLMs disponíveis, o *ChatGPT* é um modelo que se tornou popular e pode agilizar a elicitação e documentação de requisitos ao facilitar o diálogo com *stakeholders*, gerar textos coerentes e oferecer *feedback* em tempo real [12]. Apesar dos benefícios, deve ser usado como apoio à expertise humana, não como substituto [12]. No entanto, ainda são escassos os estudos que investigam como esse uso extensivo tem impactado a criatividade desses profissionais [8].

2.5 Trabalhos Relacionados

Chang *et al.* [3] propõem o *G.P.S. Framework* (*Goal, Prompt, Strategy*), um modelo para apoiar a colaboração com LLMs em sessões de *brainstorming* para estimular o pensamento criativo. A criatividade das ideias geradas foi avaliada com base no Teste Torrance de Pensamento Criativo (TTCT). Assim, os autores puderam conduzir uma análise quantitativa do *framework* proposto, evidenciando aumento em todas as dimensões avaliadas. De forma semelhante, nosso estudo também se concentra no aspecto da criatividade e utiliza o TTCT para avaliar os requisitos produzidos em sessões de ideação. No entanto, diferimos em pontos fundamentais: enquanto os autores focam no refinamento sistemático de *prompts* dentro de um *framework* estruturado, nosso estudo se concentra na elicitação de requisitos no contexto de educação e qualidade de software, não propõe nem avalia um *framework* específico e permite o uso livre do *ChatGPT-3.5* pelos participantes.

Xiaotong *et al.* [15] também exploram o uso de LLMs no contexto educacional voltado à criatividade, conduzindo estudos com estudantes universitários para compreender como essas ferramentas

podem apoiar o processo de ideação. Por meio da integração de LLMs a templates de design colaborativo amplamente utilizados, como o *Five Whys* e o *Competitive Analysis Template* (CTA), o estudo teve como objetivo apoiar os estudantes no desenvolvimento de ideias e na reflexão crítica durante o processo criativo. A pesquisa avaliou o uso desses templates e sua capacidade de orientar o pensamento dos participantes por meio de perguntas reflexivas e da organização das informações. Diferentemente desse enfoque, o presente trabalho investiga a relação entre o uso do *ChatGPT-3.5* e o estímulo à criatividade durante o processo de elicitação de requisitos de software, concentrando-se na análise do potencial criativo das ideias produzidas pelos alunos, em vez da avaliação de uma ferramenta ou *framework* de suporte.

Shaer *et al.* [13] apresentam um *framework* colaborativo de co-criação entre humanos e LLMs, voltado ao aprimoramento do processo de ideação. Baseado na técnica de *brainwriting*, o método busca integrar LLM tanto na fase de divergência, em que ideias são geradas, quanto na de convergência, em que as melhores propostas são avaliadas e selecionadas. O LLM atuou como um “membro adicional” do grupo, contribuindo com novas perspectivas e detalhes que ampliam o repertório criativo dos participantes. Os autores também desenvolveram ferramentas específicas, como motores de avaliação e reflexão baseados em *GPT-4*, que auxiliam na seleção e aprimoramento das ideias a partir de critérios como inovação, relevância e profundidade. De forma semelhante, nosso estudo aplica a técnica de *brainwriting* como ferramenta para eliciar requisitos. Contudo, diferentemente dos autores, foi conduzido um estudo comparativo entre etapas com e sem o uso do *ChatGPT-3.5* e avaliou-se quantitativa e qualitativamente a criatividade dos requisitos gerados por meio do TTCT, permitindo compreender de maneira mais direta a influência da ferramenta no processo criativo de estudantes em contextos educacionais voltados à qualidade de software.

Vale ressaltar que os artigos se alinham a essa pesquisa ao evidenciar os benefícios do uso de LLMs no processo criativo, destacando a importância de uma avaliação crítica das sugestões geradas por LLM. Esse estudo se diferencia ao focar na elicitação de requisitos de software, permitindo analisar de forma comparativa o impacto do *ChatGPT-3.5* na criatividade dos estudantes. Assim, contribui-se para o entendimento de como LLMs podem ser integradas a práticas pedagógicas, visando à qualidade de software, ampliando o conhecimento sobre seu uso em contextos educacionais.

3 Condução da Experiência

Com o objetivo de investigar a relação entre o uso de LLMs, em especial o *ChatGPT-3.5*, e o estímulo à criatividade durante a elicitação de requisitos de software, foi conduzida uma atividade com estudantes de uma disciplina introdutória de Engenharia de Software (IES) da Universidade Federal do Amazonas, com ênfase em Engenharia de Requisitos, na qual os alunos tiveram que gerar requisitos para um cenário proposto. A técnica utilizada para promover essa geração de requisitos foi o *Brainwriting*.

A seguir, são detalhados os procedimentos adotados nas etapas de planejamento, execução e análise da atividade.

3.1 Planejamento

3.1.1 Participantes. O estudo contou com a participação de 42 alunos de IES, disciplina ofertada no curso de Ciência da Computação. Todos os participantes já possuíam familiaridade com ER, adquirida por meio de atividades previamente desenvolvidas ao longo da própria disciplina. Entretanto, apenas 52,4% dos respondentes relataram ter experiência prévia com a técnica *Brainwriting*, seja para fins de geração de ideias ou para a elicitação de requisitos.

3.1.2 Instrumentos de Preparação. Durante a fase de preparação do estudo, foram desenvolvidos diferentes instrumentos com a finalidade de apoiar a realização da atividade e organizar o processo de coleta e análise dos dados. Esses materiais abrangiam desde o conteúdo apresentado aos estudantes até os recursos utilizados pelos avaliadores, garantindo alinhamento metodológico ao longo de todas as etapas do estudo. Entre eles, destacam-se:

- (i) o material da aula expositiva, composto por slides abordando os conceitos sobre *brainwriting*, criatividade e instruções da atividade;
- (ii) o Termo de Consentimento Livre e Esclarecido (TCLE), elaborado para garantir a participação ética e voluntária dos estudantes, apresentando os objetivos, procedimentos, riscos e benefícios do estudo e informando o direito recusa ou desistência sem penalidades;
- (iii) o formulário de caracterização, voltado à coleta de informações sobre a experiência prévia dos participantes referente a elicitação de requisitos e a técnica *brainwriting*;
- (iv) as instruções para aplicação do TTCT, contendo um documento com uma breve explicação do método TTCT, juntamente com as diretrizes para a avaliação das dimensões de fluência, flexibilidade, originalidade e elaboração. A fluência foi definida como a quantidade total de requisitos gerados; a flexibilidade, como o número de categorias em que esses requisitos podem ser classificados. Para isso, os avaliadores categorizaram cada requisito e, ao final, contabilizaram o total de categorias distintas. Para as dimensões de originalidade e elaboração, os avaliadores atribuíram uma nota de 0 à 3 para cada requisito com base nos critérios definidos nas instruções;
- (v) a lista de requisitos comuns elaborada pelos pesquisadores e utilizada como referência para a análise da originalidade.

Estes instrumentos estão disponíveis no material complementar no seguinte link: <https://figshare.com/s/51b15c04f5265d7d529f>

3.1.3 Instrumentos de execução. Os instrumentos utilizados durante a execução do estudo com os estudantes incluiram materiais desenvolvidos especificamente para orientar e apoiar a realização da atividade. Entre eles, destacam-se:

- (i) o documento com o escopo do sistema Portal Vida Ativa, desenvolvido exclusivamente para a atividade. O sistema foi descrito como uma plataforma digital voltada a pessoas interessadas em superar o sedentarismo e adotar hábitos mais saudáveis. O material apresentava uma breve descrição da solução, o perfil dos principais públicos-alvo e suas respectivas necessidades, servindo como base para a geração de requisitos;

- (ii) o roteiro da atividade, contendo instruções detalhadas sobre cada etapa do processo. O roteiro incluía orientações para o uso da plataforma colaborativa *online IdeaBoardz*¹, bem como diretrizes sobre a estrutura e o formato padronizado dos artefatos a serem produzidos, como a lista de requisitos;
- (iii) a própria plataforma colaborativa *online IdeaBoardz*, uma ferramenta gratuita para colaboração em grupo. A partir de quadros virtuais com colunas personalizáveis, os participantes puderam adicionar e organizar ideias de forma síncrona ou assíncrona. Na atividade, um integrante criava o quadro conforme o roteiro e compartilhava o *link* com os demais, permitindo o registro coletivo dos requisitos, posteriormente o conteúdo dos quadros pôde ser exportado em PDF para facilitar a análise.

3.1.4 Instrumentos de avaliação. Com o objetivo de analisar os resultados obtidos na atividade de elicitação de requisitos, foram utilizados dois instrumentos principais:

- (i) a lista de requisitos gerados pelas equipes, que serviu de base para a avaliação das quatro dimensões do TTCT (Teste Torrance de Pensamento Criativo);
- (ii) um questionário de *feedback* aplicado aos participantes, voltado à coleta de percepções individuais sobre a experiência com o *ChatGPT-3.5*. Ao final de cada etapa, as equipes exportaram da plataforma colaborativa um PDF com os requisitos organizados em tabela padronizada (E1 e E2); os dados foram extraídos e organizados em uma planilha do *Google Sheets*² com colunas para cada etapa e para a avaliação do TTCT. Paralelamente, um questionário no *Google Forms* captou percepções individuais sobre a atividade com o *ChatGPT-3.5*, abordando desempenho, uso, utilidades e vantagens/desvantagens. O instrumento, com 14 perguntas (10 objetivas e 4 dissertativas), está disponível como material suplementar.

3.2 Execução

O estudo foi conduzido em um laboratório de informática da universidade, com duração total de 120 minutos (vide Figura 1). Cada participante teve acesso individual a um computador para a realização das atividades. Inicialmente, foi ministrada uma aula expositiva de aproximadamente 20 minutos, abordando a técnica de *brainwriting* (seu funcionamento e aplicação), conceitos de criatividade e as orientações sobre a dinâmica da atividade de elicitação de requisitos. Em seguida, os estudantes puderam esclarecer dúvidas pontuais e, antes de prosseguir com a atividade, assinaram o TCLE, declarando sua participação voluntária na pesquisa.

Para a realização da atividade prática, os estudantes foram orientados a se organizarem em grupos de 3 a 5 integrantes, sendo a formação por conveniência. Ao todo, formaram-se 10 equipes. A atividade de elicitação foi dividida em duas etapas de aproximadamente 40 minutos cada: na E1, as equipes trabalharam sem o auxílio do *ChatGPT-3.5*; já na E2, foi incentivado o uso do *ChatGPT-3.5* como apoio para a geração de requisitos criativos.

¹<https://ideaboardz.com>

²<http://sheets.google.com>



Figura 1: Aula sobre a atividade

Cada etapa da atividade foi estruturada em dois momentos: um primeiro momento individual, com 10 minutos para que cada participante registrasse suas próprias ideias de requisitos; e um segundo momento coletivo, com 30 minutos para discussão e consolidação em grupo. Um integrante de cada equipe ficou responsável por seguir o roteiro da atividade, acessando a plataforma colaborativa *Ideaboardz* e configurando o quadro compartilhado com os demais colegas. Ao final de cada etapa, as equipes exportaram o quadro em formato PDF contendo os requisitos consolidados. Com isso, foram obtidos 20 documentos no total, dois por equipe, correspondentes às etapas E1 e E2. A Figura 2 mostra a interação dos estudantes durante a execução da atividade.



Figura 2: Grupos de estudantes executando a atividade de elicitação de requisitos

Encerradas as etapas de geração de requisitos, os alunos foram convidados a responder, individualmente, a um questionário com o objetivo de captar suas percepções sobre a atividade e, em especial, sobre o uso do *ChatGPT-3.5* na elicitação de requisitos criativos. No total, foram coletadas 42 respostas.

3.3 Análise dos Requisitos

A análise dos requisitos foi conduzida com base nas diretrizes do TTCT, conforme descrito na Subseção 2.2. O objetivo desta etapa foi avaliar a criatividade dos requisitos gerados pelas equipes, considerando as quatro dimensões propostas pelo TTCT: fluência, flexibilidade, originalidade e elaboração.

Esta análise foi realizada por três avaliadores: o primeiro autor da pesquisa, que analisou os requisitos de todas as 10 equipes, e dois colaboradores externos com experiência prévia em ER. Cada um dos avaliadores externos ficou responsável por avaliar, de forma independente, os requisitos de 5 equipes distintas. De modo que, no conjunto, todas as equipes fossem contempladas com duas avaliações. Além disso a revisão final contou com a participação da autora Tayana Conte, especialista na área. Para auxiliar na tarefa de análise, foi entregue a cada avaliador uma planilha com os requisitos de suas respectivas equipes, organizados em formato tabular conforme descrito na Subseção 3.1.4.

A dimensão de fluência, definida pela quantidade de requisitos gerados, foi computada previamente durante o processo de extração dos documentos PDF entregues pelas equipes. Por se tratar de uma simples contagem objetiva, esse critério não exigiu julgamento subjetivo.

As demais dimensões, flexibilidade, originalidade e elaboração, exigiram julgamento qualitativo por parte dos avaliadores.

- **Flexibilidade:** foi avaliada com base na diversidade de categorias identificadas nos requisitos de cada equipe. Para isso, os avaliadores atribuíram uma categoria a cada requisito, considerando diferentes tipos de funcionalidades propostas, de modo a refletir a variedade de ideias presentes no conjunto. Por exemplo, requisitos voltados à personalização do perfil do usuário podem pertencer a uma mesma categoria, enquanto funcionalidades de monitoramento de desempenho físico constituiriam outra.
- **Originalidade:** foi analisada a partir do grau de novidade dos requisitos. Para apoiar esse julgamento, os avaliadores utilizaram uma lista de requisitos considerados comuns ou não criativos, elaborada previamente com base em soluções convencionais de sistemas semelhantes. Com isso, buscou-se garantir maior uniformidade na atribuição das notas, que variaram de 0 a 3. A Tabela 1 apresenta os critérios utilizados para essa pontuação.
- **Elaboração:** considerou o nível de detalhamento expresso em cada requisito, levando em conta a clareza, completude e profundidade das descrições fornecidas pelos participantes. Os critérios de pontuação para essa dimensão também seguiram uma escala de 0 a 3, conforme descrito na Tabela 2.

Ao final do processo, cada avaliador gerou uma tabela com os seguintes dados para cada equipe: número total de requisitos (fluência), número de categorias distintas (flexibilidade), média de originalidade e média de elaboração, tanto para a E1 quanto para a E2. Em

seguida, os dados das tabelas dos avaliadores foram consolidados, calculando-se a mediana para a fluência e flexibilidade e a média para a originalidade e elaboração.

Tabela 1: Diretrizes para avaliação de originalidade

Nota	Significado	Descrição
0	Nada criativo	Se o requisito estiver na lista de requisitos comuns
1	Pouco criativo	Ideia simples, com leve variação do padrão, sem grande inovação
2	Moderadamente criativo	Ideia diferente e interessante, mas ainda dentro de um raciocínio comum
3	Muito criativo	Ideia incomum, surpreendente ou que mostra pensamento inovador

Tabela 2: Diretrizes para avaliação de elaboração

Nota	Significado	Descrição
0	Nenhuma elaboração	A ideia é vaga, genérica ou mal explicada; não traz nada além do básico
1	Pouca elaboração	A ideia é clara, mas pouco desenvolvida; não há detalhes adicionais ou justificativas
2	Elaboração moderada	A ideia é detalhada, com elementos explicativos ou melhorias funcionais
3	Elaboração elevada	A ideia é completa, estruturada, com múltiplos detalhes, aspectos técnicos e justificativas

Os resultados por equipe, assim como os valores agregados da turma, são apresentados e discutidos na Tabela 3, sendo utilizados como base para as análises quantitativas e qualitativas descritas na próxima seção.

4 Resultados

A Tabela 3 sumariza os dados obtidos com a aplicação do método TTCT sobre os requisitos gerados pelas equipes. A seguir, será apresentada uma análise quantitativa e qualitativa baseada em cada um dos fatores de criatividade apresentados na Seção 2.2:

4.1 Fluência

A fluência pode ser entendida como a capacidade de gerar uma grande quantidade de ideias a partir de um determinado problema ou estímulo. Pode-se observar na Tabela 3 um aumento nesse aspecto durante a E2, quando foi permitido o uso do *ChatGPT-3.5* como apoio à elicitação. A produção mediana da turma passou de 34 requisitos na E1 (sem o uso da ferramenta) para 38 na E2. Esse crescimento sugere que o apoio da LLM contribuiu para uma maior fluência criativa por parte dos participantes. Com exceção das Equipes 2, 6 e 8, todas as demais apresentaram aumento no número de requisitos gerados após a introdução do *ChatGPT-3.5* no processo de elicitação de requisitos.

A preocupação com a fluência também se refletiu em algumas respostas à pergunta aberta sobre como foi o uso do *ChatGPT-3.5* na E2. Alguns participantes relataram ter adotado estratégias explícitas para maximizar a produção de requisitos, como exemplificado pelo estudante P35: "[...] pedi pra ele gerar em torno de 60 ideias e filtrei apenas algumas". Já o participante P14 destacou: "[...] Quanto mais você conseguir, melhor! <O cenário disponibilizado>", esses relatos indicam que parte dos estudantes utilizou o *ChatGPT-3.5* com foco deliberado em aumentar o volume de ideias geradas.

Quando indagados sobre as principais vantagens observadas na utilização do *ChatGPT-3.5*, muitos participantes destacaram a possibilidade de gerar um número maior de requisitos em menos tempo. Essa percepção é exemplificada pelo relato de P31: "Acredito que as principais (vantagens) foram a economia de tempo e o maior número de ideias geradas, comparado à elicitação sem ChatGPT". Tal comentário reforça a ideia de que o uso da ferramenta não apenas ampliou a fluência, mas também otimizou a agilidade do processo de geração de requisitos.

Porém, em função do grande volume de requisitos gerados, as principais limitações apontadas pelos participantes foram a necessidade de uma filtragem posterior e a tendência do *ChatGPT-3.5* a se desviar do tema proposto. Como destacou P8: "Algumas ideias se mostraram muito tangentes ao tema central [...] sendo necessária uma filtragem mais atenta das respostas". De forma semelhante, P9 relatou: "A maior dificuldade é não tangenciar demais o escopo. [...] Então o filtro tem que ser bem maior do que se fôssemos nós pensando". Esses relatos indicam que, embora a geração extensiva de ideias contribua positivamente para a fluência, ela também pode exigir um esforço adicional de curadoria por parte dos estudantes, comprometendo a eficiência do processo e exigindo maior discernimento na seleção dos requisitos realmente relevantes.

Logo, no que se refere à fluência, embora tenham sido apontadas vantagens e limitações, destacou-se o uso do *ChatGPT-3.5* como ferramenta para geração em larga escala de ideias de requisitos. Essa produção inicial foi, em geral, seguida por etapas de refinamento, que incluíram filtragem de ideias repetidas ou fora de escopo, reformulação de *prompts* e detalhamento de pontos específicos. Esses relatos sugerem que os estudantes adotaram uma abordagem iterativa no uso da ferramenta, combinando volume e revisão, o que reforça o potencial do *ChatGPT-3.5* em contribuir significativamente para a dimensão de fluência do TTCT.

4.2 Flexibilidade

A flexibilidade, no contexto do TTCT, refere-se à capacidade de gerar ideias que pertençam a diferentes categorias ou abordagens, refletindo a variedade na forma de pensar e resolver problemas. No que tange a essa dimensão, ao observar a Tabela 3, nota-se que a mediana de categorias geradas pelas equipes na E1 foi de 14, enquanto na E2 esse número aumentou para 17. Quando analisado por equipe, observa-se que, com exceção da Equipe 2, todas as demais mantiveram ou ampliaram sua mediana de categorias na E2 em comparação com a E1. Esse crescimento sugere que um maior número de perspectivas, funcionalidades ou aspectos distintos do sistema foi considerado na E2 da elicitação.

Tabela 3: Resumo dos resultados por equipe (E1 e E2)

Equipe	Flue. E1	Flue. E2	Flex. E1	Flex. E2	Orig. E1	Orig. E2	Elab. E1	Elab. E2
Equipe 1	17	24	7	10	1,29	1,46	1,29	2,08
Equipe 2	39	37	13	12	1,21	1,70	1,51	1,99
Equipe 3	21	48	10	16	1,24	1,17	1,50	1,55
Equipe 4	24	27	9	18	1,06	1,81	2,04	2,30
Equipe 5	38	50	16	17	1,04	2,07	1,59	2,12
Equipe 6	39	34	16	16	1,73	1,32	1,79	1,81
Equipe 7	32	66	15	19	1,20	1,42	1,73	2,08
Equipe 8	45	41	18	18	1,36	1,60	2,04	2,17
Equipe 9	20	36	10	17	1,13	1,38	1,90	2,19
Equipe 10	36	39	17	18	1,25	2,14	1,94	2,36
Média	-	-	-	-	1,25	1,61	1,74	2,06
Mediana	34	38	14	17	-	-	-	-

Os dados qualitativos coletados por meio do questionário reforçam essa tendência observada. Quando questionados sobre a contribuição do *ChatGPT-3.5* para a geração de novos *insights*, 85,7% dos participantes relataram que a ferramenta apresentou requisitos que ainda não haviam sido considerados por suas equipes, resultado ilustrado na Figura 3. Essa sugestão de novos *insights* foi percebida também por 92,9% dos estudantes como uma das maiores utilidades do *ChatGPT-3.5* na E2, como aponta a Tabela 4, o que evidencia a percepção de valor do *ChatGPT-3.5* como suporte à flexibilidade criativa no processo de elicitação de requisitos.

Tabela 4: Principal utilidade do *ChatGPT-3.5* durante a E2 (com o uso do ChatGPT)

Utilidade	Nº de pessoas	%
Sugestão de requisitos que eu não havia considerado	39	92,9
Economia de tempo	27	64,3
Ajuda na escrita/redação dos requisitos	27	64,3
Redução do esforço cognitivo	22	52,4
Organização das ideias	13	31,0
Estímulo à criatividade	11	26,2
Outros: Adaptar requisitos de uma outra aplicação	1	2,4
Nenhuma utilidade significativa	0	0,0

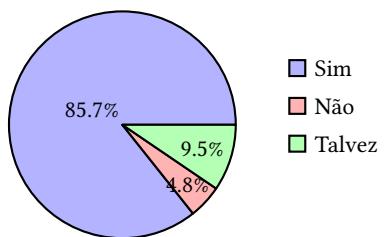


Figura 3: Novos *insights* com o uso do *ChatGPT*

Ao descreverem o modo de uso do *ChatGPT-3.5* durante a E2, alguns participantes relataram estratégias específicas adotadas para

ampliar a diversidade das ideias geradas. Entre os exemplos, destacam-se abordagens que buscaram simular diferentes pontos de vista ou personagens para enriquecer a variedade dos requisitos: P29 mencionou que “para tentar obter requisitos mais criativos e variados, pedi para que o *ChatGPT* interpretasse um personagem de série de televisão que pratica muitos esportes e motiva as pessoas a realizarem atividades físicas”; já P38 relatou: “eu descrevi o cenário pra ele, solicitei para ele criar as duas personagens do cenário e se passar por elas, e fiz uma entrevista com ambas as personagens que ele estava interpretando”. Essas estratégias revelam um uso inventivo do modelo como meio de explorar múltiplas perspectivas na geração de requisitos.

As principais vantagens e limitações do uso do *ChatGPT-3.5* para estimular a flexibilidade criativa também foram evidenciadas nas respostas dos participantes. Entre os aspectos positivos, destaca-se a capacidade da ferramenta de auxiliar na superação de bloqueios criativos, como relatado por P5: “Além de ser mais rápido, quando temos algum bloqueio nas ideias, ele pode fornecer um ou mais motivos que podem fazer mais sentido”. No entanto, também foram apontadas limitações relacionadas à qualidade e à variedade das respostas geradas. P31 observou que “o chat às vezes tinha dificuldade para entender exatamente o escopo, fugindo um pouco do tema ou dando ideias muito irrealistas para esse sistema”, enquanto P2 comentou sobre a recorrência de ideias semelhantes: “muitas ideias repetidas com as ideias geradas na rodada inicial [...]”. Esses relatos demonstram que, embora o uso do *ChatGPT-3.5* possa acelerar o processo e auxiliar em momentos de bloqueio criativo, ele ainda exige supervisão humana para garantir relevância e adequação ao escopo.

4.3 Originalidade

Originalidade refere-se à capacidade de produzir ideias incomuns ou inesperadas. Segundo os dados da Tabela 3, observou-se um aumento na média de originalidade da turma, passando de 1,25 na E1 para 1,61 na E2. Entretanto, 40,5% dos participantes relataram ter produzido requisitos mais originais sem o auxílio da ferramenta (vide Figura 4). Do ponto de vista do ensino-aprendizagem, esse resultado é bastante positivo, pois sugere que muitos estudantes percebem sua própria capacidade criativa como superior à do *ChatGPT-3.5* na elicitação. Essa percepção fortalece o pensamento crítico e a autonomia intelectual dos estudantes, indicando que o uso de

modelos generativos não substituiu o raciocínio individual, mas atuou como complemento. Ainda assim, apenas as Equipes 3 e 6 apresentaram desempenho superior em originalidade na E1. Todas as demais equipes registraram melhora nesse critério na etapa com uso do *ChatGPT-3.5*, sugerindo uma tendência positiva no apoio da ferramenta à geração de requisitos mais originais.

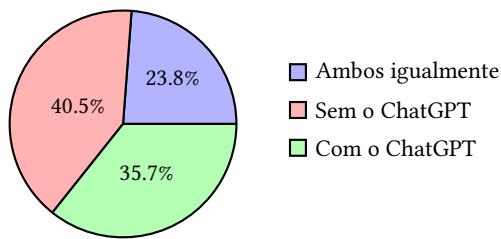


Figura 4: Etapa com requisitos mais originais

Quando questionados sobre o uso do *ChatGPT-3.5*, alguns participantes relataram utilizá-lo com foco explícito na geração de requisitos inovadores. P11, por exemplo, descreveu: “*Inicialmente, contextualizei o ChatGPT com o objetivo do sistema. Depois, perguntei sobre o que poderia ser inovador na ideia proposta. O chat deu algumas ideias [...] mas 2 me chamaram a atenção. [...] As ideias eram voltadas a um olhar mais psicológico dentro do sistema, o que eu não tinha pensado anteriormente*”. De modo semelhante, P33 afirmou: “[...] *Pedi para ele listar apenas requisitos criativos e diferenciais, e, a partir disso, consegui extrair diversos requisitos que eu ainda não tinha pensado*”. Esses relatos indicam que a ferramenta foi usada intencionalmente para estimular ideias fora do comum e ampliar o escopo de possibilidades consideradas, contribuindo diretamente para a dimensão de originalidade no processo de elicitação.

No que diz respeito às vantagens observadas, alguns estudantes relataram que o uso do *ChatGPT-3.5* contribuiu significativamente para a geração de ideias originais e criativas. P9 comentou: “*Acredito que ele conseguiu ser bem mais criativo, inclusive algumas vezes sendo bem absurdo. Assim deu pra sair bem mais da caixa*”. De forma semelhante, P15 afirmou: “*Ele conseguiu criar requisitos que não tínhamos pensado antes, bem fora da caixa mesmo [...]*”. Essas percepções reforçam a ideia de que o *ChatGPT-3.5* pode atuar como um estímulo à inovação, ao sugerir possibilidades inesperadas ou não convencionais que dificilmente surgiriam em um processo puramente humano, como destacou P22: “*Fiquei surpreso com a quantidade de ideias geradas a partir do ChatGPT, a maioria das quais dificilmente iríamos pensar, então acredito que a principal vantagem foi justamente essa de nos mostrar ideias muito inovadoras e criativas*.”

Por outro lado, algumas desvantagens ou limitações relacionadas à originalidade também foram apontadas por alguns estudantes. Entre os relatos, destaca-se o de P1: “*Uma limitação foi que ele não conseguiu trazer novas ideias; quando a minha equipe perguntou sobre ideias inovadoras, as respostas foram quase as mesmas, apenas com palavras diferentes*”. De forma semelhante, P7 relatou: “*Acho que o chat acabou gerando muitos requisitos genéricos, não aplicáveis e que não cobrem o sistema como um todo, ao contrário do primeiro brainwriting*”. Já P37 comentou: “*Senti que a capacidade dele de gerar ideias verdadeiramente criativas é limitada*”. Esses depoimentos

evidenciam que, apesar do potencial para ampliar a criatividade, o modelo ainda enfrenta desafios para gerar soluções que sejam, ao mesmo tempo, originais e viáveis dentro do escopo do sistema.

A maioria dos participantes reconheceu que o *ChatGPT-3.5* foi capaz de estimular a geração de ideias que ainda não haviam sido consideradas, como evidenciado pela fala de P6: “[...] ele [ChatGPT] gera requisitos não pensados antes, tanto básicos quanto mais específicos.” No entanto, essa originalidade foi frequentemente percebida como limitada. Muitas das ideias geradas foram descritas como genéricas, repetitivas ou pouco viáveis, o que compromete seu potencial inovador. P29 destacou que “*as ideias criadas com o uso do ChatGPT foram bem menos criativas ou inovadoras do que os requisitos criados somente com o nosso grupo*”, enquanto P25 afirmou que o modelo “[...] produziu mais do mesmo [...].” Essa percepção foi reforçada por P28: “[...] achei que eu e o grupo ficamos engessados depois de gerar os requisitos com o modelo [ChatGPT] [...].” Esses relatos indicam que, embora o *ChatGPT-3.5* possa funcionar como ponto de partida para ampliar o repertório de ideias, ele não substitui a criatividade humana na busca por soluções verdadeiramente originais e alinhadas ao problema em questão.

4.4 Elaboração

A elaboração diz respeito ao grau de detalhamento, clareza e desenvolvimento das ideias produzidas. Nesse aspecto, o uso do *ChatGPT-3.5* se destacou, conforme mostra a Tabela 3, com aumento na média geral da turma de 1,74 na E1 para 2,06 na E2. Observa-se também que todas as equipes apresentaram crescimento nesse critério, reforçando a tendência de que o modelo contribuiu para maior refinamento das ideias. Essa percepção é sustentada pelos dados da Tabela 4, segundo os quais 64,3% dos participantes apontaram que a principal utilidade do *ChatGPT-3.5* foi ajudar na escrita e redação dos requisitos, enquanto 31% destacaram a organização das ideias como um benefício relevante, ambos os aspectos diretamente relacionados à dimensão de elaboração.

De forma complementar, 42,9% dos estudantes relataram ter conseguido gerar requisitos mais completos com o auxílio do modelo, e 40,5% afirmaram que os requisitos se tornaram mais claros e organizados, conforme ilustram as Figuras 5 e 6, respectivamente.

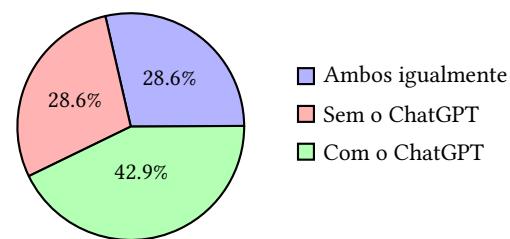


Figura 5: Requisitos mais completos

Essa utilidade também foi destacada nas respostas dos participantes sobre o uso do *ChatGPT-3.5*, como relatado por P9: “*Expliquei qual era a ideia do projeto e falei sobre algumas ideias que tive durante a primeira rodada, para aprimorá-las [...]*”. No entanto, a forma como o modelo elaborou os requisitos gerou percepções divergentes. Para alguns, esse nível de detalhamento foi um ponto positivo, como

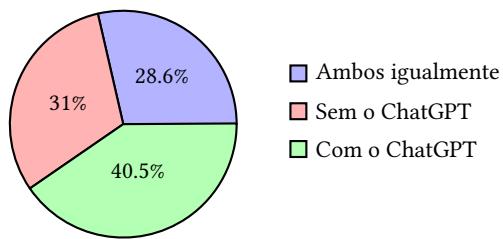


Figura 6: Requisitos mais claros e organizados

afirmou P17: "*Ele (ChatGPT) foi mais específico nos requisitos que elaborou, acredito que isso foi um ponto positivo no uso dele [...]*". Por outro lado, houve quem considerasse essa característica excessiva e contraprodutiva, como relatado por P36: "*Ele (ChatGPT) é muito prolixo e também contextualiza muito as ideias dele. Perdi muito tempo lendo a mesma coisa, apesar de ele gerar requisitos padrões e clássicos muito bem [...]*".

A percepção geral é que o *ChatGPT-3.5* contribui positivamente para a clareza textual e a organização das ideias, sendo especialmente útil na estruturação de requisitos já concebidos ou na produção de documentos mais legíveis. Como apontaram alguns participantes, P19 destacou: "[...] o *ChatGPT* auxilia muito na questão da coerência e a estruturar melhor os requisitos [...]", enquanto P9 afirmou: "[...] não sou uma pessoa que sei me explicar bem, então conseguir falar minha ideia e ter algo que me entenda e consiga explicar melhor, facilita muito [...]. Outro relato, de P36, reforça esse papel de apoio na reformulação textual: "[...] (O *ChatGPT*) seria bom para corrigir a minha ortografia [...] e também para me ajudar a expressar melhor as minhas ideias pois ele tende a usar as palavras melhor que eu".

No entanto, vários estudantes ressaltaram que essa clareza formal não implica, necessariamente, em uma elaboração conceitual aprofundada. Muitos consideraram os requisitos gerados genéricos, prolixos ou desalinhados com as necessidades específicas do sistema. Assim, embora a ferramenta seja valorizada pela sua capacidade de melhorar a forma e a estrutura textual, sua contribuição para o conteúdo significativo, a elaboração conceitual, ainda é percebida como limitada.

5 Lições Aprendidas

Com base nos dados analisados, identificou-se um conjunto de lições aprendidas que podem orientar professores interessados em aplicar o *ChatGPT-3.5* no contexto de geração de ideias criativas na fase de elicitação de requisitos.

- **O uso de LLMs na geração de requisitos pode promover reflexões educacionais relevantes:** a atividade evidenciou o potencial das LLMs, como o *ChatGPT*, para apoiar estudantes na elaboração de requisitos criativos. Muitos relataram ganhos práticos, como o apoio à estruturação de ideias e complementação de raciocínios em andamento.
- **O uso indiscriminado de LLMs pode comprometer a criatividade coletiva:** apesar dos ganhos relatados, muitos

estudantes alertaram para os riscos de depender excessivamente do *ChatGPT* nas etapas iniciais de ideação. O uso precoce da ferramenta foi associado à limitação do pensamento original e ao direcionamento prematuro das discussões.

- **O uso complementar das LLMs pode favorecer o equilíbrio entre eficiência e criatividade:** muitos estudantes recomendaram que o *ChatGPT* seja integrado após uma rodada inicial de brainstorming humano, atuando como apoio para refinar ideias ou explorar lacunas não percebidas. Essa estratégia foi vista como uma forma de preservar a criatividade original do grupo, ao mesmo tempo em que se aproveitam os benefícios da automação para ampliar e organizar os resultados.
- **O uso crítico das LLMs pode potencializar a aprendizagem em engenharia de requisitos:** a professora responsável pela disciplina destacou que a atividade permitiu aos alunos experimentar, de forma prática, os benefícios e limitações do *ChatGPT*, reforçando a importância do julgamento humano ao empregar a ferramenta. A experiência demonstrou que o uso autônomo e não mediado do modelo compromete a geração de requisitos com profundidade criativa e aderência ao contexto.
- **Antecipar o uso das LLMs pode favorecer seu papel como apoio estratégico:** a professora sugeriu que a atividade poderia ser realizada antes dos trabalhos principais de elicitação. Dessa forma, os estudantes já iniciariam suas tarefas com uma compreensão crítica do potencial e das limitações do *ChatGPT*, sabendo utilizá-lo como apoio estratégico, e não como substituto.
- **O uso reflexivo de LLMs pode contribuir para uma formação alinhada à qualidade de software:** ao explorar os limites e as oportunidades das LLMs no contexto educacional, a atividade contribui não apenas para o desenvolvimento da criatividade em engenharia de requisitos, mas também para uma formação alinhada aos princípios da qualidade de software, ao incentivar decisões conscientes, criteriosas e tecnicamente fundamentadas no uso de tecnologias emergentes.
- **O *ChatGPT* pode ser um aliado importante para ampliar a criatividade dos estudantes:** a ferramenta pode potencializar a geração de ideias, aumentando a quantidade e diversificando as categorias de requisitos, o que pode enriquecer as atividades de elicitação.
- **Supervisionar e orientar a filtragem das respostas do *ChatGPT*:** é fundamental estabelecer momentos de análise crítica, incentivando os estudantes a avaliarem a relevância, originalidade e consistência das ideias geradas, dado que o modelo pode produzir respostas irrelevantes ou superficiais.
- **Desenvolver atividades que integrem o uso do *ChatGPT* com reflexão crítica:** por exemplo, após gerar ideias com o suporte do modelo, incentivar discussões sobre sua viabilidade, inovação e alinhamento com o problema, fortalecendo as habilidades de avaliação e filtragem dos estudantes.
- **Atenção às limitações do *ChatGPT* na criatividade:** embora seja um estímulo efetivo, o modelo não substitui a criatividade humana, sendo importante que o professor destaque

a necessidade de supervisão e complementação das ideias por parte dos estudantes.

- **Integrar o uso do *ChatGPT* em metodologias ativas e colaborativas:** atividades em grupo, em que os estudantes dividam tarefas relacionadas ao uso da ferramenta, podem potencializar o aprendizado e o desenvolvimento de habilidades criativas e de trabalho em equipe.

Essas lições fornecem orientações práticas para professores que desejam incorporar o *ChatGPT* na formação de futuros engenheiros de requisitos, aproveitando seu potencial como ferramenta de estímulo criativo de forma crítica e eficaz.

6 Considerações Finais

Este estudo investigou a relação do uso de LLMs, especificamente o *ChatGPT-3.5*, no estímulo à criatividade durante a elicitação de requisitos de software. Por meio de uma atividade prática realizada com 42 estudantes de computação, foi possível comparar as produções de requisitos de duas etapas: uma sem o uso do *ChatGPT-3.5* e outra com seu apoio, considerando as dimensões de fluência, flexibilidade, originalidade e elaboração, conforme o modelo do TTCT de avaliação da criatividade.

Os resultados indicam um aumento significativo nas dimensões de fluência e elaboração com o uso do *ChatGPT-3.5*, refletindo maior quantidade e melhor estruturação dos requisitos gerados. Além disso, os participantes expressaram percepções favoráveis quanto à utilidade da ferramenta, ressaltando benefícios como a economia de tempo, a organização das ideias e a geração de novos *insights* que ampliaram o repertório criativo.

Quanto à originalidade, os dados do questionário revelam que 40,5% dos participantes se sentiram mais criativos e originais na etapa sem o uso do *ChatGPT*. Esse resultado é positivo do ponto de vista educacional, pois indica que uma parte expressiva dos estudantes reconhece sua própria capacidade de gerar ideias originais de forma autônoma. Tal percepção fortalece o pensamento crítico e a confiança criativa, habilidades fundamentais na formação de engenheiros de software inovadores. Em relação à flexibilidade, os requisitos produzidos com o apoio do *ChatGPT-3.5* apresentaram maior diversidade de categorias e perspectivas, sugerindo uma ampliação do repertório criativo promovida pela ferramenta.

Por outro lado, foram identificadas limitações importantes em relação ao *ChatGPT-3.5*, como a necessidade de uma filtragem dos requisitos gerados para garantir relevância e adequação, bem como a tendência do modelo a produzir ideias genéricas, prolixas ou que se desviam do escopo definido. Esses aspectos ressaltam a importância da supervisão humana e do uso estratégico da ferramenta, reforçando que o *ChatGPT-3.5* deve ser um complemento ao processo criativo, e não seu substituto.

Do ponto de vista educacional, a atividade proposta proporcionou uma experiência prática valiosa para que os estudantes compreendessem as potencialidades e limitações do uso de LLMs para elicitação de requisitos criativos. Essa reflexão crítica é fundamental para formar futuros engenheiros de software capazes de integrar essas ferramentas de forma consciente e eficaz, potencializando a criatividade e a qualidade dos produtos desenvolvidos.

Logo, o *ChatGPT-3.5* pode atuar como um estimulante da criatividade no contexto da engenharia de requisitos, especialmente

quando utilizado de forma estratégica e com mediação crítica por parte dos usuários. Embora não substitua a criatividade humana, o modelo demonstrou potencial para ampliar o repertório de ideias e apoiar a construção de requisitos mais elaborados. Como implicação prática, recomenda-se a incorporação orientada de LLMs em atividades educacionais e profissionais de ER, aliando suas capacidades técnicas à experiência e ao julgamento humano.

Este estudo apresenta limitações que devem ser consideradas. Primeiramente, trata-se de um estudo experimental, realizado em um contexto educacional específico, com estudantes de uma turma introdutória à Engenharia de Software, o que pode limitar a generalização dos resultados.

Outras limitações são referentes a: utilização apenas da versão gratuita do *ChatGPT*, restringindo a análise a uma única ferramenta; a avaliação da criatividade envolve certo grau de subjetividade. Para minimizar o efeito da subjetividade, utilizou-se os padrões do TTCT e a avaliação foi realizada por três avaliadores; os dados foram baseados em produções dos participantes e respostas autorrelatadas, sujeitos a possíveis vieses. Para minimizar vieses de interpretação, os dados foram analisados por dois pesquisadores.

Como trabalhos futuros, propõe-se aprofundar a investigação sobre diferentes formas de interação com LLMs no processo de elicitação de requisitos, considerando estratégias mais refinadas de *prompt engineering* para potencializar a qualidade e a criatividade das respostas geradas. Além disso, destaca-se a importância de identificar o momento mais adequado para a introdução dessas ferramentas no fluxo criativo, de modo a evitar interferências ou bloqueios na fase de ideação e preservar a originalidade das contribuições humanas. Outra linha promissora consiste na realização de estudos comparativos entre diferentes LLMs disponíveis no mercado, visando avaliar suas capacidades específicas na geração de requisitos criativos, bem como sua adaptação a diferentes perfis de usuários e contextos de aplicação. Por fim, sugere-se ampliar a análise para outros fatores que influenciam a criatividade, como a dinâmica colaborativa entre os participantes, a experiência prévia dos envolvidos e o grau de familiaridade com o domínio do problema, contribuindo assim para um entendimento mais abrangente sobre o papel das LLMs na formação de engenheiros de software criativos e críticos.

DISPONIBILIDADE DE ARTEFATO

Os artefatos utilizados estão disponíveis no material complementar no seguinte link: <https://figshare.com/s/51b15c04f5265d7d529f>

ACKNOWLEDGMENTS

Agradecemos a todos os participantes do estudo e aos membros do grupo de pesquisa USES pelo apoio. Esta pesquisa foi financiada pela CAPES - PROEX - Código de Financiamento 001, pelo CNPq, processos 314797/2023-8, 443934/2023-1 e 445029/2024-2, pelo Projeto nº 017/2024 – DIVULGA CT&I/FAPEAM, e pela Universidade do Estado do Amazonas, por meio do Programa de Produtividade Acadêmica 01.02.011304.026472/2023-87. Este trabalho também contou com apoio parcial da Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM), por meio do projeto POSGRAD 2025-2026.

REFERENCES

- [1] Zeynep G Akdemir-Beveridge, Arash Zaghi, and Connie Syharat. 2025. Understanding and Evaluating Engineering Creativity: Development and Validation of the Engineering Creativity Assessment Tool (ECAT). arXiv:2504.12481 [stat.OT] <https://arxiv.org/abs/2504.12481>
- [2] Ahmed M. Abdulla Alabbasi, Sue Hyeon Paek, Daehyun Kim, and Bonnie Crumond. 2022. What do educators need to know about the Torrance Tests of Creative Thinking: A comprehensive review. *Frontiers in Psychology* Volume 13 - 2022 (2022). doi:10.3389/fpsyg.2022.1000385
- [3] Hung-Fu Chang and Tong Li. 2024. A Framework for Collaborating a Large Language Model Tool in Brainstorming for Triggering Creative Thoughts. arXiv:2410.11877 [cs.HC] <https://arxiv.org/abs/2410.11877>
- [4] Marian Daun, Alicia M. Grubb, Viktoria Stenkova, and Bastian Tenbergen. 2022. A systematic literature review of requirements engineering education. *Requir. Eng.* 28, 2 (May 2022), 145–175. doi:10.1007/s00766-022-00381-9
- [5] Wouter Groeneveld. 2025. Leveraging Creativity as a Problem Solving Tool in Software Engineering. arXiv:2502.03280 [cs.SE] <https://arxiv.org/abs/2502.03280>
- [6] Wouter Groeneveld, Laurens Luyten, Joost Venninkens, and Kris Aerts. 2021. Exploring the Role of Creativity in Software Engineering. arXiv:2101.00837 [cs.SE] <https://arxiv.org/abs/2101.00837>
- [7] Victoria Jackson, André van der Hoek, Rafael Prikladnicki, and Kuldeep Singh. 2025. Introduction to the Special Issue on Creativity in Software Engineering. *IEEE Software* 42, 3 (2025), 27–32. doi:10.1109/MS.2025.3540536
- [8] Victoria Jackson, Bogdan Vasilescu, Daniel Russo, Paul Ralph, Malihah Izadi, Rafael Prikladnicki, Sarah D'Angelo, Sarah Inman, Anielle Lisboa, and Andre van der Hoek. 2024. Creativity, Generative AI, and Software Development: A Research Agenda. arXiv:2406.01966 [cs.SE] <https://arxiv.org/abs/2406.01966>
- [9] Ranin Khojah, Mazen Mohamad, Philipp Leitner, and Francisco Gomes de Oliveira Neto. 2024. Beyond Code Generation: An Observational Study of ChatGPT Usage in Software Engineering Practice. *Proc. ACM Softw. Eng.* 1, FSE, Article 81 (July 2024), 22 pages. doi:10.1145/3660788
- [10] Kyung Hee Kim. 2006. Can We Trust Creativity Tests? A Review of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal* 18, 1 (2006), 3–14. arXiv:https://doi.org/10.1207/s15326934crj1801_2 doi:10.1207/s15326934crj1801_2
- [11] Marcela Litcanu, Octavian Postean, Cosmin Oros, and Alin Vasile Mnerie. 2015. Brain-Writing Vs. Brainstorming Case Study For Power Engineering Education. *Procedia - Social and Behavioral Sciences* 191 (2015), 387–390. doi:10.1016/j.sbspro.2015.04.452 The Proceedings of 6th World Conference on educational Sciences.
- [12] Nuno Marques, Rodrigo Rocha Silva, and Jorge Bernardino. 2024. Using ChatGPT in Software Requirements Engineering: A Comprehensive Review. *Future Internet* 16, 6 (2024). doi:10.3390/fi16060180
- [13] Orit Shaer, Angelora Cooper, Andrew L. Kun, and Osnat Mokry. 2024. Toward Enhancing Ideation through Collaborative Group-AI Brainwriting. In *Joint Proceedings of the ACM IUI Workshops 2024 (CEUR Workshop Proceedings, Vol. 3660)*. CEUR-WS.org, Greenville, SC, USA. <https://ceur-ws.org/Vol-3660/paper13.pdf> Workshop co-located with the 29th Annual ACM Conference on Intelligent User Interfaces.
- [14] Pradeep Kashinath Waychal. 2014. Developing Creativity Competency of Engineers. In *Proceedings of the 2014 American Society for Engineering Education Annual Conference Exposition*. American Society for Engineering Education, Indianapolis, IN, USA. <https://monolith.asee.org/public/conferences/32/papers/8802/view> Paper #8802, Entrepreneurship Engineering Innovation Education Session 3.
- [15] Xiaotong (Tone) Xu, Jiayu Yin, Catherine Gu, Jenny Mar, Sydney Zhang, Jane L. E., and Steven P. Dow. 2024. Jamplate: Exploring LLM-Enhanced Templates for Idea Reflection. In *Proceedings of the 29th International Conference on Intelligent User Interfaces* (Greenville, SC, USA) (IUI '24). Association for Computing Machinery, New York, NY, USA, 907–921. doi:10.1145/3640543.3645196
- [16] Didar Zowghi and Chad Coulin. 2005. Requirements elicitation: A survey of techniques, approaches, and tools. In *Engineering and Managing Software Requirements*. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, 19–46. <https://opus.lib.uts.edu.au/bitstream/10453/11626/1/2005003295.pdf> Open Access via OPUS UTS, originally part of a Springer publication.