

A Integração de LLMs no Desenvolvimento Colaborativo de Software: Uma Análise Sociotécnica no GitHub Discussions

Raquel de Souza Zimmer¹, Leticia dos Santos Machado¹, Karina Kohl Silveira¹

¹Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

{rszimmer, leticia.machado, karina.kohl}@inf.ufrgs.br

Abstract. *Integrating Large Language Models (LLMs) into software development reshapes human-human interaction dynamics, yet the implications of externalizing these interactions in collaborative ecosystems remain underexplored. This study investigates why developers share ChatGPT links on GitHub Discussions by analyzing the DevGPT dataset through the 3C+P collaboration model (Communication, Coordination, Cooperation, and Awareness). Thematic analysis reveals four main categories: contextualizing doubts, proposing ideas, proposing solutions, and seeking validation. Results show LLMs act as cognitive mediators in pre-deliberative stages, communication mechanisms, redistributing cooperative workload, and expanding collective awareness by externalizing previously private reasoning. We argue these shared interactions function as digital Boundary Negotiation Artifacts, redefining AI's role within collaborative ecosystems.*

Resumo. *A incorporação de Large Language Models (LLMs) ao desenvolvimento de software tem reconfigurado a dinâmicas de interação humano-humano, mas a externalização pública dessas interações em ecossistemas colaborativos permanece pouco explorada. Este estudo investiga por que desenvolvedores compartilham links do ChatGPT no GitHub Discussions, analisando o dataset DevGPT sob o modelo 3C+P (Comunicação, Coordenação, Cooperação e Percepção). A análise temática revela quatro categorias: contextualização de dúvidas, proposição de ideias, proposição de soluções e busca por validação. Os resultados indicam que LLMs atuam como mediadores cognitivos em estágios pré-deliberativos, reconfigurando a comunicação, redistribuindo a carga cooperativa e ampliando a percepção coletiva ao externalizar raciocínios privados. Argumenta-se que essas interações operam como Artefatos de Negociação de Fronteira digitais, redefinindo o papel da IA em ecossistemas colaborativos.*

1. Introdução

O desenvolvimento de software é reconhecido como uma atividade intrinsecamente colaborativa [Whitehead 2007, Dabbish et al. 2012] sustentada por ecossistemas sociotécnicos que articulam pessoas, artefatos e processos distribuídos. Plataformas como o GitHub consolidaram esse modelo ao integrar versionamento, documentação e espaços públicos de deliberação, viabilizando a construção coletiva de conhecimento técnico em larga escala. Nos últimos anos, entretanto, esses ecossistemas passaram a incorporar um novo

tipo de ator: assistentes inteligentes baseados em *Large Language Models* (LLMs), como ChatGPT, GitHub Copilot, Gemini e Claude Code.

A adoção quase universal dessas tecnologias tem transformado não apenas a produtividade individual [Zhao et al. 2025], mas também as dinâmicas públicas de interação em comunidades on-line, como Stackoverflow¹, Reddit² e GitHub Discussions³. Ao mediar a formulação de perguntas, a elaboração de soluções e a argumentação técnica, os LLMs introduzem novas camadas de mediação nas práticas colaborativas, tensionando modelos tradicionalmente centrados na interação exclusivamente humano–humano [Zhao et al. 2025].

Após o lançamento do ChatGPT em 2022 [OpenAI 2022], plataformas de perguntas e respostas, do inglês *questions and answers* (Q&A) como o Stack Overflow registraram queda significativa de atividade [Burtch et al. 2024, Guzman e Lewis 2020], sugerindo uma migração de interações predominantemente humano–humano para interações humano–LLM. Tradicionalmente estruturado como um ecossistema de validação coletiva, no qual respostas são avaliadas por meio de votos e sistemas de reputação, o Stack Overflow exemplifica um modelo de construção de conhecimento fortemente baseado na curadoria comunitária de desenvolvedores de software. A emergência de interações “sintéticas”, isto é, humano–artificiais [Guzman e Lewis 2020], indica uma reconfiguração desses arranjos colaborativos, deslocando parte da mediação do conhecimento da comunidade para agentes baseados em Inteligência Artificial (IA).

A forma como essa mudança de paradigma está refletindo no processo de desenvolvimento colaborativo ainda é pouco explorada na literatura de sistemas colaborativos, uma vez que exige repensar os modelos clássicos de colaboração à luz da integração de sistemas inteligentes em espaços colaborativos da comunidade de desenvolvimento de software [Shi et al. 2025].

A motivação deste estudo reside na necessidade crítica de compreender, sob a perspectiva da área de Sistemas Colaborativos (SC), a colaboração humano–LLM não como uma interação pontual ou instrumental, mas como um processo sociotécnico integrado a ecossistemas colaborativos complexos. Em vez de tratar os LLMs apenas como ferramentas individuais de apoio à programação, este trabalho os analisa como elementos que passam a mediar práticas coletivas de comunicação, coordenação, cooperação e percepção em comunidades de desenvolvimento. Para isso, utilizamos o modelo 3C+P de colaboração [Fuks et al. 2005] como lente analítica, investigando de que maneira a adoção de LLMs reconfigura essas dimensões clássicas e quais novos desafios emergem para a área de Sistemas Colaborativos diante da incorporação de agentes sintéticos nos processos deliberativos e produtivos do desenvolvimento de software.

Assim, esta pesquisa busca compreender os objetivos que levam desenvolvedores a compartilharem suas conversas com LLMs em ambientes públicos de colaboração e como esse compartilhamento se reflete nas práticas de desenvolvimento de software. Mais especificamente, são explorados os dados públicos do GitHub Discussions [Xiao et al. 2024]. Para guiar este estudo, a seguinte pergunta de pesquisa foi definida:

¹<https://stackoverflow.com/questions>

²<https://www.reddit.com/>

³<https://github.com/features/discussions>

RQ1. Quais objetivos e funções sócio-técnicas emergem quando desenvolvedores compartilham interações com LLMs em espaços colaborativos públicos como o GitHub Discussions?

Este trabalho preenche a lacuna sobre a compreensão das novas dinâmicas de colaboração mediadas por LLMs em ecossistemas de desenvolvimento de software. Dessa forma as contribuições para a área de Sistemas Colaborativos são: (i) caracterizar empiricamente os objetivos que levam desenvolvedores a compartilharem interações com LLMs em fóruns públicos de desenvolvimento de software; (ii) reinterpretar o modelo 3C+P à luz da colaboração humano-IA; e (iii) propor o conceito de LLM compartilhado como um Artefato de Negociação de Fronteira, ampliando seu escopo para artefatos digitais gerados por LLMs.

2. Referencial Teórico

Ecossistemas de conexão em larga escala funcionam como espaços de construção de conhecimento coletivo [Mamykina et al. 2011], onde a interação social e técnica ocorre de forma especializada, unindo desenvolvedores que buscam conhecimentos específicos e detentores desses conhecimentos [Treude et al. 2011], mas distribuída, pois o acesso é universal [Kittur et al. 2013], servindo ao propósito fundamental de resolução de problemas e troca de saberes. No entanto, o funcionamento desses ecossistemas não é isento de desafios, como por exemplo, a sobrecarga de informação [Agichtein et al. 2008] e a complexidade em manter a qualidade do conteúdo gerado [Zhang et al. 2021].

Sob a ótica de Sistemas Colaborativos, a colaboração é conceituada por diversos modelos teóricos [Nardi 1996, Rogers e Ellis 1994, Lee e Paine 2015], com destaque para o Modelo 3C de Colaboração [Fuks et al. 2005]. Este modelo decompõe o processo nas dimensões interdependentes de comunicação, coordenação e cooperação, integrando ainda a percepção (awareness) como elemento transversal relevante [Gerosa et al. 2003].

Os pilares do modelo 3C+P estruturam a colaboração a partir de dimensões complementares. **Comunicação** refere-se à troca de mensagens que sustenta negociação, tomada de decisão e construção de entendimento mútuo. **Coordenação** envolve o gerenciamento de pessoas, tarefas e recursos, organizando esforços no tempo e no espaço para evitar conflitos e desperdícios. **Cooperação** corresponde à atuação conjunta no espaço compartilhado de trabalho, onde artefatos são produzidos e manipulados em direção a objetivos comuns. Por fim, **Awareness (Percepção)** funciona como elemento central e facilitador da colaboração, permitindo compreender continuamente o que ocorre no grupo e no estado dos artefatos, orientando contribuições individuais.

Conforme sintetizado na matriz apresentada na Figura 1, cada dimensão passou por uma trajetória evolutiva: da comunicação síncrona localizada a agentes de IA mediadores; da supervisão manual a sistemas inteligentes de triagem; da edição individual ao social coding e à coautoria com IA; e de indicadores rudimentares de presença ao fornecimento de contexto situacional profundo gerado por LLMs. Essa transição evidencia a passagem de modelos centralizados e estáticos para ecossistemas distribuídos e, mais recentemente, aumentados por inteligência artificial, preservando os fundamentos conceituais do 3C+P enquanto redefine suas formas de operacionalização.

A dinâmica colaborativa centrada nesse modelo, tradicionalmente focada na interação entre humanos, atravessa hoje uma transformação com a emergência da

Matriz de Evolução do 3C+P

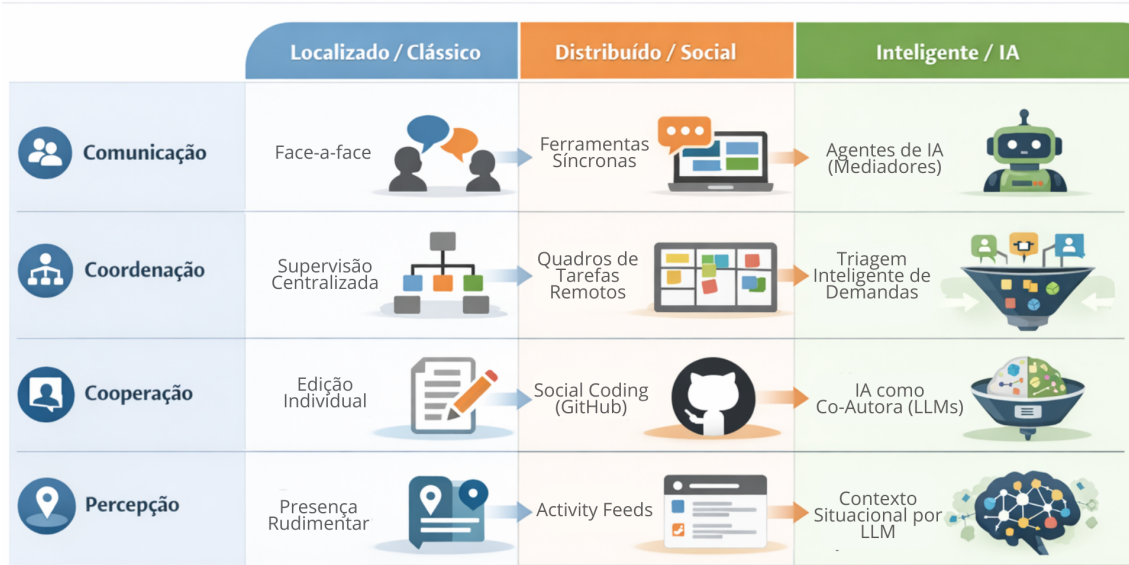


Figura 1. Matriz de evolução de ferramentas de suporte à colaboração sob os pilares de Comunicação, Coordenação, Cooperação e Percepção

colaboração humano-artificial, impulsionada pelos LLMs. A introdução dessas IAs nos ecossistemas de conexão em grande escala altera o paradigma de ferramentas passivas para agentes ativos, capazes de atuar como co-participantes na resolução de problemas e na geração de código ou conhecimento [Beganovic et al. 2023, Fan et al. 2023, Zheng et al. 2025]. Apesar de desafios amplamente documentados, como alucinações e vulnerabilidades [Gao et al. 2025], os LLMs vêm sendo progressivamente incorporados às práticas de desenvolvimento.

3. Trabalhos Relacionados

A colaboração em desenvolvimento de software tem sido caracterizada como um processo sociotécnico sustentado por artefatos compartilhados, transparência social e mecanismos de coordenação distribuída [Whitehead 2007, Dabbish et al. 2012]. Em plataformas como o GitHub, a visibilidade de ações, discussões e artefatos desempenha papel central na organização do trabalho coletivo [Treude e Storey 2011]. Esses estudos, contudo, assumem predominantemente interações humano–humano como base da colaboração.

A literatura recente passou a investigar LLMs como assistentes de produtividade e suporte cognitivo como assistentes inteligentes integrados a Ambientes de Desenvolvimento Integrados (IDEs) [Hou et al. 2024, Zhao et al. 2025]. Neste trabalho é explorada a externalização de interações entre desenvolvedores e LLMs, no caso o ChatGPT e analisada as novas dinâmicas de colaboração.

Sob a perspectiva de SC, artefatos desempenham papel importante na mediação da coordenação e na construção de entendimento compartilhado. Como parte dos artefatos observados na área de SC, Artefatos de Negociação de Fronteira, do inglês *Boundary Negotiation Artifact* - (BNA), são objetos tradicionalmente desenvolvidos por humanos e utilizados para resolver incertezas técnicas entre atores em contextos distribuídos [Herbsleb

2007]. Este estudo propõe que os LLMs operem como BNAs digitais, fundamentando-se no fato de que os *outputs* gerados por LLMs agora fazem parte do ciclo de colaboração.

4. Metodologia

Esta pesquisa caracteriza-se como um estudo qualitativo de natureza exploratória, fundamentado na análise de dados secundários provenientes do dataset *DevGPT* [Xiao et al. 2024], produzido para o *Mining Software Repositories (MSR) Challenge 2024*. O *DevGPT* compreende um conjunto de 17.955 entradas coletadas entre julho e outubro de 2023, consistindo em links de conversas entre desenvolvedores e o ChatGPT compartilhados na plataforma online de gestão e colaboração em projetos de software - GitHub. A estrutura geral do dataset abrange fontes que incluem o *GitHub Commits, Pull Requests, Issues, Discussions, Files* e *Hacker News Threads*.

Neste trabalho o foco de análise concentrou-se nas threads do GitHub Discussions totalizando 398 entradas. O GitHub Discussions é um espaço integrado aos repositórios da plataforma GitHub destinado a conversas abertas e colaborativas entre membros da comunidade. A escolha desse ambiente como unidade de análise justifica-se por seu papel central nas dinâmicas de comunicação e coordenação entre desenvolvedores.

4.1. Coleta de Dados

Do total de entradas do dataset *DevGPT* [Xiao et al. 2024], o subconjunto do GitHub Discussions foi submetido à etapa de "Pré-processamento dos dados" (Figura 2), resultando em 54 discussões elegíveis. O refinamento consistiu na remoção de duplicatas, preservando apenas a versão mais recente e singular de cada entrada; na exclusão de discussões em idiomas diferentes do inglês, identificadas pela biblioteca *Lingua* [Stahl 2026]; e no descarte de links que se tornaram indisponíveis ou foram removidos.

Os dados pré-processados foram organizados através de uma seleção aleatória em um subconjunto que viabilizasse a análise na íntegra das informações, resultando em 60 entradas onde (i) 30 entradas foram referentes às conversas entre desenvolvedores e o ChatGPT, e as outras (ii) 30 entradas compreenderam discussões do GitHub Discussions nas quais pessoas desenvolvedoras compartilharam com a comunidade links de conversas com o ChatGPT sobre diferentes assuntos (figura 2). Neste artigo serão apresentados os resultados da Análise Temática e associação teórica realizadas no conjunto de dados referente a troca de conversas entre desenvolvedores no canal coletivo - GitHub Discussions (ii), etapa ilustrada na figura 2 pela área demarcada como "Presente Estudo".

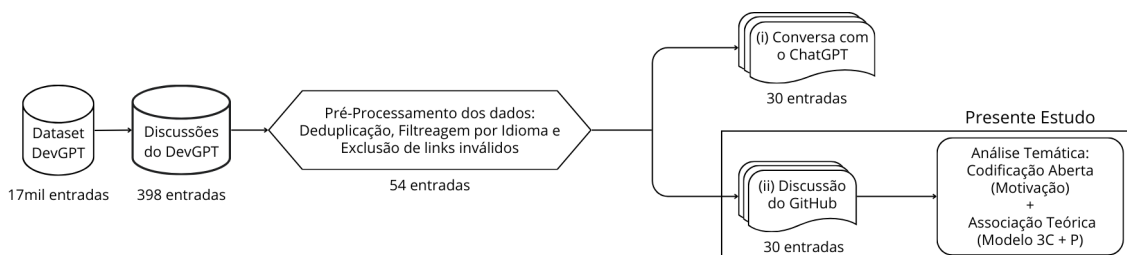


Figura 2. Visão geral do processo metodológico

4.2. Procedimentos de Análise Qualitativa

A análise dos dados seguiu uma abordagem de Análise Temática [Braun e Clarke 2006], adotando um procedimento sistemático, iterativo e realizado por meio da anotação manual [Braun e Clarke 2006] do compartilhamento dos links para conversas com o ChatGPT no GitHub Discussions.

O processo foi apoiado pelo conteúdo da conversa com o ChatGPT, pelo conteúdo das discussões do GitHub Discussions, pelo local de compartilhamento do link (**Descrição**: seção de texto que inicia uma discussão; **Comentário**: seção de texto que complementa uma descrição; **Resposta**: seção de texto vinculada a um comentário) e pelos papéis atribuídos pelo GitHub Discussions aos desenvolvedores participantes da thread (**Mantenedor**: um membro oficial do repositório com permissões administrativas; **Colaborador**: usuário não-membro do repositório com permissão para contribuir; **Usuário**: usuário sem vínculo com o repositórios; e **Autor**: usuário que inicia uma discussão). A análise pode ser dividida em três fases descritas a seguir:

Fase 1: Codificação Aberta e Indutiva. Inicialmente, realizou-se uma etapa manual de codificação aberta e indutiva [Braun e Clarke 2006] das postagens no GitHub Discussions, conduzida pelas três pesquisadoras de forma independente para criar uma lista inicial de códigos, identificando diálogos relevantes nos dados analisados. Divergências foram discutidas e o codebook foi refinado iterativamente até a estabilização das categorias. Esta etapa resultou na identificação das intencionalidades do compartilhamento de conversas realizadas no ChatGPT para a comunidade GitHub de desenvolvedores.

Fase 2: Associação dos Códigos com o modelo 3C+P. Na segunda fase, os códigos obtidos foram interpretados sob as lentes do modelo de colaboração 3C+Percepção [Fuks et al. 2005]. Esse processo foi feito com o cruzamento dos conceitos de Comunicação, Cooperação, Coordenação e Percepção apresentados na seção 2 com os códigos extraídos na fase 1 e informações pertinentes às discussões (Quem compartilhou o link para conversa com o ChatGPT; Onde o link para conversa com o ChatGPT foi compartilhado).

Fase 3: Criação das Categorias. As interpretações encontradas foram refinadas, nomeadas e descritas para formalizar as categorias de análise. Esta etapa culminou no refinamento do *codebook*. O *codebook*, juntamente com as definições de códigos, categorias e termos utilizados, está disponível para acesso⁴.

5. Resultados

A codificação inicial do subconjunto de 30 entradas gerou nove códigos. Ao serem interpretados à luz do modelo 3C+P [Fuks et al. 2005], os códigos derivaram um conjunto de quatro categorias, cada uma atrelada a um conceito do modelo. Essa construção é apresentada na Tabela 5.

5.1. Modelo 3C+P: Comunicação

A dimensão de Comunicação do modelo 3C+P foi representada pela categoria *Proposição de Ideias*, que visa englobar o esforço humano de divulgar em comunidades de desenvolvimento de software ideias pessoais que foram construídas ou refinadas com ajuda do ChatGPT.

⁴<https://doi.org/10.5281/zenodo.18804019>

Tabela 1. Categorização do compartilhamento das interações com o ChatGPT no GitHub Discussions fundamentada no modelo colaborativo 3C+P.

Categoria	Descrição	Códigos	Modelo 3C+P	Quote
Contextualização de Dúvidas	Expor o histórico de tentativas e a lógica de um problema não resolvido para facilitar a ajuda e o diagnóstico coletivo.	Exploração de Possíveis Problemas Transparência de log de erros	Awareness	<i>[AUTOR] "GPT writing functions: [link...] Summary: Message 1 = Environment description... Note: GPT originally hallucinated a wrong answer..."</i>
Proposição de Ideias	Compartilhar conhecimentos, respostas ou processos desenvolvidos em conjunto com o ChatGPT com objetivo de contribuir com a comunidade.	Refinamento Prévio Apoio na Explicação Nivelamento	Comunicação	<i>[MANTENEDOR] "You can't do that sort of query with our client... [USUÁRIO] "I really like this idea! ChatGPT tells me this is not possible :)"</i>
Proposição de Solução	Apresentar uma possível resolução para o problema discutido, utilizando o ChatGPT para gerar o código, diagnóstico ou explicação técnica...	Agilidade de Detalhamento Compartilhamento de Artefatos Práticos	Cooperação	<i>[AUTOR] "How do I complete the following code?... [MANTENEDOR] "I asked ChatGPT4 for a simple example... here's what it wrote..."</i>
Busca por Validação	Usar a conversa com o ChatGPT como evidência técnica para sustentar um argumento e buscar a confirmação ou segunda opinião da comunidade.	Exploração Antecipada Revisão de Qualidade	Coordenação	<i>[USUÁRIO] "Full disclosure, I used ChatGPT to help figure out a sensible plan of action... I hope these instructions might form a basis..."</i>

Observa-se dessa forma, uma possível intenção do desenvolvedor em compartilhar assuntos técnicos no GitHub Discussions para que os demais membros da comunidade possam interagir sobre o tema postado, trazendo novos comentários e fomentando a troca de ideias. Essa categoria foi desdobrada através da Análise Temática para os seguintes códigos: Refinamento Prévio, Apoio na Explicação e Nivelamento.

Refinamento Prévio diz respeito ao papel do ChatGPT na comunicação entre desenvolvedores como um auxiliar para organizar e refinar ideias. O quote

[AUTOR] "Improved UX and Security: Switching to Time-based One Time Passwords (TOTPs)! And a very interesting ChatGPT 4 conversation [about it] here: [link ChatGPT]" (ID 13)

apresenta que o autor da discussão trouxe para a comunidade melhorias sobre "Improved UX and Security" que podem auxiliar outros desenvolvedores com esse assunto. Isso

indica que o LLM atua como mediador prévio de formulação, sustentando a transição da interação humano-LLM para o debate coletivo.

Apoio na Explicação aborda o uso do ChatGPT como um validador de ideias. O quote de **ID 22**

[AUTOR] "What do you prefer as the name for a UUID where all bits are set to 1? Max UUID or Omni UUID" [USUÁRIO] "I don't think it is the best option to only make a poll between the proposed "Omni" option and the old "Max" option... I tried to find some more [options]... You can read the chat here: [link ChatGPT]" (ID 22)

expõe uma situação na qual um desenvolvedor, cujo papel é o de usuário não vinculado ao repositório, discorda das proposições do autor que iniciou a discussão. Tal usuário, utiliza uma conversa com o ChatGPT como uma forma de contraponto à proposição do autor. Essa situação revela como é possível utilizar LLMs para reforçar e diversificar a argumentação dos desenvolvedores, atribuindo maior profundidade à comunicação.

Nivelamento consiste em utilizar o ChatGPT como organizador de ideias e tradutor técnico. O quote

[AUTOR] "given the importance of the topic, it would be very useful to add this in the landing page how to fully / partially revert file(s) [...] ps: this might be a starting point [link ChatGPT]" (ID 2)

faz referência a como o autor da discussão, que é um usuário não vinculado ao repositório, deseja propor uma melhoria na *landing page* do projeto, mas devido ao seu conhecimento limitado sobre os comandos utilizados no repositório, precisou do ChatGPT para formalizar a ideia em uma linguagem que pudesse ser transmitida para a comunidade. Isso mostra como LLMs podem ser utilizados para dar consistência técnica aos conhecimentos ainda não formalizados de novos desenvolvedores, facilitando a comunicação entre diferentes níveis de conhecimento.

5.2. Modelo 3C+P: Cooperação

A dimensão de Cooperação do modelo 3C+P foi associada à categoria *Proposição de Solução*, que descreve o ato de apresentar resoluções para problemas discutidos na comunidade utilizando o ChatGPT para gerar código, diagnósticos ou explicações técnicas. O objetivo é colaborar com o progresso do projeto, independentemente do resultado final da solução ser positivo ou negativo. Essa categoria foi traduzida para os códigos: Agilidade de Detalhamento e Compartilhamento de Artefatos Práticos.

Agilidade de Detalhamento refere-se ao uso do ChatGPT para produzir uma explicação ou alternativa técnica que responde a uma dúvida ou problema proposto. O quote

[MANTENEDOR] "I don't know what ALB is but I think this can easily be done with NGINX: [link ChatGPT]" (ID 4)

mostra um mantenedor que, apesar de não conhecer uma ferramenta (ALB), utiliza o ChatGPT para fornecer prontamente uma solução equivalente usando NGINX. O uso do LLM nesse contexto evidencia sua função como acelerador de resposta técnica, redistribuindo esforço cognitivo sem deslocar a responsabilidade pela validação final.

Compartilhamento de Artefatos Práticos envolve a divulgação de processos de conversão ou geração de scripts feitos pelo ChatGPT para que outros possam testar ou utilizar a solução. No exemplo a seguir

[MANTENEDOR] "I asked ChatGPT to convert our powerwall.yml docker-compose file into a script with separate docker runs. I then asked it if it could convert it to a kubernetes manifest... To be fair, I'm sure there is some errors here... See the chat session: [link para conversa com o ChatGPT]" (ID 5)

o mantenedor apresenta uma conversa na qual pediu ao ChatGPT para converter arquivos de configuração (*docker-compose* para *kubernetes*), com objetivo de ajudar o autor a prosseguir com o próprio desenvolvimento. Mesmo sem garantia de perfeição do código desenvolvido, o compartilhamento da conversa pode servir como um ponto de partida técnico que encoraja a cooperação com teste coletivo de novas abordagens e integração de conhecimentos práticos.

5.3. Modelo 3C+P: Coordenação

A dimensão de Coordenação do modelo 3C+P foi aplicada à categoria *Busca por Validação*, que foca no uso do ChatGPT como uma evidência técnica para sustentar argumentos ou buscar uma segunda opinião da comunidade. Essa prática visa alinhar o entendimento técnico entre os participantes da Discussão e da comunidade. Essa categoria foi organizada sob os códigos: Exploração Antecipada e Revisão de Qualidade.

Exploração Antecipada Diz respeito ao emprego da IA para investigar conceitos ou arquiteturas antes de apresentá-los abertamente à comunidade. O quote

[AUTOR] "This may be silly, but I had a fantastic conversation with ChatGPT about this [comparison between SON Schema, JSON-LD and RDF], curious if this tracks with your thinking –¿ [link ChatGPT]" (ID 21)

apresenta o autor de uma discussão compartilhando uma comparação entre diferentes padrões técnicos (JSON-LD, RDF) feita com o ChatGPT para saber se o raciocínio está alinhado com o pensamento da comunidade. O compartilhamento opera como mecanismo de alinhamento coletivo, permitindo antecipar objeções e estabilizar interpretações antes da tomada de decisão.

Revisão de Qualidade Faz referência ao uso do ChatGPT como uma fonte de informação que necessita de curadoria da comunidade. No quote a seguir:

[AUTOR] "Hi all, I have a usecase where I want to trigger a method when something is added to my ref, or the ref didn't change for x seconds. I asked ChatGPT for a suggestion: [link GPT]... Would it be useful to add this to vueuse, or not?" (ID 1)

o autor apresenta à comunidade um caso de uso desenvolvido com o auxílio do ChatGPT a fim de validar sua utilidade e viabilidade de implementação no repositório. Esse cenário ilustra o papel recentemente atribuído aos desenvolvedores como curadores de informações geradas sinteticamente, uma função essencial para garantir que o conhecimento comunitário permaneça correto e coordenado.

5.4. Modelo 3C+P: Percepção (Awareness)

A dimensão de Percepção do modelo 3C+P foi instanciada para a categoria *Contextualização de Dúvidas*, que visa expor o histórico de tentativas e a lógica de um problema para facilitar o diagnóstico coletivo. O uso do ChatGPT aqui serve para dar visibilidade ao que já foi testado, permitindo que a comunidade perceba o estado atual da dúvida. Essa categoria foi mapeada para os códigos: Exploração de Possíveis Problemas e Transparência de log de erros.

Exploração de Possíveis Problemas Refere-se ao ato de compartilhar reflexões construídas em conjunto com o ChatGPT sobre riscos ou falhas potenciais em uma implementação. O quote:

[AUTOR] "Should we worry about imports performance in handlers? My dialogue with chatgpt about it: [link ChatGPT]" (ID 7)

mostra o autor da discussão questionando a comunidade sobre performance e anexando seu diálogo com o ChatGPT sobre o tema. Essa prática possibilita aumentar a consciência da comunidade sobre possíveis problemas ou fragilidades do repositório, permitindo que todos acompanhem a linha de raciocínio preventiva desenvolvida pelo desenvolvedor juntamente com LLMs.

Transparência de log de erros Faz referência ao compartilhamento de erros de execução acompanhados do contexto fornecido pelo ChatGPT para ajudar na solução. No quote

[AUTOR] "Getting the following error: [Command Failed] when trying to execute: Browsershot::url('https://www.google.com/') - ¿save('example.png'); // Specify the output file. AI chat for context: [link ChatGPT]" (ID 8)

o autor da discussão apresenta um erro específico de comando falho e anexa a conversa com o ChatGPT para fornecer um contexto relacionado com as tentativas de solução anteriores. Ao anexar o diálogo com o ChatGPT, o autor externaliza tentativas prévias de resolução, ampliando a percepção coletiva sobre o estado do problema.

6. Discussão

Os resultados indicam que o compartilhamento de links para conversas com o ChatGPT no GitHub Discussions segue padrões associados aos papéis sociais dos participantes. O uso do LLM distribui-se de forma distinta entre autores, usuários e mantenedores, evidenciando sua incorporação às dinâmicas colaborativas do projeto.

Embora amplie a capacidade técnico-explicativa das interações, o LLM não substitui a deliberação coletiva. As decisões permanecem ancoradas na agência humana, especialmente nos mantenedores, reforçando que a IA atua como suporte cognitivo e mediador, e não como instância decisória. Esse movimento dialoga com a transição observada na literatura de uma lógica de desconfiança para produtividade assistida [Hou et al. 2024].

6.1. Autoridade Técnica e LLM como Suporte Cognitivo

Um achado central refere-se ao papel dos mantenedores como legitimadores simbólicos do uso de IA. Ao compartilharem links do ChatGPT para sustentar argumentos, contribuem para normalizar a ferramenta no ecossistema do projeto, reforçando mecanismos de transparência social descritos por Dabbish et al. (2012).

Além disso, observa-se que autores utilizam o ChatGPT predominantemente na formulação inicial de propostas, especialmente na categoria Proposição de Ideias. Esse padrão sugere que o LLM opera como mecanismo de preparação cognitiva ou scaffolding [Yilmaz e Karaoglan Yilmaz 2023], mitigando barreiras de entrada amplamente documentadas em projetos open source [Steinmacher et al. 2015].

Nesse contexto, o LLM sustenta o amadurecimento prévio de contribuições sem substituir expertise, contribuindo para elevar o nível técnico das discussões iniciais.

6.2. Os LLMs como Artefato de Negociação de Fronteira

Quando o link para a conversa com o ChatGPT é tornado público, o LLM deixa de operar apenas como suporte individual e passa a desempenhar função social distinta. Nesses casos, o conteúdo gerado atua como artefato de negociação de fronteira [Herbsleb 2007], mediando a articulação de argumentos em contextos de incerteza técnica.

Nos casos classificados como Busca por Validação e Proposição de Ideias, o compartilhamento do link funciona como âncora informacional [Treude e Storey 2011], estruturando a discussão e reduzindo ambiguidades interpretativas. Ao mobilizar o LLM como referência externa, o autor redistribui parcialmente a carga cognitiva da argumentação.

Esse movimento sugere também a redução de assimetrias de senioridade, ao aproximar contribuições iniciais das expectativas técnicas dos mantenedores. Ainda assim, a validação final permanece sob autoridade humana, preservando a centralidade da deliberação coletiva.

Os resultados sugerem não apenas a instância do modelo 3C+P, mas sua reconfiguração operacional [Fuks et al. 2005, Gerosa et al. 2003]. A presença de LLMs desloca parte da Comunicação para um estágio pré-deliberativo privado, redistribui a carga da Cooperação e altera os mecanismos de Awareness ao externalizar raciocínios previamente invisíveis.

7. Ameaças à Validade

Apesar do rigor metodológico adotado, este estudo apresenta limitações que devem ser consideradas na interpretação dos resultados.

Primeiramente, por tratar-se de uma investigação qualitativa sem pretensão de generalização estatística, o subconjunto de discussões foi reduzido para 30 (n=30). Os achados refletem padrões do dataset *DevGPT*, o que introduz o viés de modelo, visto que os dados se limitam ao ChatGPT e não consideram alternativas como Gemini e Claude. Além disso, os resultados focam no *GitHub Discussions*, podendo variar em outros contextos colaborativos (*Pull Requests, Issues*) ou projetos com diferentes culturas de governança (viés de plataforma). Adicionalmente, os dados analisados correspondem exclusivamente a interações nas quais desenvolvedores optaram por compartilhar publicamente links de conversas com LLMs. Assim, o estudo captura apenas o uso externalizado dessas ferramentas, excluindo interações privadas e não compartilhadas (viés de visibilidade). Como a análise se baseia em dados secundários públicos, não foi possível confirmar diretamente as intenções subjetivas dos participantes. As categorias interpretativas derivam da análise temática do conteúdo textual e refletem significados negociados nas interações observáveis.

Outro aspecto refere-se à predominância de interações com o ChatGPT no período analisado. Embora representativo do momento estudado, o foco em um modelo específico pode limitar a compreensão de dinâmicas associadas a outros LLMs (dependência de modelo). Por fim, os dados foram coletados entre julho e outubro de 2023, período inicial de consolidação do uso de LLMs em contextos colaborativos. Dado o ritmo acelerado de evolução tecnológica e de adaptação social, os resultados devem ser compreendidos como um retrato situado de uma fase específica da adoção dessas tecnologias.

8. Conclusões e Trabalhos Futuros

Este trabalho investigou os objetivos e os significados associados ao compartilhamento de links para conversas entre desenvolvedores e o ChatGPT em comunidades de desenvolvimento de software, com foco no GitHub Discussions. A partir da análise do dataset *DevGPT*, foi possível identificar que esses compartilhamentos, embora motivados por diferentes objetivos explícitos, produzem impactos involuntários nos processos colaborativos dessas comunidades. Entre esses impactos, destacam-se a reconfiguração das dinâmicas de argumentação técnica, o uso do ChatGPT como artefato mediador de comunicação e de negociação de significado, e a externalização de raciocínios que antes permaneceriam implícitos ou individuais.

Como trabalhos futuros, este estudo prevê a ampliação do *corpus* analisado, de modo a acompanhar a evolução contínua dos LLMs e suas formas de apropriação pelas comunidades de desenvolvimento. Pretende-se, ainda, realizar análises comparativas entre diferentes tipos de plataformas colaborativas, como fóruns e sistemas de perguntas e respostas, com o objetivo de compreender como distintas arquiteturas de interação influenciam o papel do ChatGPT nos processos colaborativos. Adicionalmente, será explorado o design de ferramentas de *social coding* que ofereçam suporte nativo à integração de IA, com ênfase na rastreabilidade das interações e na preservação do conhecimento coletivo gerado por meio dessas práticas.

Em síntese, os resultados indicam que os LLMs não são utilizados apenas como ferramentas individuais de apoio à programação, mas também como elementos ativos nas dinâmicas colaborativas das comunidades de desenvolvimento de software, contribuindo para a construção, negociação e disseminação do conhecimento técnico necessário à produção de software.

Agradecimento

Este trabalho foi apoiado pela FAPERGS, Processo nº 25/2551-0000883-2.

Sobre o uso de Inteligência Artificial Generativa

Este trabalho utilizou ferramentas de Inteligência Artificial Generativa para a produção da figura 1 a partir de instruções elaboradas pelos autores e verificação de ortografia, sintaxe e coesão por meio do envio de trechos textuais. Os modelos utilizados foram Gemini 3 e GPT-5.2. Todas as imagens e correções passaram por revisão humana antes de sua inclusão no artigo.

Referências

- Agichtein, E., Castillo, C., Donato, D., Gionis, A. e Mishne, G. (fev. de 2008) Finding high-quality content in social media. Em: *Proceedings of the 2008 International Conference on Web Search and Data Mining*. WSDM '08. New York, NY, USA: Association for Computing Machinery, pp. 183–194. ISBN: 9781595939272. DOI: 10.1145/1341531.1341557. URL: <https://dl.acm.org/doi/10.1145/1341531.1341557>.
- Beganovic, A., Jaber, M. A. e Abd Almisreb, A. (2023) Methods and applications of chatgpt in software development: A literature review. Em: *Southeast Europe Journal of Soft Computing* 12.1, pp. 08–12.
- Braun, V. e Clarke, V. (2006) Using thematic analysis in psychology. Em: *Qualitative Research in Psychology* 3.2, pp. 77–101. DOI: 10.1191/1478088706qp063oa.
- Burtch, G., Lee, D. e Chen, Z. (mai. de 2024) The consequences of generative AI for online knowledge communities. en. Em: *Scientific Reports* 14.1, p. 10413. ISSN: 2045-2322. DOI: 10.1038/s41598-024-61221-0. URL: <https://www.nature.com/articles/s41598-024-61221-0>.
- Dabbish, L., Stuart, C., Tsay, J. e Herbsleb, J. (fev. de 2012) Social coding in GitHub: transparency and collaboration in an open software repository. Em: *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. CSCW '12. New York, NY, USA: Association for Computing Machinery, pp. 1277–1286. ISBN: 9781450310864. DOI: 10.1145/2145204.2145396. URL: <https://doi.org/10.1145/2145204.2145396>.
- Fan, A., Gokkaya, B., Harman, M., Lyubarskiy, M., Sengupta, S., Yoo, S. e Zhang, J. M. (2023) Large language models for software engineering: Survey and open problems. Em: *2023 IEEE/ACM International Conference on Software Engineering: Future of Software Engineering (ICSE-FoSE)*. IEEE, pp. 31–53.
- Fuks, H., Raposo, A. B., Gerosa, M. A. e Lucena, C. J. (2005) Applying the 3C model to groupware development. Em: *International Journal of Cooperative Information Systems* 14.2–3, pp. 299–328. DOI: 10.1142/S0218843005001171. URL: <https://www.scopus.com/pages/publications/21144453963>.
- Gao, C., Hu, X., Gao, S., Xia, X. e Jin, Z. (mai. de 2025) The Current Challenges of Software Engineering in the Era of Large Language Models. Em: *ACM Trans. Softw. Eng. Methodol.* 34.5, 127:1–127:30. ISSN: 1049-331X. DOI: 10.1145/3712005. URL: <https://dl.acm.org/doi/10.1145/3712005>.
- Gerosa, M. A., Fuks, H. e Lucena, C. (2003) Analysis and Design of Awareness Elements in Collaborative Digital Environments: A Case Study in the AulaNet Learning Environment. en. Em: *Journal of Interactive Learning Research* 14.3. ERIC Number: EJ822238, pp. 315–332.
- Guzman, A. L. e Lewis, S. C. (2020) Artificial intelligence and communication: A Human–Machine Communication research agenda. Em: *New Media Society* 22.1, pp. 70–86. ISSN: 1461-7315. DOI: 10.1177/1461444819858691.
- Herbsleb, J. D. (mai. de 2007) Global Software Engineering: The Future of Socio-technical Coordination. Em: *Future of Software Engineering (FOSE '07)*, pp. 188–198. DOI: 10.1109/FOSE.2007.11. URL: <https://ieeexplore.ieee.org/abstract/document/4221620>.

- Hou, X., Zhao, Y., Liu, Y., Yang, Z., Wang, K., Li, L., Luo, X., Lo, D., Grundy, J. e Wang, H. (dez. de 2024) Large Language Models for Software Engineering: A Systematic Literature Review. Em: *ACM Trans. Softw. Eng. Methodol.* 33.8. ISSN: 1049-331X. DOI: 10.1145/3695988. URL: <https://doi.org/10.1145/3695988>.
- Kittur, A., Nickerson, J. V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., Lease, M. e Horton, J. (fev. de 2013) The future of crowd work. Em: *Proceedings of the 2013 conference on Computer supported cooperative work*. CSCW '13. New York, NY, USA: Association for Computing Machinery, pp. 1301–1318. ISBN: 9781450313315. DOI: 10.1145/2441776.2441923. URL: <https://dl.acm.org/doi/10.1145/2441776.2441923>.
- Lee, C. P. e Paine, D. (fev. de 2015) From The Matrix to a Model of Coordinated Action (MoCA): A Conceptual Framework of and for CSCW. Em: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work Social Computing*. CSCW '15. New York, NY, USA: Association for Computing Machinery, pp. 179–194. ISBN: 9781450329224. DOI: 10.1145/2675133.2675161. URL: <https://dl.acm.org/doi/10.1145/2675133.2675161>.
- Mamykina, L., Manoim, B., Mittal, M., Hripcsak, G. e Hartmann, B. (mai. de 2011) Design lessons from the fastest qa site in the west. Em: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. New York, NY, USA: Association for Computing Machinery, pp. 2857–2866. ISBN: 9781450302289. DOI: 10.1145/1978942.1979366. URL: <https://doi.org/10.1145/1978942.1979366>.
- Nardi, B. A. (1996) Activity theory and human-computer interaction. Em: *Context and consciousness: Activity theory and human-computer interaction* 436.1, pp. 7–16.
- OpenAI (mar. de 2022). pt-BR. URL: <https://openai.com/pt-BR/index/chatgpt/> (acesso em 08/02/2026).
- Rogers, Y. e Ellis, J. (1994) Distributed Cognition: An Alternative Framework for Analysing and Explaining Collaborative Working. EN. Em: *Journal of Information Technology* 9.2, pp. 119–128. ISSN: 0268-3962. DOI: 10.1177/026839629400900203. URL: <https://doi.org/10.1177/026839629400900203>.
- Shi, M., Wang, G., Zhang, S., Jiang, J. e Xiang, Y. (2025) AISCW: Artificial Intelligence Supported Cooperative Work The New Frontier of CSCW. en. Em: *Computer Supported Cooperative Work and Social Computing*. Ed. por Sun, H., Fan, H., Gao, Y., Wang, X., Liu, D., Du, B. e Lu, T. Singapore: Springer Nature, pp. 3–17. ISBN: 9789819623730. DOI: 10.1007/978-981-96-2373-0_1.
- Stahl, P. (2026). Python Library: Lingua - Language Detector. URL: <https://pemistahl.github.io/lingua-py> (acesso em 26/02/2026).
- Steinmacher, I., Graciotto Silva, M. A., Gerosa, M. A. e Redmiles, D. F. (mar. de 2015) A systematic literature review on the barriers faced by newcomers to open source software projects. Em: *Information and Software Technology* 59, p. 2015. ISSN: 0950-5849. DOI: 10.1016/j.infsof.2014.11.001. URL: <https://www.sciencedirect.com/science/article/pii/S0950584914002390>.

- Treude, C., Barzilay, O. e Storey, M.-A. (mai. de 2011) How do programmers ask and answer questions on the web? (NIER track). Em: *Proceedings of the 33rd International Conference on Software Engineering. ICSE '11*. New York, NY, USA: Association for Computing Machinery, pp. 804–807. ISBN: 9781450304450. DOI: 10.1145/1985793.1985907. URL: <https://dl.acm.org/doi/10.1145/1985793.1985907>.
- Treude, C. e Storey, M.-A. (2011) Effective communication of software development knowledge through community portals. Em: *Proceedings of the 19th ACM SIGSOFT symposium and the 13th European conference on Foundations of software engineering. ESEC/FSE '11*. New York, NY, USA: Association for Computing Machinery, pp. 91–101. ISBN: 9781450304436. DOI: 10.1145/2025113.2025129. URL: <https://dl.acm.org/doi/10.1145/2025113.2025129>.
- Whitehead, J. (2007) Collaboration in software engineering: A roadmap. Em: *Future of Software Engineering (FOSE'07)*. IEEE, pp. 214–225.
- Xiao, T., Treude, C., Hata, H. e Matsumoto, K. (2024) Devgpt: Studying developer-chatgpt conversations. In *2024 IEEE/ACM 21st International Conference on Mining Software Repositories (MSR)*.
- Yilmaz, R. e Karaoglan Yilmaz, F. G. (ago. de 2023) Augmented intelligence in programming learning: Examining student views on the use of ChatGPT for programming learning. Em: *Computers in Human Behavior: Artificial Humans 1.2*, p. 100005. ISSN: 2949-8821. DOI: 10.1016/j.chbah.2023.100005. URL: <https://www.sciencedirect.com/science/article/pii/S2949882123000051>.
- Zhang, H., Wang, S., Chen, T.-H. e Hassan, A. E. (nov. de 2021) Reading Answers on Stack Overflow: Not Enough! Em: *IEEE Transactions on Software Engineering* 47.11, pp. 2520–2533. ISSN: 1939-3520. DOI: 10.1109/TSE.2019.2954319. URL: <https://ieeexplore.ieee.org/abstract/document/8906075>.
- Zhao, W., Shang, W. e Liu, Y. (2025) From Code Completion to Autonomous Pipeline Orchestration: How LLM-Powered Developer Tools Are Reshaping Software Engineering Workflows. en. Em: *American Journal Of Big Data* 6.05, pp. 111–139. ISSN: 2688-9994. DOI: 10.71465/ajbd3444. URL: <https://australiansciencejournals.com/bigdata/article/view/3444>.
- Zheng, Z., Ning, K., Zhong, Q., Chen, J., Chen, W., Guo, L., Wang, W. e Wang, Y. (2025) Towards an understanding of large language models in software engineering tasks. Em: *Empirical Software Engineering* 30.2, p. 50.