**Ficha da Pesquisa**

| **Título do Trabalho** | Quão acostumado o chat gpt está a perguntas de desenvolvedores ? |
| --- | --- |
| **Autor(res) e instituição** | João Vitor Durso Ferraz, Igor Wiese, Lucas Silva Couto |
| **Email do autor(res)** | [pg405541@uem.br](mailto:pg405541@uem.br) , pg908455@uem.br |
| **Orientador(es)** | Prof.Dra.Lynnie |
| **Tipo** | ( ) TCC ( \* ) Só para cumprir a disciplina |
| **Contexto** | O trabalho atual pretende ser uma oportunidade de explorar algumas técnicas de manipulação de data sets que podem vir a ser proveitosas para o escopo principal. |
| **Objetivo da pesquisa** | Analisar o dataset e definir quais os tipos de perguntas mais comuns feitas por desenvolvedores |
| **Referencial teórico** | Usando como base o artigo Analyzing Developer Use of ChatGPT, onde o mesmo analisa os erros mais comuns e como o prompt influencia o erro. De maneira simplificada se usou de técnicas de análise de texto para categorizar as conversas. O trabalho atual seria estender essa pesquisa por texto por todo o dataset, aprofundando em classes mais específicas de pergunta. |
| **Trabalhos relacionados** | * Analyzing Developer Use of ChatGPT Generated Code in Open Source GitHub Projects , fala bastante a respeito de como as pessoas usam o código do chat para resolver o problema com pequenas alterações. * Write me this Code: An Analysis of ChatGPT Quality for Producing Source Code , como o uso do chat ajuda os desenvolvedores a criar. * Investigating the Utility of ChatGPT in the Issue Tracking System: An Exploratory Study, fala sobre como utilizam o chat para lidar com erros. * ChatGPT in Action: Analyzing Its Use in Software Development, fala sobre a importância do prompt com relação ao tempo da conversa * The role of library versions in Developer-ChatGPT conversations, mostra alguns pontos importantes que acontecem de possíveis erros quanto às bibliotecas. |
| **Problema de pesquisa** | Descobrir qual o estado atual do chat gpt e sua confiança quando se trata de resolver problemas em códigos. |
| **Justificativa e relevância** | Uma das grandes vantagens da popularização das IAs é o auxílio que elas dão aos programadores iniciantes, entretanto é sempre importante analisar onde se encontram as maiores dificuldades do ambiente. Esse tipo de levantamento ajuda a montar o cenário atual da programação, encontrando onde está a maior dificuldade dos jovens programadores. Tal levantamento auxiliará professores e desenvolvedores a onde se encontram as lacunas do cenário da programação. |
| **Evidências do problema** | O dataset é montado justamente entre interações dos desenvolvedores e o chat, justamente com o objetivo de resolverem problemas. A própria existência do dataset evidencia que as pessoas usam ele para respostas, portanto deve-se analisar onde se encontram as maiores lacunas desse processo. Usar um código do chat sem entender o processo pode ser prejudicial, pois podem conter vulnerabilidades de segurança, são partes tiradas de um contexto maior, ele pode não cobrir cenários diferentes do descrito no prompt. Portanto é de suma importância entender quais as deficiências do cenário, para que o auxílio do chat seja melhor aproveitado. |
| **Hipótese/Questão** | A hipótese é de que o chat gpt é útil para resolver problemas desde que o prompt seja bem descrito. |
| **Proposta de Solução** | Para responder a pergunta será usado um dataset montado com várias conversas entre desenvolvedores no chat. Usando técnicas de mineração pretende-se levantar os perfis de problemas mais comuns e ver geralmente qual a demora ou quantas iterações são necessárias até o usuário chegar em uma resposta satisfatória. |
| **Projeto de Avaliação** | Para avaliar o projeto os resultados serão comparados com outros artigos da área e comparar se as respostas entre eles está muito discrepante. Obviamente que uma semelhança não é esperada uma vez que a pergunta principal é diferente. |
| **Falseamento** |  |
| **Abordagem** | ( ) Teórica ( \* ) Empírica  Se empírica: ( \* ) Quantitativa ( ) Qualitativa ( ) Mista |
| **Finalidade** | ( ) Explanatória ( ) Exploratória ( \* ) Descritiva |
| **Método a ser seguido na pesquisa** | ( ) Experimento ( \* ) Levantamento/survey ( ) Estudo de caso ( ) Grounded Theory  ( ) Etnografia ( ) Pesquisa-ação ( ) Outro: \_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| **Técnicas de coleta de dados** | ( ) Medição ( ) Questionário ( ) Entrevista ( ) Grupo focal  ( ) Observação direta ( \* ) Coleta de Documentos ( ) Benchmark  ( ) Outras: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| **Dados** | Os dados fornecidos pelo dataset são conversas entre desenvolvedores e o chat gpt |
| **Técnicas de análise dos dados coletados** | ( ) Estatística descritiva ( \* ) Estatística inferencial ( ) Análise do discurso  ( ) Análise de conteúdo ( ) Codificação de dados ( ) Outra: \_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| **Validade**  Como você auto-avalia a validade da pesquisa? | Validade interna (o quão robusto é o estudo em relação à possível introdução de erros ou vieses)? Péssimo ( ) 0 ( ) 1 ( ) 2 (\* ) 3 ( ) 4 ( ) 5 Excelente  Validade externa (o quão generalizável ou aplicável em contextos diferentes são os resultados)? Péssimo ( ) 0 ( ) 1 ( \* ) 2 ( ) 3 ( ) 4 ( ) 5 Excelente |
| **Contribuições científicas** | Analisar o cenário atual da inteligência artificial mais popular em termos de número de usuários. |
| **Contribuições tecnológicas** |  |
| **Estágio do trabalho** | ( ) Proposta não defendida ( \* ) Proposta já defendida |
|  | |
| **Outros comentários** |  |

**Referências citadas neste documento.**