

Vergleich zwischen Erklärbarkeitsanforderungen von LLMs und Anforderungsanalysten

Kommentiert [M1]: Erfordert kritische Reflexion, mehr als bei anderen Themen

Hintergrund

Im Rahmen unseres Projekts [softXplain](#) haben wir Erklärungsbedarf von Nutzern in App-Reviews identifiziert [1]. Dieser Erklärungsbedarf entsteht häufig, wenn Nutzer Fragen zur Software haben, insbesondere bei unerwartetem Systemverhalten. Um diesen Bedarf zu decken, müssen Anforderungen an die Erklärbarkeit formuliert und darauf basierend Erklärungen entwickelt werden. Ein vielversprechender Ansatz zur (semi-)automatisierten Unterstützung dieses Prozesses sind generative LLMs (Large Language Models) wie ChatGPT oder Google Gemini. Diese Modelle können basierend auf Erklärungsbedarf automatisch Anforderungen und Erklärungen generieren. Allerdings muss untersucht werden, wie präzise und korrekt die von LLMs generierten Anforderungen und Erklärungen im Vergleich zu denen von Anforderungsanalysten sind.

Aufgabe

In dieser Arbeit sollen Sie Anforderungen und Erklärungen für bereits identifizierten Erklärungsbedarf aus einem Datensatz von App-Reviews formulieren lassen. Zunächst lassen Sie Anforderungsanalysten in einem strukturierten Workshop Anforderungen und Erklärungen entwickeln. Anschließend sollen Sie mithilfe von LLMs (z.B. ChatGPT) ähnliche Anforderungen und Erklärungen generieren, wobei Sie geeignete Prompts einsetzen. Schließlich vergleichen Sie die manuell und automatisch generierten Ergebnisse in einer Studie. Diese Arbeit gliedert sich in die folgenden Schritte:

1. Einarbeiten in die Grundlagen von Erklärbarkeit und des Requirements Engineering
2. Einarbeiten in den App-Review-Datensatz [1]
3. Erstellung von Konzepten:
 - a. Entwicklung eines Workshops zur Formulierung von Erklärbarkeitsanforderungen und Erklärungen durch Anforderungsanalysten und LLMs
 - b. Vergleich der automatisch von einem LLM generierten Anforderungen und Erklärungen mit denen der Anforderungsanalysten
4. Durchführung des Workshops (mindestens 6 Teilnehmer)
 - a. Formulierung von (mindestens) 200 Erklärbarkeitsanforderungen und Erklärungen durch Anforderungsanalysten.
5. Generierung von Erklärbarkeitsanforderungen und Erklärungen mithilfe eines LLMs (z.B. ChatGPT) durch gezieltes Prompt-Finetuning
6. Durchführung einer Studie, in der Anforderungsanalysten und Nutzer darüber abstimmen, welche Anforderungen und Erklärungen sie bevorzugen
7. Optional: Evaluation eines Tools zur automatischen Formulierung von Anforderungen und Erklärungen
 - a. Einarbeiten in ein vorhandenes Tool [2]
 - b. Durchführung der Interviewstudie (mindestens 10 Teilnehmern)
8. Evaluation und Diskussion der Ergebnisse

Organisatorisches

Betreuer: Martin Obaidi

Prüfer: Prof. Dr. Kurt Schneider

Beginn: ab sofort möglich

Literaturempfehlung

[1] [Automatisierte Detektion von Erklärungsbedarf in Nutzerfeedback zu Software](#)

[2] [Automatische Erstellung von Erklärbarkeitsanforderungen und Erklärungen](#)