

Reforge

Natalie Stricker
Fakultät Elektrotechnik, Medien und Informatik
Ostbayerische Technische Hochschule Amberg-Weiden
Amberg, Deutschland

Einleitung

Die Einleitung behandelt die zunehmende Bedeutung von KI in verschiedenen Lebensbereichen, wie autonome Fahrzeuge und Bildung. Es wird diskutiert, wie KI unterstützend eingesetzt werden kann, aber auch potenzielle Probleme wie Bias und Abhängigkeit werden angesprochen. Die Arbeit zielt darauf ab, einen webbasierten Berichtsgenerator namens zu entwickeln, um Abschlussarbeiten zu analysieren und zusammenzufassen. Die Arbeit ist in Kapitel unterteilt, die von theoretischen Grundlagen über Forschungsstand bis zur Implementierung und Evaluierung führen.

Grundlagen

In dem Kapitel "Grundlagen" geht es um die automatisierte Textgenerierung und die Verwendung eines Sprachmodells, insbesondere des -3.5-Turbo Modells von OpenAI. Es wird erklärt, wie das Modell trainiert wird und wie es auf Benutzereingaben reagiert. Des Weiteren werden reguläre Ausdrücke und deren Anwendung zur Mustererkennung in Texten erläutert. Im Hinblick auf Webdesign wird über die Gestaltung von Webseiten für eine benutzerfreundliche Erfahrung gesprochen, einschließlich responsivem Webdesign und visueller Hierarchie. Außerdem wird die Wahl von TypeScript für die Programmierung und von React als Frontend-Framework erklärt, sowie die Verwendung von APIs und RESTful Web-Services zur Kommunikation zwischen Frontend und Backend.

Forschungsstand und verwandte Arbeiten

Der Abschnitt beschreibt den aktuellen Stand der Technologien, die für das Verständnis und die Umsetzung des Projekts wichtig sind, wie Textzusammenfassung, LaTeX und Web-Entwicklung. Es wird erwähnt, dass Textzusammenfassung bei einer Kompressionsrate von 15-30\ am besten funktioniert und extraktive sowie abstrakte Zusammenfassungsmethoden verwendet werden können. Es wird diskutiert, dass automatisierte Zusammenfassungen leistungsfähig sind, aber sorgfältig überprüft werden müssen, um Halluzinationen zu vermeiden. Es wird auch darauf hingewiesen, dass automatisierte Texte noch nicht mit menschlichem Schreiben mithalten können. Die Sicherheit von Textgenerierungssystemen wie s ist wichtig, da diese anfällig für Manipulationen sind. LaTeX wird als leistungsstarkes Textsatzsystem für wissenschaftliche Dokumente erwähnt, das präzise Kontrolle über das Layout und die Formatierung bietet. Zudem wird auf die Integration von Copilot in erwähnt, um den Schreibprozess zu unterstützen. Die Bedeutung der Frontend-Entwicklung mit JavaScript-Frameworks wie React wird betont. Schließlich wird das vorliegende Projekt als innovative Lösung zur Generierung technischer Berichte aus LaTeX- oder -Dokumenten dargestellt.

Konzeption und Entwurf von Reforge

Die Anforderungsanalyse für das Softwareprojekt Reforge erfolgt nach dem Modell von Kleuker. Funktionale Anforderungen werden in "muss", "soll" und "wird" klassifiziert, wobei "muss" essentiell für den Projekterfolg ist. Akzeptanzkriterien werden definiert, um Anforderungen genauer zu bestimmen. Es gibt Anforderungen zur Textextraktion, inneren Struktur, Zusammenfassungsgenerator, Textaufteilung, Schnittstelle für OpenAI, Exportfunktionen, Sprachauswahl, Benutzeroberfläche und Fehlerbehandlung. Nicht-funktionale Anforderungen werden nach dem FURPS-Modell kategorisiert. Ein Wireframe und die Wahl des OpenAI-Modells sind wichtige Schritte. Die Systemarchitektur basiert auf RESTful Kommunikation und externen Systemen wie OpenAI und DeepL. Eine Datenbank wird bewusst nicht verwendet, und die Speicherung erfolgt temporary. Ein Sequenzdiagramm visualisiert den Datei-Upload-Prozess.

Implementierung von Reforge

Das Backend von Reforge verarbeitet hochgeladene Dateien und generiert technische Berichte. Die Kommunikation zwischen Frontend und Backend für den Datenaustausch und den Berichtsdownload erfolgt über einen RESTful API-Request. Der Ablauf beginnt, wenn der Nutzer im Frontend einen LaTeX-ZIP-Ordner hochlädt und auf "Start the generation" klickt. Der Code im Frontend ruft die RESTful API auf, um Dateien