

EINSATZ VON LLM TOOLS IM SOFTWARE ENGINEERING PROZESS

BACHELORARBEIT
ZUR ERLANGUNG DES AKADEMISCHEN GRADES
BACHELOR OF ENGINEERING (B.ENG.)

Jonas Gundlach

Betreuerin:
Prof. Dr. rer. nat. Andrea Baumann

Tag der Abgabe: XX.XX.XXXX

eingereicht bei
Universität der Bundeswehr München
Fakultät für Elektrotechnik und Technische Informatik



Universität der Bundeswehr München

Institut für
Software Engineering

Neubiberg, Juli 2024

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst, noch nicht anderweitig für Prüfungszwecke vorgelegt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, insbesondere keine anderen als die angegebenen Informationen.

Der Speicherung meiner Bachelorarbeit zum Zweck der Plagiatsprüfung stimme ich zu. Ich versichere, dass die elektronische Version mit der gedruckten Version inhaltlich übereinstimmt.

Neubiberg, den XX.XX.XXXX

Jonas Gundlach

Zusammenfassung

TODO!!!!!!!!!!!!!!

Ein Abstract ist ein Text, der kurz und knapp die wichtigsten Punkte deiner Bachelorarbeit wiedergibt. Er soll die Fragestellung, Methoden, Quellen und Ergebnisse zusammenfassen, damit die Lesenden schnell die Relevanz der Bachelorarbeit für ihr Interessensgebiet beurteilen können.

Inhaltsverzeichnis

Erklärung	III
1 Einleitung	1
1.1 Motivation	1
1.2 Zielsetzung der Arbeit	1
1.3 Struktur der Arbeit	1
2 Grundlagen	3
2.1 LLM Tools	3
2.1.1 ChatGPT	4
2.1.2 Google Gemini	6
2.1.3 Le Chat	8
2.2 Software Engineering Prozess	8
2.2.1 Besprechungsprotokoll	11
2.2.2 Projekthandbuch	11
2.2.3 Risikoliste	12
2.2.4 Anforderungsspezifikation	13
2.2.5 Architekturdokument	15
2.2.6 Testspezifikation	16
3 Anwendung von LLM Tools im Software Engineering	17
3.1 Besprechungsprotokoll	17
3.2 Projekthandbuch	17
3.3 Risikoliste	19
3.4 Anforderungsspezifikation	19
3.5 Architekturdokument	21
3.6 Testspezifikation	21
4 Praxisergebnisse und Vergleich	23
4.1 Bewertung	23
4.2 Besprechungsprotokoll	24
4.3 Projekthandbuch	25
4.4 Risikoliste	27
4.5 Anforderungsspezifikation	30
4.6 Architekturdokument	42

5 Herausforderungen und Limitationen	47
5.1 Technische Herausforderungen	47
5.2 Limitationen von LLM Tools im Software Engineering Kontext	47
6 Zusammenfassung und Ausblick	49
6.1 Zusammenfassung der Ergebnisse	49
6.2 Implikationen für die Praxis	49
6.3 Ausblick und zukünftige Entwicklungen	49
Abbildungsverzeichnis	VII
Tabellenverzeichnis	IX
Quellcodeverzeichnis	XI
Stichwortverzeichnis	XIII
Literaturverzeichnis	XV

1 Einleitung

TODO!!!!!!!!!!

1.1 Motivation

In einer Welt, die von immer komplexeren und dynamischeren Softwareanwendungen geprägt ist, wird die effektive Nutzung von Technologien zunehmend entscheidend für den Erfolg von Softwareprojekten. In diesem Zusammenhang gewinnen Werkzeuge des Natural Language Processing eine immer größere Bedeutung. Insbesondere die jüngsten Fortschritte im Bereich der Large Language Models (LLM) eröffnen neue Möglichkeiten für die Verbesserung des Software Engineering Prozesses.

Die Motivation hinter dieser Arbeit liegt in der Erkundung des Einsatzes von LLM-Tools im Software Engineering und der Analyse ihrer Auswirkungen auf verschiedene Phasen des Entwicklungszyklus. Die Frage, wie diese fortschrittlichen LLM-Tools genutzt werden können, um den Softwareentwicklungsprozess effizienter, präziser und insgesamt erfolgreicher zu gestalten, steht im Mittelpunkt dieser Untersuchung.

TODO!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

1.2 Zielsetzung der Arbeit

Todo!!!!!!!!!!!!!!!!!!!!!!

1.3 Struktur der Arbeit

Todo!!!!!!!!!!!!!!!!!!!!!!

2 Grundlagen

In diesem Kapitel wird erläutert, was LLM Tools sind, wie sie trainiert werden und wie sie funktionieren. Anschließend werden die einzelnen Tools ChatGPT, Google Gemini und Le Chat vorgestellt und deren Entwicklung beschrieben.

Im zweiten Teil des Kapitels wird der Software Engineering Prozess erklärt und erläutert, was dieser umfasst. Dabei wird als Softwareprozess ausschließlich das Wasserfallmodell betrachtet, da in dieser Arbeit nur mit diesem Modell gearbeitet wird.

2.1 LLM Tools

Unter LLM-Tools versteht man Sprachmodelle, die auf einem Large Language Model (LLM) basieren. Ein LLM ist ein Deep-Learning-Algorithmus, der mit sehr großen Datensätzen trainiert wird. Diese Modelle finden häufig Anwendung im Bereich des Natural Language Processing (NLP), wo sie verwendet werden, um Abfragen in natürlicher Sprache zu beantworten oder Ergebnisse zu liefern. LLMs können neue Inhalte verstehen, zusammenfassen, generieren und vorhersagen. Durch das Training sammeln LLMs Milliarden von Parametern, bei denen es sich um Variablen handelt, die im Modell angepasst werden, um neue Inhalte abzuleiten.

LLMs basieren auf einem Transformer-Modell, das Eingaben in Token umwandelt und dann gleichzeitig mathematische Gleichungen ausführt, um Beziehungen zwischen den Token zu ermitteln. Dadurch kann der Computer Muster erkennen, die auch ein Mensch wahrnehmen würde, wenn ihm die gleiche Frage gestellt wird. Zudem verwenden Transformer-Modelle Selbstaufmerksamkeitsmechanismen, die es dem Modell ermöglichen, schneller zu lernen als herkömmliche Modelle. Somit kann das Transformer-Modell verschiedene Teile der Sequenz oder den gesamten Kontext eines Satzes berücksichtigen, um Vorhersagen zu generieren.

Grundsätzlich bestehen LLMs aus vier neuronalen Netzwerkschichten: der wiederkehrenden Ebene, der Einbettungsebene, der Feedforward-Ebene und der Aufmerksamkeitsebene. Diese Schichten arbeiten zusammen, um den Eingabetext zu verarbeiten und Ausgabeinhalte zu generieren.

Die wiederkehrende Ebene dient dazu, die Wörter des Eingabetextes der Reihe nach zu interpretieren und die Beziehungen zwischen den Wörtern in einem Satz zu erfassen. Die Einbettungsebene erfasst die semantische und syntaktische Bedeutung der Eingabe, sodass das Modell den Kontext verstehen kann. Die Feedforward-Schicht besteht aus mehreren vollständig verbundenen Schichten, die die Eingabeeinbettungen transformieren. Dadurch ermöglichen diese Schichten dem Modell, Abstraktionen auf höherer Ebene zu verstehen und somit die Absicht des Benutzers mit der Texteingabe zu erfassen. Die Aufmerksamkeitsebene ermöglicht es dem Sprachmodell, sich auf einzelne Teile des Eingabetextes

zu konzentrieren, die für die Aufgabe relevant sind, und dadurch die genauesten Ausgaben zu generieren. Damit ein großes Sprachmodell Texteingaben empfangen und eine Ausgabevorhersage generieren kann, muss es zunächst allgemein geschult und anschließend feinabgestimmt werden, um spezifische Aufgaben ausführen zu können. Für die Schulung werden riesige Datenmengen im Petabyte-Bereich benötigt. Das Training verläuft mehrstufig und beginnt in der Regel mit einem unbeaufsichtigten Lernansatz, bei dem das Modell mit unstrukturierten und unbeschrifteten Daten trainiert wird, da diese in größeren Mengen verfügbar sind. In dieser Phase leitet das Modell Beziehungen zwischen verschiedenen Wörtern und Konzepten ab.

Anschließend erfolgt die Schulung und Feinabstimmung durch eine Form des selbstüberwachten Lernens. Dabei wird eine Datenkennzeichnung durchgeführt, durch die das Modell verschiedene Konzepte besser und genauer identifizieren kann. Im nächsten Schritt durchläuft das LLM den transformatorischen neuronalen Netzwerkprozess im Rahmen des Deep Learning. Die Transformer-Modellarchitektur ermöglicht es dem LLM, mittels eines Selbstaufmerksamkeitsmechanismus Beziehungen und Verbindungen zwischen Wörtern und Konzepten zu erkennen, indem bestimmten Elementen (Token) Bewertungen zugewiesen werden, um die Beziehungen festzulegen.

Entscheidend über die Leistungsfähigkeit und die Qualität des Sprachmodells sind die Datensätze, welche zum trainieren benutzt werden. Im folgenden wird auf die drei unterschiedliche LLM-Tools eingegangen, welche in dieser Arbeit näher untersucht werden [1].

2.1.1 ChatGPT

ChatGPT wurde von OpenAI am 30. November 2022 veröffentlicht und entwickelte sich aufgrund seiner fortschrittlichen generativen KI-Fähigkeiten rasch zu einer internationalen Sensation. Wenn man ChatGPT fragt, ob es einen Infotext zu sich selbst erstellen kann, erhält man die Antwort in Prompt 2.1 Die Beta-Version der Plattform war zunächst kostenlos verfügbar, und OpenAI bot den Nutzern weiterhin eine kostenlose Basisversion an. Im Februar 2023 führte OpenAI ein kostenpflichtiges Abonnement ein, das unbegrenzten Zugriff auf ChatGPT sowie eine erweiterte Palette von Funktionen und Diensten ermöglichte.

Trotz der begeisterten Resonanz löste ChatGPT auch einige Kontroversen aus. Es lieferte teilweise fehlerhafte Antworten, insbesondere bei Aufgaben wie Schlussfolgerungen ziehen, Nuancen analysieren, Meinungen bewerten oder Vorhersagen treffen. Pädagogen weltweit äußerten Bedenken hinsichtlich des Einsatzes von ChatGPT im Bildungsbereich, auch an Universitäten wurde der Einsatz kritisch diskutiert. Trotz dieser Kontroversen schien ChatGPT bereit zu sein, die Erstellung von Inhalten im Internet zu revolutionieren. Webentwickler und Content-Ersteller nutzten ChatGPT zunehmend als Recherche- und Schreibhilfe für Website-Texte.

Im März 2023 brachte OpenAI eine neue Version heraus, bekannt als GPT-4. Ein bedeutender Unterschied zu früheren Versionen bestand darin, dass GPT-4 Eingaben sowohl in Text- als auch in Bildformaten akzeptierte. Dies ermöglichte dem Chatbot, Daten aus Diagrammen, Grafiken und Screenshots zu interpretieren. Zudem machte GPT-4 deutlich weniger Denkfehler und sachliche Fehler, auch "Halluzinationen" genannt, als das zuvor verwendete GPT-3.5-Modell. Am Ende des Jahres hatte sich ChatGPT weiterentwickelt und konnte beispielsweise Bilder analysieren, Sprachansagen verarbeiten

Prompt: Erstell mir einen Infotext über dich für meine Bachelorarbeit

ChatGPT ist ein fortschrittlicher KI-gestützter Sprachassistent, der auf der GPT-4-Architektur von OpenAI basiert. Entwickelt im Rahmen der Forschung und Entwicklung im Bereich Künstliche Intelligenz, ist ChatGPT in der Lage, natürliche Sprache zu verstehen und darauf zu antworten. Dies macht ihn zu einem vielseitigen Werkzeug für eine Vielzahl von Anwendungen, von der Beantwortung alltäglicher Fragen bis hin zur Unterstützung bei komplexeren Aufgaben wie dem Schreiben von Texten, der Generierung von Ideen und der Bereitstellung von Informationen.

Die GPT-4-Architektur, auf der ChatGPT basiert, verwendet tiefes maschinelles Lernen, um menschliche Sprache zu verarbeiten und zu erzeugen. Durch die Analyse großer Mengen an Textdaten kann das Modell Muster und Zusammenhänge erkennen, die es ihm ermöglichen, kontextuell relevante und kohärente Antworten zu geben. Diese Fähigkeiten machen ChatGPT zu einem nützlichen Instrument in Bereichen wie Bildung, Forschung, Kundenservice und mehr.

Ein weiterer bemerkenswerter Aspekt von ChatGPT ist seine Fähigkeit, sich an eine Vielzahl von Themen anzupassen und Informationen auf klare und präzise Weise zu vermitteln. Dies wird durch kontinuierliches Training und Verbesserung des Modells erreicht, wobei sowohl technische Fortschritte als auch Nutzerfeedback berücksichtigt werden.

Zusammengefasst ist ChatGPT ein leistungsstarker und vielseitiger KI-Assistent, der durch seine fortschrittliche Technologie und seine breite Anwendbarkeit einen wertvollen Beitrag in vielen Bereichen leistet.

Prompt 2.1: Infotext ChatGPT, Quelle: ChatGPT

und gesprochene Antworten geben.

Im Mai 2024 kündigte OpenAI ein aktualisiertes Modell namens ChatGPT-4o an, das schnellere Ergebnisse sowie verbesserte Text- und Audiofunktionen versprach. Dieses Modell umfasst auch ein neues Konversations-KI-System namens "Sky" [2].

2.1.2 Google Gemini

Im Dezember 2023 stellte Google eine generative Künstliche Intelligenz mit dem Namen Google Gemini vor. Sie wurde von Google DeepMind entwickelt und ist laut Aussage von Google das größte, leistungsstärkste und vielseitigste große Sprachmodell (LLM), das bisher von Google entwickelt und realisiert wurde. In der ursprünglichen Ankündigung wurde das KI-Sprachmodell unter dem Namen Gemini vorgestellt. Im Februar 2024 entschied Google, auch den bisher als Bard bekannten Chatbot in Gemini umzubenennen. Damit ist Gemini mittlerweile sowohl der Name des Google Chatbots als auch des zugrundeliegenden Sprachmodells.

Bei Nachfrage an Gemini, ob es einen Infotext über sich erstellen kann, erhält man die Antwort in Prompt 2.2 und Prompt 2.3.

Gemini 1.0 wurde in drei verschiedenen Größen optimiert: Gemini Ultra als das größte und leistungsstärkste Modell für hochkomplexe Aufgaben, Gemini Pro als das beste Modell zur Anwendung in einem breiten Aufgabenspektrum und Gemini Nano als das effizienteste Modell für Aufgaben, die direkt auf dem Gerät verarbeitet werden.

Die Gemini-Modelle sind von Grund auf multimodal konzipiert, was bedeutet, dass sie mit verschiedenen Arten von Informationen umgehen können. Gemini ist in der Lage, Texte, Programmcode, Bilder, Videos, Audioinformationen und deren Kombinationen zu analysieren, zu verstehen, zu kombinieren und zu generieren. Da Gemini auch mit Skizzen und handgeschriebenem Text umgehen kann, eignet es sich für den Einsatz in Fachbereichen wie zum Beispiel Mathematik oder Physik. Im Programmierbereich werden gängige Programmiersprachen wie Python, Java, C++ oder Go unterstützt. Gemini kann Programmcode analysieren, kommentieren und auch generieren. Dadurch steht Gemini in Konkurrenz zu anderen LLMs wie den GPT-Sprachmodellen von Open AI. Gemini soll jedoch die Konkurrenzprodukte in vielen Benchmarks deutlich übertreffen.

Gemini baut auf den ebenfalls von Google entwickelten Sprachmodellen LaMDA und PaLM 2 auf. Wie üblich für LLMs (Large Language Models) basiert auch Gemini auf einer Transformer-Architektur. Google hat die Architektur des Modells zur Skalierung der Trainingsfähigkeiten optimiert und mit einem Multi-Query-Aufmerksamkeitsmechanismus ausgestattet. Die unterstützte Kontextlänge beträgt 32.000 Tokens und in der Version 1.5 bis zu einer Million Tokens. Gemini wurde mit Daten verschiedener Modalitäten trainiert, und als Trainingsplattform verwendete Google Rechner-Cluster und Tensor Processing Units (TPUs) der neuesten Generation v4 und v5 als KI-Beschleuniger.

Google hat begonnen, die KI in zahlreiche eigene Produkte zu integrieren, um die Fähigkeiten der Websuche, des Chrome-Browsers, des Gmail-Services, verschiedener Entwicklungstools und anderer Produkte und Services zu verbessern und zu erweitern. Mitte Februar 2024 stellte Google Gemini 1.5 und 1.5 Pro vor, welche bis zu eine Million Tokens verarbeiten können. Diese neuen Versionen sollen die Benchmark-Ergebnisse der Vorgängerversion deutlich übertreffen [3].

Prompt: Erstell mir einen Infotext über dich für meine Bachelorarbeit**Was ist Gemini?**

Gemini ist ein Large Language Model (LLM) von Google AI, das mit einem riesigen Datensatz aus Text und Code trainiert wurde. Es kann Text generieren, Sprachen übersetzen, verschiedene Arten von kreativen Inhalten schreiben und Ihre Fragen auf informative Weise beantworten. Gemini befindet sich noch in der Entwicklung, hat aber bereits gelernt, viele Arten von Aufgaben auszuführen, einschließlich:

- Befolgen Sie Anweisungen und erledigen Sie Aufgaben sorgfältig.
- Beantworten Sie Ihre Fragen umfassend und informativ, auch wenn sie offen, herausfordernd oder seltsam sind.
- Generieren Sie verschiedene kreative Textformate, wie Gedichte, Codes, Skripte, Musikstücke, E-Mails, Briefe usw. Ich werde mein Bestes tun, um alle Ihre Anforderungen zu erfüllen.

Technische Details

Gemini basiert auf der Transformer-Architektur und wird mit einem massiven Datensatz aus Text und Code trainiert. Es verwendet eine Vielzahl von Techniken, um seine Leistung zu verbessern, einschließlich:

- **Aufmerksamkeit:** Ermöglicht es Gemini, sich auf die relevantesten Teile einer Eingabe zu konzentrieren.
- **Beam Search:** Ermöglicht Gemini, die besten möglichen Ausgaben zu generieren.
- **Multitasking:** Ermöglicht Gemini, mehrere Aufgaben gleichzeitig zu lernen und auszuführen.

Anwendungen

Gemini kann für eine Vielzahl von Anwendungen eingesetzt werden, einschließlich:

- **Generierung von kreativen Inhalten:** Gemini kann verwendet werden, um Gedichte, Codes, Skripte, Musikstücke, E-Mails, Briefe usw. zu generieren.
- **Sprachübersetzung:** Gemini kann verwendet werden, um Text von einer Sprache in eine andere zu übersetzen.
- **Beantwortung von Fragen:** Gemini kann verwendet werden, um Ihre Fragen umfassend und informativ zu beantworten, auch wenn sie offen, herausfordernd oder seltsam sind.
- **Zusammenfassen von Texten:** Gemini kann verwendet werden, um lange Textpassagen zu zusammenfassen.
- **Schreiben verschiedener Arten von kreativen Textformaten:** Gemini kann verwendet werden, um verschiedene Arten von kreativen Textformaten zu schreiben, wie Gedichte, Codes, Skripte, Musikstücke, E-Mails, Briefe usw.

Prompt 2.2: Infotext Gemini 1, Quelle: Gemini

Ethische Überlegungen

Es ist wichtig, Gemini verantwortungsvoll und ethisch zu nutzen. Zu den ethischen Überlegungen, die bei der Verwendung von Gemini zu beachten sind, gehören:

- **Vermeidung von Vorurteilen:** Gemini wird mit einem riesigen Datensatz aus Text und Code trainiert. Dieser Datensatz kann Vorurteile enthalten, die sich in den Ausgaben von Gemini widerspiegeln können. Es ist wichtig, sich dieser Vorurteile bewusst zu sein und Maßnahmen zu ergreifen, um sie zu mindern.
- **Vermeidung von Schadenszufügung:** Gemini kann verwendet werden, um schädlichen Text zu generieren, z. B. Hassreden oder Gewaltandrohungen. Es ist wichtig, Gemini nicht zu verwenden, um anderen Schaden zuzufügen.
- **Respekt vor der Privatsphäre:** Gemini kann verwendet werden, um personenbezogene Daten zu generieren. Es ist wichtig, Gemini nicht zu verwenden, um die Privatsphäre anderer zu verletzen.

Fazit

Gemini ist ein leistungsstarkes Tool, das für eine Vielzahl von Anwendungen eingesetzt werden kann. Es ist jedoch wichtig, Gemini verantwortungsvoll und ethisch zu nutzen.

Prompt 2.3: Infotext Gemini 2, Quelle: Gemini

2.1.3 Le Chat

In Prompt 2.4 sieht man die Antwort die Le Chat gibt, wenn man ihn nach einem Infotext über sich fragt. Le Chat ist ein, von dem französischen KI-Startup Mistral AI veröffentlichter Chatbot. Mistral AI startete mit seinen frei verwendbaren Sprachmodellen, also auf der Grundlage von Open Source, erfolgreich durch. Nun hat Mistral AI sein bislang größtes Modell 'Mistral Large' veröffentlicht. Diesmal allerdings nicht auf der Basis von Open Source, sondern ausschließlich über die eigene Webseite und der KI-Infrastruktur Microsoft Azure. Es lassen sich allerdings API-Keys für Programmierschnittstellen erstellen, um z.B. Mistral Large über seinen eigenen Server laufen zu lassen und so für andere User auf der eigenen Homepage verfügbar zu machen. Mit Mistral Large wurde auch Le Chat veröffentlicht, welcher aktuell kostenfrei verwendet werden kann. Le Chat bietet derzeit noch sehr wenige Funktionen an. Es stehen lediglich Texteingabe und -ausgabe zur Verfügung. Die Datengrundlage reicht nur bis 2021, weshalb es auch hier, für die Jahre 2022 bis heute, zu der Problematik der Halluzination kommen kann. Grundsätzlich kann man zwischen drei Sprachmodellen auswählen: Large, Next und Small. Large bietet überlegene Denkfähigkeit, Next ist ein Prototyp-Modell für erhöhte Kürze und Small arbeitet schnell und kosteneffektiv [4].

2.2 Software Engineering Prozess

Software-Engineering ist eine technische Disziplin, die sich umfassend mit der Erstellung von Software beschäftigt, beginnend bei der Konzeption, über den Betrieb, bis hin zur Wartung. Die wesentlichen Aktivitäten im Software-Engineering umfassen die Softwarespezifikation, -entwicklung, -validierung und

Prompt: Erstell mir einen Infotext über Le Chat für meine Bachelorarbeit Einleitung

Le Chat ist ein fortschrittliches Sprachmodell, das auf künstlicher Intelligenz (KI) und maschinellem Lernen basiert. Entwickelt von einem Team aus Experten für KI und Sprachverarbeitung, hat Le Chat die Fähigkeit, menschliche Sprache zu verstehen, zu interpretieren und darauf zu reagieren, was es zu einem wertvollen Werkzeug in verschiedenen Branchen macht.

Funktionsweise

Le Chat nutzt eine Kombination aus Techniken des Natural Language Processing (NLP) und des Deep Learning. NLP ermöglicht es dem Modell, die Struktur und den Kontext der menschlichen Sprache zu verstehen, während Deep Learning-Algorithmen verwendet werden, um Muster in den Daten zu erkennen und die Genauigkeit der Vorhersagen zu verbessern.

Anwendungsbereiche

Le Chat kann in einer Vielzahl von Anwendungen eingesetzt werden. Einige der wichtigsten Anwendungsbereiche sind:

- **Kundenservice:** Le Chat kann als virtueller Assistent eingesetzt werden, um Kundenanfragen zu beantworten, Produktempfehlungen auszusprechen und den Kunden bei der Lösung von Problemen zu unterstützen.
- **Bildung:** Le Chat kann als Tutor-System eingesetzt werden, um Schülern und Studenten bei der Lösung von Aufgaben und beim Verständnis von komplexen Konzepten zu helfen.
- **Gesundheitswesen:** Le Chat kann als virtueller Gesundheitsassistent eingesetzt werden, um Patienten bei der Verwaltung ihrer Medikamente zu unterstützen, Gesundheitstipps zu geben und in Notfällen Erste Hilfe-Anweisungen bereitzustellen.

Vorteile

Le Chat bietet eine Reihe von Vorteilen gegenüber herkömmlichen Sprachmodellen. Einige der wichtigsten Vorteile sind:

- **Verbesserte Genauigkeit:** Le Chat nutzt fortschrittliche Deep Learning-Algorithmen, um die Genauigkeit der Vorhersagen zu verbessern und eine bessere Benutzererfahrung zu bieten.
- **Skalierbarkeit:** Le Chat kann leicht skaliert werden, um eine große Anzahl von Anfragen gleichzeitig zu bearbeiten, was es zu einer idealen Lösung für Unternehmen mit hohem Kundenaufkommen macht.
- **Kosteneffizienz:** Le Chat kann als kosteneffiziente Alternative zu menschlichen Kundenservice-Mitarbeitern eingesetzt werden, was zu erheblichen Einsparungen für Unternehmen führen kann.

Fazit

Le Chat ist ein innovatives Sprachmodell, das die Fähigkeit hat, die Art und Weise, wie wir mit Maschinen interagieren, zu revolutionieren. Mit seiner fortschrittlichen Technologie und seinen vielfältigen Anwendungsmöglichkeiten ist Le Chat ein wichtiger Meilenstein in der Entwicklung der KI und ein Werkzeug, das in Zukunft eine noch größere Rolle spielen wird.

Dokument	Inhalte
01 Besprechungsprotokoll	Protokolle
02 Projekthandbuch	Projekthandbuch
03 Risikoliste	Risikoliste
04 Anforderungsspezifikation	Einleitung
	Benutzeranforderungen
	Systemübersicht
	Fachliches Datenmodell
	Anwendungsfälle
	AF
	Benutzungsschnittstelle
	Dialog
	Nichtfunktionale Anforderungen
05 Architekturdokument	Technische Infrastruktur
	Komponentenmodell
	Datenmodell
	Detaillierte Schnittstellenbeschreibung
	AF
06 Testspezifikationen	Testkonzept
	AF
	Testprotokoll
	Testdaten

Abbildung 2.1: Dokumente im Software Engineering

-weiterentwicklung [5]. Während des Prozesses werden mehrere Dokumente erstellt und im Verlauf angepasst. Diese umfassen Besprechungsprotokolle, eine Risikoliste, ein Projekthandbuch, eine Anforderungsspezifikation, ein Architekturdokument und eine Testspezifikation. In der Abbildung 2.1 sieht man welche Inhalte die einzelnen Dokumente umfassen. Außerdem wird im folgenden zu jedem Dokument und dessen Inhalte ein Überblick gegeben.

2.2.1 Besprechungsprotokoll

Alle Team- und Kundengespräche werden systematisch in Besprechungsprotokollen dokumentiert. Diese Protokolle enthalten Aufträge, Beschlüsse, Feststellungen und Informationen, die in Form von Texten, Bildern, Diagrammen usw. festgehalten werden. Aufträge müssen mit einem Fälligkeitsdatum und einem Verantwortlichen versehen sein, während Beschlüsse klar und unmissverständlich formuliert werden sollten. Im Rahmen der Anforderungsanalyse werden in Beschlüssen die Benutzeranforderungen erfasst. Feststellungen hingegen sind Beschlüsse, die keine Abstimmung benötigen, können aber ebenfalls Anforderungen an das System enthalten. Informationen bieten den Projektmitgliedern wichtige Hinweise. Dabei ist zu beachten, dass ein Besprechungsprotokoll keine Erlebniserzählung sein soll.

Der Aufbau eines Besprechungsprotokolls gestaltet sich wie folgt: Am Anfang stehen Datum, Thema, Verfasser, Teilnehmer und Verteiler. Anschließend wird eine Tabelle mit den Spalten Nummer, Art, Beschreibung, Termin und Verantwortlicher erstellt. Die Nummer stellt eine eindeutige ID im Protokoll dar. Die Art spezifiziert, ob es sich um einen Auftrag (=A), einen Beschluss (=B), eine Feststellung (=F) oder eine Information (=I) handelt. Die Beschreibung fasst den Auftrag, den Beschluss, die Feststellung oder die Information kurz und präzise zusammen. Der Termin gibt an, bis wann der Auftrag erledigt sein muss. Der Verantwortliche benennt die Person, die den Auftrag ausführen muss, den Beschluss gefasst hat oder die Feststellung bzw. die Information geliefert hat.

2.2.2 Projekthandbuch

Das Projekthandbuch dient dazu, die Projektmitglieder umfassend über das Projekt zu informieren, einschließlich der Vorgeschichte, Vertragsbasis, Grundlagen, Ziele, Ergebnisse und mehr. Es bietet zudem eine Übersicht über die Projektorganisation, wie z.B. die Teamaufstellung, das Berichtswesen und die Dokumentenablage, sowie über wesentliche Projektabläufe wie Änderungsverfahren, Risikomanagement und Aufwandsverfolgung. Ziel des Projekthandbuchs ist es, neuen Teammitgliedern den Einstieg zu erleichtern und auf wichtige Dokumente hinzuweisen, die für die Projektabwicklung notwendig sind. Damit fungiert das Projekthandbuch als zentrale Anlaufstelle für das Projekt.

Das Dokument beginnt mit einer Einleitung, die die Rahmenbedingungen klärt. Dieser Abschnitt enthält den Zweck des Dokuments, Hinweise zur Redaktion, in denen festgelegt wird, wer für das Dokument verantwortlich ist, sowie Informationen zum Verteiler, also wer bei Änderungen informiert werden muss.

Im Anschluss folgt die Projektbeschreibung, die die Rahmenbedingungen des Projekts darlegt. Dieser Teil enthält einen Abschnitt zur Vorgeschichte, in dem beschrieben wird, wie es zu dem Projekt kam.

Ein weiterer Abschnitt bietet eine inhaltliche Kurzdarstellung des Projekts. Danach werden in einem Abschnitt zur Vertragsbasis der zeitliche Projektrahmen, die Anzahl der Teammitglieder und der Umfang der Bearbeiter-Stunden festgelegt. Es folgt eine Beschreibung des Projektergebnisses, in der festgelegt wird, welche Produkte geliefert werden müssen. Ein weiterer Abschnitt zu sonstigen Besonderheiten beschreibt informelle Ziele.

Ein nachfolgendes Kapitel verweist auf die Risikoliste.

Daraufhin folgt ein Kapitel zur Projektorganisation, das die Teamstruktur und die Kommunikationswege beschreibt. Dieses Kapitel beginnt mit dem Teamaufbau, in dem die Aufgabenverteilung und die E-Mail-Adressen der einzelnen Teammitglieder vermerkt sind. Anschließend wird die Zusammenarbeit mit den Kunden definiert. Danach folgt ein Abschnitt über regelmäßige Pflichtbesprechungen für die Teammitglieder.

Das folgende Kapitel widmet sich der Planung und behandelt verschiedene Planungsaspekte. Der erste Abschnitt befasst sich mit dem Projektplan. Hier werden die festgelegten Meilensteine beschrieben, einschließlich der Startzeiten der einzelnen Aktivitäten, der parallelen Abläufe, der Abschlusszeiten der verschiedenen Tätigkeiten, der Entstehungszeitpunkte der Produkte und der möglichen Entscheidungspunkte.

Anschließend folgt ein Kapitel zur Dokumentation und Ablage, in dem beschrieben wird, wie und wo Produkte bzw. Dokumente abgelegt werden. Es beginnt mit einem Abschnitt zum Dokumenteneingang und -ausgang, in dem festgelegt wird, wie mit extern eingehenden Informationen umgegangen wird und wie ausgehende Informationen protokolliert werden. Im zweiten Abschnitt wird die Ablage und Archivierung geregelt, wobei erläutert wird, wie die Daten im Projektarchiv zu organisieren sind und in welchen Verzeichnissen die jeweiligen Daten abgelegt werden.

2.2.3 Risikoliste

Die Risikoliste ist ein zentrales Instrument, das dem gesamten Team einen umfassenden Überblick über bekannte Risiken und die entsprechenden Maßnahmen bietet. Der Projektleiter muss diese Risiken stets im Auge behalten, um eine fundierte Projektplanung sicherzustellen. Zusätzlich kann die Risikoliste dem Anforderungsanalysten bei der Identifikation von Anforderungen, dem Systemarchitekten bei der Wahl geeigneter technischer Lösungen und dem Testmanager bei der Entwicklung von Testfällen nützlich sein.

Der Inhalt der Risikoliste umfasst die Autoren des Dokuments, die Historie des Dokuments mit einer Zusammenfassung der Änderungen, die Liste der Risiken sowie eine Liste der möglichen Maßnahmen für jedes identifizierte Risiko. Die Historie enthält Spalten für Datum, Autor*in und Änderungen, um später vorgenommene Modifikationen leichter nachvollziehbar zu machen.

Die Risikoliste selbst setzt sich aus mehreren Attributen zusammen. Die Identifikationsnummer ist eine eindeutige Nummer zur Identifizierung des Risikos, während die Risikobezeichnung eine kurze Beschreibung des Risikos bietet. Eine ausführliche Darstellung des Risikos findet sich in der Spalte für die Risikobeschreibung. Das Datum beschreibt, wann das Risiko identifiziert wurde, und der Autor gibt an, wer das Risiko gemeldet hat. In der Spalte für Auswirkungen werden die möglichen Folgen des Risikos beschrieben. Die Risikowahrscheinlichkeit ist eine Schätzung der Eintrittswahrscheinlichkeit des Risikos, und der Risikoschaden oder das Risikoausmaß beschreibt das potenzielle Schadensausmaß, einschließlich eventueller Vertragsstrafen. Das Risikomaß berechnet sich aus der Multiplikation von Risikowahrscheinlichkeit und Risikoschaden. Die Risikoklasse ist eine Priorisierung der Risiken, typischerweise in Kategorien wie tolerierbar, unerwünscht, kritisch oder katastrophal. Schließlich wird für jedes Risiko ein Status festgelegt, der zwischen aktiv, eingetragen und geschlossen unterscheiden kann.

Die Maßnahmen zur Risikobewältigung umfassen ebenfalls mehrere Attribute. Zunächst wird der Typ der Maßnahme festgelegt, wobei zwischen Risiko verhindern, Risiko lindern oder minimieren, Risiko übertragen oder teilen und Risiko akzeptieren unterschieden werden kann. Eine detaillierte Beschreibung der Maßnahme folgt, und falls das Risiko den Status "geschlossen" erhält, wird hier die Begründung eingetragen. Wenn eine Maßnahme nicht sofort eingeleitet werden soll, wird das Ereignis, das die Maßnahme auslöst, in der Trigger-Spalte definiert. Die verantwortliche Person, die für die Umsetzung der Maßnahme zuständig ist, wird ebenfalls eingetragen. Der geplante Termin gibt an, bis wann die Maßnahme abgeschlossen sein soll, und der Ist-Termin beschreibt den voraussichtlichen Abschluss basierend auf aktueller Einschätzung. Der geplante Aufwand gibt die geschätzten Kosten der Maßnahme an, während der Ist-Aufwand den aktuellen Aufwand widerspiegelt. Auch für die Maßnahme gibt es einen Status, bei dem man zwischen geplant, aktiv oder beendet wählen kann. Das letzte Attribut umfasst die Restrisikowahrscheinlichkeit, den Restrisikoschaden, das Restrisikomaß und die Restrisikoklasse. Dies sind die geschätzte Wahrscheinlichkeit, der geschätzte Schaden, das geschätzte Maß und die geschätzte Klasse des Restrisikos nach Durchführung der Maßnahmen.

2.2.4 Anforderungsspezifikation

Die Anforderungsspezifikation definiert formell, welche Leistungen das zu entwickelnde System erbringen muss. Sie beschreibt detailliert, was das System leisten soll, ohne jedoch festzulegen, wie diese Leistungen erbracht werden. Das "Wie" wird im Architekturdokument behandelt. In der Praxis sind Anforderungen und Systementwurf eng miteinander verknüpft. Die Anforderungsspezifikation ist oft Teil eines Vertrags und muss daher so vollständig und eindeutig wie möglich formuliert sein.

Diese Spezifikation dient mehreren Zwecken: Sie hilft Kunden und Endbenutzern zu beurteilen, ob das System die gewünschten Leistungen erbringen wird und wie es später zu bedienen ist. Projektmanager nutzen die Spezifikation zur Erstellung des Angebots und zur Planung des detaillierten Entwicklungsprozesses. Administratoren können damit entscheiden, wie das System verwaltet werden muss. Für Systemarchitekten dient sie als Grundlage für ihr Architekturdokument, und Softwareentwickler nutzen sie, um zu verstehen, welches System entwickelt werden soll. Zusammen mit dem Architekturdokument

bildet sie die Basis für die Implementierung. Testmanager verwenden die Spezifikation, um Validierungstests und Testfallspezifikationen zu entwickeln.

Der Inhalt einer Anforderungsspezifikation ist folgendermaßen aufgebaut:

Zunächst werden der Name des Softwareprodukts und des Herstellers genannt. Es folgt ein Vorwort, das die erwartete Leserschaft definiert, eine Versionshistorie des Dokuments mit Begründungen für neue Versionen und eine Zusammenfassung der Änderungen enthält sowie die Autoren des Dokuments nennt.

Die Einleitung beschreibt den Zweck und Umfang des Dokuments. Dies umfasst eine Beschreibung der Notwendigkeit des Systems, eine Kurzbeschreibung der Funktionalität und der Nachbarsysteme sowie eine Definition der Bedeutung des Systems für die gesamtwirtschaftlichen oder strategischen Ziele des Unternehmens. Es folgen Verweise auf weitere Ressourcen sowie Definitionen von Begriffen und Abkürzungen. Eine Übersicht des Dokuments schließt die Einleitung ab.

Anschließend werden die Benutzeranforderungen definiert. Diese enthalten eine Beschreibung der für den Benutzer bereitgestellten Dienste und der nichtfunktionalen Anforderungen in verständlicher Sprache, Diagrammen oder anderen für den Kunden nachvollziehbaren Notationen.

Es folgt ein Überblick über die Systemarchitektur, der die erwartete Systemarchitektur und die Einordnung des Systems in die Systemlandschaft des Kunden beschreibt.

Daraufhin wird die Spezifikation der Systemanforderungen behandelt, die sowohl funktionale als auch nichtfunktionale Anforderungen beschreibt. Hier können auch Schnittstellendefinitionen zu externen Systemen enthalten sein.

Abschließend werden die Systemmodelle vorgestellt, zu denen Klassen-, Datenfluss- und semantische Modelle gehören, die die Beziehungen zwischen Systemkomponenten sowie zwischen dem System und seiner Umgebung darstellen.

Ein spezielles Modell ist das fachliche Datenmodell, das aus Entitätstypen besteht. Eine Entität ist ein individuelles und identifizierbares Exemplar von Dingen, Personen oder Begriffen der realen oder Vorstellungswelt, beschrieben durch ihre Eigenschaften. Entitätstypen besitzen Attribute und Beziehungen, jedoch keine Fähigkeiten. Das fachliche Datenmodell wird mit Hilfe eines UML Klassendiagrammes mit ergänzenden Beschreibungen bzw. Einschränkungen spezifiziert. Dieses enthält alle Entitätstypen mit deren Eigenschaften, Beziehungen und Einschränkungen.

Zu den Systemmodellen gehören auch Anwendungsfälle, die die funktionalen Anforderungen an das System beschreiben. Ein Anwendungsfalldiagramm gibt eine Übersicht über alle Anwendungsfälle. Ein Anwendungsfall besteht aus mehreren Szenarien, die konkrete Abläufe beschreiben, wie das System genutzt werden kann. Anwendungsfälle können textuell beschrieben oder durch Aktivitäts- oder Sequenzdiagramme ergänzt werden.

Ein Anwendungsfall hat mehrere Bestandteile: Der Titel dient als eindeutige Identifikation, die Kurzbeschreibung definiert das Thema, und das Ziel legt fest, was erreicht werden soll. Akteure beschreiben, wer den Anwendungsfall auslöst, und der Auslöser definiert den Grund für die Ausführung. Die Vorbedingung legt fest, welche Bedingungen zu Beginn vorherrschen müssen, das Szenario beschreibt

den Ablauf, und das Ergebnis erläutert die gelieferten Informationen. Die Häufigkeit gibt eine grobe Schätzung, wie oft der Anwendungsfall ausgeführt wird, und nicht-funktionale Anforderungen werden ebenfalls beschrieben. Schließlich können Bemerkungen hinzugefügt werden.

In der Anforderungsspezifikation gibt es dazu eine Übersicht aller Anwendungsfälle im UML Use-Case-Diagramm, welches alle Akteure, Anwendungsfälle, Beziehungen zwischen Akteur und Anwendungsfällen und eventuell include- und extend-Beziehungen zwischen den Anwendungsfällen darstellt. Ebenfalls gibt es eine Spezifikation jedes einzelnen Anwendungsfalls mit Hilfe von UML Aktivitätsdiagrammen und ergänzenden Beschreibungen. Diese enthalten eine genaue Ablaufbeschreibung, den Akteur, die Aufrufhäufigkeit, die Vorbedingung und den Auslöser sowie eine Dokumentation des Datenflusses.

Neben den Anwendungsfällen müssen auch die Dialoge und Schnittstellen beschrieben werden, über die die Anwendungsfälle ausgeführt werden. Eine Dialogbeschreibung enthält alle Details, einschließlich der Feldlängen und der zugeordneten Aktionen für jeden Button. In den Schnittstellen müssen die Parameter genau festgelegt werden.

Im Dokument werden alle Dialoge mit Bildern und strukturierter und detaillierter Beschreibung der Bestandteile spezifiziert. Auch alle Schnittstellen zu den Nachbarsystemen werden Spezifiziert.

Zum Schluss werden die nichtfunktionalen Anforderungen in natürlicher Sprache inklusiver Abnahmebeschreibung aufgelistet.

2.2.5 Architekturdokument

Das Architekturdokument, auch bekannt als Softwarearchitektur-Dokument, beschreibt die Umsetzung des zu entwickelnden Systems. Die Anforderungsspezifikation [2.2.4] legt fest, „was“ das System leisten soll, jedoch nicht „wie“ dies erreicht wird. Diese Umsetzung wird im Architekturdokument detailliert beschrieben. In der Praxis sind die Anforderungsspezifikation und das Architekturdokument eng miteinander verzahnt.

Die Zielgruppe des Architekturdokuments umfasst hauptsächlich Softwareentwickler und Testmanager. Softwareentwickler nutzen das Dokument, um zu verstehen, wie das in der Anforderungsspezifikation beschriebene System entwickelt werden soll. Zusammen mit der Anforderungsspezifikation bildet es die Grundlage für die Implementierung. Testmanager verwenden das Architekturdokument, um Komponententests für das System zu entwickeln.

Der Inhalt des Architekturdokuments gliedert sich in drei Hauptabschnitte: Technische Architektur, Anwendungsarchitektur und IT-Konzept.

Die Technische Architektur umfasst zwei Hauptkomponenten. Die Technische Infrastruktur (TI-Architektur) beschreibt die Hardware- und Produkthanforderungen. Beide Komponenten werden durch erklärende Texte und Grafiken dargestellt.

Die Anwendungsarchitektur besteht ebenfalls aus zwei Hauptteilen: dem Komponentenmodell und dem Datenmodell. Das Komponentenmodell beschreibt die fachliche Struktur der Lösung auf einer groben Ebene mittels eines UML-Komponenten-Diagramms. Das Datenmodell hingegen beschreibt

die Struktur der persistenten Daten und wird entweder durch ein UML-Komponenten-Diagramm oder Entity-Relationship-Diagramme dargestellt.

Das IT-Konzept enthält die detaillierte Schnittstellenbeschreibung. Diese beschreibt die Semantik und Syntax der Komponentenschnittstellen ausführlich und wird durch Code oder Code-Dokumentation dokumentiert.

Darüber hinaus beinhaltet das IT-Konzept dynamische Beschreibungen exemplarischer Anwendungsfälle im Komponentenmodell, die mittels Aktivitätsdiagrammen oder Sequenzdiagrammen veranschaulicht werden.

2.2.6 Testspezifikation

Die Testfallspezifikation oder auch Software Test Dokumentation dient dazu die Testdurchführung zu planen, die Testfälle zu Spezifizieren und die Testergebnisse zu protokollieren. Mit Hilfe dieser ist es möglich die Anwendung systematisch und strukturiert zu testen. Die Protokollierung der Testergebnisse ermöglicht es eine Aussage über die Qualität zu geben. Nach einer Fehlerbehebung können die selben Tests erneut durchgeführt werden. Dadurch wird nichts vergessen.

Die Testspezifikation besteht aus einem Testkonzept, einer Testfall-Spezifikation und einem Testprotokoll. Das Testkonzept legt die Vorgehensweise, die verwendeten Mittel und den Ablauf der Testaktivitäten fest. In diesem Rahmen muss festgelegt werden, ob es irgendwelche Anforderungen an die Umsetzung gibt oder ob irgendwelche Dummy-Komponenten entwickelt werden müssen.

Die Testfall-Spezifikation listet die Testfälle auf. Ein Testfall besteht aus einer Kurzbeschreibung, der Testvoraussetzung, den Eingabewerten inkl. dem Testablauf, und den erwarteten Ausgabewerte.

Im Testprotokoll wird die Testdurchführung protokolliert. Dies beinhaltet wann welcher Test durchgeführt wurde und ob der Test erfolgreich war oder nicht.

3 Anwendung von LLM Tools im Software Engineering

Dieses Kapitel untersucht ausführlich die potenziellen Anwendungsfälle von LLM-Tools im Software Engineering. Dabei wird jedes Dokument einzeln betrachtet und aufgezeigt, wie LLM-Tools bei der Erstellung und Verwaltung dieser Dokumente eingesetzt werden können. Die Abbildung 3.1 bietet eine Übersicht darüber, in welchen Dokumenten und bei welchen Inhalten LLM-Tools unterstützend eingesetzt werden können.

3.1 Besprechungsprotokoll

Im Unterabschnitt 2.2.1 wurde bereits erläutert, welche Inhalte in einem Besprechungsprotokoll enthalten sind. Nun stellt sich die Frage, wie LLM-Tools bei der Erstellung eines solchen Dokuments unterstützen können.

Eine Möglichkeit wäre, das Besprechungsprotokoll aufzeichnen zu lassen und anschließend automatisch daraus das Protokoll erstellen zu lassen. Da aktuell jedoch nur die kostenfreie Version ChatGPT-4o mit Dokumenten arbeiten kann, ist dies nur in begrenztem Umfang und ohne Vergleiche testbar.

Eine weitere Möglichkeit ist es, das Meeting zu verschriftlichen und damit als Texteingabe zu arbeiten. Mit dem schriftlichen Gespräch könnten die Tools dann automatisch das Besprechungsprotokoll erstellen. Sollte dies nicht möglich sein, könnte man versuchen, sich die Aufgaben aus dem Gespräch zusammenfassen zu lassen. Dadurch wäre es einfacher, das Besprechungsprotokoll selbst zu erstellen. Falls auch dies nicht funktioniert, kann das Gespräch zusammengefasst und um unwichtige Themen, wie etwa Abschweifungen, gekürzt werden.

3.2 Projekthandbuch

Das Projekthandbuch ist ein zentrales Dokument im Software Engineering Prozess, das alle wesentlichen Informationen und Richtlinien eines Projekts zusammenfasst. Die spezifischen Inhalte dieses Dokuments wurden bereits im Unterabschnitt 2.2.2 dargelegt. Doch wie können LLM-Tools bei der Erstellung des Projekthandbuches unterstützend wirken?

Eine Möglichkeit besteht darin, das gesamte Projekthandbuch oder einzelne Kapitel von diesen Tools erstellen zu lassen. Dafür benötigen die Tools jedoch die relevanten Informationen. Um die Kapitel präzi-

Dokument	Inhalte	Mögliche KI Unterstützung
01 Besprechungsprotokoll	Protokolle	(Audio zu Text)
		Schriftliches Gespräch in Protokoll umformen
		automatische Identifizierung der Aufgaben Gespräch zusammenfassen
02 Projekthandbuch	Projekthandbuch	Text schreiben (oder ausformulieren)
		Bei Änderungen Aktualisieren
03 Risikoliste	Risikoliste	Risiken/Maßnahmen -liste erstellen lassen
		Risiken/Maßnahme erstellen lassen, Attribute selber festlegen
		Risiken/Maßnahmen aktualisieren
04 Anforderungsspezifikation	Einleitung	Text schreiben (oder ausformulieren)
	Benutzeranforderungen	Text schreiben (oder ausformulieren)
	Systemübersicht	Komponenten & Beziehungen beschreiben
	Fachliches Datenmodell	Text schreiben (oder ausformulieren)
		FDM vollständig erstellen
		Entitäten erstellen lassen
		Attribute erstellen lassen
		Beziehungen erstellen lassen
		Beschreibung erstellen lassen
		Einschränkungen erstellen lassen
	Anwendungsfälle	Rollen + Anwendungsfälle beschreiben
	AF	Textuelle Beschreibung
	Benutzungsschnittstelle	Textuelle Beschreibung
	Dialog	Zuordnung Dialoge zu Anwendungsfällen
		Dialog erstellen
		Felder des Dialogs erklären
	Nichtfunktionale Anforderungen	Wann der Dialog auftreten soll
		Erstellen
05 Architekturdokument	Technische Infrastruktur	Zugehörigen Test erstellen
	Komponentenmodell	Diagramm als Text beschreiben
		Erklärenden Text formulieren (oder ausformulieren)
	Datenmodell	Diagramm als Text beschreiben
		Erklärenden Text formulieren (oder ausformulieren)
	Detaillierte Schnittstellenbeschreibung	Gleich FDM? Wo unterschiede?
	AF	Beschreibung erstellen Sequenzdiagramme erstellen Vorbedingungen erklären
06 Testspezifikationen	Testkonzept	Testablauf erstellen
	AF	Tabelle erstellen
		Tests zu den Anwendungsfällen erstellen (Testablauf & -ergebnis)
		(Tabelle erstellen)/ Rahmen erstellen
		Tabelle erstellen

Abbildung 3.1: Mögliche LLM Tool Unterstützung

ser zu gestalten, könnte man diese zunächst in Form von Stichpunkten vorschreiben und anschließend von den LLM-Tools in ausformulierte Texte umwandeln lassen.

Dabei sollte man sich überlegen, ob es sinnvoll ist, bestimmte Kapitel automatisiert erstellen zu lassen. Ein Beispiel hierfür ist die Projektorganisation. In diesem Kapitel werden der Teamaufbau mit den Namen, Rollen und Erreichbarkeiten der einzelnen Teammitglieder dokumentiert. In diesem Fall würde es länger dauern, den Prompt für die automatisierte Erstellung zu schreiben, als die Daten selbst in die Tabelle einzutragen.

Ein weiterer unterstützender Ansatz wäre, dass das Projekthandbuch bei Änderungen, wie beispielsweise personellen Veränderungen, automatisch aktualisiert wird.

3.3 Risikoliste

Nachdem im Unterabschnitt 2.2.3 die Bestandteile der Risikoliste beschrieben wurden, stellt sich auch hier die Frage, wie LLM Tools bei der Erstellung und Pflege der Risikoliste mitwirken und unterstützen können.

Eine erste Möglichkeit besteht darin, die gesamte Risikoliste einschließlich der zugehörigen Maßnahmen-tabelle von LLM Tools erstellen zu lassen. Dabei sollten alle Attribute für die Risiken festgelegt werden. Die Werte der einzelnen Risikoattribute sollten zudem zueinander in einem konsistenten Verhältnis stehen.

Falls dies nicht möglich ist, können LLM-Tools zur Erstellung von Risiken und Maßnahmen herangezogen werden, während die Festlegung der Attribute manuell erfolgt.

Eine weitere Möglichkeit, wie LLM-Tools bei der Risikoliste unterstützen können, liegt in der Pflege der Liste. Die Tools könnten zu bestimmten Zeitpunkten während des Projekts abgefragt werden, ob und wie sich die Risikoliste verändert hat. Sollten sich der Status oder andere Attribute der Risiken verändert haben oder sollten Risiken eingetreten sein, können die Tools die Risikoliste entsprechend aktualisieren.

3.4 Anforderungsspezifikation

Das Unterabschnitt 2.2.4 beschreibt die einzelnen Inhalte der Anforderungsspezifikation. Doch wie können LLM-Tools bei der Erstellung dieser unterstützen?

Zunächst könnte die Einleitung des Dokuments entweder komplett oder kapitelweise automatisch erstellt werden. Da der Kontext hierbei jedoch entscheidend ist, wäre es vermutlich sinnvoller, die Kapitel in Stichpunkten selbst zu verfassen und diese dann von den LLM-Tools in zusammenhängende Texte umwandeln zu lassen.

Im nächsten Schritt folgen die Benutzeranforderungen. Auch hier besteht die Möglichkeit, den gesamten

Text von den Tools erstellen zu lassen oder, um eine präzisere Formulierung zu erreichen, die Inhalte zunächst in Stichpunkten zu notieren und diese dann ausformulieren zu lassen.

Danach kommt die Systemübersicht. Da außer ChatGPT-4o keiner der Chatbots eigenständig Dokumente erstellen kann, muss die Zeichnung manuell erstellt werden. Es ist jedoch möglich, die Beschreibung der Komponenten und deren Beziehungen von den Tools erstellen zu lassen und diese Beschreibung dann in eine Zeichnung umzusetzen. Ebenso kann der Text, der das System beschreibt und in die Systemlandschaft einordnet, automatisch generiert werden. Auch hier kann es hilfreich sein, vorab Stichpunkte zu erstellen und diese dann in einen flüssigen Text umwandeln zu lassen.

Ähnlich verhält es sich beim fachlichen Datenmodell. Die kostenfreien LLM-Tools können zwar kein UML-Diagramm erstellen, aber sie können die Entitäten sowie deren Attribute und Beziehungen beschreiben. Die Diagrammerstellung muss daher manuell erfolgen. Die dazugehörigen Beschreibungen und Einschränkungen können hingegen von den Tools generiert werden. Falls dies nicht zufriedenstellend funktioniert, können die einzelnen Inhalte wie Entitäten, Attribute, Beziehungen sowie Beschreibungen und Einschränkungen jeweils nacheinander separat erstellt werden.

Anschließend folgt das Anwendungsfalldiagramm und eine detaillierte Beschreibung der einzelnen Anwendungsfälle mit Aktivitätsdiagramm. Das Anwendungsfalldiagramm muss wieder manuell erstellt werden, jedoch können die LLM-Tools die Rollen, deren Verbindungen und die zugehörigen Anwendungsfälle auflisten. Falls zusätzliche Bemerkungen zu den Anwendungsfällen erforderlich sind, können diese auch von den Tools erstellt werden. Das Aktivitätsdiagramm für die Anwendungsfälle muss ebenfalls manuell erstellt werden, wobei eine genaue und detaillierte Beschreibung erforderlich ist. Die Beschreibung des Anwendungsfalls kann hingegen vollständig von den Tools generiert werden.

Darauf folgt die Benutzungsschnittstelle, die ebenfalls aus einem Diagramm besteht, das zeigt, von welchem Dialog aus man zu anderen navigieren kann. Hier kann lediglich eine textuelle Beschreibung von den Tools erstellt werden, während das Diagramm manuell erstellt werden muss. Zusätzliche Bemerkungen und die Zuordnung der Dialoge zu den Anwendungsfällen können ebenfalls von den LLM-Tools übernommen werden. Die einzelnen Dialoge können textuell beschrieben werden, wobei die Tools die einzelnen Eingabefelder und Buttons erläutern sowie definieren können, wann der Dialog aufgerufen wird und was beim Schließen des Dialogs passiert.

Zuletzt werden die nichtfunktionalen Anforderungen mitsamt den dazugehörigen Tests aufgelistet. Die Anforderungen können von den Tools basierend auf den vorherigen Dokumenten erstellt werden. Falls die Tools die Anforderungen nicht eigenständig erstellen können, sollten zumindest die Tests von ihnen aufgestellt werden können.

3.5 Architekturdokument

Das Architekturdokument beschreibt die grundlegende Struktur und die Designentscheidungen eines Softwareprojekts. Die genauen Inhalte wurden bereits im Unterabschnitt 2.2.5 erläutert. Doch wie können LLM-Tools bei der Erstellung dieses Dokuments unterstützen?

Die technische Infrastruktur besteht aus einer Grafik und einem erklärenden Text. Bei der Erstellung der Grafik stellt sich das Problem, dass nur ChatGPT-4o in der Lage ist, Grafiken zu generieren. Daher muss man sich die einzelnen Hardwarekomponenten mit den entsprechenden Versionen und deren Beziehungen zueinander beschreiben lassen und die Grafik selbst erstellen. Den dazugehörigen erklärenden Text kann man automatisch generieren lassen. Falls dieser nicht präzise genug formuliert ist, kann es hilfreich sein, den Text zunächst in Stichpunkten vorzuschreiben und anschließend ausformulieren zu lassen.

Anschließend folgt das Komponentenmodell. Auch dieses kann nicht grafisch erstellt werden; lediglich die Komponenten und deren Schnittstellen zueinander können definiert werden. Hierzu gibt es einen begleitenden Text, der die Komponenten und Schnittstellen beschreibt und deren Funktion erläutert. Dieser Text kann ebenfalls vollständig generiert oder in Stichpunkten vorgeschrieben und anschließend ausformuliert werden.

Das Datenmodell entspricht häufig dem fachlichen Datenmodell. Hier kann überprüft werden, ob dies der Fall ist und, falls nicht, wo die Unterschiede liegen bzw. wie das Datenmodell tatsächlich aussieht.

Danach folgt die detaillierte Schnittstellenbeschreibung. Diese ist abhängig von der Projektstruktur der Implementierung. LLM-Tools könnten basierend auf den vorherigen Dokumenten eine Projektstruktur erstellen und anschließend die Schnittstellenbeschreibung generieren. Wenn jedoch bereits ein Verzeichnis mit den passenden Ordnern existiert, sollte die Schnittstellenbeschreibung manuell durchgeführt werden.

Nach der Schnittstellenbeschreibung folgt die dynamische Beschreibung exemplarischer Anwendungsfälle. Diese besteht aus einem Sequenzdiagramm und den für den Anwendungsfall benötigten Vorbedingungen. Das Sequenzdiagramm kann nur als Text beschrieben werden und muss anschließend selbstständig erstellt werden. Die Vorbedingungen sollten ohne Schwierigkeiten generiert werden können.

3.6 Testspezifikation

Im Unterabschnitt 2.2.6 wurde bereits erklärt, welche Inhalte die Testspezifikation umfasst. Doch wie können LLM-Tools bei der Erstellung dieses Dokuments unterstützen?

Zunächst kommt das Testkonzept. Dieses kann automatisch generiert werden. Falls das generierte Konzept nicht den eigenen Vorstellungen entspricht, kann man zunächst Stichpunkte selbst schreiben und sich dann den Text daraus erstellen lassen. Auch die dazugehörige Tabelle sollte sich von den LLM-Tools erstellen lassen.

Anschließend folgen die Tests für die einzelnen Anwendungsfälle. Die Testfälle kann man vollständig von den Tools erstellen lassen. Die Testdaten, die bei diesen Tests benötigt werden, werden in einer Tabelle zusammengefasst. Auch diese Tabelle könnten die Tools automatisch generieren.

Der Rahmen der Tabelle mit dem Testprotokoll kann ebenfalls von den LLM-Tools erstellt werden. Die Ergebnisse könnten die Tools auch generieren; allerdings stellt sich hier die Frage, ob dies zu einer Zeitersparnis oder einer Arbeitserleichterung führt.

4 Praxisergebnisse und Vergleich

TODO!!!!!!!

Generell: - Schwierigkeiten in einem großen/langen Chat, die Tools von einer Formulierung weg zu bringen. - Le Chat erstellt immer Link mit Bild aber /Xyxyxyxy ? - Le Chat warnt manchmal vor Eingabe: "Der Inhalt kann gefährliche oder sensible Themen beinhalten" für Eingabe: "Formuliere die Beschreibung der Aktivitäten etwas genauer. Beschreibe welcher Wert wo gespeichert wird mithilfe des Fachlichen Datenmodells."

4.1 Bewertung

Um die Ausgaben der Tools besser zu vergleichen und ein nachvollziehbares Fazit zu ziehen, wurde ein Bewertungssystem entwickelt, das fünf Stufen umfasst. Diese Bewertungsstufen ermöglichen eine strukturierte und transparente Methode zur Beurteilung der Qualität der Ausgaben von LLM Tools. Die fünf Bewertungsstufen sind wie folgt definiert:

“Sehr gut” wird vergeben, falls der Inhalt der Ausgabe fachlich richtig ist und auch das Format den Vorgaben entspricht. Also, dass die Ausgabe genauso in das entsprechende Dokument eingefügt werden kann.

Die Ausgabe wird als “Gut” eingestuft, wenn der Inhalt etwas fehlerhaft ist, man aber mit ein paar wenigen Eingaben den Inhalt korrigieren kann und das Format den Vorgaben entspricht. Eine Ausgabe wird auch als “Gut” bewertet, wenn der Inhalt richtig ist, aber das Format etwas fehlerhaft ist. Man sollte aber den richtigen Inhalt einfach aus der Ausgabe kopieren können und in das entsprechende Dokument mit dem richtigen Format einfügen können.

Als “Befriedigend” wird eine Ausgabe bewertet, wenn der Inhalt und das Format etwas fehlerhaft sind, sich diese aber mit ein paar wenigen Eingaben korrigieren lassen oder man den richtigen Inhalt aus der Ausgabe kopieren kann.

“Ausreichend” wird vergeben, wenn in der Ausgabe inhaltlich Teile fehlen, man aber mit ganz genauen Eingaben, von dem was man will, die Ausgabe etwas verbessert, aber diese immer noch nicht vollständig ist.

Zuletzt wird eine Ausgabe als “Mangelhaft” bewertet, wenn der Inhalt der Ausgabe falsch oder nicht verständlich ist und es sich auch nicht korrigieren lässt.

Außerdem fließen auch spezifische Aspekte, die nur für das jeweilige Aspekt wichtig sind, in die Bewertung mit ein. Diese werden in den entsprechenden Kapiteln für die Dokumente im Fazit erläutert.

Erstell mir aus folgendem Gespräch ein erstes Besprechungsprotokoll für ein Projekt. Das Protokoll soll nur die wichtigen Punkte in Form einer Tabelle mit den Spalten "Nummer: was eine eindeutige Nummer zur Identifikation ist, Ärt: eine Auswahl ob es eine Information, ein Auftrag, eine Feststellung oder eine Beschluss ist, "Beschreibung: Was den Punkt kurz und präzise zusammenfasst, "Termin: ein genaues Datum bis wann der Auftrag erledigt sein muss und "Verantwortlich: Welches Teammitglied verantwortlich ist, enthalten. Dabei müssen Aufträge mit einem genauen Fälligkeitsdatum und einem Verantwortlichen versehen sein. Beschlüsse müssen klar und unmissverständlich formuliert werden. Feststellungen sind Beschlüsse, die keine Abstimmung benötigen und Informationen bieten den Projektmitgliedern wichtige Hinweise: [Hier folgte das Gespräch]

Auf die Spalte "Termin" musste ebenfalls ein besonderes Augenmerk gelegt werden. Wenn im Prompt nicht explizit angegeben war, dass ein genaues Datum erforderlich ist, arbeiteten die Tools oft mit relativen Angaben wie "+2 Wochen". Selbst mit der Klarstellung, dass der Termin ein genaues Datum sein sollte, traten bei Le Chat Probleme auf: Es wurden immer Termine gewählt, die in der Vergangenheit lagen. Dazu wurde angegeben, dass die Termine beispielhaft gewählt und auf das aktuelle Datum, den

29.03.2023, bezogen seien.

Google Gemini hatte, wie bereits erwähnt, erhebliche Schwierigkeiten, nur die wichtigsten Punkte aus dem Gespräch herauszufiltern. Häufig wurden Besprechungsprotokolle mit etwa 30 Punkten erstellt, wobei jede einzelne Information separat und auch unwichtige Informationen aufgeführt wurden. Gemini fügte zudem eine neue Art, "Frage", selbstständig hinzu, wodurch im Besprechungsprotokoll teilweise mehrere Punkte für eine Information erstellt wurden. Bei erneuten Eingaben variierte die Länge des Besprechungsprotokolls stark. Trotz des Hinweises, nur die wichtigen Punkte in das Protokoll aufzunehmen, neigte Gemini weiterhin dazu, sehr detaillierte und kleinteilige Protokolle zu erstellen.

Le Chat hingegen hielt sich häufig zu knapp. Dadurch wurden immer wieder wichtige Anforderungen nicht in das Besprechungsprotokoll aufgenommen, und es musste besonders darauf geachtet werden, ob alle relevanten Informationen enthalten waren. Besonders, dass mehrere Spiele gleichzeitig gespielt werden können und das ansprechende Animationen verwendet werden sollen, wurden nicht mit aufgeführt.

Zusammenfassend lässt sich sagen, dass die Erstellung des Besprechungsprotokolls durch die LLM Tools grundsätzlich schon ganz gut funktioniert. Besonders ChatGPT hat wenig Schwankungen in seinen generierten Ausgaben und hat eigentlich immer die gleichen Punkte im Besprechungsprotokoll, formuliert diese nur ein bisschen unterschiedlich. Daher wird ChatGPT mit "Sehr gut" bewertet. Die anderen beiden haben damit etwas mehr Schwierigkeiten, doch wenn man sich das Besprechungsprotokoll immer wieder neu generieren lässt, kommt irgendwann ein anständig erstelltes Protokoll. Da die Zusammenfassung des Gesprächs auf die wesentlichen Punkte, die Hauptaufgabe eines Besprechungsprotokolls ist, erhalten Le Chat und Gemini lediglich die Bewertung "Ausreichend". Das Le Chat und Gemini nicht alle Anforderungen in das Protokoll aufnehmen, kann, falls dies nicht auffällt, im Nachhinein das Projekt verzögern und die Kosten in die Höhe treiben. Geminis erstellte Protokolle mit etwa 30 Punkten verfehlen dahingegen komplett den Sinn eines Besprechungsprotokolls.

4.3 Projekthandbuch

Das Projekthandbuch wurde in zwei Schritten erstellt. Zunächst wurde die Einleitung mit dem Zweck des Dokuments, der Redaktion und dem Verteiler verfasst. Hierzu wurde der folgende Prompt im selben Chat eingegeben, in dem auch das Besprechungsprotokoll erstellt wurde:

Erstell mir für dieses Projekt die Einleitung für das Projekthandbuch. Die Einleitung besteht aus einem Abschnitt für den Zweck des Dokuments, einen Abschnitt zur Redaktion, in welchem geklärt wird, wer für das Dokument verantwortlich ist und einen Abschnitt zu dem Verteiler, also wer bei Änderungen zu informieren ist. Verantwortlich für das Dokument ist der Projektleiter und über Änderungen wird das gesamte Team informiert. Dazu wird eine entsprechende Nachricht in den Discord Channel geschrieben.

Prompt 4.2: Prompt Einleitung Projekthandbuch

Für die Einleitung sind spezifische Informationen erforderlich, die im Prompt angegeben werden müssen. In diesem Projekt sind das die Verantwortlichkeit des Projektleiters und der Verteiler über den Discord-Channel. Es fiel lediglich auf, dass ChatGPT im Abschnitt "Verteiler" alle Teammitglieder nannte, was nicht unbedingt nötig ist. Ansonsten wurden die Abschnitte gut erstellt [Verweis!!!!!!!!!!!!!!!!!!!!!!!!!!!!].

Der zweite Teil betrifft das Kapitel "Projektdefinition" mit den Abschnitten "Vorgeschichte" und "Inhaltliche Kurzdarstellung". Dazu wurde der folgende Prompt verwendet, welcher nach der Ausgabe der Einleitung im Chat eingegeben wurde:

Erstelle mir nun das Kapitel "Projektdefinition" des Projekthandbuches. Der erste Abschnitt soll die Vorgeschichte des Projekts beschreiben und anschließend soll ein Abschnitt eine inhaltliche Kurzdarstellung beschreiben.

Prompt 4.3: Prompt Projektdefinition Projekthandbuch

Diese zweite Eingabe erforderte keine weiteren Informationen, da diese im Gespräch bereits vorgegeben waren und die Tools darauf zugreifen können sollten. Auch hier wurden die Abschnitte gut erstellt, und die Tools konnten die benötigten Informationen aus dem Gespräch gut extrahieren [Verweis!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!].

Bei der Erstellung der Einleitung zeigte sich, dass eine genaue Beschreibung der benötigten Abschnitte entscheidend war. Ohne diese klare Vorgabe neigten die Tools dazu, eigene Strukturen und Inhalte zu erstellen, die nicht den Anforderungen entsprachen.

Außerdem traten bei Gemini teilweise Fehler in den Namen auf. Einmal wurde beispielsweise geschrieben, dass Frau Schmidt (Teamchefin) mit Herrn Müller (Kunde) Kontakt aufnahm, was eine fehlerhafte Zuordnung darstellt.

Abschließend lässt sich sagen, dass die Abschnitte von allen drei Tools gut waren. Der Fehler von Gemini zeigt allerdings, dass man sich die Ausgaben trotzdem nochmal achtsam durchlesen sollte. Aufgrund der Bewertungskriterien wird Gemini durch seinen Fehler mit "Gut" bewertet. Le Chat und

Gemini werden beide mit "Sehr gut" benotet.

4.4 Risikoliste

Auch die Risikoliste wurde in zwei Schritten erstellt. Zunächst wurde nur die Risikotabelle erstellt und im zweiten Schritt die Tabelle mit den Maßnahmen. Die beiden Ergebnisse sind [Verweis!!!!!!!!!!!!!!].

Der Prompt für die Risikotabelle

Erstell mir für diese Projekt eine Risikoliste. Diese soll aus mehreren Spalten bestehen: ID für eine eindeutige Identifikationsnummer, "Beschreibung" für eine ausführliche Beschreibung des Risikos und der Auswirkungen, "Datum" für den Zeitpunkt, wann das Risiko identifiziert wurde, "Äutor" für die Person die das Risiko gemeldet hat, "Wahrscheinlichkeit (in%)" für einen Schätzwert der Eintrittswahrscheinlichkeit des Risikos, "SSchaden (in €)" für eine Schätzung wie groß der Schaden ist, "Maß (in €)" was das Produkt aus Wahrscheinlichkeit und Schaden ist, "Risikoklasse" für eine Priorisierung der potentiellen Risiken wo zwischen Tolerierbar, Unerwünscht, Kritisch und Katastrophal Unterschieden wird und "SStatus" wo zwischen aktiv, eingetreten und geschlossen unterschieden wird. Das Risiko ist Tolerierbar wenn das Risikomaß geringer als 0,1% des Projektvolumen ist, Unerwünscht wenn es größer als 0,1% ist, Kritisch wenn es größer als 1% ist und Katastrophal wenn es größer als 10% ist. In der Risikoliste sollen Team-, Technische-, Methodische-, Kunden-, Fachliche-, Produkt-, Management- und Planungsrisiken betrachtet werden. Diese sollen mit einer leer Zeile getrennt werden, in welchen die Oberbegriffe stehen. Das Projektvolumen beträgt 500000€.

Prompt 4.4: Prompt Risikotabelle

ist sehr umfangreich formuliert. Damit die Tabelle mit den richtigen Spalten erstellt wird, wurde im Prompt jede einzelne Spalte aufgezählt und beschrieben. Ebenfalls ist es wichtig, die möglichen Werte für die Risikoklasse zu definieren, da hier sonst immer "Hoch", "Mittel" und "Niedrig" von den Tools verwendet wird. Es musste auch festgelegt werden, wann die einzelnen Risikoklassen auftreten, da die Zuweisung ansonsten recht schwammig ausfällt und schwerwiegende Risiken eine geringere Risikoklasse erhalten als eher unwichtige Risiken. Außerdem sollte beschrieben werden, welche Arten (Team-, Technische-, Methodische-, Kunden-, Fachliche-, Produkt-, Management- und Planungsrisiken) von Risiken betrachtet werden sollen und dass diese Arten in einer leer Zeile stehen, welche die dazugehörigen Risiken von den anderen Arten trennen. Ansonsten kann es passieren, dass die Risiken durcheinander geschrieben werden und damit nicht den Arten zuzuordnen sind. Damit der Risikoschaden zwischen den Tools vergleichbar ist, sollte das Projektvolumen definiert werden. Ansonsten wird auch dies von Tools festgelegt und führt zu Ungenauigkeiten im Vergleich der einzelnen Ausgaben.

Grundsätzlich war die Erstellung der Risiken kein Problem. Schwierig war es eher die gewünschte Formatierung zu erhalten, sowie eine richtige Zuweisung der restlichen Attribute. Besonders Gemini

und Le Chat haben dabei Schwierigkeiten. Außerdem war bei allen drei Tools auffällig, dass der Autor immer die Person ist, zu dem das Risiko in den Tätigkeitsbereich fällt. Also die Person, die auch der Verantwortliche ist. Fraglich ist dabei teilweise, wenn Kundenrisiken vom Kunden, also Herr Müller, erstellt werden, da dieser an der Erstellung der Risikoliste überhaupt nicht beteiligt ist.

ChatGPT hat die Tabelle sehr gut erstellt. Lediglich die Zuweisung der Risikoklasse war falsch. Nach einem Hinweis diesbezüglich wurden diese jedoch, mit einer Ausnahme, richtig korrigiert.

Ein Problem bei Gemini ist, dass das Risikomaß nicht richtig berechnet wird und auch die Risikoklassen nicht korrekt zugewiesen werden. Das Komma der Werte ist um eine Stelle zu weit links. Es muss alles einmal $\cdot 10$ gerechnet werden, damit die Werte stimmen. Kritisch bei Gemini ist, dass die Tabelle nicht in einem Zug erstellt werden kann. Es wird während der Generierung einfach aufgehört. Auch wenn man Gemini fragt, ob die vollständige Tabelle generiert werden kann, wird mitten drin aufgehört. Man muss explizit nur nach den noch offenen Risikoarten fragen. Diese werden dann erstellt, jedoch soll man den Schaden und das Maß selbst eintragen und berechnen. Nach anschließender Frage, ob er den Risikoschaden und das Schadensmaß festlegen kann, sagt er, dass er detaillierte Informationen über das Projekt und die potenziellen Risiken benötigt um genaue Werte festlegen zu können.

Le Chat hatte ein ähnliches Problem wie bereits im Besprechungsprotokoll [Abschnitt 4.2], dass das Datum immer auf den 01.04.2023 gesetzt wird. Außerdem hatte Le Chat Probleme damit, die Tabelle in das gewünschte Format zu bringen, dass die Risikoarten in einer Zeile stehen und die dazugehörigen Risiken darunter aufgelistet werden. Sie wurden lediglich mit 1.x bis 8.x beschriftet wodurch Sie sich zu den Arten zuweisen ließen. Auch hier wurden die Risikoklassen falsch bestimmt und auch nach einem Hinweis wurden diese nicht korrigiert. Jedoch wurde dabei der Risikoschaden der einzelnen Risiken geändert. Auch hier ist die Zuweisung der Risikoklassen nicht richtig.

Die Maßnahmentabelle wurde mit folgendem Prompt erstellt:

Erstelle mir nun eine dazu passende Maßnahmentabelle. Auch diese besteht aus mehreren Spalten: "Typ" beschreibt ob die Maßnahme das Risiko verhindert, lindert oder überträgt. In "Beschreibung" wird die Maßnahme beschrieben. Falls ein Risiko als nicht mehr relevant eingestuft wird, wird in der Spalte "Beschreibung" eine Begründung eingetragen und die Maßnahme auf "beendet" gesetzt. "Trigger" ist das Ereignis, das den Start der Maßnahme veranlasst, falls diese nicht sofort eingeleitet werden soll. "Verantwortlicher" ist die zuständige Person für die Durchführung der Maßnahme. "Status" unterscheidet zwischen geplant, aktiv und beendet. Anschließend gibt es jeweils eine Spalte für "Restwahrscheinlichkeit (in %)", "Restschaden (in €)", "Restmaß (in €)" und "Restklasse", was die geschätzte Wahrscheinlichkeit, geschätzter Schaden, Maß und Klasse des Restrisikos, nach Durchführung der Maßnahme entsprechen. Für jedes Risiko sollen zwei Maßnahmen erstellt werden. Dazu wird über die zwei Maßnahmen eine Zeile mit der ID von dem Risiko beschrieben, auf die sich die Maßnahmen beziehen.

Prompt 4.5: Prompt Maßnahmentabelle

Der Prompt für die Maßnahmen ist ähnlich wie der, für die Risikotabelle. Es wird jede Spalte der Tabelle einmal beschrieben, damit das Format der Tabelle mit den Vorgaben übereinstimmt. Anschließend wird gesagt, dass die ID des dazugehörigen Risikos mit in die Tabelle übernommen werden soll, damit man die Maßnahmen den Risiken zuordnen kann. Die Probleme sind hier ähnlich zu denen bei der Risikotabelle, jedoch erstellen alle drei Tools grundsätzlich anständige Maßnahmen für die Risiken. Bei allen dreien fühlt sich allerdings die Zuweisung der Attribute zu den Maßnahmen repetitiv an. Häufig wechseln sich die auswählbaren Parameter ab und auch die Wahrscheinlichkeiten und der Restschaden sind meistens immer wieder die gleichen Zahlen.

ChatGPT erstellt auch die Maßnahmentabelle sehr gut. Lediglich die Zuweisung der Restklasse stimmt nicht überein.

Bei Gemini wird für die Begründung, warum ein Risiko als nicht mehr relevant eingestuft wird, eine eigene Spalte erstellt. Auch wenn man die Beschreibung dafür im Prompt ändert, wird die Spalte erstellt. Außerdem ist die Formatierung bei Gemini manchmal nicht so gut, da bei jeder 2. Zeile alle Daten ab der Spalte "Begründung" um eine Stelle nach links verschoben wird. Lässt man sich die Antwort neu generieren und auch wenn man Gemini sagt er soll die Formatierung korrigieren bleibt das Problem bestehen. Ein weiteres Problem ist, dass die Spalte "Trigger" nicht gefüllt wird und auch das Problem, dass die Tabelle nicht ganz vollständig generiert wird sondern mitten drin aufhört, tritt wieder auf. Teilweise ist das Restmaß falsch berechnet und auch die Zuweisung der Restklasse ist nicht immer richtig.

Besonders aufgefallen bei Le Chat ist, dass die Attribute für die Maßnahmen sehr repetitiv festgelegt wurden. Die Restwahrscheinlichkeit beträgt für jede Maßnahme 5% und auch der Restschaden ist häufig für die Maßnahmen für eine Risikoart gleich. Auch hier ist die Restklasse nicht entsprechend zugewiesen worden sondern ist, bis auf bei der ersten Maßnahme, auf Unerwünscht festgelegt. Auch

die Festlegung des Typs der Maßnahmen wirkt sehr repetitiv, da hier immer zwischen “verhindern” und “lindern” gewechselt wird.

Zusammenfassend lässt sich sagen, dass die Erstellung von Risiken und dazu gehörigen Maßnahmen gut funktioniert, allerdings die Festlegung der zusätzlichen Attribute nicht den Vorgaben entspricht und häufig zu Fehlern führt. Auch die richtige Formatierung der Tabelle ist häufig etwas schwierig, lässt sich aber in den meisten Fällen korrigieren. Da ChatGPT deutlich weniger Probleme bei der Erstellung hatte und auch die Zuweisung der Attribute, abgesehen von der Risikoklasse, und die Formatierung gut funktioniert hat, wird es mit “Gut” bewertet.

Le Chat hatte etwas mehr Probleme gemacht. Bei der Risikotabelle war die Formatierung insofern kein Problem, da man die einzelnen Spalten einfach in eine eigene Tabelle übernehmen konnte und die Risikoarten selbst dazwischen schreibt. Erst bei der Maßnahmentabelle hat LeChat Probleme mit der Zuweisung der Attribute bekommen. Daher wird Le Chat mit “Befriedigend” bewertet.

Gemini hat am meisten Probleme gemacht. Allein schon die Tatsache, dass Gemini mitten in der Generierung diese als fertig betrachtet und dadurch die Tabellen unvollständig erstellt werden, sorgt dafür, dass Gemini für solche Erstellungen eher ungeeignet ist. Die anderen Probleme wie die fehlerhafte Formatierung und dass die Attribute falsch berechnet und festgelegt werden kommen dabei noch hinzu. Dies sorgt dafür, dass Gemini hier die Bewertung “Mangelhaft” erhält.

Wenn man allerdings sich nur die Erstellung von Risiken mit dazugehörigen Maßnahmen anschaut, erhalten alle drei Tools die Bewertung “Sehr gut”. Die erstellten Risiken und Maßnahmen waren zwar alle eher welche, die man vermutlich als erstes benennt und auch Projekt unabhängig sind, wurden aber dennoch sehr gut erstellt. Auch bei Gemini werden einzelne Risiken mit dazugehörigen Maßnahmen in einem Rutsch erstellt, weshalb man dieses Problem bei der Erstellung der Risiken und Maßnahmen nicht ankreiden kann.

4.5 Anforderungsspezifikation

Für die Anforderungsspezifikation wurde für jeden Inhalt ein eigener Prompt erstellt. Daher wird im folgenden die Erstellung der Einleitung, der Systemarchitektur, des fachlichen Datenmodells, des Anwendungsfalldiagramms, der detaillierten Anwendungsfall-Beschreibung mit Aktivitätsdiagramm, der Benutzungsschnittstelle mit der Dialog Navigation, der detaillierten Ausarbeitung von Dialogen und der nichtfunktionalen Anforderungen einzeln beschrieben und anschließend bewertet.

Einleitung

Für die Einleitung wurde der folgende Prompt verwendet:

Erstell mir für die Anforderungsspezifikation eine Einleitung. Diese soll kurz, mit 2-3 Sätzen, in das Dokument einführen. Anschließend soll ein Abschnitt zum Zweck und Umfang des Dokuments kommen, welcher eine Beschreibung der Notwendigkeit des Systems für den Kunden sowie eine Kurzbeschreibung der Funktionalität und der Nachbarsysteme beinhaltet. Danach wichtige Begriffe und Abkürzungen erklärt werden. Zum Schluss kommt ein Abschnitt "Verweise auf sonstige Ressourcen", wo auf die Spielregeln verwiesen wird.

Prompt 4.6: Prompt Einleitung Anforderungsspezifikation

Bei der Erstellung des Prompts musste darauf geachtet werden, den Abschnitt "Zweck und Umfang des Dokuments" genauer zu beschreiben, da hier ansonsten die generierte Ausgabe sehr ausschweifend und unspezifisch erstellt werden. Grundsätzlich hat die Erstellung gut geklappt.

Bei ChatGPT ist lediglich auffällig, dass er beim Abschnitt "Zweck und Umfang des Dokuments" nicht den Zweck der Anwendung beschreibt, sondern den Zweck des Dokuments. Die Kurzbeschreibung der Funktionalität wurde dafür gut erstellt. Bei der Kurzbeschreibung der Nachbarsysteme wurden die einzelnen Teile der Anwendung erläutert, also das Spielsystem, das Kommunikationsprotokoll für RMI und die Benutzeroberfläche, und keine anderweitigen System, da die Anwendung alleinstehend ist.

Bei Gemini []

Le Chat hat die Beschreibung gut erstellt. Der einzige Aspekt der auffällig war ist, dass beim Abschnitt "Verweis auf sonstige Ressourcen" ein Link auf eine Webseite für die Regeln von Mensch ärgere dich nicht eingefügt wird.

Bewertet wird ChatGPT mit "Gut", da hier der Zweck des Dokuments beschrieben wurde, obwohl im Prompt explizit nach einer Beschreibung der Notwendigkeit des Systems für den Kunden gefragt wurde. Le Chat wird mit "Sehr gut" benotet, da hier keine inhaltlichen Fehler vorhanden sind und soweit alles gepasst hat.

Systemarchitektur

Die Systemarchitektur wurde mit folgenden Prompt erstellt:

Erstell mir für die Anforderungsspezifikation die Systemarchitektur. Diese beinhaltet einen groben Überblick über die erwartete Systemarchitektur und Einordnung des Systems in die Systemlandschaft des Kunden. Falls Schnittstellen zu Nachbarsystemen bestehen, dann müssen diese abgebildet werden.

Prompt 4.7: Prompt Systemarchitektur

Auch hier muss beschrieben werden, was genau erstellt werden soll. Wichtig hier ist, dass man in den selben Chat schreibt, in dem man die Dokumente vorher schon erstellt hat. Dadurch wissen die Tools was zu erstellen ist.

Bei ChatGPT wurde sogar ein Diagramm in den Chat "gezeichnet". Die Beschreibung und das Diagramm ist jedoch für die Systemarchitektur schon viel zu genau. Es hätte gereicht den Client, den Server und die Datenbank darzustellen und diese richtig miteinander zu verbinden. Bei der ersten Erstellung wurden die Verbindungen zwischen Client und Server und Server und Datenbank unidirektional dargestellt. Nach einer weiteren Eingabe, ob das so gewollt ist, wurde dies jedoch korrigiert und die Verbindung wurde bidirektional eingezeichnet.

Gemini[[[[[[[[[[[[[[[[[[[[[[[[]]]]]]]]]]]]]]]]]]]]]]

Nach der Eingabe kam bei Le Chat eine kurze Textuelle Beschreibung der Client- und Server-Anwendung sowie der Datenbank. Außerdem wurde beschrieben, dass Schnittstellen zu Nachbarsystemen ebenfalls mit aufgenommen werden sollen. Diese Beschreibungen haben soweit gut gepasst. Auf nachfrage, wie denn dann die Abbildung aussehen soll, wurde beschrieben, dass der Client, der Server und die Datenbank jeweils eigenständige Blöcke sein sollen. Dies ist soweit richtig. Es wurde jedoch nicht genau Beschrieben wie die Beziehungen zwischen den Komponenten aussehen. Daher wurde nochmal eine Eingabe getätigt, wie denn die Beziehung zwischen den Komponenten aussieht. Hier wurde richtig dargestellt, dass der Client und der Server kommunizieren und der Server mit der Datenbank, jedoch nicht der Client mit der Datenbank. Die restliche Beschreibung, wie diese Komponenten kommunizieren, ist wieder etwas zu ausführlich. Die Beschreibung reicht jedoch, um die Skizze selber zu erstellen, was kein negatives Kriterium ist, da bei den kostenfreien Versionen nicht davon ausgegangen wird, dass diese das können.

Zusammenfassend lässt sich hier sagen das alle [zwei???] Tools die Systemarchitektur etwas zu genau erzeugen. Jedoch lassen sich die richtigen Inhalte aus der Ausgabe kopieren und die zu detaillierten Beschreibungen können einfach rausgelöscht werden, weshalb die Tools mit der Note "Gut" bewertet werden.

Fachliches Datenmodell

Das erste größere und anspruchsvollere Dokument ist das Fachliche Datenmodell. Dieses wurde mit folgendem Prompt erstellt:

Erstell mir nun für die Anforderungsspezifikation das Fachliche Datenmodell. Dieses soll mit Hilfe eines UML Klassendiagrammes mit ergänzenden Beschreibungen bzw. Einschränkungen spezifiziert werden. Das Modell soll alle Entitätstypen mit deren Eigenschaften, Beziehungen und Einschränkungen besitzen.

Prompt 4.8: Prompt Fachliches Datenmodell

Dabei muss explizit erwähnt werden, dass ein UML-Klassendiagramm gewünscht ist, damit die Ausgabe das richtige Format besitzt. Außerdem sollte erwähnt werden, dass ergänzende Beschreibungen mit erstellt werden soll, damit diese Vorgabe auch direkt mit erfüllt ist und die Tools ihr Diagramm einmal näher erläutern.

Bei ChatGPT war direkt auffällig, dass für jede Klasse eine ID erstellt wurde. Dies führt zu einem dazu, dass die Anwendung sehr Datenbanklastig ist und zum anderen ist dies auch einfach bei manchen Klassen nicht nötig. Zum Beispiel beim User könnte man, je nach logIn-Daten, den Namen oder die E-Mail-Adresse einzigartig machen und damit den User identifizieren. Des Weiteren haben die Beziehungen nicht immer gepasst. Zum Beispiel die 1:1 Beziehung zwischen "Spielfigur" und "Feld". Es muss natürlich nicht auf jedem Feld eine Spielfigur stehen sondern auf einem Feld kann (0..1) eine Spielfigur stehen. Dieses Problem zieht sich auch durch die gesamten Versuche, das fachliche Datenmodell zu korrigieren. Häufig passiert es auch, dass wenn ein anderer Fehler korrigiert wird, eine richtige Beziehung zu etwas falschem geändert wird. Bei der erstellten Skizze waren ebenfalls die Beziehungen falsch eingezeichnet. Außerdem zeigten Beide Pfeile zwischen zwei Klassen in die gleiche Richtung. Bei Nachfrage, ob die Beziehung nicht in beide Richtungen zeigen sollten, wurde gesagt, dass das stimmt und die Beziehungen bidirektional sein sollten, jedoch wurde an der Zeichnung nichts geändert. Des Weiteren ist aufgefallen, dass bei der Klasse "Feld" ein Attribut gefehlt hat, das beschreibt, ob das Feld belegt ist oder nicht. Außerdem wurde in einem String in "Spiel" das Spielbrett gespeichert. Auf Nachfrage wie dies implementiert werden soll, wurde geantwortet, dass dies nicht die beste Lösung ist, sondern das Spielbrett dynamisch in der Anwendung generiert und angezeigt werden sollte. Weitere Probleme hatte ChatGPT mit der Speicherung der Farbe und der Position. Manchmal wurden diese Attribute in "Spieler", manchmal in "Spielfigur" und teilweise auch in beiden gespeichert. Damit es später einfacher zu realisieren ist, an mehreren Spielen teilnehmen zu können, wurde die Eingabe, dass es mehr Sinn macht, den "Spieler" in "User" und "Spieler" aufzuteilen, getätigt. ChatGPT hat dies dann genau wie beschrieben gemacht. Des Weiteren wurde der Würfel vergessen, welcher jedoch nach Hinweis diesbezüglich in "Spiel" hinzugefügt wurde. Nach der letzten Ausgabe des fachlichen Datenmodells wurde, nach Aufforderung, ein PlantUML-Diagramm erstellt. In diesem wurde an den Assoziationen Leseunterstützung angefügt. Diese Leseunterstützung sind jedoch nur von oben nach unten hilfreich. Zum Beispiel "Feld steht auf Spielfigur" ist keine hilfreiche Beschriftung. Des Weiteren sind alle Festlegungen der Vielfachheit auf der falschen Seite. Als Beispiel muss an der Assoziation zwischen "Spieler" und "Spielfigur" die "1" auf der Seite von "Spieler" sein und die "1..4" auf der Seite von "Spielfigur". Des Weiteren würden Pfeile an den Assoziationen helfen, die Flussrichtung zu verdeutlichen.

Die zu dem Dokument gehörige ergänzende Beschreibungen und Einschränkungen wurden erst auf Nachfrage erstellt, doch ab da an passten diese ganz gut. Auch, dass beim Format direkt die Länge der Strings begrenzt wurden, zeigt das ChatGPT hier auch Sicherheitstechnische Aspekte mit einbezieht.

Gemini[[[[[[[[]]]]]]]]

Auch Le Chat hatte zu Beginn nur eine Klasse "Spieler". Für eine Anwendung, wo ein Spieler an mehreren Spielen teilnehmen soll, macht es jedoch Sinn, diese Klasse auf zwei Klassen, "User" und "Player", aufzuteilen. Desweiteren wurde bei Le Chat, wie es eigentlich in einem UML-Klassendiagramm gehört, die Typen der Attribute nicht festgelegt. Außerdem ist die Frage, wenn die E-Mail-Adresse eines Spielers eindeutig sein muss, warum dann trotzdem eine eindeutige ID benötigt wird. Daneben wurde auch die Anforderung des Kunden, dass ein Punktesystem eingebunden werden soll, ignoriert, da es kein

Attribut zum Speichern der Punktzahl gibt. Die Bezeichnung der Beziehungen zwischen den Klassen ist, in der Textuellen Beschreibung für die Beziehung, eher an einem ER-Modell orientiert und nicht an einem UML-Klassendiagramm. Die genaue Bezeichnung für die Vielfachheit muss aus der Beschreibung selbst erstellt werden. Diese war jedoch auch teilweise fehlerhaft. Zum Beispiel, dass ein Player an einem oder mehreren Spielen teilnehmen kann, macht kein Sinn, da die Klasse "Player" die Eigenschaften eines Spielers im Spiel beschreibt. Diese kann schlecht in mehreren Spielen gleich sein. Auch einfachere Sachen wie die Zuweisung der Spielfiguren war ein Problem. Zum Beispiel bei der Beschreibung der Beziehung zwischen Player und Spielfigur wurde definiert, dass ein Player eine oder mehrere Spielfiguren besitzen kann. Dies entspricht aber nicht den normalen Regeln von Mensch ärgere dich nicht, wo jeder Spieler immer genau vier Spielfiguren besitzt. Daneben wurde auch hier der "Player"-Klasse ein Attribut für die Position auf dem Spielbrett festgelegt. Nach der Eingabe [Verweis!!!! 05_03_LeChat] wurden genau diese Beziehungen richtig geändert, jedoch waren immer noch Fehler enthalten wie zum Beispiel, dass ein Spiel eine oder mehrere Spielfiguren enthalten kann. Anschließend wurde gefragt ob der Würfel auch in das Fachliche Datenmodell mit aufgenommen werden sollte, woraufhin sowohl eine eigene Klasse "Dice" als auch ein Attribut "dice: Dice" in der "Game"-Klasse erstellt wurde. Anschließend wurde gefragt ob das Fachliche Datenmodell in PlantUML erstellt werden kann. Hier fällt besonders auf, dass die Verbindungen zwischen den Klassen alles Kompositionen sind, was aber soweit auch Sinn ergibt. Außerdem fällt auch auf, dass hier bereits die Primär- und Fremdschlüssel dargestellt werden. Dies ist für das Fachliche Datenmodell allerdings zu spezifisch. Ein weiterer Fehler ist die Beziehung zwischen "Game" und "Figure". Hier wird festgelegt, dass ein Spiel aus 16 Figuren besteht. Allerdings gibt es auch ein Attribut "max_players" und es wurde bereits festgelegt, dass ein Spiel aus 1..4 Spieler besteht. Daher kann es auch sein, falls nicht immer auf 4 Spieler mit KI-Gegner aufgefüllt wird, dass weniger als 16 Figuren auf dem Spielfeld sind. Hier lässt sich allerdings eh kritisch hinterfragen ob die Komposition zwischen "Game" und "Figure" überhaupt notwendig ist, da durch die Verbindung zwischen "Figure" und "Player" und der zwischen "Game" und "Player" die Figuren sowieso genau einem Spiel zugeordnet sind.

Le Chat erstellt das Format und die Beschreibung der Attribute auch erst auf Nachfrage. Jedoch hält sich die Ausgabe in Grenzen. Bei dem Format wird lediglich der Typ der Attribute beschrieben und es gibt keine Eingrenzung der Größe und der Länge der Werte.

Alle drei Tools schaffen es, ein grobes Grundgerüst für das Fachliche Datenmodell zu erstellen. Bei [ChatGPT und Le Chat] war dabei auffällig, dass für jede Klasse eine eindeutige ID erstellt wurde. Das Hauptproblem bei ChatGPT stellt die richtige Darstellung der Assoziationen zwischen den Klassen dar. Besonders die Beschreibung der Vielfachheit stellt ChatGPT vor Schwierigkeiten. Dies funktioniert bei Le Chat etwas besser, aber immer noch nicht fehlerfrei. Dafür hatte Le Chat mit mehreren kleineren generellen Problemen zu kämpfen. Es ist außerdem auffällig dass ab einem gewissen Punkt, die Erstellung des Prompts und das analysieren der neuen Antwort länger dauert als die Fehler selbstständig anzupassen. Daher ist es nicht allzu schlimm wenn ein Attribut irgendwo doppelt vorkommt oder eine Assoziation falsch beschriftet ist. Deshalb und aufgrund der Komplexität dieses Diagrammes erhalten sowohl ChatGPT als auch Le Chat die bewertung "Gut". Das Grundgerüst kann hier ohne Probleme übernommen werden und anschließend an die eigenen Bedürfnisse und Anforderungen angepasst

werden.

Anwendungsfalldiagramm

Erstell mir nun für die Anforderungsspezifikation ein Anwendungsfalldiagramm. Dieses soll alle Anwendungsfälle beinhalten und den Rollen zuweisen, die für die Applikation benötigt werden.

Prompt 4.9: Prompt Anwendungsfalldiagramm

Mit diesem Prompt wurde das Anwendungsfalldiagramm erstellt. Der Prompt ist sehr einfach gehalten und beschreibt lediglich kurz den Inhalt des Diagrammes.

ChatGPTs erste Ausgabe enthielt zwei Rollen, den "Spieler" und den "User". Der "User" hatte den Anwendungsfall "Spiel erstellen". Dadurch kam die Frage auf, ob man als Spielersteller besondere Rechte haben sollte. ChatGPT hat darauf hin die neue Rolle "Gastgeber" hinzugefügt. Fraglich ist allerdings die Einzeichnung in das PlantUML-Diagramm. Dem zufolge erbt der "Gastgeber" von "Spieler" und dieser von "User". Das ist allerdings nur Teilweise richtig. Der "Gastgeber" muss von "Spieler" erben, da auch der "Gastgeber" die zum Spielen benötigten Anwendungsfälle benötigt. Allerdings sollte "Spieler" nicht von "User" erben, da die Anwendungsfälle auf die der "User" zugreifen kann, während eines Spiels nicht benötigt werden. Des Weiteren hatte "Spieler" die Anwendungsfälle "Zug machen" und "Spielfigur bewegen". Dadurch war zunächst unklar welche Tätigkeiten "Zug machen" beinhaltet. Auf Nachfrage antwortete ChatGPT allerdings, dass zu "Zug machen" alle Tätigkeiten beinhaltet, die man während einem Zug durchführen muss (Würfeln, Auswählen und Bewegen der Spielfigur, Schlagen gegnerischer Figuren und Beenden des Zuges). Daher ist der Anwendungsfall "Spielfigur bewegen" unnötig. Auf einen Hinweis diesbezüglich entfernt ChatGPT den Anwendungsfall aus dem Diagramm. Zuletzt konnte der Anwendungsfall "Chat nutzen" mit der Rolle "User" durchgeführt werden. Dies entspricht allerdings nicht den Anforderungen, da nur während eines Spiels ein Chat vorhanden sein soll. Auch hier wird der Anwendungsfall nach einem Hinweis von ChatGPT wieder entfernt. Grundsätzlich wurde das Anwendungsfalldiagramm so passen, allerdings muss man sich die Frage stellen, ob man noch eine Rolle für Administrative Zwecke benötigt, welcher direkt Daten, wie zum Beispiel User, aus der Datenbank löschen und hinzufügen kann. Positiv herausheben sollte man, dass ChatGPT bei jeder Ausgabe automatisch Code für ein Diagramm bei PlantUML erzeugt hat.

[illegible]

Eine erste Ausgabe mit direkt vier Rollen erstellt dahingegen Le Chat. Hier wird eine eigene Rolle "Gast" erstellt, welche ein Benutzer erhält der noch nicht eingeloggt ist, und es wird direkt ein Rolle "Admin" mit angegeben. Ein unterschied zu [ChatGPT] ist hier, dass wenn ein Spieler am Zug ist, jede Aktion ein eigener Anwendungsfall ist. Dabei ist zu beachten, dass "Figur setzen" und "Figur schlagen" nicht getrennt werden dürfen, da diese Aktionen zusammengehören. Wenn eine Figur gesetzt wird muss

automatisch überprüft werden ob sich auf dem Feld bereits eine Figur befindet, wenn ja muss diese automatisch geschlagen werden. Außerdem sind die Anwendungsfälle "Benutzer verwalten" und "Profil bearbeiten" zu ungenau definiert. Hier muss man für jede Aktion einen eigenen Anwendungsfall erstellen. Diese beiden Punkte wurden Le Chat als Eingabe gegeben und anschließend anständig korrigiert. Auf Nachfrage wurde hier auch Code für ein PlantUML-Diagramm erzeugt. Dieses ist jedoch sehr unübersichtlich, da für jeden Anwendungsfall ein "Männchen" erstellt wird und diese einfach untereinander aufgelistet werden. Die Versuche, ihn davon wegzubringen und das Diagramm ähnlich wie ChatGPT zu erstellen waren alle vergebens.

Zusammenfassend lässt sich sagen, dass [ChatGPT und Le Chat] das Anwendungsfalldiagramm schon ganz gut beschreiben. Ein paar Ungenauigkeiten gibt es dennoch und aus diesem Grund werden sowohl ChatGPT als auch Le Chat mit "Gut" bewertet. Dass Le Chat das PlantUML-Diagramm nicht übersichtlich erstellen kann fließt hier nicht negativ in die Bewertung mit ein, da das Erstellen des Diagramms keine Anforderung war sondern es reicht wenn die Diagramme anständig beschrieben werden, so dass man diese selbst erstellen kann. Dies ist bei der Beschreibung von Le Chat gegeben.

Anwendungsfälle

Nun geht es um die Erstellung der detaillierten Beschreibung für die einzelnen Anwendungsfälle. Dazu wurden zwei Anwendungsfälle ausgesucht. Einmal "Anmelden" als leichteren Anwendungsfall und "Zug machen" [ChatGPT] bzw. "Figur setzen oder schlagen" [Le Chat] als komplexeren Anwendungsfall. Dazu wurde mit zwei Prompts gearbeitet. Für den Anwendungsfall "Anmelden" wurde folgende Eingabe benutzt:

Erstell mir nun für die Anforderungsspezifikation, basierend auf dem Anwendungsfalldiagramm, eine detaillierte Beschreibung für den Anwendungsfall „Anmelden“. Dabei soll zunächst der Anwendungsfall in einem Satz beschrieben werden, was dieser macht. Anschließend soll der Akteur, der Auslöser, die Vorbedingung und die Häufigkeit beschrieben werden. Danach soll ein Aktivitätsdiagramm erstellt werden. Dabei werden die Aktivitäten dem Client oder dem Server zugewiesen, wo diese bearbeitet werden. In dem Diagramm sollen keine Dialog beschreibung, wie Fenster mit Fehlermeldung anzeigen, beschrieben werden. Es sollen lediglich Dienste beschrieben werden, die der Server anbietet. Zuletzt werden die Aktivitäten aus dem Diagramm kurz beschrieben.

Prompt 4.10: Prompt AF Anmelden

Der Prompt beschreibt die Struktur des Dokuments mit den Inhalten. Außerdem wird der Aufbau des Aktivitätsdiagramms beschrieben und dass keine Dialog Beschreibungen in das Diagramm gehören. Einige Sachen werden hier direkt ganz gut erstellt. Auffällig ist, dass bei der "Häufigkeit" keine grobe Schätzung angegeben wird sondern Aussagen wie "Jedes Mal, wenn der Benutzer die Anwendung nutzen möchte.". Dies ist keine Sinnvolle Beschreibung der Häufigkeit.

Die erste Ausgabe wurde mit ChatGPT-4o erstellt. Dabei ist vor allem aufgefallen, dass die Beschreibung nicht zu dem PlantUML-Diagramm passt. Die Aktivitäten sind anders benannt und auch die Anzahl der Aktivitäten stimmt nicht überein. Außerdem wird sowohl in der Beschreibung als auch in dem Diagramm bereits viel zu genau beschrieben was passiert. Zum Beispiel die Aktivität "Session erstellen" ist für die Beschreibung eines Anwendungsfalls viel zu genau. Dies beschreibt bereits "Wie" der Anwendungsfall umgesetzt wird und gehört somit in das Architekturdokument. Aufgrund, dass 4o hier zuviel erstellt hat, wurde die Nachricht mit ChatGPT-3.5 nochmal generiert. Man sieht schonmal direkt, dass deutlich weniger Aktivitäten und keine Beschreibung der Implementierung erstellt wurden. Die Aktivitäten passen zum Anwendungsfall und beschreiben diesen gut. Wenn man jedoch die Diagramme von 4o und von 3.5 vergleicht, fallen einem einige Fehler auf, die beide machen. Zum einen wird die Aktivität, die überprüft wird, in den Entscheidungsknoten reingeschrieben. Diese muss jedoch als eigene Aktivität oben drüber stehen. Des weiteren sollten die Beschriftungen an den Pfeilen des Entscheidungsknoten in eckige Klammern. Auch das die Pfade, die aus dem Entscheidungsknoten entstehen, wieder zusammenführen, ist falsch. Bei dem Diagramm der 3.5er Version fehlt außerdem der Start- und Endknoten und auch bei der 4o Version sollten zwei Endknoten vorhanden sein. Einmal für "Anmeldung fehlgeschlagen" und einmal für "Anmeldung erfolgreich" [01&02 GPT Anmelden]. Wenn man die 3.5er Version auf diese Hinweise allgemein hinweist [siehe 2. Prompt 03 GPT Anmelden] führen zwar die beiden Pfade nach dem Entscheidungsknoten nicht mehr zusammen, jedoch endet der eine ohne Endknoten und im Entscheidungsknoten steht immernoch die Aktivität. Wenn man ihm genau sagt, was er mit welcher Aktivität machen soll, wird das Aktivitätsdiagramm komplett fehlerhaft mit unterbrochenen Verbindungen und einer alleinestehenden Aktivität erstellt.

Gibt man den groben Hinweis der 4o-Version [05 GPT Anmelden] setzt dieser die Vorgaben richtig um. Lediglich die Dialog Interaktion "Benutzer klickt auf "Anmelden"" ist noch vorhanden, welche jedoch nach einem Hinweis auch entfernt wird [06 GPT Anmelden]. Die letzten Kleinigkeiten die noch auffallen sind zum einen, dass im Entscheidungsknoten nur "ja" steht. Hier ist die Frage ob man in PlantUML einen leeren Entscheidungsknoten ohne Text erstellen kann. Ähnlich verhält es sich an den Endknoten. Diese können vermutlich auch nicht beschriftet werden, weshalb direkt davor noch Aktivitäten mit der Beschreibung erstellt werden.

Gemini []

Le Chat erstellt wieder nur auf Nachfrage ein PlantUML-Diagramm, welches vom Format her auch nicht ganz passend ist, da hier Client und Server übereinander stehen und nicht nebeneinander. Außerdem gibt es hier die selben Probleme wie bei ChatGPT, dass im Entscheidungsknoten die Aktivität steht und die Pfade danach wieder zusammengeführt werden. Die Textuelle Beschreibung von der ersten Ausgabe [01 Le Chat Anmelden] ist ganz gut und man könnte daraus, mit einbisschen mitdenken, ein richtiges Aktivitätsdiagramm erstellen. Bei der zweiten Ausgabe, die mit der selben Eingabe erstellt worden ist, fällt auf das trotz des Hinweises, dass keine Dialog Eigenschaften beschrieben werden sollen, die erste Aktivität "Anmeldeformular anzeigen" genau das tut.

Der Anwendungsfall “Zug machen” oder “Spielfigur setzen oder schlagen” wurde direkt nach dem

Anwendungsfall "Anmelden" erstellt. Dadurch wurde beim Prompt die Beschreibung, was erstellt wurde weggelassen. Stattdessen wurde beschrieben, dass wenn in einer Aktivität ein Wert gespeichert werden soll, auf den Wert im Fachliche Datenmodell referenziert werden muss. Außerdem wurde der Name des Anwendungsfalls bei den Tools entsprechend angepasst. Für ChatGPT sieht der Prompt zum Beispiel wie folgt aus:

Erstell mir nun eine detaillierte Beschreibung für den Anwendungsfall SZug machen". Wenn bei einer Aktivität ein Wert gespeichert wird, sollte in der Beschreibung der Aktivität beschrieben werden, um welchen Wert es sich aus dem Fachlichen Datenmodell handelt. Als Beispiel sollte der gewürfelte Wert in Spiel.wuerfel gespeichert werden.

Prompt 4.11: Prompt AF Zug machen

Die Ausgabe von ChatGPT scheint zunächst sehr überwältigend, doch beim näheren betrachten ist diese eigentlich schon ziemlich gut. Das Format des PlantUML-Diagrammes zeigt die selben Fehler auf wie bereits beim Anwendungsfall "Anmelden", lassen sich aber alle korrigieren. Eine Aktivität die etwas erschreckend klingt ist Aktivität 6 "Berechne möglichen Züge". Grundsätzlich ist dies nicht grade eine effiziente herangehensweise, was bei einem Mensch ärgere dich nicht Spiel allerdings nicht ganz so schlimm ist, da sich hier die möglichen Züge auf maximal vier beschränkt. Es fällt auch auf, dass das Würfeln auf Server-Seite stattfindet. Dies ist ansich kein gutes Spielerlebnis, wenn man als Spieler nicht würfelt. Das Würfeln sollte am besten ein eigener Anwendungsfall sein, da das Würfelergebnis wieder auf Client-Seite gebracht werden muss und dies innerhalb eines Anwendungsfalls schwierig zu implementieren ist. Auch die Verweise auf das Fachliche Datenmodell halten sich mit zwei Verweisen in grenzen und die Klasse "Spielzug" wird überhaupt nicht betrachtet. Als Hilfe wurde daher nochmal eine Eingabe, mit dem Fachliche Datenmodell [12 GPT FDM] und der Anweisung sich dieses zu merken, erstellt. Anschließend wurde der Prompt 4.11 erneut eingegeben. Diese Ausgabe schafft es etwas besser die Verweise auf das Fachliche Datenmodell zu beschreiben. Lediglich "Spiel.status" und "Feld.belegt" werden nicht erwähnt. Das Würfeln wird auf die "Client"-Seite geschoben. Jedoch ist dieser Ablauf, wie bereits beschrieben, schwierig zu Implementieren und man sollte Würfeln als eigenen Anwendungsfall beschreiben. Es entfällt auch die Aktivität die Überprüft, ob ein Spieler an der Reihe ist. Ebenfalls negativ ist, dass die "Client"-Seite im Aktivitätsdiagramm zu "Spieler" umbenannt wurde. Dies wird jedoch nach einem Hinweis wieder zurück geändert. Ein Fehler ist auch, dass die Position der Spielfigur zweimal gespeichert wird. Einmal in Aktivität 7 und einmal in Aktivität 9 [04 GPT Zug machen]. Was nach einem Hinweis aber auch korrigiert wird.

[illegible]

Bei Le Chat war zu Beginn die Beschreibung etwas kurz und es gab auch, außer "Spiel.wuerfel", was explizit im Prompt genannt wurde, keine Referenz aufs Fachliche Datenmodell. Erst nach der nächsten Eingabe [LeChat 02 ZugMachen] wurden die Beschreibungen etwas ausführlicher und es gab Referenzen aufs Fachliche Datenmodell. Diese Referenzen waren jedoch falsch und entsprachen nicht dem zuvor erstellten Fachlichen Datenmodell. Hier wurde nämlich häufig auf die Position der Figuren und

der Felder über “.x” und “.y” Werte zugegriffen. Grundsätzlich ist das natürlich eine gute Art und Weise, die Position leichter zu definieren. Jedoch wurde im Fachlichen Datenmodell über einfache Int-Werte gespeichert. Außerdem ist die Beschreibung der Aktivitäten nicht entsprechend der Regeln von Mensch ärgere dich nicht. In der Aktivität “Figur setzen oder schlagen” wird beschrieben, dass eine geschlagene Figur aus der Liste der Figuren des Spielers entfernt wird. Dies ist natürlich nicht richtig, sondern die Figur sollte einfach auf ein Startfeld der entsprechenden Farbe bzw. des entsprechenden Spielers gesetzt werden. Außerdem wurde Würfeln als eine Aktivität hier aufgeführt. Im Anwendungsfalldiagramm wurde allerdings für Würfeln ein eigener Anwendungsfall erstellt, weshalb die Aktivität hier unnötig und falsch ist. Nach einem Hinweis diesbezüglich, hat Le Chat ausgegeben, dass das stimmt und die Aktivität entfernt werden muss. Anschließend wurde, auf Nachfrage, eine Beschreibung des Aktivitätsdiagramms erstellt. Dabei wurde einmal die Aktivität “Figur setzen oder schlagen” in zwei Aktivitäten aufgeteilt. Einmal “Figur setzen” und “Figur schlagen” wobei mit einem Entscheidungsknoten entschieden wird, in welche Aktivität gegangen wird. Die Beschreibung der Flusslinien ist sehr sporadisch und sagt sogar eigentlich aus, dass die erst die Aktivitäten miteinander verbunden sind und diese dann erst mit den Entscheidungsknoten und den Endpunkten verbunden werden. Dies sollte aber nicht so sein, da die Entscheidungsknoten ja zwischen den Aktivitäten stehen müssen. Dazu wird beschrieben, dass es zwei Endpunkte gibt. Für den Fall, dass erfolgreich die Figur gesetzt oder geschlagen wird und das Verweigern des Zugs. Ebenfalls negativ auffällig ist, dass bei den ersten beiden Aktivitäten “Figur auswählen” und “Zielfeld auswählen”, jeweils die Information an den Server übermittelt wird. Besser und einfacher wäre es aber, die Infos zusammen in einem Zug an den Server zu senden. Damit würde es reichen eine Funktion zu erstellen und an den Server zu senden. Der “Trick”, das zuletzt erstellte Fachliche Datenmodell als Eingabe zu nutzen, hat nicht wie bei ChatGPT funktioniert. Wenn man Le Chat nach dem aktuellen Fachlichen Datenmodell fragt, auf welches er sich referenziert, wird auch ein völlig anderes Datenmodell ausgegeben [LeChat 05 ZugMachenFDM].

Benutzungsschnittstelle

Erstell mir nun für die Anforderungsspezifikation die Benutzungsschnittstelle. Diese zeigt, von welchem Dialog aus man zu anderen navigieren kann. Es sollen alle Dialoge enthalten sein, mit all deren Verbindungen zueinander.

Prompt 4.12: Prompt Benutzungsschnittstelle

Mit diesem Prompt wird die Benutzungsschnittstelle beschrieben die zeigt, wie man zwischen den Dialogen navigieren kann.

Bei ChatGPT wurde zwei mal die Eingabe getätigt.

Die erste Eingabe hat im Chat ein “gezeichnetes” Diagramm enthalten. Dort waren jedoch sowohl die Verbindungen zwischen Dialogen als auch die Einzeichnung der Navigation etwas schwierig, da hier nur

in eine Richtung navigiert wird. Außerdem passte dieses auch nicht richtig zu der textuellen Beschreibung die mit generiert wurde. Bei der Beschreibung ist aufgefallen, dass ein "Passwort-zurücksetzen-Dialog" enthalten ist. Allerdings ist in dem Anwendungsfalldiagramm kein Anwendungsfall für das Passwort zurücksetzen erstellt worden, weshalb man entweder den Dialog aus dem Diagramm löscht oder den Anwendungsfall "Passwort zurücksetzen" nachträglich erstellen muss. Ähnlich verhält es sich mit den Anwendungsfällen "Punkte anzeigen" und dem Dialog "Anzeige der Spieler-Liste". Außerdem ist fraglich ob ein noch nicht angemeldeter Nutzer die Buttons "Spiel erstellen" und "Spiel beitreten" sehen soll. Hier wäre es angebrachter einen eigenen Dialog für nicht-angemeldete Nutzer zu erstellen und zusätzlich ein Hauptmenü.

Nach der zweiten Eingabe erhielt man ebenfalls eine textuelle Beschreibung von welchem Dialog man wohin navigieren kann, einen PlantUML-Code für ein UML-Diagramm sowie eine Beschreibung der Dialoge. Das PlantUML-Diagramm hatte zunächst einen Syntax Fehler, welcher jedoch nach einem Hinweis korrigiert wurde. Jedoch ist das Diagramm recht unübersichtlich, da zum einen die Pfeile nur in eine Richtung zeigen und zum anderen die Struktur etwas durcheinander ist. Die Beschreibung ist dafür abgesehen von einem Punkt sehr gut. Wenn man einem Spiel beigetreten oder ein Spiel gestartet hat gelangt man zu dem Dialog "Spielbildschirm". Wenn man dort direkt das Spielfeld sieht, was bei einem Hauptbildschirm während des Spiels wohl der Fall ist, muss man den Anwendungsfall "Spiel anzeigen" aus dem Anwendungsfalldiagramm wieder entfernen.

Gemini[[[[[[]]]]]]

Le Chat hat bei der Erstellung der Benutzungsschnittstelle probiert, diese in die Ausgabe mit zu "Zeichnen". Die Abbildung war jedoch völlig durcheinander und unübersichtlich. Die Verbindung zwischen den Dialogen ist unverständlich und ergeben keinen Sinn. Auch mit Eingaben, dass die Verbindung doch überarbeitet werden soll, wurde dies nicht besser. Dafür kann man sagen, dass die erstellten Dialoge soweit zu den Anwendungsfällen passen und die meisten enthalten sind. Jedoch sind auch einige Dialoge erstellt worden, für die man eigentlich keinen eigenen Dialog erstellen muss. Zum Beispiel "Spiel verlassen". Da würde ein Button reichen, mit dem man zurück zum Hauptmenü gelangt. Auch die Textuelle Beschreibung ist nicht vollständig. Bei der Beschreibung des Verlaufes werden Anwendungsfälle, wie zum Beispiel "Registrieren" vergessen. Außerdem stellt sich die Frage, was man beim "Spiel beitreten" Dialog sieht. In dem Anwendungsfalldiagramm wurden die Anwendungsfälle "Spiel suchen" und "Spiel beitreten" erstellt. Daher ist etwas ungenau ob die Anwendungsfälle über den selben Dialog gehen oder ob einer der Anwendungsfälle unnötig ist. Dies sollte bei einer näheren Betrachtung der Anwendungsfälle festgelegt werden. Bei der Beschreibung fehlt auch, wie der Admin zu seinen Verfügbaren Dialogen navigieren kann.

Dialog

Für die Erstellung der Dialoge wurden ebenfalls zwei recht einfache Prompts verwendet, die lediglich beschreiben, wie das Dokument aufgebaut sein soll. Einmal für den Anwendungsfall "Anmelden" und einmal für den Dialog "Spielbildschirm":

Erstell mir nun für die Anforderungsspezifikation einen passenden Dialog für den Anwendungsfall „Anmelden“. Dabei wird zunächst beschrieben, wann der Dialog geöffnet wird und wozu er dient. Anschließend folgt eine Skizze des Dialoges. Zuletzt werden die einzelnen Elemente des Dialogs beschrieben und definiert, was dort eingetragen wird bzw. was beim drücken passiert.

Prompt 4.13: Prompt Dialog Anmelden

ChatGPT hat in der Ausgabe wieder eine „gezeichnete“ Skizze erstellt. Bei der ersten Ausgabe für den „Anmelden“ Anwendungsfall wurde noch ein Button für „Passwort vergessen?“ erstellt. Ansich macht dies natürlich Sinn hier einzufügen, allerdings wurde im Anwendungsfalldiagramm kein Anwendungsfall dafür erstellt. Positiv auffällig ist, dass die erste Ausgabe [GPT 01 Anmelden] in dem Chat erstellt wurde, wo GPT-3.5 den Anwendungsfall „Anmelden“ erstellt hat und dort festgelegt hat, dass sich der User mit einem Benutzernamen und einem Passwort anmeldet [GPT 01 Anmelden] und die zweite Ausgabe [GPT 02 Anmelden] in dem Chat erstellt wurde, wo GPT-4o sich vorher beim Anwendungsfall für eine Email und ein Passwort entschieden hat. Bemängeln kann man dabei, dass bei der Beschreibung der Elemente nicht auf das Fachliche Datenmodell verwiesen wurde. Nach einer groben Eingabe, dass doch auf die Werte aus dem Fachlichen Datenmodell referenziert werden soll, wurde dies dann auch richtig gemacht. [GPT 03 Anmelden].

Bei der Erstellung des Dialogs „Spielbildschirm“ wurde auch wieder eine textuelle Beschreibung und eine Skizze erstellt. Die Skizze umfasst allerdings nur einen Vermerk in eckigen Klammern, wo das Spielbrett graphisch dargestellt werden soll. Außerdem wurde auch hier nochmal die Eingabe gemacht, dass auf das Fachliche Datenmodell referenziert werden soll. Interessant ist, dass ein extra Button fürs Würfeln erstellt wurde und einer um einen Zug zumachen. Bei dieser Implementierung ist es nahelegender, dass für beide Button ein eigener Anwendungsfall gestartet wird. Bei Nachfrage, zu beschreiben, wenn ein Anwendungsfall gestartet wird, gibt ChatGPT aus, dass Würfel ein eigener Anwendungsfall ist. Die Restlichen Verweise auf das Fachliche Datenmodell und beschreibung welcher Anwendungsfall gestartet wird passen soweit.

- Gemini:

Beim „Anmelden“-Dialog hat Le Chat wieder nur eine Textuelle Beschreibung erstellt. Diese ist soweit passend. Fürs Anmelden hat sich Le Chat auf die Email-Adresse und ein Passwort festgelegt. Im Fachlichen Datenmodell wurde allerdings nur gesagt, dass die ID eindeutig ist, das heißt das entweder die Email-Adresse ebenfalls eindeutig gemacht werden muss oder sich diese nicht eignet fürs Anmelden. Ansonsten könnten zwei Benutzer mit derselben Email-Adresse und dem selben Passwort nicht unterschieden werden. Außerdem fehlt ein Button um zum „Registrieren“-Dialog zukommen, welcher zumindest laut dem Anwendungsfalldiagramm vorhanden sein sollte. Passen würde es jedoch zur Benutzungsschnittstelle, da hier kein Dialog fürs Registrieren eingebunden wurde.

Bei der Beschreibung des Dialogs „Spielbrett“ werden die Inhalte gut beschrieben. Es fehlt lediglich ein Button oder ähnliches um den Spielstand anzuzeigen. Dies ist laut dem Anwendungsfalldiagramm ein eigener Anwendungsfall, wobei jedoch die Frage ist, was dieser beschreiben soll oder ob dieser durch

das dauerhafte anzeigen des Spielbrettes hinfällig geworden ist. Bei Nachfrage, ob Le Chat beschreiben kann wann ein Anwendungsfall gestartet wird, gibt dieser eine zu dem Dialog passende Beschreibung aus. Auch hier fehlen allerdings die Anwendungsfälle "Spielstand anzeigen" und "Spielergebnis anzeigen".

Nichtfunktionale Anforderungen

Erstell mir nun für die Anforderungsspezifikation die nichtfunktionalen Anforderungen und jeweils einen dazu gehörigen Test.

Prompt 4.14: Prompt nichtfunktionale Anforderungen

Dieser sehr einfach gehaltene Prompt wurde für die Erstellung der nichtfunktionalen Anforderungen verwendet. Es musste lediglich dazu erwähnt werden, dass zu jeder Anforderung ein Test erstellt werden soll. Bei [ChatGPT und Le Chat] werden gute Anforderungen erstellt und die Tests passen auch sehr gut zu diesen. Hier werden die [beiden] Tools mit "Sehr gut" bewertet.

- Gemini:

4.6 Architekturdokument

Da auch das Architekturdokument ein umfangreiches Dokument ist, werden hier ebenfalls die Ausgaben der Technische Infrastruktur, des Komponentenmodell und der dynamischen Beschreibung exemplarischer Anwendungsfälle im Komponentenmodell jeweils einzeln beschrieben und bewertet.

Technische Infrastruktur

Da die Technische Infrastruktur auf die Systemübersicht aufbaut und alle drei Tools Schwierigkeiten hatten sich an dieses zu erinnern, wurde die Textuelle Beschreibungen der jeweiligen Tools ihnen nochmal als Eingabe mitgegeben. Dazu beschreibt der Prompt, was die Technische Infrastruktur ist und was dort beschrieben ist. So kommt folgender Prompt zustande:

Erstell mir nun, basierend auf folgender Systemarchitektur, ein Diagramm und/oder eine Textuelle Beschreibung der Technischen Infrastruktur. Diese Listet die Hardware- und Produktanforderungen auf: Hardware, Betriebssysteme, Produkte, Nachbarsysteme. Dabei sollen die einzelnen Betriebssysteme mit Versionsnummer und die Verbindungen zwischen den Komponenten mit der Verbindungsart beschrieben werden: [Textuelle Beschreibung der Systemübersicht des jeweiligen Tools]

Prompt 4.15: Prompt Technische Infrastruktur

ChatGPT erstellt hier wieder eine Skizze in die Ausgabe mit. Diese ist auch ganz gut, lediglich die Datenbank ist etwas verrutscht. Negativ aufgefallen ist die Beschreibung der Verbindung von der KI-Komponente. Hier wird beschrieben, dass diese in die Spiel-Logik auf der Server-Seite integriert ist und daher keine externe Verbindung erforderlich ist. In der Beschreibung der Systemübersicht wird die Spiel-Logik allerdings nur auf der Client-Seite erstellt. Den Anwendungsfällen entsprechend ist die Spiel-Logik auf der Server-Seite allerdings korrekt. Fragt man ChatGPT danach, wird ausgegeben, dass die Spiel-Logik beim Server hinzugefügt werden soll. Ansonsten wurde die Technische Infrastruktur von ChatGPT gut erstellt.

Gemini [[[]]]

Auch Le Chat versucht eine Skizze in die Ausgabe mit einzufügen. Dies klappt hier allerdings nicht so richtig, da die Komponenten mit den Verbindungen nicht richtig verbunden werden. Während ChatGPT die Datenbank extern, als eigene Komponente definiert, ist bei Le Chat diese Teil des Servers. Außerdem werden bei den Nachbarsystemen ein Externes Authentifizierungssystem sowie ein Zahlungssystem erstellt, was allerdings nicht Teil der Anforderungen war. Daher sind diese hier falsch. Jedoch passt die Zuweisung der Versionen für die Betriebssysteme sowie der Produkte ganz gut.

Komponentenmodell

- 1) Erstell mir nun für das Architekturdokument das Komponentenmodell. Das Komponentenmodell beschreibt die fachliche Struktur der Lösung auf einer groben Ebene mittels eines UML-Komponenten-Diagramms.
- 2) Beim Komponentenmodell handelt es sich nicht um die Technische Infrastruktur. Hier sollen die Komponenten und deren Schnittstellen erstellt werden, die später für die Programmierung verwendet werden.

Prompt 4.16: Prompt Komponentenmodell

Mithilfe des ersten Promptes wurde versucht das Komponentenmodell zu erstellen. Dies hat jedoch bei [ChatGPT und Le Chat] eher weniger funktioniert. Beide Tools haben sich zu sehr an die Technische Infrastruktur gehalten und sind von diesem Format nicht weggekommen. Dadurch wurde bei [beiden] Tools mit dem zweiten Prompt versucht diese davon etwas wegzubringen.

Bei ChatGPT hat dies dann auch schon ganz gut funktioniert, allerdings wird sich trotzdem noch stark an der Technischen Infrastruktur orientiert. Außerdem sind keine Schnittstellen zwischen den Komponenten beschrieben worden. Es fällt auch auf, dass "Game-Logic", welche die Spielregeln implementiert und lokale Spielzüge validiert, nur auf Client-Seite ist, was nicht den Anwendungsfällen entspricht, wo festgelegt wurde dass die Validierung der Spielzüge auf Server-Seite geschieht. Nach einer Eingabe, wo

ChatGPT dies erklärt wird, wird "Game-Logic" auf Server-Seite geschoben. Außerdem werden die für die RMI nötigen Komponenten "Client Communication" und "Server Communication" hinzugefügt, welche vorher durch die "Communication"-Komponente zwischen Client und Server falsch dargestellt wurde. Da nun allerdings auf der Server-Seite "Game Logic" und "Game Manager" vorhanden waren, wurde ChatGPT gefragt ob man diese nicht zusammenfassen könnte. Dies wurde bejaht und die Komponenten wurde zu "Game Manager/Logic" zusammengefasst. Da der vom Kunden angeforderte Chat nirgends erwähnt wurde, wurde ChatGPT diesbezüglich nochmal gefragt ob dieser in einer Komponente enthalten sei oder eine eigene Komponente für diesen benötigt wird. Auch hier wurde dann die Komponente "Chat Manager" richtig erstellt und in das Diagramm eingefügt. Zusätzlich könnte man noch eine "Authentication"-Komponente in das Diagramm mit aufnehmen, welche sich um die Anmeldungen und Registrierungen und ähnliche Anwendungsfälle kümmert. Da der Chat ohne Probleme in das Diagramm mit aufgenommen wurde, wird das mit einer "Authentication"-Komponente ähnlich sein, weshalb dies nicht als extra Eingabe getätigt wurde. Zum Schluss wurde ChatGPT nochmal nach den Schnittstellen zwischen den Komponenten gefragt. Hier versucht ChatGPT allerdings die Funktionen, die dort benötigt werden, zu erstellen. Diese sind zum einen nicht vollständig und zum anderen an dieser Stelle falsch. Auch mit weiteren Eingaben lassen sich die Schnittstellen nicht richtig darstellen.

Gemini[[[]]]

Wie zu Beginn bereits erwähnt hatte Le Chat dasselbe Problem wie ChatGPT, dass dieser sich zu nah an der Technischen Infrastruktur orientiert hat. Allerdings hat man Le Chat auch nach mehreren Eingaben nicht davon weggebracht. Die Beschreibung der Komponenten war eher bedürftig und es wurden keine Komponenten für die RMI Verbindung erstellt. Erst nach einer expliziten Eingabe, dass eine "ClientCommunication"- und eine "ServerCommunication"-Komponente erforderlich ist, wurden diese erstellt. Das daraufhin erstellte Diagramm war allerdings eher wieder in der Form der Technischen Infrastruktur. Die Komponenten besaßen Interfaces, welche allerdings nicht richtig gepasst haben. Erst nach einer expliziten Eingabe, wie das Komponentendiagramm auszusehen hat, wurde dieses Anständig erstellt. Allerdings haben auch hier die Schnittstellen gefehlt. Auf Nachfrage diesbezüglich wurden diese sogar einigermaßen ok erstellt. Allerdings werden auch hier schon die Methoden der Schnittstellen aufgelistet. Auch hier sind diese allerdings an der falschen Stelle erstellt worden und sind ebenfalls nicht vollständig. Dazu wurde beschrieben, dass die "Data"-Komponente auf andere Schnittstellen zugreift. Dies ist allerdings falsch, da die Datenbank ja keine Funktionen von alleine aufruft, sondern lediglich einen Return liefern soll. Nach einem Hinweis korrigiert Le Chat dies allerdings auch. Ein ähnliches Verhalten sieht man bei den anderen Komponenten auch, dass die Komponenten auf der Server-Seite in Richtung Client-Seite Funktionen schicken können. Hier ist die Frage wie zum Beispiel eine Spielfeld aktualisierung umgesetzt werden soll. Wenn der Server die aktuellen Positionen der Spielfiguren von alleine an den Client schicken soll, müssen die Komponenten auf der Server-Seite Richtung Client-Seite kommunizieren können. Wenn man allerdings eine Funktion wie "Spielfeld aktualisieren()" auf der Client-Seite hinzufügt, könnte man die Daten über einen Return-Wert zurück schicken. In diesem Fall müssten die Komponenten auf der Server-Seite nicht in Richtung Client-Seite kommunizieren können. Der letzte Auffällige Punkt ist, dass die Komponente "ClientLogic" die Schnittstelle "ClientLogicInterface" besitzt, auf

welche von der Komponente "Client-Anwendung" zugegriffen wird. Nach dem Komponentendiagramm ist die "ClientLogic"-Komponente allerdings die letzte auf der Client-Seite. Hier ist vermutlich die Anweisung, "Als erstes gibt es eine ClientLogic-Komponente welche mit einer ClientCommunication-Komponente verbunden ist." ungeeignet gewesen, da hier interpretiert wird, dass die "ClientLogic" nur die Logik auf der Client-Seite umsetzt, nicht aber die Visuelle Darstellung. Daher hätte man die Komponente in der Eingabe besser "View" oder ähnlich nennen sollen.

5 Herausforderungen und Limitationen

TODO!!!!!!!

5.1 Technische Herausforderungen

TODO!!!!!!!!!!!!!!

5.2 Limitationen von LLM Tools im Software Engineering Kontext

6 Zusammenfassung und Ausblick

TODO!!!!!!!!!!

6.1 Zusammenfassung der Ergebnisse

TODO!!!!!!!!!!!!

6.2 Implikationen für die Praxis

Alles Produkte von Unternehmen, welche Gewinnorientiert sind. Daher Diebstahl von Infos und Herangehensweisen der Projekte wenn man diese mit dem Tool zusammenfassen lässt. Grade im militärischen Bereich.

Lokale LLM Tools

Immer unterschiedliche Ausgaben trotz gleicher Eingabe. Erfahrung hat gezeigt, das falls mal nur mist rauskommt, einfach selbe eingabe immer wieder machen, bis was anständiges rauskommt.

Gemini merkt sich infos aus dem gleichen Chat nicht. Wenn man am nächsten Tag weiter arbeitet, kann Gemini nicht mehr auf die Infos vorher zugreifen.

Le Chat sagt, dass es ein Bild zu den Diagrammen erstellen kann. Jedoch sieht man zunächst den Link nicht und wenn man ihn danach fragt liefert er ein Link zu imgur.com, wo error 404 angezeigt wird.

Le Chat mit der Zeit auch Probleme die eigentlichen Anforderungen umzusetzen. Erstellt immer eigene neue Anforderungen.

Ich hätte vor der Untersuchung festlegen sollen, wieviele eingaben gemacht werden soll. So ist es bei jedem Tool etwas unterschiedlich was dazu führt dass das Endergebnis, bei Tools mit mehr Eingaben, vielleicht etwas besser ist als bei Tools mit weniger eingaben.

TODO!!!!!!!!!!!!

6.3 Ausblick und zukünftige Entwicklungen

TODO!!!!!!!!!!!!

Anhang

****Kunde (Herr Müller):**** Guten Tag, ich bin Herr Müller. Wie geht es ihnen?

****Entwickler (Frau Schmidt):**** Hallo, ich bin Frau Schmidt. Mir geht es gut. Wie geht es ihnen?

****Kunde (Herr Müller):**** Mir geht es auch gut. Ich möchte ein digitales "Mensch ärgere dich nichtSpiel entwickeln lassen und brauche Ihre Expertise dafür.

****Entwickler (Frau Schmidt):**** Es freut mich, dass Sie sich für unsere Dienstleistungen interessieren. Können Sie mir bitte mehr über Ihr Projekt erzählen? Welche Funktionen und Eigenschaften sollte das Spiel haben?

****Kunde:**** Natürlich. Zunächst möchte ich, dass das Spiel als JavaFX Anwendung auf dem Computer läuft. Die Anwendung soll eine Server-Client-Applikation, die mit Hilfe von Java realisiert wird, sein. Server und Client sollen über RMI kommunizieren. Es sollte das klassische "Mensch ärgere dich nichtSpielbrett und die entsprechenden Regeln enthalten. Außerdem soll es auch möglich sein, mehrere Spiele gleichzeitig zu spielen.

****Entwickler:**** Verstehe. Sollen die Spieler gegen den Computer, gegen andere Spieler online oder beides spielen können?

****Kunde:**** Beides wäre ideal. Es sollte einen Einzelspielermodus gegen KI-Gegner sowie einen Mehrspielermodus geben, in dem man online gegen Freunde oder zufällige Gegner spielen kann.

****Entwickler:**** Gut, dann werden wir einen Multiplayermodus integrieren. Wie viele Spieler sollen maximal teilnehmen können?

****Kunde:**** Wie im klassischen Spiel sollten es bis zu vier Spieler sein.

****Entwickler:**** Perfekt. Haben Sie Vorstellungen bezüglich des Designs? Soll das Spiel ein klassisches, traditionelles Aussehen haben oder möchten Sie ein modernes Design?

****Kunde:**** Ein klassisches Design mit einer ansprechenden, modernen Benutzeroberfläche wäre gut. Es sollte an das traditionelle Spiel erinnern, aber gleichzeitig frisch und ansprechend wirken.

****Entwickler:**** Verstanden. Gibt es besondere Funktionen oder Features, die Ihnen wichtig sind? Zum Beispiel Chat-Funktionalitäten, Punktesysteme oder spezielle Animationen?

****Kunde:**** Ein Chat während des Spiels wäre nett. Ein Punktesystem könnte die Motivation erhöhen, also wäre das auch eine gute Idee. Animationen sollten einfach, aber ansprechend sein - nichts Übertriebenes.

****Entwickler:**** Das klingt alles machbar. Wie sieht es mit dem Zeitrahmen aus? Haben Sie eine bestimmte Deadline, bis wann das Spiel fertig sein soll?

****Kunde:**** Ich denke, ein Zeitraum von sechs Monaten sollte machbar sein. Glauben Sie, dass das realistisch ist? Darf ich außerdem Fragen wie ihr Team aufgestellt ist?

****Entwickler:**** Ja, sechs Monate sind ein realistischer Zeitraum für dieses Projekt, vorausgesetzt, wir haben regelmäßige Feedback-Schleifen und Sie können uns zeitnah Ihre Anforderungen und Rückmeldungen geben. Wir sind ein Team aus 4 Personen. Ich bin die Projektleiterin, der Herr Schneider ist unser Anforderungsanalyst und fachlicher Chefdesigner, die Frau Fischer ist unsere Systemarchitektin und der Herr Becker ist der Testmanager.

****Kunde:**** Das klingt gut. Was wären die nächsten Schritte?

****Entwickler:**** Wir würden zunächst einen detaillierten Projektplan erstellen, der alle Funktionen und Meilensteine festlegt. Danach beginnen wir mit der Entwicklung des Prototyps. Während der gesamten Entwicklung würden wir regelmäßige Meetings abhalten, um den Fortschritt zu besprechen und eventuelle Anpassungen vorzunehmen.

****Kunde:**** Das klingt nach einem guten Plan. Wie sind sie eigentlich dazu gekommen, Entwickler zu werden?

****Entwickler:**** Ich habe mich schon als Kind für Computerspiele interessiert und auch wie diese im Hintergrund funktionieren. Daher habe ich recht früh angefangen das Programmieren zu lernen. Das hat mir immer viel Spaß gemacht und so habe ich nach der Schule eine Ausbildung in der Branche angefangen.

****Kunde:**** Ah ich verstehe. Sie sind jetzt mittlerweile Projektleiterin oder?

****Entwickler:**** Ja, genau.

****Kunde:**** Fehlt ihnen die Arbeit als Programmiererin? Als Projektleitern machen Sie ja wahrscheinlich nicht mehr ganz so viel in dem Bereich oder?

****Entwickler:**** Ne, das stimmt. Aber man ist ja trotzdem noch damit in Berührung. Wenn mal irgendwo Schwierigkeiten bei der Implementierung auftreten, kann ich auch helfen diese zu lösen. Ein wenig fehlen tut es mir schon, aber meine jetzige Tätigkeit als Projektleiterin macht mir auch sehr viel Spaß.

****Kunde:**** Alles klar. Ich freue mich darauf, mit Ihnen zusammenzuarbeiten.

****Entwickler:**** Vielen Dank, Herr Müller. Wir freuen uns auch auf die Zusammenarbeit. Wir werden Ihnen in Kürze einen detaillierten Projektplan und ein Angebot zukommen lassen.

****Kunde:**** Perfekt, ich warte auf Ihre Nachricht. Einen schönen Tag noch!

****Entwickler:**** Ihnen auch, Herr Müller. Bis bald!

Prompt 6.2: Kundengespräch

Abbildungsverzeichnis

2.1	Dokumente im Software Engineering	10
3.1	Mögliche LLM Tool Unterstützung	18

Tabellenverzeichnis

Quellcodeverzeichnis

Stichwortverzeichnis

Anforderungsspezifikation, 13, 19, 30	Zusammenfassung der Ergebnisse, 49
Anwendung von LLM Tools im Software Engineering, 17	Zusammenfassung und Ausblick, 49
Architekturdokument, 15, 21, 42	
Ausblick und zukünftige Entwicklungen, 49	
Besprechungsprotokoll, 11, 17, 24	
Bewertung, 23	
ChatGPT, 4	
Einleitung, 1	
Google Gemini, 6	
Grundlagen, 3	
Herausforderungen und Limitationen, 47	
Implikationen für die Praxis, 49	
Le Chat, 8	
Limitationen von LLM Tools im Software Engineering Kontext, 47	
LLM Tools, 3	
Motivation, 1	
Praxisergebnisse und Vergleich, 23	
Projekthandbuch, 11, 17, 25	
Risikoliste, 12, 19, 27	
Struktur der Arbeit, 1	
Technische Herausforderungen, 47	
Testspezifikation, 16, 21	
Zielsetzung der Arbeit, 1	

Literaturverzeichnis

- [1] Milfad Sadar. *Large Language Models - Grundlagen KI-getriebener Kommunikation*. Zugriff am 27. Mai 2024. 2023. URL: <https://weissenberg-group.de/was-ist-ein-large-language-model/#:~:text=Zum%20IPA%20Whitepaper-,Wie%20funktioniert%20ein%20LLM%3F,zwischen%20den%20Token%20zu%20ermitteln> (siehe Seite 4).
- [2] Jim Greene. *ChatGPT(software)*. Zugriff am 28. Mai 2024. 2024. URL: <https://eds.p.ebscohost.com/eds/detail/detail?vid=3&sid=1983b266-2f53-4090-a0f2-1aef00f51d1f%40redis&bdata=Jmxhbm9ZGUmc2l0ZT1lZHMtbGl2ZQ%3d%3d#AN=163030996&db=ers> (siehe Seite 6).
- [3] Dipl.-Ing. (FH) Stefan Luber. *Was ist Google Gemini?* Zugriff am 29. Mai 2024. 2024. URL: <https://www.bigdata-insider.de/was-ist-google-gemini-a-f855ef105c02f4533c18e913d4643007/> (siehe Seite 6).
- [4] Matthias Kindt.
Mistral AI: Französisches KI-Startup stellt Chatbot "Le Chat" und Kooperation mit Microsoft Azure vor. Zugriff am 16. Mai 2024. 2024. URL: <https://www.unidigital.news/mistral-ai-franzoesisches-ki-startup-stellt-chatbot-le-chat-und-kooperation-mit-microsoft-azure-vor/> (siehe Seite 8).
- [5] Ian Sommerville. *Software Engineering*. 9. Auflage. Harlow, England: Addison-Wesley, 2010. ISBN: 978-0-13-703515-1 (siehe Seite 11).

