

BISE Student

<https://bise-student.io>

MASTER'S THESIS

Nutzung der Leistungsfähigkeit generativer künstlicher Intelligenz zur Steigerung der Qualität und Produktivität im Use Case des Requirements Engineering

Publication Date: 2024-11-27

Author

Oliver OLIVER GATNAR

University of Hamburg

Hamburg, Germany

oliver.gatnar@gmail.com

0xadd10AFB85574351737CAeBa1Aeb226D80C99481

Abstract

This study examines the role of generative artificial intelligence (GenAI) in the field of requirements engineering (RE). The aim is to increase quality and productivity in the RE process and provide a solution to existing RE challenges. An artifact for generating user stories is developed and evaluated, based on the design science research (DSR) methodology. The methodology includes a literature review to provide an adequate problem definition. Furthermore, eleven interviews with experts from fields such as RE and GenAI, as well as an additional literature search, are conducted to develop design principles for the resulting artifact. Finally, the resulting artifact is evaluated with ten RE experts. The results show that compared to manual methods the use of GenAI in the RE process increases productivity by enabling a more efficient creation of user stories. Additionally, the evaluation results suggest that a hybrid approach can improve the quality of the generated user stories. Furthermore, it is demonstrated that...

Keywords: Generative Artificial Intelligence

Methods: Design Science Research

Submission Date: 2024-11-04

Submission Contract: 0xe5dE9B3bA79F2e109F21A23eA96De8B0A53Ee8f0



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

MASTERARBEIT

Nutzung der Leistungsfähigkeit generativer künstlicher Intelligenz zur Steigerung der Qualität und Produktivität im Use Case des Requirements Engineering

vorgelegt von
Oliver Gatnar

Fakultät:	Wirtschaftsinformatik, Sozio-Technische Systemgestaltung (WISTS)
Studiengang:	IT- Management und -Consulting
Matrikelnummer:	7029695
Erstgutachterin:	Prof. Dr. Eva Bittner
Zweitgutachterin:	Izabel Cvetkovic, M. Sc.
Betreuer Universität Hamburg:	Lucas Memmert, M. Sc.
Betreuer adesso SE:	Julius Haas, M. Sc.

vorgelegt am:
Hamburg, den 30. Mai 2024

Abstract

This study examines the role of generative artificial intelligence (GenAI) in the field of requirements engineering (RE). The aim is to increase quality and productivity in the RE process and provide a solution to existing RE challenges. An artifact for generating user stories is developed and evaluated, based on the design science research (DSR) methodology. The methodology includes a literature review to provide an adequate problem definition. Furthermore, eleven interviews with experts from fields such as RE and GenAI, as well as an additional literature search, are conducted to develop design principles for the resulting artifact. Finally, the resulting artifact is evaluated with ten RE experts. The results show that compared to manual methods the use of GenAI in the RE process increases productivity by enabling a more efficient creation of user stories. Additionally, the evaluation results suggest that a hybrid approach can improve the quality of the generated user stories. Furthermore, it is demonstrated that GenAI can be employed to address existing RE challenges. This research contributes to the literature by investigating the integration of GenAI in the RE process and by highlighting practical implications. It provides insights for the further development of GenAI applications in RE and emphasizes the importance of balancing the use of human and artificial intelligence.

Keywords: Generative Artificial Intelligence, Requirements Engineering, User Stories, Hybrid Intelligence, Design Science Research

Zusammenfassung

Diese Forschungsarbeit untersucht die Rolle generativer künstlicher Intelligenz (GenAI) im Requirements Engineering (RE). Ziel ist es, die Qualität und Produktivität im RE-Prozess zu steigern und eine Möglichkeit zur Lösung bestehender RE-Herausforderungen bereitzustellen. Basierend auf der Design Science Research (DSR) Methodik wird ein Artefakt zur Generierung von User Stories entwickelt und evaluiert. Die Methodik umfasst ein Literatur-Review zur Aufstellung der Problemstellung. Darüber hinaus werden elf Experteninterviews im Bereich RE und GenAI sowie eine zusätzliche Literaturrecherche zur Entwicklung von Design Prinzipien für das Ergebnisartefakt durchgeführt. Abschließend wird das Ergebnisartefakt mit zehn Experten aus dem Bereich des RE evaluiert. Die Ergebnisse zeigen, dass der Einsatz von GenAI im RE-Prozess die Produktivität erhöht, indem eine effizientere Erstellung von User Stories im Vergleich zu manuellen Methoden ermöglicht wird. Zusätzlich deuten die Evaluationsergebnisse darauf hin, dass ein hybrider Ansatz die Qualität der generierten User Stories verbessern kann. Darüber hinaus wird aufgezeigt, dass GenAI zur Lösung bestehender RE-Herausforderungen eingesetzt werden kann. Diese Forschungsarbeit trägt zur Literatur bei, indem sie die Integration von GenAI im RE untersucht und praktische Implikationen aufzeigt. Es werden Erkenntnisse für die Weiterentwicklung von GenAI-Anwendungen im RE geboten und die Bedeutung einer ausgewogenen Nutzung von menschlicher und künstlicher Intelligenz betont.

Stichworte: Generative Künstliche Intelligenz, Requirements Engineering, User Stories, Hybride Intelligenz, Design Science Research

Genderhinweis

In dieser Arbeit wird aus Gründen der besseren Lesbarkeit das generische Maskulinum verwendet. Weibliche und anderweitige Geschlechteridentitäten werden dabei ausdrücklich mitgemeint, soweit es für die Aussage erforderlich ist.

Inhaltsverzeichnis

Abstract	I
Zusammenfassung	II
Abbildungs- und Tabellenverzeichnis	V
Abkürzungsverzeichnis	VII
1 Einleitung	1
1.1 Problemdefinition und Motivation	1
1.2 Zielsetzung der Forschungsarbeit	2
1.3 Forschungsfragen	3
1.4 Struktur der Forschungsarbeit	4
2 Grundlagen	5
2.1 Requirements Engineering	5
2.2 Generative Künstliche Intelligenz	6
2.2.1 Künstliche Intelligenz	6
2.2.2 Prompt Engineering	7
2.2.3 Hybride Intelligenz	8
2.2.4 Einsatzbereiche und Potenziale im Kontext hybrider Intelligenz	8
2.2.5 Risiken und Limitationen	9
2.2.6 Abgrenzung	10
2.3 Generative Künstliche Intelligenz im Requirements Engineering	10
3 Methodik	12
3.1 Literaturrecherche	13
3.2 Experteninterviews - Allgemein	16
3.2.1 Experteninterviews - RE	17
3.2.2 Experteninterviews - GenAI	20
3.3 Entwicklung der Lösung	21
3.4 Evaluation der Webanwendung	21
4 Beschreibung des Ergebnisartefakts	25
4.1 Problembeschreibung	25
4.1.1 RE-Herausforderungen - Ergebnisse Literaturrecherche	25
4.1.2 RE-Herausforderungen - Ergebnisse RE-Experteninterviews	29
4.1.3 RE-Qualitätskriterien - Ergebnisse Literaturrecherche	31
4.1.4 RE-Qualitätskriterien - Ergebnisse RE-Experteninterviews	33
4.1.5 Produktivität	35

4.2	Design Prinzipien	35
4.2.1	Bewältigung der RE-Herausforderungen mithilfe von GenAI	37
4.2.2	Herausforderungen von GenAI	40
4.3	Lösungsgestaltung	41
4.3.1	Entwicklung der Prompt	42
4.3.2	Aufbau der Prompt	44
4.3.3	Webanwendung - Benutzeroberfläche	45
4.3.4	Technische Implementierung	47
4.3.5	Abgrenzung	48
5	Evaluationsergebnisse	49
5.1	Produktivität	50
5.2	Qualität	51
5.3	Wahrgenommene Vorteile	52
5.4	Wahrgenommene Herausforderungen	54
5.5	Beobachtungen während der Simulation	56
6	Diskussion	58
6.1	Interpretation der Ergebnisse	58
6.2	Beitrag zur Literatur	59
6.3	Beitrag zur Praxis	60
6.4	Kritische Reflexion und Limitationen	61
6.5	Ausblick	62
7	Fazit	64
A	Literaturverzeichnis	VIII
B	Anhang	XXI
B.1	AIS eLibrary Suchstring Problemdefinition:	XXI
B.2	Ergebnisartefakt - Prompt zur Erstellung von User Stories	XXI
B.3	Ergebnisse der Experteninterviews	XXVIII
B.4	Leitfaden RE-Experteninterviews	XXXVI
B.5	Leitfaden GenAI-Experteninterviews	XXXIX
B.6	Evaluation	XLI
B.6.1	Leitfaden Nachevaluation	XLI
B.6.2	Agenda und Aufgabenstellung	XLII
B.6.3	Beispiel-Projektkontext	XLIV

Abbildungsverzeichnis

Abbildung 1:	DSR-Vorgehen nach Kuechler und Vaishnavi (2008) (Eigene Darstellung in Anlehnung an Meth et al. (2015))	12
Abbildung 2:	Design Prinzipien und Design Features des Ergebnisartefakts (Eigene Darstellung in Anlehnung an Meth et al. (2015))	36
Abbildung 3:	Aufbau der Prompt (Eigene Darstellung)	45
Abbildung 4:	Ausgabe Webanwendung - Teil 1	46
Abbildung 5:	Ausgabe Webanwendung - Teil 2	47
Abbildung 6:	Ausgabe Webanwendung - Teil 3	48
Abbildung 7:	Generierte User Stories innerhalb der Evaluation (Eigene Darstellung)	50
Abbildung 8:	Zuordnung der wahrgenommenen Vorteile auf die identifizierten RE-Herausforderungen (Eigene Darstellung)	52

Tabellenverzeichnis

Tabelle 1:	Suchergebnisse für GenAI und KI im RE (Stand: 08.12.2023) . .	14
Tabelle 2:	Interviewpartner RE	18
Tabelle 3:	Interviewpartner GenAI	20
Tabelle 4:	Evaluationsteilnehmer	23
Tabelle 5:	Teilartefakt: Konzeptmatrix - RE-Herausforderungen in der Literatur (Eigene Darstellung in Anlehnung an Webster und Watson (2002))	26
Tabelle 6:	Kategoriensystem und Häufigkeiten der Kodierungen der RE-Experteninterviews (Eigene Darstellung)	29
Tabelle 7:	Teilartefakt: Zuordnung der genannten RE-Herausforderungen aus den RE-Experteninterviews mit der Literatur	30
Tabelle 8:	Teilartefakt: Neu identifizierte RE-Herausforderungen aus den RE-Experteninterviews	31
Tabelle 9:	Teilartefakt: Konzeptmatrix - RE-Qualitätskriterien in der Literatur (Eigene Darstellung in Anlehnung an Webster und Watson (2002))	32
Tabelle 10:	Teilartefakt: Zuordnung RE-Qualitätskriterien aus den RE-Experteninterviews mit der Literatur	34
Tabelle 11:	Kategoriensystem und Häufigkeiten der Kodierungen der GenAI-Experteninterviews (Eigene Darstellung)	38
Tabelle 12:	Auswahl der genutzten Richtlinien für Prompt Engineering im RE nach Arvidsson und Axell (2023)	43
Tabelle 13:	Kategoriensystem und Häufigkeiten der Kodierungen der Evaluationsinterviews (Eigene Darstellung)	49

Tabelle 14:	Wahrgenommene Vorteile in der Evaluation	53
Tabelle 15:	Wahrgenommene Herausforderungen in der Evaluation	55
Tabelle 16:	Technisch lösbare wahrgenommene Herausforderungen in der Evaluation	55
Tabelle 17:	Aufgestellte Hypothesen in der Evaluation	56
Tabelle B.1:	Teilartefakt: Abgleich der RE-Herausforderungen der Literatur mit RE-Experten (Eigene Darstellung in Anlehnung an Webster und Watson (2002))	XXVIII
Tabelle B.2:	Teilartefakt: Abgleich der RE-Qualitätskriterien der Literatur in den RE-Experteninterviews (Eigene Darstellung in Anlehnung an Webster und Watson (2002))	XXX
Tabelle B.3:	Teilartefakt: Einschätzung bezüglich des Einsatzes von GenAI zum Lösen der RE-Herausforderungen in den GenAI-Experteninterviews (Eigene Darstellung in Anlehnung an Webster und Watson (2002))	XXXI
Tabelle B.4:	Teilartefakt: GenAI-Vorteile aus den GenAI-Experteninterviews .	XXXIV
Tabelle B.5:	Teilartefakt: GenAI-Herausforderungen aus den GenAI-Experten interviews	XXXV
Tabelle B.6:	Interviewleitfaden RE - Fragen des ersten Abschnitts	XXXVII
Tabelle B.7:	Interviewleitfaden RE - Fragen des zweiten Abschnitts	XXXVII
Tabelle B.8:	Interviewleitfaden RE - Fragen des dritten Abschnitts	XXXVII
Tabelle B.9:	Interviewleitfaden RE - Fragen des vierten Abschnitts	XXXVIII
Tabelle B.10:	Interviewleitfaden GenAI - Fragen des ersten Abschnitts	XL
Tabelle B.11:	Interviewleitfaden GenAI - Fragen des zweiten Abschnitts	XL
Tabelle B.12:	Interviewleitfaden Evaluation - Fragen des ersten Abschnitts . . .	XLII
Tabelle B.13:	Interviewleitfaden Evaluation - Fragen des zweiten Abschnitts . .	XLII

Abkürzungsverzeichnis

AIRE Workshop zu künstlicher Intelligenz im Requirements Engineering

ASD Agile Software Development

DL Deep Learning

DSR Design Science Research

Ei Experteninterview

Et Evaluationsteilnehmer

FEDS Framework for Evaluation in Design Science Research

GenAI Generative Künstliche Intelligenz

KI Künstliche Intelligenz

LLM Large Language Model

ML Machine Learning

RE Requirements Engineering

UX User Experience

1 Einleitung

1.1 Problemdefinition und Motivation

In der gegenwärtigen Unternehmenslandschaft ist der Einsatz von künstlicher Intelligenz (KI) weit verbreitet und nimmt weiter zu. Laut einer Studie von McKinsey (2023) nutzen 55 % der Unternehmen KI in zumindest einem ihrer Geschäftsbereiche. Im Vergleich dazu lag dieser Anteil im Jahr 2017 bei 20 % (McKinsey 2022). Ein Teilgebiet der KI ist die generative künstliche Intelligenz (GenAI), welche vor allem in den letzten eineinhalb Jahren einen Aufschwung erlebt hat (Dencik et al. 2023). Diese Entwicklung eröffnet wirtschaftliche Potenziale: Mit GenAI kann jährlich ein weltweiter Umsatz von bis zu 4,4 Billionen US-Dollar erwirtschaftet werden (Chui und Yee 2023). Zudem wird erwartet, dass GenAI bis 2025 einen Anteil von 30 % am gesamten KI-Markt ausmacht (BCG 2023).

Der Einsatz von GenAI hat sich seit der Veröffentlichung von ChatGPT im November 2022 weit verbreitet (Chen, Jiao et al. 2024). Dies unterstreicht den disruptiven Charakter der Technologie und lässt darauf schließen, dass ihr Einsatz Veränderungen in den Geschäftsprozessen und -strukturen von Unternehmen bewirken wird. Daraus lässt sich ein gestiegenes Forschungsinteresse begründen. Dellermann et al. (2019) nennen ein Beispiel für das gestiegene Forschungsinteresse, indem sie die Betrachtung des Einsatzes von KI zur Lösung komplexer Problemstellungen fordern. Inayat et al. (2015) sowie Hoy und Xu (2023) zeigen durch ihre Auflistungen von Herausforderungen im Bereich des Requirements Engineering (RE), dass dieses als eine komplexe Problemstellung beschrieben werden kann. Das RE ist ein Prozess zum systematischen Erfassen und Dokumentieren von Anforderungen an neu zu entwickelnde Systeme (Schön, Thomaschewski et al. 2017).

KI-Systeme vereinfachen die Mustererkennung in großen Datensätzen (Dellermann et al. 2019). Für komplexe Herausforderungen sind allerdings nach wie vor menschliche Fähigkeiten entscheidend (Dellermann et al. 2019). Dazu gehören unter anderem Kreativität und Empathie (Dellermann et al. 2019). Beides begünstigt eine effiziente Lösungsfindung (Dellermann et al. 2019). Die Kombination menschlicher und künstlicher Intelligenz kann die Problemlösung bisher unlösbarer Probleme ermöglichen (Akata et al. 2020). Diese Kombination, bekannt als hybride Intelligenz (HI), vereint menschliche Kreativität mit der Rechenkraft und Effizienz von KI (Akata et al. 2020).

HI findet Anwendung in verschiedenen Bereichen. Im Bildungswesen kommen beispielsweise Roboter zum Einsatz, um Lernprogramme zu erstellen. Indem sie das menschliche Fachwissen mit den eigenen Beobachtungen kombinieren, können sie den Lernfortschritt von Kindern überwachen (Akata et al. 2020). In der Wissenschaft werden HI-Assistenten eingesetzt, um umfangreiche Datenmengen zu analysieren und schnell an Informationen

zu gelangen (Akata et al. 2020). Auch im RE sind viele Informationen notwendig, die aus Daten generiert werden können (Maalej et al. 2016), weshalb HI-Ansätze in diesem Bereich genutzt werden könnten, um die Herausforderungen im RE zu bewältigen.

In der Auseinandersetzung mit offenen Problemstellungen erweisen sich andere KI-Ansätze als unzureichend (von Krogh 2018). GenAI bietet in diesem Kontext Vorteile beim Bearbeiten von sprachgestützten Aufgaben (Brown et al. 2020; Lester et al. 2021; Wu et al. 2022), wie beispielsweise dem RE. Dabei kann insbesondere HI Unterstützung leisten (Akata et al. 2020; Dellermann et al. 2019; Dwivedi et al. 2023). GenAI-Systeme sind bereits seit einigen Jahren in der wissenschaftlichen Gemeinschaft bekannt und Gegenstand verschiedener Forschungen (Brown et al. 2020; Vaswani et al. 2017). Die Einführung benutzerfreundlicher Anwendungen dieser Technologie, insbesondere durch ChatGPT im Jahr 2022, ermöglicht auch Nutzern ohne technischen Hintergrund von den Vorteilen von GenAI zu profitieren (Jiang et al. 2022). Seitdem ist sowohl das Interesse als auch der Druck gestiegen, GenAI in Unternehmen zu implementieren (Dencik et al. 2023).

1.2 Zielsetzung der Forschungsarbeit

In dieser Forschungsarbeit wird der Mehrwert von GenAI im RE untersucht. Das Ziel besteht darin, die Produktivität im RE-Prozess zu steigern und die Qualität von Anforderungsdokumenten zu erhöhen. Das wissenschaftliche Vorgehen besteht aus der Analyse des aktuellen Stands der wissenschaftlichen Literatur und der alltäglichen Praxis im RE. Dabei sind die Herausforderungen im RE das Ziel dieser Untersuchung.

Im Kontext der Qualität fassen Heck und Zaidman (2018) die bestehende Forschung zu Qualitätskriterien zusammen und bestätigen, dass der Einsatz von Tools Qualitätskriterien automatisch überprüfen kann. Sie empfehlen, dass in der Forschung neue Tools zur automatischen Sicherstellung der Qualität in Anforderungsdokumenten entwickelt und evaluiert werden sollen (Heck und Zaidman 2018).

In Bezug auf die Produktivität prognostizieren Dwivedi et al. (2023) eine Produktivitätssteigerung durch den Einsatz von GenAI-Systemen. Die Produktivität wird im Rahmen dieser Arbeit anhand des Zeitfaktors überprüft.

Anhand des Zeitfaktors und der Wahrnehmung der Qualität der generierten Ausgaben soll ein Vergleich zwischen der herkömmlichen Methodik zur Anforderungsdokumentation gegenüber einer hybriden Herangehensweise stattfinden. Es wird die These aufgestellt, dass die HI eine erhöhte Produktivität und Qualität erbringt.

1.3 Forschungsfragen

Parikh (2023) empfiehlt am Ende seines systematischen Literatur-Reviews zukünftige Forschungsaktivitäten auf die Rolle von GenAI im Software-Produktmanagement, also auch dem RE, zu konzentrieren. Zusätzlich soll die Analyse der Auswirkungen auf die Effizienz und die allgemeine Produktqualität untersucht werden (Parikh 2023). Daraus ergibt sich die folgende zentrale Forschungsfrage dieser Forschungsarbeit, die nachfolgend schrittweise anhand der Teilforschungsfragen erläutert wird:

Welchen Einfluss hat der Einsatz generativer künstlicher Intelligenz im Kontext hybrider Intelligenz auf die Qualität und Produktivität der Anforderungserhebung im Requirements Engineering?

Teilforschungsfragen:

(TF 1) *Welche Herausforderungen gibt es im Requirements Engineering im Kontext der agilen Softwareentwicklung?*

→ Zur Beantwortung der Teilforschungsfrage 1 wird eine Übersicht über die Herausforderungen im RE erarbeitet. Dafür wird die wissenschaftliche Literatur mit der alltäglichen Praxis verglichen. Die Betrachtung der RE-Herausforderungen dient als Ausgangspunkt für die Problembeschreibung.

(TF 2) *Wie kann generative künstliche Intelligenz genutzt werden, um die Herausforderungen im Requirements Engineering zu adressieren?*

→ Die Beantwortung der Teilforschungsfrage 2 adressiert mehrere Forschungslücken. Akata et al. (2020) und Dellermann et al. (2019) fordern die Untersuchung der Nutzung von HI zur Bewältigung komplexer Herausforderungen. Lester et al. (2021) und Wu et al. (2022) erachten die Generierung spezifischer Inhalte als erforschenswert. Anhand der Zusammenführung von Ergebnissen aus der Literaturrecherche mit den Ergebnissen von Experteninterviews sollen diese Forschungslücken adressiert werden. Die Einblicke in die Umsetzbarkeit und Effektivität von GenAI im RE zeigen, wie GenAI im Rahmen von HI genutzt werden kann. Dabei wird untersucht, ob Herausforderungen des RE durch GenAI gelöst werden können.

(TF 3) *Wie nehmen Experten und Stakeholder im Bereich des Requirements Engineering die Unterstützung durch generative künstliche Intelligenz wahr und welche Vorteile oder Herausforderungen identifizieren sie?*

→ Die Beantwortung der Teilforschungsfrage 3 zielt ebenfalls darauf ab, eine bestehende Forschungslücke zu adressieren. Jiang et al. (2022), Lester et al.

(2021) und Wu et al. (2022) empfehlen die Untersuchung der Interaktion zwischen Menschen und GenAI-Systemen. Die Ergebnisse der Evaluation zeigen wie die Unterstützung durch GenAI wahrgenommen wird und bringen Erkenntnisse über die Vorteile und Herausforderungen, die aus der Nutzung entstehen. Beides trägt zum besseren Verständnis der Interaktion zwischen GenAI-Systemen und Menschen bei.

1.4 Struktur der Forschungsarbeit

Der Aufbau dieser Forschungsarbeit berücksichtigt die Phasen des Design Science Research (DSR) nach Kuechler und Vaishnavi (2008). Im einleitenden Kapitel wird die Problemstellung und die Zielsetzung motiviert. Darauf folgt eine Erläuterung grundlegender Konzepte und Begriffe, wie zum Beispiel dem RE oder von GenAI. Danach wird die Methodik dieser Forschungsarbeit erläutert. Das darauf folgende Kapitel beinhaltet die Beschreibung der Teilartefakte und des Ergebnisartefakts, wobei Design Prinzipien und Design Features nach Meth et al. (2015) vorgestellt werden. Anschließend werden die Evaluationsergebnisse hinsichtlich der Wirksamkeit und des Nutzens des Ergebnisartefakts vorgestellt. Im vorletzten Kapitel werden die Ergebnisse diskutiert und kritisch reflektiert. Dabei werden die Limitationen dieser Forschungsarbeit aufgegriffen und Perspektiven für zukünftige Forschung skizziert. Abschließend werden die zentralen Erkenntnisse zusammengefasst.

2 Grundlagen

In diesem Kapitel werden wesentliche theoretische Konzepte und praxisorientierte Ansätze vorgestellt, die für das Verständnis der Inhalte dieser Forschungsarbeit wichtig sind. Kapitel 2.1 führt das Thema RE ein. Kapitel 2.2 definiert und erläutert die für diese Arbeit relevanten Aspekte von GenAI. In diesem Zusammenhang wird die Rolle von Large Language Modellen (LLMs) erörtert und verschiedene Prompting-Techniken diskutiert, um die Potenziale und Einschränkungen dieser Technologie zu untersuchen. Abschließend betrachtet Kapitel 2.3 die Anwendung von GenAI im Kontext des RE. Dieses behandelt den aktuellen Stand der Forschung sowie praktische Umsetzungen und skizziert, wie GenAI das Feld des RE aktuell beeinflusst.

2.1 Requirements Engineering

Im Kontext der agilen Softwareentwicklung bildet das RE die Basis für die Entstehung von Softwareprodukten. Dieser Prozess umfasst die Kernaktivitäten *Ermittlung*, *Dokumentation*, *Validierung*, *Verhandlung* und *Verwaltung von Anforderungen* für ein zu entwickelndes System und wird iterativ durchgeführt (Schön, Thomaschewski et al. 2017), (Pohl und Rupp 2015, vgl. S. 3 f.). Eine Anforderung lässt sich als ein Bedarf oder Wunsch definieren, der von einem Stakeholder zu einem bestimmten System geäußert wird (Glinz 2022, vgl. S. 16). Dies schließt auch Funktionalitäten, die ein System aufweisen soll, sowie dessen Dokumentation ein (Glinz 2022, vgl. S. 12). Anforderungen lassen sich grundsätzlich in zwei verschiedene Kategorien einteilen: *funktionale Anforderungen* und *nicht-funktionale Anforderungen* (auch Qualitätsanforderungen genannt) (Glinz 2022, vgl. S. 12, 14, 16). Während funktionale Anforderungen das erwartete Ergebnis oder Verhalten einer Systemfunktion spezifizieren, umfassen nicht-funktionale Anforderungen kritische Leistungsmerkmale, die über die grundlegenden Funktionalitäten eines Systems hinausgehen (Glinz 2022, vgl. S. 14, 16). Dies beinhaltet beispielsweise die *Wartbarkeit*, *Zuverlässigkeit* und *Sicherheit* eines Systems (Glinz 2022, vgl. S. 13, 16, 19).

Ein häufig verwendetes Artefakt zur Dokumentation von Anforderungen innerhalb des RE in der agilen Softwareentwicklung sind *User Stories* (Schön, Thomaschewski et al. 2017). User Stories illustrieren spezifische Beschreibungen von Funktionalitäten von Dienstleistungen oder Produkten (Schön, Thomaschewski et al. 2017). Diese Beschreibungen werden aus der Perspektive der Endnutzer formuliert, welche aus den betreffenden Funktionen Mehrwert generieren können (Cohn 2004). Typischerweise umfassen User Stories sowohl eine textuelle Darstellung als auch zugehörige Akzeptanzkriterien zur Evaluation ihrer Erfüllung (Hallmann 2020; Lucassen, Dalpiaz, van der Werf et al. 2016a; Schön, Thomaschewski et al. 2017).

User Stories werden in der Praxis häufig nach einem bestimmten Schema formuliert. In einer Studie mit 182 Teilnehmer bevorzugten 59 % die folgende Struktur für das Verfassen von User Stories:

Als [Rolle] möchte ich [Ziel], so dass [Nutzen] . . .

(Lucassen, Dalpiaz, Werf et al. 2016b).

Weitere 10 % der Studienteilnehmer wählen dasselbe Schema, jedoch ohne den Abschnitt des [Nutzens] (Lucassen, Dalpiaz, Werf et al. 2016b).

Dieses Schema gliedert sich in drei grundlegende Elemente:

- Die *Rolle* definiert einen Systemnutzer, der mit dem System bestimmte Ziele erreichen will (Amna und Poels 2022).
- Das *Ziel* konkretisiert die Aktionen des Systems, um den Nutzer bei seinen Zielen zu unterstützen (Amna und Poels 2022).
- Der *Nutzen* begründet, warum diese Aktion für den Nutzer wichtig ist (Amna und Poels 2022).

Die Überprüfung, ob ein System oder eine Software die aufgestellten Anforderungen und Akzeptanzkriterien erfüllt, erfolgt durch den Einsatz von Testszenarien (Mollah und van den Bos 2023). Diese Testszenarien lassen sich aus User Stories ableiten (Mollah und van den Bos 2023) und eignen sich dabei als Möglichkeit eine User Story zu ergänzen (Cohn 2004). Zudem können Testszenarien auch als eigene Anforderungsdokumente genutzt werden (Bjarnason et al. 2016).

Um Anforderungsdokumente, wie beispielsweise User Stories, in ihrer Qualität zu kontrollieren werden in der Praxis verschiedene Qualitätskriterien genutzt. Diese werden im Rahmen dieser Arbeit zusammengefasst und in Kapitel 4.1.3 vorgestellt.

2.2 Generative Künstliche Intelligenz

In diesem Kapitel werden die Grundlagen von GenAI erläutert. Der Fokus liegt hierbei auf Sprachmodellen, die auf Textgenerierung ausgerichtet sind. Es werden unterschiedliche Aspekte dieser Technologie, von ihrem Ursprung bis zu ihrer Anwendung behandelt. Zusätzlich wird die Integration von menschlicher und künstlicher Intelligenz betrachtet.

2.2.1 Künstliche Intelligenz

Im Laufe der Zeit haben sich die Definitionen von KI verändert. McCarthy et al. (1955) beschreiben das Ziel der KI-Entwicklung vor etwa 70 Jahren darin, Maschinen zu entwickeln,

die als intelligent angesehen werden können. Eine aktuelle Definition beschreibt KI als eine Software mit der Fähigkeit, Aufgaben auszuführen, die traditionell menschliche Intelligenz erfordern (Chui, Hazan et al. 2023). Um diese Fähigkeit zu realisieren, werden im Bereich der KI verschiedene Technologien angewendet. Im Bereich der KI umfasst das Spektrum unter anderem die Teilgebiete *Machine Learning* (ML) und dessen Unterbereich *Deep Learning* (DL) (Chui, Hazan et al. 2023; Lund et al. 2023). ML befasst sich mit Mustererkennung sowie dem Erlernen von Vorhersagen und der Abgabe von Empfehlungen durch die Verarbeitung von Daten (Chui, Hazan et al. 2023). DL nutzt unter anderem neuronale Netzwerke, um aus Daten zu lernen (Bommasani et al. 2022; Chui, Hazan et al. 2023). Diese Technologie wird beispielsweise in der Sprachverarbeitung eingesetzt (Hecker et al. 2017).

GenAI nutzt sogenannte *Foundation Models*. Diese werden durch ihre Fähigkeit charakterisiert, Inhalte zu erzeugen (Bommasani et al. 2022). Foundation Models werden auf einer umfangreichen Datenbasis trainiert (Bommasani et al. 2022). Innerhalb der Foundation Models bilden *Large Language Models* eine Unterkategorie, die als Grundlage für viele GenAI-Anwendungen dient (Bommasani et al. 2022). LLMs basieren auf der Transformer-Architektur, welche einen Teilbereich von DL darstellt (Vaswani et al. 2017). Die Transformer-Architektur ermöglicht es, das Verständnis und die Verarbeitung natürlicher Sprache durch Sprachmodelle zu verbessern (Vaswani et al. 2017). Ein Beispiel für LLMs ist die Serie der GPT-Modellen: Während GPT im Jahr 2018 etwa 110 Millionen Lernparameter nutzte, erweiterte GPT-2 diese Zahl 2019 auf 1,5 Milliarden (Floridi und Chiriatti 2020). GPT-3 erreichte schließlich 175 Milliarden Lernparameter (Floridi und Chiriatti 2020). Diese Skalierung ermöglicht es LLMs, aus denselben Prompts neue und qualitativ hochwertige Ausgaben zu generieren (Benbya et al. 2024).

Zu weiteren bekannten Beispielen für LLMs zählen OpenAI's *GPT-4* (OpenAI et al. 2023), *BERT* (Devlin et al. 2019) sowie *Gemini 1.0 Pro* (Gemini-Team et al. 2023) von Google und *RoBERTa* (Liu et al. 2019) von Facebook.

2.2.2 Prompt Engineering

Die Kommunikation mit einem Sprachmodell erfolgt durch sogenannte *Prompts* (Bommasani et al. 2022). Diese dienen als Aufgabenbeschreibungen und sind zentral für das Steuern der Generierung von Ausgaben von GenAI-Systemen (Bommasani et al. 2022; White et al. 2023; Ye et al. 2023). Das Feld des *Prompt Engineering* widmet sich der Entwicklung und Optimierung dieser Prompts, um die bestmögliche Qualität der generierten Antworten zu erreichen (Chui, Hazan et al. 2023; Ye et al. 2023). Modelle, wie zum Beispiel Gemini 1.0 Pro oder GPT-4o, unterstützen verschiedene Modalitäten als Eingaben, darunter Texte, Bilder und Videos, was eine umfangreiche Aufstellung der Prompt ermöglicht (C. Lu et al.

2024). White et al. (2023) zeigen, dass Prompt Engineering zur Verbesserung der Qualität der Ausgaben beitragen kann.

In diesem Kontext sind unterschiedliche Techniken wie *Fine-Tuning*, *Zero-Shot*-, *One-Shot*- und *Few-Shot-Learning* von Bedeutung (Brown et al. 2020). Fine-Tuning trainiert ein bereits vortrainiertes Modell für eine spezielle Aufgabe (Brown et al. 2020; Chui, Hazan et al. 2023). Dabei handelt es sich um ein Trainingsverfahren, bei dem das Modell nachhaltig aus bereitgestellten Daten lernt und einen Teil seiner Parameter entsprechend anpasst (Brown et al. 2020). Im Gegensatz dazu bleiben die Modellparameter bei Zero-Shot-, One-Shot- und Few-Shot-Learning unverändert. Stattdessen werden den Sprachmodellen kontextuelle Informationen und Beispiele im Rahmen einer Prompt zur Verfügung gestellt (Brown et al. 2020). Bei der Verwendung von Zero-Shot-Learning wird dem Modell nur die Aufgabenbeschreibung zur Verfügung gestellt, ohne dass Beispiele mitgeliefert werden (Brown et al. 2020). Beim Einsatz von One-Shot-Learning wird dem Modell zusätzlich ein einzelnes Beispiel zur Orientierung bereitgestellt (Brown et al. 2020). Beim Few-Shot-Learning werden mehrere Beispiele genutzt, um das Modell an die Aufgabenstellung anzupassen (Brown et al. 2020). Diese Ansätze verringern den Bedarf an großen, spezifischen Datensätzen und erleichtern die schnelle Anpassung von Modellen an neue Aufgaben (Brown et al. 2020).

2.2.3 Hybride Intelligenz

HI repräsentiert ein Konzept, das darauf abzielt, die komplementären Stärken von menschlicher und künstlicher Intelligenz zu vereinen (Akata et al. 2020; Dellermann et al. 2019). Dies dient dem Zweck menschliche Fähigkeiten zu erweitern, anstatt sie zu ersetzen (Akata et al. 2020) und ermöglicht Lösungen für Probleme, die weder allein von Menschen noch durch Maschinen lösbar sind (Akata et al. 2020).

Diese Zusammenarbeit von menschlicher und künstlicher Intelligenz, auch als das Konzept *machines as teammates* bekannt, wird immer relevanter (Akata et al. 2020; Seeber et al. 2020; D. Wang et al. 2020). Mit den Fortschritten innerhalb der GenAI erhöht sich die Relevanz der HI (Kar et al. 2023). GenAI wird zunehmend in Interaktionen mit Menschen eingesetzt, um spezifische, kontextbezogene Aufgaben zu erfüllen (Goel et al. 2023; Marczak-Czajka und Cleland-Huang 2023).

2.2.4 Einsatzbereiche und Potenziale im Kontext hybrider Intelligenz

Die Integration von GenAI in den Kontext der HI eröffnet Möglichkeiten zur Steigerung von Leistung und Effizienz in verschiedenen Anwendungsbereichen. Ein Vorteil im Einsatz von GenAI liegt in der Unterstützung von Prozessen in frühen Entwicklungsphasen, wie Goel

et al. (2023) beim Einsatz von GPT-3 bei der Erstellung von Personas zeigen. Empfohlen wird der Einsatz von GenAI zur Generierung erster Entwürfe (Goel et al. 2023). Dabei kann durch die Beantwortung von Folgefragen auf initial generierte Ausgaben die Qualität der Ergebnisse verbessert werden (Goel et al. 2023). Weiterhin kann GenAI Nutzer effektiv bei spezifischen Aufgaben unterstützen (Goel et al. 2023). Im Fall der Persona-Erstellung mithilfe von GenAI werden von unerfahrenen Designern Ergebnisse generiert, die mit der Anwendung manueller Methoden vergleichbar sind (Goel et al. 2023). Allerdings werden die mit GPT-3 erstellten Inhalte auch in wichtigen Aspekten wie Nutzungsbereitschaft, Glaubwürdigkeit und Empathie niedriger bewertet (Goel et al. 2023).

Ein weiteres wichtiges Potenzial von GenAI ist die Erleichterung der Wissensbeschaffung und der Synthese von Wissen (Alavi et al. 2024). GenAI hat die Fähigkeit, sich die Konversation mit einem Nutzer zu merken und darauf basierend zugeschnittene Antworten zu liefern (Alavi et al. 2024). Zudem ermöglicht GenAI das Herstellen von Verbindungen innerhalb vorgegebener Daten. Diese können genutzt werden, um zuvor nicht lösbare Probleme zu lösen (Orrù et al. 2023).

Im Bereich des Brainstormings demonstriert die HI ein hohes Einsatzpotenzial (Memmert und Tavanapour 2023). In einer Studie von Memmert und Tavanapour (2023) wird beobachtet, dass ein Großteil der Teilnehmer die Zusammenarbeit mit GenAI der individuellen Arbeit vorzieht. GenAI trägt an dieser Stelle durch das Einbringen von Ideen bei (Memmert und Tavanapour 2023).

Zusammenfassend bietet die Anwendung von GenAI im Rahmen von HI Vorteile in der Wissensbeschaffung und -synthese sowie in der Förderung der Kreativität. Der Einsatz dieser Technologie ermöglicht es, die Kapazitäten menschlicher Intelligenz zu erweitern und die Effizienz in der Informationsverarbeitung und Ideengenerierung zu steigern.

2.2.5 Risiken und Limitationen

Der Einsatz von GenAI ist jedoch auch mit Risiken und Limitationen verbunden. Diese können die Authentizität und Zuverlässigkeit der von GenAI generierten Ergebnisse beeinträchtigen (Lund et al. 2023).

Ein auftretendes Problem sind sogenannte *Halluzinationen*, bei denen generative Modelle falsche oder irrelevante Informationen produzieren (Alavi et al. 2024; Budhwar et al. 2023). Darüber hinaus besteht bei Halluzinationen die Gefahr, dass GenAI potenziell schädliche oder anstößige Inhalte erzeugt, wie bei der Nutzung von GPT-3 beobachtet wurde (Budhwar et al. 2023).

Zusätzlich trägt *Voreingenommenheit (Bias)* zur Verzerrung der Ergebnisse bei, da GenAI

nicht immer in der Lage ist, Fehlinformationen zu erkennen und korrekt zu verarbeiten (Ferrara 2023; Lund et al. 2023). Die *Missinterpretation* mehrdeutiger oder zusammengesetzter Begriffe birgt eine weitere Gefahr (Lund et al. 2023). Ein mit den zuvor genannten Problemen verbundenes Risiko, auch im Kontext von HI, stellt sich durch das *übermäßige Vertrauen von Nutzern in die Ausgaben von GenAI* dar (Vasconcelos et al. 2023).

Darüber hinaus sind die für den Betrieb von GenAI-Systemen erforderlichen Algorithmen und Datenspeicher, wie bei OpenAI beobachtet, mit einem erheblichen Energieverbrauch verbunden (Zhou et al. 2021). Diese Aspekte verdeutlichen die Notwendigkeit einer sorgfältigen Überwachung und Weiterentwicklung der Rahmenbedingungen im Umgang mit GenAI.

2.2.6 Abgrenzung

Der Fokus dieser Arbeit liegt auf der Implementierung einer Lösung, die mithilfe von GenAI auf die RE-Herausforderungen eingeht. Des Weiteren wird die Interaktion zwischen Menschen und dieser Lösung betrachtet. Technische Implementierungsaspekte, wie unter anderem Guardrails (Rebedea et al. 2023), Reranking (Ma et al. 2023) sowie agentbasierte Ansätze (Park et al. 2023) werden in dieser Arbeit nicht vertieft behandelt. Dennoch tragen diese Techniken zur Sicherheit, Effizienz und Anwendbarkeit von GenAI-Systemen bei. Eine detaillierte Betrachtung dieser technischen Aspekte, einschließlich ihrer Potenziale für künftige Forschungsarbeiten, ist in Kapitel 6.5 beschrieben.

2.3 Generative Künstliche Intelligenz im Requirements Engineering

In einer wissenschaftlichen Forschungsarbeit von Wagner et al. (2019) zum aktuellen Stand im RE gibt die Mehrheit der Befragten an, dass sie Anforderungen hauptsächlich in textueller Form dokumentieren. Dies geschieht sowohl in freier Form als auch innerhalb bestimmter Vorlagen, wie beispielsweise dem in Kapitel 2.1 beschriebenen Schema für User Stories (Wagner et al. 2019). Halbformale Methoden, wie unter anderem UML-Diagramme, werden dabei überwiegend für Anwendungen wie Datenmodelle verwendet (Wagner et al. 2019). Diese Tendenz zur textbasierten Anforderungsdokumentation erklärt bisherige Versuche GenAI im RE einzusetzen.

Ein Meilenstein in diesem Bereich ist der internationale Workshop zu KI und RE (AIRE), welcher seit 2014 jährlich stattfindet (Abualhaija et al. 2023). Diese Konferenz zielt darauf ab, Wechselwirkungen zwischen KI und RE zu erkunden. AIRE konzentriert sich darauf, komplexe RE-Probleme zu identifizieren, die von der Anwendung von KI-Technologien profitieren können und umgekehrt. Dies trägt dazu bei, RE-Herausforderungen durch den

Einsatz von KI-Tools zu adressieren. Im Rahmen der AIRE Konferenz werden verschiedene Forschungsarbeiten präsentiert, die den Einsatz von GenAI im Bereich des RE illustrieren. Görer und Aydemir (2023) untersuchen beispielsweise, wie LLMs zur automatischen Generierung von Interviewskripten für den Prozess der Anforderungserhebung genutzt werden können. Sie entwickeln eine Methode zur Erstellung dieser Skripte durch Prompt Engineering (Görer und Aydemir 2023). Marczak-Czajka und Cleland-Huang (2023) erforschen die Verwendung von ChatGPT zur Generierung von User Stories, die menschliche Werte widerspiegeln. Diese dienen als Kreativitätsauslöser, um menschliche Werte für ein Produkt zu spezifizieren (Marczak-Czajka und Cleland-Huang 2023).

Unabhängig von der AIRE Konferenz wird ebenfalls am Einsatz von GenAI im RE geforscht. In ihrer Studie analysieren Goel et al. (2023) den Einsatz von GenAI zur Unterstützung von unerfahrenen Designern bei der Persona-Erstellung. Ajagbe und L. Zhao (2022) präsentieren BERT4RE, ein auf Anforderungstexten trainiertes BERT-Modell, welches für RE Aufgaben optimiert ist. Es dient unter anderem der Klassifizierung von Anforderungen, der Erkennung von Sprachproblemen (Ajagbe und L. Zhao 2022).

Auch im Bereich der KI-gestützten Tools im RE werden Beiträge geleistet. Ezzini, Abualhaija und Sabetzadeh (2022) stellen *WikiDoMiner* vor. Dabei handelt es sich um ein Tool zur automatischen Generierung von domänenspezifischen Textkorpora durch das Durchsuchen von Wikipedia, das RE-Aufgaben unterstützt (Ezzini, Abualhaija und Sabetzadeh 2022). Ezzini, Abualhaija, Arora et al. (2023) präsentieren *QAssist*. Bei *QAssist* handelt es sich um einen Ansatz zur automatisierten Unterstützung bei der Analyse von natürlichsprachlichen Anforderungen durch ein Frage-Antwort-System (Ezzini, Abualhaija, Arora et al. 2023). *REInDetector* automatisiert die Entdeckung von Inkonsistenzen zwischen Anforderungen und unterstützt eine effiziente Problembehandlung (Nguyen et al. 2012).

Zum Abschluss folgt ein praxisnahes Beispiel, das im Rahmen dieser Forschungsarbeit in Interviews¹ mit RE-Experten aufgenommen wird. Ein Interviewteilnehmer berichtet von einem Projekt, in dem juniorige Mitarbeiter in Abstimmung mit dem Kunden Use Cases und weitere Anforderungsdokumente erstellen sollen (Ei3 - Z. 65 ff.). Bei einer nachfolgenden Überprüfung zeigt sich jedoch, dass die Qualität der erstellten Anforderungen nicht den Erwartungen entspricht (Ei3 - Z. 72 f.). Angesichts der zu diesem Zeitpunkt kürzlich erfolgten Markteinführung von ChatGPT beschließt das Team, das Potenzial dessen zu testen (Ei3 - Z. 75 ff.). Der Einsatz von ChatGPT führte innerhalb einer vergleichsweise kurzen Zeit zu einer signifikanten Verbesserung der Ergebnisse (Ei3 - Z. 78 ff.). Besonders hervorgehoben wird die verbesserte strukturelle Aufbereitung und die inhaltliche Vollständigkeit der Dokumente (Ei3 - Z. 79 ff.).

¹ Diese werden methodisch in Kapitel 3.2.1 vorgestellt.

3 Methodik

In diesem Kapitel wird das wissenschaftliche Vorgehen dieser Arbeit beschrieben, das auf der Durchführung von Design Science Research (DSR) nach Kuechler und Vaishnavi (2008) basiert. Das Vorgehen von Kuechler und Vaishnavi (2008) orientiert sich an den sieben Richtlinien für DSR von Hevner et al. (2004) und gliedert sich in fünf Phasen. Die Methodik sowie die Forschungsaktivitäten und Ergebnisse dieser Arbeit sind in Abbildung 1 dargestellt.

Das Ziel von DSR ist die Entwicklung eines Artefakts, das praktisch anwendbar ist und einen Mehrwert liefert (Hevner et al. 2004). Im Rahmen dieser Forschungsarbeit wurde ein solches Artefakt entwickelt, um mithilfe von GenAI RE-Herausforderungen zu bewältigen ((TF 1) und (TF 2)). Darüber hinaus sollte das Artefakt so gestaltet werden, dass es im Unternehmenskontext eingesetzt werden kann (TF 3). Die Anwendung des DSR eignet sich hierfür, da es eine iterative Weiterentwicklung des angestrebten Ergebnisartefakts durch die Integration theoretischer Erkenntnisse, praktischer Anwendungen und kontinuierlicher Evaluation ermöglicht (Hevner et al. 2004). Dies gewährleistet, dass die entwickelte Lösung relevant für die Forschung und in der Praxis anwendbar ist (Hevner et al. 2004).



Abbildung 1: DSR-Vorgehen nach Kuechler und Vaishnavi (2008) (Eigene Darstellung in Anlehnung an Meth et al. (2015))

Im Folgenden werden die Phasen des DSR-Vorgehens erläutert. In der ersten Phase wurde die zugrundeliegende Problemstellung präzisiert, die die Herausforderungen im Bereich des RE betrachtet. Weiterhin wurde der Wert der Lösung dieser Herausforderungen argumentiert (Kuechler und Vaishnavi 2008). Mittels einer unstrukturierten Literaturrecherche wurde das Forschungsgebiet zunächst abgegrenzt. Anschließend wurde ein Literatur-Review nach Webster und Watson (2002) für die Erweiterung des bestehenden Wissens und die Erhebung des aktuellen Forschungsstands genutzt (Kuechler und Vaishnavi 2008).

In der darauffolgenden Phase wurde die Problemstellung durch Experteninterviews aus der Praxis analysiert. Zudem wurden Design Prinzipien von Experten im Bereich des RE und GenAI identifiziert. Design Prinzipien werden als Prinzipien einer Lösungsgestaltung definiert (Meth et al. 2015). Die Synthese von literaturbasierten Erkenntnissen und praktischem Expertenwissen trug zur Erarbeitung des Ergebnisartefakts bei.

In der dritten Phase wurde eine GenAI-gestützte Webanwendung zur Erstellung von User Stories entwickelt. Dabei wurden die Design Prinzipien und die daraus abgeleiteten Design Features berücksichtigt.

In der vierten Phase wurde das erstellte Ergebnisartefakt im Unternehmenskontext von Experten evaluiert (Kuechler und Vaishnavi 2008). Die Evaluation orientierte sich dabei an Konzepten des Framework for Evaluation in Design Sciences (FEDS) nach Venable et al. (2016). Dieses Framework ermöglicht die Auswertung der Nützlichkeit, Qualität und Effektivität eines Artefakts (Venable et al. 2016). Die *Thinking-Aloud* Methode nach Lewis (1982) ermöglicht Einblicke in die Benutzererfahrung. Anhand dieser Einblicke erfolgte die Evaluation des Ergebnisartefakts. Ein zentrales Element des Evaluationsvorgehens war die Untersuchung der Produktivität des Prozesses und Qualität der erstellten Inhalte. Zusätzlich wurde die Interaktion der Evaluationsteilnehmer mit der Webanwendung analysiert.

In der abschließenden Phase wurden die Ergebnisse durch Präsentation und Dokumentation der Forschungsergebnisse kommuniziert (Kuechler und Vaishnavi 2008).

3.1 Literaturrecherche

Es wurden zwei Literatursuchen durchgeführt. Zunächst erfolgte anhand einer unstrukturierter Literaturrecherche die Erschließung des Forschungsbereichs. Anschließend wurde ein Literatur-Review nach Webster und Watson (2002) durchgeführt, um Herausforderungen und Qualitätskriterien im Kontext des RE zu identifizieren.

Im Rahmen der unstrukturierter Literaturrecherche wurde auch der aktuelle Stand der Praxis im Einsatz von KI und GenAI im RE untersucht. Hierfür wurden Suchstrings entwickelt, die auf Stichwörtern relevanter Artikel basieren. Die Suchergebnisse, die sich aus den jeweiligen Suchstrings zum aktuellen Stand von GenAI und KI im RE ergeben haben, sind in Tabelle 1 dargestellt.

Im Rahmen des Literatur-Reviews nach Webster und Watson (2002) wurden zwei separate Konzeptmatrizen erstellt: Die erste Konzeptmatrix umfasst RE-Herausforderungen, die die Problemstellung dieser Arbeit konkretisieren. Die zweite Konzeptmatrix enthält RE-Qualitätskriterien. Beide Konzeptmatrizen stellen den aktuellen wissenschaftlichen Forschungsstand dar und wurden anschließend von RE-Experten in Bezug auf ihre

Suchstrings	Web of Science	Google Scholar
<i>“Requirements Engineer*“ AND (“Generativ* AI“ OR “Generativ* Artificial* Intelligen*“ OR “AIRE“) OR “Artificial Intelligence for Requirements Engineering“</i>	6	107
<i>“Requirements Engineer*“ AND (“Generativ* AI“ OR “Generativ* Artificial* Intelligen*“ OR “Artificial* Intelligen*“ OR “AIRE“) OR “Artificial Intelligence for Requirements Engineering“</i>	76	106
<i>“Requirements Engineer*“ AND (“Hybrid Intelligen*“ OR “Collaborative AI“ OR “Machines as teammates“ OR “Human-computer collaboration“)</i>	2	6

Tabelle 1: Suchergebnisse für GenAI und KI im RE (Stand: 08.12.2023)

Praxisrelevanz ausgewertet.

Die Suche nach relevanten Artikeln erfolgte durch eine Online-Recherche innerhalb der Web of Science-Datenbank, welche Zugriff auf verschiedene Quellen wie ScienceDirect und die IEEE Xplore Library bietet. Die Web of Science-Datenbank erwies sich als effizientes Werkzeug, um die Relevanz einzelner Artikel zu überprüfen. Der Titel und die Zusammenfassung der Artikel werden anhand der im Suchstring verwendeten Begriffe hervorgehoben. Dies ermöglichte einen schnellen Überblick. Ergänzend dazu wurden die AIS eLibrary und Google Scholar durchsucht. Die Stichwörter, die zur Erstellung des Suchstrings genutzt wurden, waren Ergebnis der unstrukturierten Literaturrecherche. Der Suchstring, welcher für die Suche in den Datenbanken verwendet wurde lautet wie folgt:

Title, Abstract, Author Keywords (“Requirements Engineer“ OR “User Stor*“ OR “Product Owner“ OR “Business Analys*“) AND (“Challeng*“ OR “Difficult*“ OR “Issue*“ OR “Problem*“ OR “Complex*“ OR “Quality Criteria“) NOT (“Estimat*“)*

Mit diesem Suchstring wurden in der Web of Science-Datenbank 2.278 Publikationen identifiziert. Die AIS eLibrary verzeichnete 4.180 Treffer und auf Google Scholar wurden 83 Treffer gefunden; ohne den Ausschlussbegriff *“Estimat*“* ergaben sich 17.200 Treffer. (Stand: 12.12.2023)

Der Suchstring wurde an verschiedenen Stellen modifiziert. Zunächst erfolgte eine Anpassung des ersten Suchbegriffs zu *“Requirement* Engineer*“*, motiviert durch die Auffindung einer relevanten Arbeit in der vorherigen unstrukturierten Recherche. Diese Arbeit, welche im Titel *“Requirement Engineering“* enthält, hebt Herausforderungen im RE hervor. Des Weiteren ergab die Verwendung des Stichworts *“Business Analys*“* keine relevanten Treffer.

Aus diesem Grund wurde der Suchbegriff auf *“Business Analyst“* angepasst, um eine Auswahl an Artikeln zu erhalten, die sich auf Herausforderungen innerhalb dieser Rolle beziehen.

Das Vorgehen bei der Literatursuche in der Web of Science Datenbank gestaltete sich wie folgt: Die Artikel wurden nach ihrer Relevanz sortiert. Um die Forschungsgüte zu erhöhen, wurde nach gereviewten Artikeln gefiltert. Dadurch verringerte sich die Anzahl von 2.278 auf 55 Artikel. Die Titel dieser Artikel wurden analysiert. Sofern sie relevant erschienen und die erforderlichen Schlüsselbegriffe enthielten, wurde die Zusammenfassung gelesen.

Die Entscheidung, ob die Schlussfolgerung und der Hauptteil des Artikels näher betrachtet werden, basierte darauf, ob der Artikel RE-Herausforderungen oder -Qualitätskriterien behandelte. Artikel, die diese Informationen enthielten, wurden zunächst in die Literatursammlung aufgenommen. Anschließend wurden sie in Konzeptmatrizen nach Webster und Watson (2002) eingeordnet, um die Gemeinsamkeiten, Unterschiede und Besonderheiten der Artikel in Bezug auf RE-Herausforderungen und -Qualitätskriterien festzuhalten. Diese Einordnung erleichterte die Analyse und den Vergleich der Literatur. Die Konzeptmatrizen sind in Tabelle 5 und Tabelle 9 dokumentiert. Nach der Zuordnung eines Artikels zur Literatursammlung wurde eine Vorwärts- sowie Rückwärtssuche durchgeführt. Die Rückwärtssuche diente der Überprüfung der Relevanz der von der Literatursammlung zitierten Quellen. Die Vorwärtssuche untersuchte die Relevanz der Quellen, die die Literatursammlung zitierten.

Eine Literatursuche in der AIS eLibrary umfasste 4.180 Artikel (Stand 12.12.2023). Im Interesse der Forschungsgüte wurde nach gereviewten Artikeln gefiltert. Aus diesem Grund wird der Suchstring nach Titeln, die das Wort *“Review“* enthalten, erweitert. Diese abgeänderte Suche (siehe Kapitel Anhang B.1) ergab neun Ergebnisse, die jedoch nicht relevant für die weitere Recherche waren.

Die Erweiterung der Suche auf Google Scholar mittels des oben stehenden Suchstrings ergab ebenfalls keine zielführenden Ergebnisse. Die Kombination der Schlüsselwörter *“Requirements Engineering“*, *“Review“*, *“Challenges“* und *“Quality Criteria“* führte jedoch zu mehr als vier Millionen Einträgen. Innerhalb dieser Suche wurden der Titel und die Zusammenfassung der Artikel analysiert. Dabei zeigte sich, dass die Suchergebnisse in Google Scholar zum Teil bereits bekannte Artikel aus der Web of Science Datenbank enthielten. Da innerhalb dieser Suche nur ein weiterer relevanter Artikel in den ersten fünf Ergebnisseiten gefunden wurde, wurde die Suche an dieser Stelle beendet.

3.2 Experteninterviews - Allgemein

Es wurden zwei Interviewrunden durchgeführt. In Kapitel 3.2.1 wird die Durchführung der RE-Experteninterviews beschrieben. Anschließend wird in Kapitel 3.2.2 die Durchführung der GenAI-Experteninterviews erläutert. Zur besseren Übersichtlichkeit werden die Gemeinsamkeiten in diesem Abschnitt zusammengefasst.

Das Vorgehen für die Interviews orientierte sich an Gläser und Laudel (2010), welche drei Hauptphasen beschreiben: die *Vorbereitung der Experteninterviews*, die *Datenerhebung* und die *Datenauswertung*. In dieser Arbeit wurden die Interviews einmalig durchgeführt, was als qualitative Querschnittsanalyse bezeichnet wird (Kuenzel 2022; Wilde und Hess 2007).

Im Rahmen der Vorbereitung erfolgte zunächst die Auswahl der Experten (Gläser und Laudel 2010). Dabei wurde die theoretische Definition eines Experten berücksichtigt. Ein Experte ist eine Person, die Wissen über ein Forschungsfeld verfügt, welches nicht allgemein zugänglich ist (Meuser und Nagel 2009). Des Weiteren wurden unterschiedliche Perspektiven der Experten berücksichtigt, um eine Triangulation zu schaffen (Gläser und Laudel 2010). Die Auswahl der Experten basierte auf der Abdeckung den folgenden Perspektiven: Vielfalt der Projekteinsätze und die Rollen im Projektkontext. Die Suche nach Interviewpartnern wurde frühzeitig angestoßen, um terminliche Herausforderungen, die zu Verzögerungen führen könnten, zu adressieren. Nach der Auswahl der Experten erfolgte die Terminvereinbarung (Gläser und Laudel 2010).

Die Experteninterviews wurden semistrukturiert durchgeführt. Dies bedeutet, dass die Konstruktion eines Interviewleitfadens notwendig ist (Gläser und Laudel 2010). Eine semistrukturierte Durchführung impliziert, dass der erstellte Interviewleitfaden verwendet werden kann und gleichzeitig auch Raum für offene Nachfragen besteht. Dabei ist es wichtig die Offenheit bis zu einem gewissen Grad zu beschränken, um den Fokus auf das Forschungsinteresse nicht zu verlieren. Gleichzeitig muss genug Spielraum für unvorhergesehene, jedoch relevante Informationen gelassen werden (Helfferich 2014).

Zu Beginn des Interviews wurden die zentralen Konzepte wie zum Beispiel Einsatzmöglichkeiten von GenAI im RE erläutert. Anschließend wurden der Forschungskontext und die Ziele des Interviews vorgestellt. Den Interviewteilnehmern wurde die vertrauliche Behandlung ihrer Antworten zugesichert. Die Ergebnisse der Interviews wurden pseudonymisiert. Am Ende der Interviews wurden die Interviewpartner gefragt, ob sie einer weiteren Kontaktaufnahme für eventuelle Nachfragen zustimmen. Die Interviews wurden mit Bild und Ton aufgezeichnet. Die Transkription erfolgte KI-gestützt. Die Transkriptionen der Interviews sowie die Transkriptionsregeln sind in der separat angehängten Datei dokumentiert.

Die Auswertung der Interviewergebnisse erfolgte anhand einer qualitativen Inhaltsanalyse nach Mayring (2010). Dabei wurde ein Kategoriensystem zur systematischen Auswertung der Experteninterviews genutzt. Dieses Kategoriensystem wurde zunächst deduktiv entwickelt. Dafür wurden die Kategorien anhand von Definitionen, in diesem Fall dem Interviewleitfaden, vorab festgelegt (Mayring und Fenzl 2019). Anschließend erfolgte eine induktive Erweiterung der Kategorien. Das bedeutet, dass anhand des vorhandenen Materials weitere Kategorien aus den Textstellen formuliert wurden, um neue Kategorien zu erzeugen (Mayring und Fenzl 2019).

Für die Analyse wurde die Grundtechnik der *Strukturierung* gewählt. Diese hat das Ziel „bestimmte Aspekte aus dem Material herauszufiltern und unter vorher festgelegten Ordnungskriterien einen Querschnitt durch das Material zu legen oder das Material unter bestimmten Kriterien einzuschätzen“ (Mayring 1994, S. 164). Folglich war diese Methode geeignet, um getätigte Aussagen den Kategorien des Leitfadens zuzuweisen.

3.2.1 Experteninterviews - RE

Diese Forschungsarbeit ergänzt eine praktische Perspektive zu den RE-Herausforderungen und -Qualitätskriterien aus der Literatur (siehe Kapitel 3.1). Dies erfolgte durch die Durchführung von Interviews mit Experten, die im RE tätig sind. Die Erkenntnisse aus diesen Gesprächen wurden im Nachgang analysiert und mit den Ergebnissen der Literaturrecherche synthetisiert. Zur Beantwortung der Teilforschungsfrage (TF 1) wurde in den Interviews die Wahrnehmung von RE-Herausforderungen und -Qualitätskriterien in der Praxis analysiert. Anschließend wurde auch die Bewältigung jener Herausforderungen betrachtet. Daraus entstand die Möglichkeit der Erarbeitung eines GenAI-Tools zur Unterstützung der Arbeit von RE-Experten.

Vorbereitung der Expertinterviews

Die Interviewpartner wurden über eine Anfrage in einer RE-Community und über Kontakte im Unternehmen gewonnen. Es wurden fünf Experteninterviews terminiert, welche sich über einen Zeitraum von einer Woche erstreckten. Alle interviewten Personen waren bereits an unterschiedlichen Projekten in verschiedenen Rollen im Kontext von RE beteiligt. Detaillierte Beschreibungen der befragten Experten sind in Tabelle 2 zu finden. Aus Datenschutzgründen werden Pseudonyme verwendet. Der Identifikator „Ei“ steht für „Experteninterview“, gefolgt von einer fortlaufenden Nummer zur Unterscheidung der Interviews.

Datenerhebung

Der Leitfaden für die RE Experteninterviews (siehe Kapitel Anhang B.6.1) wurde in Form eines Fragenkatalogs aufgebaut. In diesem wurden 18 Fragen in vier Kategorien gegliedert.

Identifikator	Rolle im RE	Beschreibung
Ei1	Product Owner, Key/Power User, Projektleiter	Ei1 ist seit über zehn Jahren im RE tätig. Mit GenAI ist Ei1 sowohl im beruflichen, als auch im privaten Kontext vertraut. Ei1 nutzt speziell ChatGPT als erweiterte Suchmaschine und zur Erstellung von Texten auf Basis bestimmter Schlagwörter.
Ei2	Business Analyst, Fachanalyst, IT-Arbeitspaketverantwortlicher	Ei2 ist seit über elf Jahren im Bereich des RE aktiv. Hinsichtlich GenAI hat Ei2 berufliche Erfahrungen vorwiegend im Bereich der Übersetzungen gesammelt, während im privaten Umfeld Erfahrung im Generieren von Texten gesammelt wurde.
Ei3	Requirements Engineer, Product Owner, Business Analyst	Ei3 arbeitet seit 1981 in verschiedenen Rollen innerhalb der Softwareentwicklung, insbesondere in der Anforderungsaufnahme. Seit Anfang 2023 sammelt Ei3 zudem Erfahrung im Einsatz von GenAI in Kunden- und internen Projekten.
Ei4	Product Owner, Business Analyst, Projektleitung (Koordination des RE)	Ei4 ist seit sieben Jahren im RE aktiv. Davor hat Ei4 bereits berufsbegleitende Erfahrungen während eines dualen Studiums der Wirtschaftsinformatik berufsbegleitende Erfahrung gesammelt. Im Bereich der GenAI hat Ei4 erste Erfahrung, speziell in der Erstellung von Texten, gesammelt.
Ei5	Usability Engineer, Product Owner, Business Analyst	Ei5 arbeitet seit 2018 im RE. Ei5 hat im wissenschaftlichen Bereich an einer Hochschule schon Erfahrung mit GenAI gesammelt.

Tabelle 2: Interviewpartner RE

Die erste Kategorie diente als Einstieg in das Gespräch und zur Ermittlung der *Vorerfahrung im RE und GenAI* der Interviewteilnehmer. Die Frage nach der Erfahrung zielte darauf ab zu überprüfen, wie lange die Interviewteilnehmer im Bereich des RE tätig sind.

In der zweiten Kategorie wurde die *Vorgehensweise im RE* erfragt. Dies schloss vor allem Fragen nach in der Praxis genutzten Anforderungsdokumenten und Methoden zur Erstellung dieser ein. Diese Fragen richteten sich nach der von Schön, Thomaschewski

et al. (2017, S. 80) gestellten Forschungsfrage „Was sind die gängigen Methoden für das Anforderungsmanagement im ASD (Agile Software Development)?“.

Die dritte Kategorie - *Qualitätskriterien und Herausforderungen im RE* - bildete den Hauptteil der Interviews. Die in der Literatur identifizierten RE-Herausforderungen und RE-Qualitätskriterien wurden mit der Praxis verglichen. Dafür wurden zunächst RE-Herausforderungen und -Qualitätskriterien in offenen Fragen erörtert, die bei den Interviewten in der Praxis bereits aufgetreten sind. Dafür richtete sich die Frage nach den Qualitätskriterien an bereits in der Wissenschaft gestellten Forschungsfragen, wie beispielsweise:

- „*Welche sind die bekannten Qualitätskriterien für die agile Anforderungsspezifikation?*“ (Heck und Zaidman 2018, S. 80) und
- „*Was ist User Story Qualität?*“ (Lucassen, Dalpiaz, Van Der Werf et al. 2015, S. 1).

Dieser Ansatz wird ebenfalls für die Fragen nach den RE-Herausforderungen und deren Bewältigung gewählt. Dabei bezieht sich die Frage zu den Herausforderungen explizit auf die Teilforschungsfrage (TF 1):

- „*Was sind die praktischen Herausforderungen des agilen RE?*“ (Inayat et al. 2015, S. 924),
- „*Was sind die Herausforderungen der agilen Methodik für das RE?*“ (Elghariani und Kama 2016, S. 511),
- „*Was sind die Hauptherausforderungen im agilen RE?*“ (Schön, Winter et al. 2017, S. 38) und
- „*Wie können wir die ermittelten zentralen Herausforderungen bewältigen?*“ (Schön, Winter et al. 2017, S. 38).

Daraufhin folgte der Abgleich der in der Literatur identifizierten RE-Qualitätskriterien (Tabelle 9) und RE-Herausforderungen (Tabelle 5) mit der Praxis durch die Experten.

Die letzte Kategorie richtete sich, je nach Vorerfahrung der interviewten Person, auf eine erste Einschätzung des Einsatzes von *Einsatz von GenAI im RE*. Es wurde ein Szenario angenommen, in dem das in dieser Arbeit zu entwickelnde Ergebnisartefakt (siehe Kapitel 4.3) bereits existiert. Zur Beantwortung der Teilforschungsfrage (TF 3) wurden die Experten zu den Vor- und Nachteilen einer Webanwendung zur GenAI-gestützten Anforderungserhebung befragt. Dies diente der Erhebung potenzieller Nutzungsmöglichkeiten sowie der Aufnahme möglicher zusätzlicher Funktionalitäten eines solchen Systems. Abschließend wurde das Interesse der Interviewpartner an einer Evaluation des Ergebnisartefakts eruiert.

3.2.2 Experteninterviews - GenAI

Das Ziel der GenAI-Interviews war die Analyse der Potenziale von GenAI-Systemen, die RE-Herausforderungen zu adressieren. Zusätzlich wurden Risiken, die mit einem solchen Einsatz verbunden sein können, untersucht. Die Interviews wurden mit sechs Experten aus dem Bereich der GenAI in einem Zeitraum von zwei Wochen durchgeführt.

Vorbereitung der Experteninterviews

Ein Unterschied zu den RE-Experteninterviews (siehe Kapitel 3.2.1) lag in der Auswahl der Experten. Diese wurden sowohl über das firmeninterne Netzwerk als auch durch den vom Unternehmen bereitgestellten Betreuer akquiriert. Alle interviewten Experten waren bereits in verschiedenen Rollen und Projekten im Kontext von GenAI involviert. Eine pseudonymisierte Übersicht der interviewten Personen ist in Tabelle 3 dargestellt. Da es sich ebenfalls um Experteninterviews handelte, wurde der Identifikator beibehalten. Lediglich die fortlaufende Nummerierung wurde entsprechend angepasst.

Identifikator	Rolle im Bereich GenAI	Erfahrung in GenAI
Ei6	Software Developer	Seit einem Jahr.
Ei7	Software Engineer	Seit einem Jahr.
Ei8	(Managing) Consultant	Seit einem Jahr.
Ei9	Team-Lead	Seit zwei Jahren.
Ei10	(Senior) Consultant	Seit einem Jahr.
Ei11	(Senior) Manager	Seit einem Jahr.

Tabelle 3: Interviewpartner GenAI

Datenerhebung

Im Vordergrund dieser Interviews stand die Beantwortung der Teilforschungsfrage (TF 2). Im Rahmen dieser sollte die Anwendbarkeit von GenAI auf die bereits identifizierten RE-Herausforderungen eingeschätzt werden. Der Interviewleitfaden (siehe Kapitel Anhang B.5) war ebenfalls in Form eines Fragenkatalogs strukturiert. Dieser umfasste neun Fragen, welche in zwei Kategorien unterteilt wurden.

Die erste Kategorie *Allgemein GenAI/KI* diente als Einstieg in das Interview und zielte darauf ab, die Vorerfahrung der Interviewteilnehmer zu ermitteln. Zusätzlich diente die Frage nach der bisherigen Erfahrung dazu, die bisher gesammelten Erkenntnisse und das Wissen im Bereich der KI und GenAI zusammenzufassen, um die Expertise der befragten Personen zu bestimmen.

In der zweiten Kategorie lag der Fokus auf dem *Einsatz von GenAI im RE*. Hierbei lag der Fokus darauf zu beantworten, inwiefern der Einsatz von GenAI geeignet ist, die RE-Herausforderungen zu minimieren oder zu lösen. Gleichzeitig wurden in diesem Zusammenhang die zu berücksichtigenden Risiken bei der Verwendung von GenAI sowie die Möglichkeiten ihrer technischen Bewältigung betrachtet.

3.3 Entwicklung der Lösung

Das Ergebnisartefakt dieser Arbeit ist ein Prompt innerhalb der bereits genannten Webanwendung. Dieser Prompt wird mit Informationen gefüllt, die von den Nutzern eingegeben werden, und an eine GenAI übermittelt. Diese generiert basierend auf den erhaltenen Daten eine User Story. Die Webanwendung gibt diese an den Nutzer aus.

Zur Entwicklung dieser Lösung wurden insbesondere die Erkenntnisse aus der Literaturrecherche und die Design Prinzipien aus den RE-Experteninterviews sowie den GenAI-Experteninterviews einbezogen. Dazu gehörten auch die Empfehlungen zur Bewältigung von Risiken im Einsatz von GenAI-Systemen.

Die Zusammenführung der Erkenntnisse der Experteninterviews mit der wissenschaftlichen Literatur bildeten die Grundlage für die Erstellung des Prompts. Ein wesentliches Instrument für die Entwicklung des Prompts waren die Richtlinien von Arvidsson und Axell (2023), die auf das Ergebnisartefakt angewendet wurden. Diese Richtlinien wurden im Rahmen dieser Arbeit genutzt, da sie speziell auf den Prozess im RE ausgerichtet sind (Arvidsson und Axell 2023).

Zusätzlich wurde auf Strategien und Praktiken führender Unternehmen wie Google und OpenAI zurückgegriffen, da diese zur Verbesserung der Qualität der Ergebnisse von Sprachmodellen beitragen können (Google 2024; OpenAI 2024). Die Einbindung dieser bewährten Methoden und Richtlinien in Kombination mit der Verwendung der in dieser Arbeit erarbeiteten Design Prinzipien unterstützte die zielgerichtete Entwicklung des Prompts.

3.4 Evaluation der Webanwendung

Der Nutzen, die Qualität und die Wirksamkeit eines Ergebnisartefakts können durch Evaluation nachgewiesen werden (Hevner et al. 2004). Im Interesse der hohen Forschungsgüte wurden Elemente von FEDS, einschließlich Evaluationsstrategien und Evaluationsziele, gemäß Venable et al. (2016) berücksichtigt. Venable et al. (2016) identifizieren die Evaluation eines Ergebnisartefakts als Schlüsselaktivität im Prozess des DSR. Um den Anforderungen einer Evaluation nach Venable et al. (2016) gerecht zu werden, wären eigentlich mehrere

Evaluationsepisoden nötig. Da innerhalb dieser Forschungsarbeit nur eine Evaluation durchgeführt wurde, orientiert sich die Methodik lediglich an den Konzepten von FEDS.

Im Hinblick auf den Zeitpunkt der Evaluation erläutern Venable et al. (2016) *ex ante* und *ex post* Evaluation. *Ex ante* Evaluation findet statt, bevor Design und Konstruktion der Artefakte beginnen (Venable et al. 2016). *Ex post* Evaluation hingegen wird nach dem Design und der Konstruktion durchgeführt, um die Auswirkungen und den Wert des Systems zu beurteilen (Venable et al. 2016). Im Rahmen dieser Arbeit wurde die Durchführung einer *ex post* Evaluation gewählt, da das Ergebnis der Artefakterstellung bewertet werden sollte.

Venable et al. (2016) differenzieren bei den Gründen für eine Evaluation zwischen einem formativen und einem summativen Vorgehen. Formative Evaluationen sind darauf ausgerichtet, das Ergebnisartefakt zu verbessern, indem eine Entscheidungsgrundlage geschaffen wird (Venable et al. 2016). Im Kontrast dazu zielen summative Evaluationen darauf ab, zu beurteilen, inwieweit das Ergebnis den Erwartungen entspricht (Venable et al. 2016). Sie werden am Ende des Entwicklungsprozesses durchgeführt, um das Artefakt zu bewerten, während eine formative Evaluation vor und während der Artefakterstellung stattfindet (Köller 2009; Scriven 1972; Venable et al. 2016). Das Ergebnisartefakt basiert auf zuvor entwickelten Design Prinzipien und daraus abgeleiteten Design Features und die Gesamtheit der Einarbeitung dieser in das Ergebnisartefakt sollte bewertet werden. Aus diesem Grund handelte es sich im Rahmen dieser Forschungsarbeit um eine *summative* Evaluation.

Venable et al. (2016) differenzieren zwischen naturalistischer und künstlicher Evaluation. Die künstliche Evaluation wird für die Überprüfung von Teilen des Artefakts genutzt (Venable et al. 2016). Die naturalistische Evaluation hingegen bewertet die Leistung eines Artefakts in einem realen Kontext, beispielsweise innerhalb einer Organisation und unter Einbeziehung realer Nutzer (Venable et al. 2016). Im Rahmen der Evaluation wurden die Teilnehmer einem realitätsnahen Szenario unterzogen, was dem Ansatz der *naturalistischen* Evaluation entspricht.

Die Evaluation verfolgte das Ziel der Sicherstellung der *Forschungstreue* und *Wirksamkeit* des Vorhabens. Für die Demonstration von Forschungstreue im Rahmen der Evaluation war es erforderlich zwei Aspekte zu belegen:

1. Zum einen war es notwendig sicherzustellen, dass die beobachteten Ergebnisse ausschließlich auf das Artefakt zurückzuführen sind, indem unabhängige Faktoren oder besondere Umstände ausgeschlossen wurden (Venable et al. 2016). Dies diente der Bestätigung der Wirksamkeit des Artefakts (Venable et al. 2016).
2. Zum anderen musste die Effektivität des Artefakts unter realen Anwendungsbe-

dingungen demonstriert werden (Venable et al. 2016). Damit wurde zudem eine naturalistische Prüfung der Einsatzfähigkeit der Webanwendung in der Praxis verfolgt (Venable et al. 2016).

Für die Evaluation der Webanwendung wurden zehn Personen befragt, die im RE tätig sind. In Tabelle 4 sind die Evaluationsteilnehmer aufgeführt. Da die Teilnehmer Et1 bis Et5 mit den Interviewteilnehmern Ei1 bis Ei5 identisch sind, wird auf eine erneute Auflistung an dieser Stelle verzichtet. Informationen dieser Teilnehmer finden sich in Tabelle 2.

Identifikator	Rolle im RE	Qualifikation / Beschreibung
Et6	Requirements Engineer, Product Owner, Lead Frontend Entwickler	Et6 hat sechs Jahre Erfahrung im RE. Et6 sammelt seit 2019 Erfahrung mit GenAI, insbesondere in der Anwendung von AI Dungeon, GPT-2 und weiteren Sprachmodelle zur Codegenerierung.
Et7	Requirements Engineer, Product Owner, Business Designer	Et7 ist seit sieben Jahren im RE tätig. Mit GenAI hat Et7 bislang begrenzt Erfahrung gesammelt.
Et8	Requirements Engineer, fachliche Anforderungsanalyse	Seit dreieinhalb Jahren ist Et8 im RE aktiv. Et8 besitzt konzeptionelle und methodische Kenntnisse in GenAI, die hauptsächlich während des Studiums erworben wurden.
Et9	Product Owner, Anforderungsanalyse, Projektmanager	Et9 arbeitet seit drei Jahren im Bereich des RE und hat sich bereits intensiv mit GenAI auseinandergesetzt. Et9 hat bereits wissenschaftliche Artikel verfasst, neuronale Netze trainiert und nutzt GenAI zur Erstellung von Anforderungsspezifikationen und User Stories.
Et10	Requirements Engineer, objektorientierter Analytiker, Projektkoordinator	Et10 hat zehn Jahre Erfahrung im Bereich des RE. Mit GenAI hat Et10 bisher noch wenig Erfahrung gesammelt.

Tabelle 4: Evaluationsteilnehmer

Den Teilnehmern wurde eine Situation präsentiert, die eine Aufgabenstellung aus dem Berufsalltag simulierte. Die Aufgabenstellung, deren konkrete Formulierung im Anhang B.6.2 zu finden, beinhaltete die Erstellung von User Stories zu einem Anwendungsfall. Innerhalb der Evaluation sollten die Experten die User Stories manuell validieren, bei Bedarf anpassen und dann übernehmen, wenn die von ihnen gewünschte Qualität erreicht wurde. Dabei war

die Qualität der User Stories wichtiger als die Quantität. Die Anpassung simulierte die reale Anwendung von GenAI im RE der Teilnehmer im gegebenen Anwendungsfall. Es wurde erwartet, dass alle erstellten User Stories in einem einheitlichen Dateiformat abgegeben werden, um als wesentlicher Bestandteil der Evaluationsauswertung betrachtet werden zu können. Insgesamt standen den Evaluationsteilnehmern 25 Minuten zur Ausführung dieser Aufgabe zur Verfügung. Der Umgang mit den generierten User Stories diente als Ausgangspunkt für die Beantwortung der Teilforschungsfrage (TF 3).

Der Anwendungsfall mit dem die Teilnehmer arbeiteten simuliert die Digitalisierung in der Gastronomiebranche. Konkret sollte ein System entwickelt werden, das anfallende Aufgaben in der Gastronomie erfassen und ausführen kann. Dieser Anwendungsfall ist im Anhang B.6.3 dokumentiert. Es wurde eine Benutzeroberfläche bereitgestellt, um die Zugänglichkeit zum Prompt und somit zum Sprachmodell zu maximieren. Dabei wurde innerhalb der mitgegebenen Aufgabenstellung explizit klargestellt, dass das Hauptaugenmerk nicht auf einer vollumfänglich ausgearbeiteten Benutzerfreundlichkeit, sondern der Erstellung von qualitativ hochwertigen User Stories lag.

In dieser Evaluation wurde die *Thinking-Aloud* Methode nach Lewis (1982) genutzt. Bei dieser Methode wird eine Testperson angewiesen, während der Interaktion mit einem System ihre Gedanken auszusprechen (Lewis 1982). Dies ermöglichte Einblicke in die Fragen, Bedenken, Überlegungen und kognitiven Prozesse der Testperson (Lewis 1982). Außerdem wurden eine Beobachtung und Analyse der Benutzererfahrung in Echtzeit möglich (Lewis 1982). Im Kontext der Webanwendung wurden die Eingaben und der Umgang mit der Ausgabe der Testpersonen zur Erfüllung der Aufgabe beobachtet und analysiert. Dies diente der Beantwortung der Teilforschungsfrage (TF 3), um die Wahrnehmung der Unterstützung durch GenAI im RE während der Nutzung des Tools zu untersuchen. Um eine Auswertung zu ermöglichen wurden die Simulationen aufgezeichnet.

Um die Integrität der Evaluationsdurchführung zu gewährleisten und Interferenzen durch externe Anfragen zu vermeiden, wurde ein Authentifizierungssystem implementiert. Für jede Evaluation wurde ein einzigartiger Token generiert. Dieser wurde vorab mittels einer Testabfrage in der Webanwendung überprüft, um dessen Funktionalität sicherzustellen. Zur Gewährleistung konsistenter Evaluationsbedingungen wurde zudem ein Timer eingesetzt, welcher unmittelbar nach Überprüfung der Testabfrage aktiviert wurde.

Im Anschluss an die Evaluation wurden Interviews geführt, um Feedback zu sammeln. Der Fokus der Interviews wurde dabei auf die wahrgenommene Qualität der erstellten User Stories, die Produktivität im Erstellungsprozess und die Interaktion der Teilnehmer mit der Webanwendung gesetzt. Diese Interviews wurden transkribiert und mittels qualitativer Inhaltsanalyse nach Mayring (2010) ausgewertet. Diese Transkriptionen sind in der separat angehängten Datei dokumentiert.

4 Beschreibung des Ergebnisartefakts

In diesem Kapitel wird der Weg von den Teilartefakten bis hin zum finalen Ergebnisartefakt vorgestellt. Beim Ergebnisartefakt handelt es sich um einen Vorschlag RE-Herausforderungen mittels GenAI zu lösen oder zu mindern. Dieser Prozess orientiert sich an der in Abbildung 1 dargestellten DSR-Methodik (Kuechler und Vaishnavi 2008).

Das Kapitel beginnt mit der Darstellung des Problembewusstseins in Kapitel 4.1. Anschließend werden in Kapitel 4.2, unter anderem basierend auf den Empfehlungen aus den Experteninterviews, Design Prinzipien abgeleitet. Diese Design Prinzipien bilden die Grundlage für die Entwicklung des Ergebnisartefakts, welches in Kapitel 4.3 mithilfe der aufgestellten Design Features beschrieben wird.

4.1 Problembeschreibung

Dieser Abschnitt beginnt mit der Darstellung der Problembeschreibung, die das Bewusstsein für die bestehenden RE-Herausforderungen schärft. Im Mittelpunkt steht dabei die Beantwortung der Teilforschungsfrage (TF 1), um eine Liste an identifizierten RE-Herausforderungen bereitzustellen. Zusätzlich werden auch RE-Qualitätskriterien innerhalb der Literaturrecherche erarbeitet. Letztere definieren, was eine qualitativ hochwertige Erhebung und Verschriftlichung von Anforderungen ausmacht. Diese RE-Herausforderungen und -Qualitätskriterien gehen aus einer Literaturrecherche und einem darauffolgenden Abgleich im Rahmen der RE-Experteninterviews hervor.

4.1.1 RE-Herausforderungen - Ergebnisse Literaturrecherche

Zuerst werden die aus der Literaturrecherche identifizierten RE-Herausforderungen dargestellt. Insgesamt werden in elf Artikeln 26 verschiedene RE-Herausforderungen ermittelt, die in einer Konzeptmatrix nach Webster und Watson (2002) in Tabelle 5 ausgewertet sind.

Tabelle 5: Teilartefakt: Konzeptmatrix - RE-Herausforderungen in der Literatur (Eigene Darstellung in Anlehnung an Webster und Watson (2002))

RE-Herausforderung	Artikel										
	Amna und Poels (2022)	Inayat et al. (2015)	Hoy und Xu (2023)	Rasheed et al. (2021)	Shah und Patel (2014)	Elghariani und Kama (2016)	Alam et al. (2017)	Schön, Winter et al. (2017)	Curcio et al. (2018)	Soares et al. (2015)	L. Cao und Ramesh (2008)
Unbestimmtheit	X			X					X		
Inkonsistenz	X						X				
Unzulänglichkeit	X										
Duplikate	X										
Mehrdeutigkeit	X			X	X		X			X	
Minimale Dokumentation		X	X	X		X			X	X	X
Vernachlässigung nicht-funktionaler Anforderungen		X	X		X	X	X		X	X	X
Bewältigung von Anforderungsänderungen (und ihrer Bewertung)		X	X	X	X	X	X	X		X	
(Ungeeignete Methode der) Prioritätensetzung			X	X	X					X	X
Schlecht formulierte Anforderungen			X	X	X						
Weniger erfahrenes und qualifiziertes Team				X							
Fehlende Anforderungen				X	X		X				
Kundenverfügbarkeit		X	X		X	X	X	X		X	
Schätzung von Budget und Zeitplan		X	X			X	X				X
Ungeeignete Architektur		X	X			X					

Fortsetzung auf der nächsten Seite

Tabelle 5: Teilartefakt: Konzeptmatrix - RE-Herausforderungen in der Literatur (Eigene Darstellung in Anlehnung an Webster und Watson (2002)) - Fortsetzung

RE-Herausforderung	Artikel										
	Amna und Poels (2022)	Inayat et al. (2015)	Hoy und Xu (2023)	Rasheed et al. (2021)	Shah und Patel (2014)	Elghariani und Kama (2016)	Alam et al. (2017)	Schön, Winter et al. (2017)	Curcio et al. (2018)	Soares et al. (2015)	L. Cao und Ramesh (2008)
Unfähigkeit und Zustimmung des Kunden		X		X							
Vertragliche Beschränkungen und Volatilität der Anforderungen		X								X	
Fehlendes (Kunden-)wissen			X							X	
Kommunikation (-smethoden)			X	X	X		X			X	
Prototyping Problematik				X							X
Weniger Vorplanung, Konzentration und keine anfängliche Beteiligung des Teams				X							
Implizites Wissen				X							
Funktionale und technische Abhängigkeiten zwischen Teams								X			
Stakeholderakzeptanz, dass Teams eigene Entscheidungen treffen								X			
Wahrung Gesamtüberblick								X			
Zusammenarbeit mit den Endnutzern								X		X	

Mehrdeutigkeit wird in verschiedenen Arbeiten als RE-Herausforderung thematisiert (Alam et al. 2017; Amna und Poels 2022; Rasheed et al. 2021; Shah und Patel 2014; Soares et al. 2015). Mehrdeutigkeit kann dabei als „Problem bei der Formulierung von Anforderungen in Form von User Stories, die zu zweifelhaften, ungenauen und mehrfachen

Interpretationen dieser Anforderungen führen“ definiert werden (Amna und Poels 2022, S. 2). Diese RE-Herausforderung lässt sich in vier Problembereiche aufteilen: *Unbestimmtheit*, *Inkonsistenz*, *Unzulänglichkeit* und *Duplikate* (Amna und Poels 2022). Diese Problembereiche werden auch von einzelnen wissenschaftlichen Arbeiten entsprechend als eigene RE-Herausforderung gelistet (Alam et al. 2017; Rasheed et al. 2021; Soares et al. 2015). Aus diesem Grund werden diese auch als eigene RE-Herausforderungen in der Konzeptmatrix gelistet. In der Literatur gibt es zudem Ansätze, die die Mehrdeutigkeit als RE-Herausforderung bestätigen und sich zusätzlich mit Möglichkeiten zur Adressierung dieser Herausforderung befassen (Bano 2015).

Hoy und Xu (2023) liefern in ihrer Literaturübersicht eine detaillierte Auflistung, wie häufig bestimmte RE-Herausforderungen genannt werden. Zu den häufigsten Nennungen gehören: *Vernachlässigung nicht-funktionaler* (25 Nennungen), *minimale Dokumentation* (20 Nennungen) und *Bewältigung von Anforderungsänderungen* (23 Nennungen). Die Berücksichtigung in der Literatur unterstreicht die Bedeutung der RE-Herausforderungen.

Zusätzlich zu Herausforderungen zur Interpretation und Verschriftlichung von Anforderungen werden auch Herausforderungen betrachtet, die sich auf die Praktiken und Prozesse im RE beziehen (Inayat et al. 2015). Beispiele hierfür sind *Kundenverfügbarkeit* sowie die *Schätzung von Budget und Zeitplan*.

Weitere spezifische RE-Herausforderungen, die im Detail erläutert werden, umfassen die *Protootyping Problematik*, eine *ungeeignete Architektur* und ein *weniger erfahrenes und qualifiziertes Team*. Diese Erläuterung ist notwendig, da aus dem Namen der Herausforderung nicht unmittelbar alle Aspekte und Probleme ersichtlich sind. Beispielsweise kann eine frühe Bereitstellung eines Prototypen im agilen RE zu Fehleinschätzungen der Entwicklungsgeschwindigkeit durch Stakeholder führen (Rasheed et al. 2021). Frühe Architekturentscheidungen können beispielsweise durch das Aufkommen neuer Anforderungen obsolet werden, was zu Umgestaltungskosten führen kann (Hoy und Xu 2023; Inayat et al. 2015). Ein weniger erfahrenes und qualifiziertes Team kann Schwierigkeiten bei der adäquaten Formulierung von User Stories haben (Rasheed et al. 2021). Dies kann Missinterpretationen und Kommunikationsproblemen verursachen, was letztlich die Qualität von Projektergebnissen beeinträchtigen kann (Rasheed et al. 2021).

Diese Vielfalt an Herausforderungen macht das RE zu einer komplexen Problemstellung. Die Lösung oder Minderung dieser RE-Herausforderungen kann in verschiedenen Aspekten Mehrwerte bringen und bietet die Grundlage für die Entwicklung von Design Prinzipien für einen Lösungsvorschlag (Meth et al. 2015).

4.1.2 RE-Herausforderungen - Ergebnisse RE-Experteninterviews

Aus den RE-Experteninterviews ergeben sich ebenfalls Erkenntnisse zu Herausforderungen im RE. Zunächst wird die grundlegende Auswertung dieser Interviews präsentiert, indem das nach Mayring (2010) entwickelte Kategoriensystem beschrieben wird.

Die erste Kategorie, K1 *Expertise*, dient der Einordnung der Interviewteilnehmer als Experten für RE und der optionalen Vorerfahrung mit GenAI. Dabei werden vor allem in der RE Praxis verwendete Anforderungsdokumente, Methoden und Tools untersucht. Zusätzlich wird die bisherige Erfahrung der Teilnehmer mit GenAI im beruflichen sowie im privaten Kontext analysiert.

Kategorie K2 betrachtet die *Praxis im RE* der Teilnehmer, einschließlich der bereits genutzten RE-Qualitätskriterien, der bereits begegneten RE-Herausforderungen und deren Bewältigungsstrategien.

Die dritte Kategorie, K3 *Einsatzmöglichkeiten von GenAI*, erörtert die Potenziale und Grenzen von GenAI im Kontext des RE. Dabei wird zwischen den adressierten RE-Herausforderungen (Subkategorie K3.1 *Potenziale zur Bewältigung Herausforderungen*) und sonstigen allgemeinen Einsatzmöglichkeiten unterschieden.

Insgesamt werden 260 Segmente aus fünf Transkriptionen für die Kodierung in Tabelle 6 verwendet. Dabei kann es vorkommen, dass einzelne Segmente mehreren Kategorien zugeordnet werden. Die Aufschlüsselung der Zuordnung der Textsegmente ist der angehängten Datei zu entnehmen.

Hauptkategorie	Subkategorie 1. Ebene	Subkategorie 2. Ebene
K1 Expertise - 0	K1.1 Beruflicher Hintergrund im RE - 18	K1.1.1 Art der Anforderungsdokumente - 27
		K1.1.2 Einsatz von Methoden und Tools - 20
	K1.2 Erfahrung in GenAI - 0	K1.2.1 Privater Kontext - 8
		K1.2.2 Beruflicher Kontext - 12
K2 Praxis im RE - 0	K2.1 Qualitätskriterien - 27	-
	K2.2 Herausforderungen - 34	-
	K2.3 Bewältigungsstrategien für Herausforderungen - 16	-
K3 Einsatzmöglichkeiten von GenAI - 28	K3.1 Potenziale zur Bewältigung der genannten Herausforderungen - 20	
	K3.2 Integration von GenAI in die eigene Arbeitsweise - 5	K3.2.1 Nachteile - 13
		K3.2.2 Vorteile - 13
	K3.3 Grenzen - 7	-
K3.4 Funktionalitäten in neuer Anwendung - 12	-	
Summe: 28	Summe: 141	Summe: 93

*Notationsweise: KX: Hauptkategorie KX.Y: Subkategorie KX.Y.Z: Subsubkategorie - #: Häufigkeit

Tabelle 6: Kategoriensystem und Häufigkeiten der Kodierungen der RE-Experteninterviews (Eigene Darstellung)

Zunächst wurden die Experten gefragt, welchen Herausforderungen sie in ihrem beruflichen Kontext im RE bereits begegnet sind. Die dabei genannten RE-Herausforderungen wurden mit denen aus der Literatur (siehe Tabelle 5) abgeglichen. Eine Zuordnung dieser ist in Tabelle 7 dargestellt. Eine Übersicht über die in den Experteninterviews neu identifizierten RE-Herausforderungen ist in Tabelle 8 aufgeführt.

RE-Herausforderung	Referenz
Unbestimmtheit	Ei5 - Z. 171 ff.
Inkonsistenz	Ei3 - Z. 209 f.
Duplikate	Ei1 - Z. 85 ff., 140 ff.
Minimale Dokumentation	Ei2 - Z. 144 Ei4 - Z. 121
Vernachlässigung nicht-funktionaler Anforderungen	Ei3 - Z. 147 ff. Ei4 - Z. 123 ff.
Ungeeignete Methode der Prioritätensetzung	Ei3 - Z. 203 ff.
schlecht formulierte Anforderungen	Ei1 - Z. 87 ff.
fehlende Anforderungen	Ei1 - Z. 178 ff. Ei4 - Z. 120 f.
Kundenverfügbarkeit	Ei2 - Z. 140 ff. Ei4 - Z. 118 f.
Fehlendes (Kunden-)wissen	Ei5 - Z. 156 ff., 180 ff.
Kommunikation (-smethoden)	Ei3 - Z. 180 ff.
Prototyping Problematik	Ei2 - Z. 143 f.

Tabelle 7: Teilartefakt: Zuordnung der genannten RE-Herausforderungen aus den RE-Experteninterviews mit der Literatur

Daraufhin werden in den Interviews die in der Literatur identifizierten RE-Herausforderungen mit den RE-Experten direkt abgeglichen. Der Abgleich dient der Überprüfung, ob die RE-Herausforderungen auch in der Praxis auftreten. Während dieses Prozesses werden die einzelnen Herausforderungen den Experten direkt vorgelesen und bei Bedarf erläutert. Dies kann ein mögliches Bias und dementsprechend eine Überschätzung der tatsächlich auftretenden Herausforderungen zur Folge haben. Die Ergebnisse werden daher als Tendenzaussagen betrachtet, um aufzuzeigen, dass die literaturbasierten RE-Herausforderungen grundsätzlich in der Praxis erkennbar sind. Zusätzlich handelt es sich hierbei um eine Auswertung, deren begrenzter Stichprobenumfang keine allgemeingültigen Schlüsse zulässt. Dieser Abgleich ist im Anhang in Tabelle B.1 dokumentiert.

RE-Herausforderung	Referenz
Kunde kennt seine Anforderungen nicht	Ei1 - Z. 92 ff., 188 ff., Ei4 - Z. 112 ff.
Kunde denkt eher in Lösungen, als in Anforderungen	Ei2 - Z. 107 ff., 132 ff.
Fachabteilungen können Use Cases nicht formulieren	Ei2 - Z. 134 f.
Fachabteilungen können die Prozesse nicht beschreiben	Ei2 - Z. 138 ff.
Fehlende Nachvollziehbarkeit von früher umgesetzten Anforderungen	Ei1 - Z. 90 ff., Ei2 - Z. 145 f., Ei3 - Z. 222 ff., Ei4 - Z. 130 ff.
Zu späte Einbindung Softwareentwicklerteams	Ei2 - Z. 146 ff.
Anforderungen werden aus der falschen Rolle formuliert	Ei3 - Z. 204 ff.
Unreflektierte Prozessübernahme	Ei3 - Z. 210 f., 218ff.

Tabelle 8: Teilartefakt: Neu identifizierte RE-Herausforderungen aus den RE-Experteninterviews

4.1.3 RE-Qualitätskriterien - Ergebnisse Literaturrecherche

In der durchgeführten Literaturrecherche werden insgesamt 19 RE-Qualitätskriterien für die Formulierung von Anforderungsdokumenten in vier Artikeln ermittelt. Diese sind in einer Konzeptmatrix nach Webster und Watson (2002) zusammengetragen und in Tabelle 9 dargestellt.

In der Praxis werden zur Qualitätskontrolle von User Stories oder anderen Anforderungsdokumenten verschiedene Modelle genutzt. Hervorzuheben ist das INVEST-Modell nach Wake (2003), da sich einige der Kriterien den in der Literatur vorkommenden RE-Qualitätskriterien zuordnen lassen. Wake (2003) schlägt im Rahmen des INVEST-Modells folgende Kriterien für eine qualitativ hochwertige User Story vor: *unabhängig*, *verhandelbar*, *wertvoll*, *schätzbar*, *klein* und *prüfbar*. Dieses Modell wird zudem als gängiges Qualitätskriterium integriert und ist nicht nur für User Stories, sondern auch auf andere Anforderungsdokumente anwendbar (Heck und Zaidman 2018). Heck und Zaidman (2018) klassifizieren in diesem Zusammenhang Qualitätskriterien für agile Anforderungsspezifikationen in drei Hauptkategorien:

1. *Vollständigkeit*: Diese Kategorie bezieht sich auf strukturelle Elemente einer Anforderung wie eine Identifikationsnummer und die Priorität eines Anforderungsdokuments.
2. *Einheitlichkeit*: Hierbei wird ein standardisierter Stil von Anforderungsdokumenten berücksichtigt.
3. *Konsistenz und Korrektheit*: Diese umfasst Aspekte wie die Verständlichkeit und die Einzigartigkeit von Anforderungen, ergänzt durch das INVEST-Modell.

RE-Qualitätskriterium	Artikel			
	(Lucassen, Dalpiaz, Van Der Werf et al. 2015)	(Hallmann 2020)	(Heck und Zaidman 2018)	(Vieira et al. 2012)
Syntax	X			
Atomarität	(X)			X
Prägnanz	(X)			X
Semantik	X	X		
Konsistenz und Korrektheit	(X)		X	X
Klarheit	(X)			X
Pragmatik	X			
Vollständigkeit	(X)		X	
Einheitlichkeit	(X)		X	
Einzigartigkeit	(X)			X
Lexikalität		X		
Formalität		X		
Sättigung		X		
INVEST		X		
Card, Conversation, Confirmation		X		
Cohn's Guideliness		X		
Durchführbarkeit				X
Abstraktheit				X
Überprüfbarkeit				X

Tabelle 9: Teilartefakt: Konzeptmatrix - RE-Qualitätskriterien in der Literatur (Eigene Darstellung in Anlehnung an Webster und Watson (2002))

Zudem wird die Entwicklung und Implementierung von Tools zur Überprüfung der Qualität von Anforderungsdokumenten ebenfalls von Heck und Zaidman (2018) diskutiert. Solche Tools können dazu beitragen, die Einhaltung der definierten Qualitätskriterien zu automatisieren.

Lucassen, Dalpiaz, Van Der Werf et al. (2015) führen übergeordnete Qualitätskriterien der Syntax, Semantik und Pragmatik für User Stories ein. Diesen sind jeweils Unterkriterien zugeordnet. Einige der Unterkriterien werden auch in anderen Forschungsarbeiten aufgelistet. Diese sind dabei mit (X) in der ersten Spalte der Matrix gekennzeichnet. Die Unterkriterien, die in anderen Arbeiten nicht explizit genannt werden, werden in der Konzeptmatrix nicht detaillierter aufgelistet, um eine Übersichtlichkeit zu wahren. Die ermittelten Unterkriterien werden jedoch zur vollständigen Dokumentation erörtert:

Die *Syntax* bezieht sich auf die formale Struktur von User Stories. Eine User Story sollte eine Funktionalität beschreiben (*Atomarität*), sich auf die Rolle, die Funktionalität und den Nutzen beschränken (*Prägnanz*) und eine *wohlgeformte Aussage* bilden (Lucassen, Dalpiaz, Van Der Werf et al. 2015).

In der *Semantik* wird die Bedeutung der in den User Stories verwendeten Sätze und Begriffe überprüft. Sie sollten *klar und konsistent* sein, keine Konflikte mit anderen User Stories aufweisen (*konfliktfrei*) und sich auf das Problem konzentrieren (*problemorientiert*), nicht auf die Lösung. Ein weiteres Ziel ist es, *Mehrdeutigkeit zu vermeiden*, damit jede User Story eine Funktion und deren Begründung korrekt ausdrückt (Lucassen, Dalpiaz, Van Der Werf et al. 2015). An dieser Stelle besteht eine direkte Verbindung zu den RE-Herausforderungen, insbesondere in Bezug auf die Mehrdeutigkeit und Duplikate.

Die *Pragmatik* betrachtet die Praxistauglichkeit von User Stories. Diese sollten *vollständig* sein und *keine unnötigen Abhängigkeiten* zu anderen User Stories aufweisen. Das bedeutet, dass sie handhabbar sein sollten, um zu gewährleisten, dass die daraus resultierenden Anwendungen bzw. deren Funktionalitäten effektiv geplant und umgesetzt werden können (Lucassen, Dalpiaz, Van Der Werf et al. 2015).

4.1.4 RE-Qualitätskriterien - Ergebnisse RE-Experteninterviews

In Analogie zu den RE-Herausforderungen werden die Experten ebenfalls zu RE-Qualitätskriterien befragt. Obwohl keine neuen Qualitätskriterien identifiziert werden, ist eine Zuordnung auf bestehende RE-Qualitätskriterien möglich. Diese Zuordnung ist in Tabelle 10 dokumentiert.

Ein Interviewteilnehmer betont, dass eine User Story einen Mehrwert liefern muss (Ei2 - Z. 109 f.). Das Ziel einer User Story ist es, einen Mehrwert für die im Dokument definierte Zielgruppe zu liefern (Cohn 2004). Daher wird dies nicht als separates Qualitätskriterium

Qualitätskriterien	Referenz
Atomarität	Ei4 - Z. 85 f. Ei5 - Z. 132 ff.
Konfliktfreiheit	Ei2 - Z. 111 f. Ei4 - Z. 86
Problemorientiert	Ei2 - Z. 107 f. Ei3 - Z. 179 ff.
Konsistenz und Korrektheit	Ei4 - Z. 86 ff. Ei5 - Z. 142 f.
Konzeptionell stimmig	Ei1 - Z. 61, 68 Ei2 - Z. 114 ff. Ei4 - Z. 90 ff. Ei5 - Z. 135 ff.
Vollständigkeit	Ei1 - Z. 60 Ei2 - Z. 101 ff.
Einheitlichkeit	Ei3 - Z. 173 ff.
Klarheit	Ei1 - Z. 60, 68 Ei2 - Z. 99 Ei5 - Z. 128 ff.
Prägnanz	Ei4 - Z. 84 f. Ei5 - Z. 132
Überprüfbarkeit	Ei1 - Z. 57 f., 68 Ei5 - Z. 138 f.

Tabelle 10: Teilartefakt: Zuordnung RE-Qualitätskriterien aus den RE-Experteninterviews mit der Literatur

betrachtet. Ein anderer Teilnehmer nennt die Zugänglichkeit von User Stories für alle Stakeholder als Qualitätskriterium (Ei5 - Z. 137 f.). Da dies jedoch nicht die Formulierung von Anforderungsdokumenten betrifft, wird es nicht als neues Qualitätskriterium aufgenommen. Die Unsicherheit des Interviewteilnehmers bezüglich der Einstufung dieses Aspekts als Qualitätskriterium bestätigt diese Entscheidung ebenfalls (Ei5 - Z. 138).

Anschließend erfolgt ein Abgleich der in der Literatur identifizierten RE-Qualitätskriterien mit der Praxis. Dafür werden den Interviewteilnehmern die Qualitätskriterien nacheinander vorgelesen und bei Bedarf erläutert. Die Auswertung dieses Abgleichs ist im Anhang unter Tabelle B.2 dokumentiert. Auch an dieser Stelle muss das mögliche Bias und eine daraus mögliche resultierende Überschätzung des Einsatzes dieser Qualitätskriterien berücksichtigt werden. Die begrenzte Stichprobengröße dieser Auswertung schränkt ebenfalls

allgemeingültige Schlüsse in Bezug auf die Verwendung spezifischer RE-Qualitätskriterien ein. Dennoch bieten die ermittelten Ergebnisse eine Tendenz zur praktischen Anwendbarkeit der RE-Qualitätskriterien.

4.1.5 Produktivität

In der initialen, unstrukturierten Literaturrecherche konnten keine spezifischen Kriterien zur Produktivität im Kontext des RE identifiziert werden. Diese Beobachtung bestätigte sich auch im anschließenden Literatur-Review. Aufgrund dessen wird Produktivität in dieser Arbeit anhand des zeitlichen Aspekts betrachtet.

4.2 Design Prinzipien

Nach der Darstellung der Problemstellung werden nun die Design Prinzipien vorgestellt und deren Herleitung gemäß Meth et al. (2015) erläutert. In der Analyse werden weiterhin die Erkenntnisse aus RE-Experteninterviews einbezogen. Dabei werden auch die Strategien zur Bewältigung der RE-Herausforderungen durch den Einsatz von GenAI betrachtet. Zudem fließen Einsichten aus der unstrukturierten Literaturrecherche und dem Literatur-Review mit ein. Zuletzt werden auch Erkenntnisse aus den GenAI-Experteninterviews miteinbezogen. Hierbei werden einerseits Einschätzungen der GenAI-Experten hinsichtlich des Einsatzes von GenAI im Kontext der bereits identifizierten RE-Herausforderungen betrachtet. Andererseits wird auch ein Design Prinzip identifiziert, das sich aus den Problemen bei der Implementierung eines GenAI-gestützten Systems ergibt.

Chandra Kruse et al. (2015) unterscheiden drei Arten von Design Prinzipien:

1. *Handlungsorientierte Design Prinzipien* definieren Aktionen zur Erzielung bestimmter Ergebnisse.
2. *Materialitätsorientierte Design Prinzipien* beschreiben physische und technologische Aspekte eines Lösungsansatzes, einschließlich der Eigenschaften und Konfigurationen der verwendeten Materialien und Technologien. Dazu gehört deren Einfluss auf die Funktionalität und Leistung des Systems.
3. *Kombinierte handlungs- und materialitätsorientierte Design Prinzipien* integrieren beide Ansätze, um effektive und effiziente Ergebnisse zu erzielen und bieten eine holistische Sicht auf den Lösungsansatz.

Design Prinzipien, die sowohl auf Aktionen als auch auf Materialität abzielen, erlangen an Bedeutung, wenn die Interaktion zwischen sozialen und technologischen Elementen im Fokus liegt (Chandra Kruse et al. 2015). In dieser Arbeit wird die Interaktion zwischen

Menschen und GenAI-Systemen betrachtet. Aus diesem Grund werden kombinierte Design Prinzipien verwendet.

Die Auflistung der innerhalb dieser Arbeit entwickelten Design Prinzipien und deren abgeleiteten Design Features wird in Abbildung 2 präsentiert. Ergänzt wird dies durch die Beschreibung der zugrundeliegenden Vorgehensweisen.

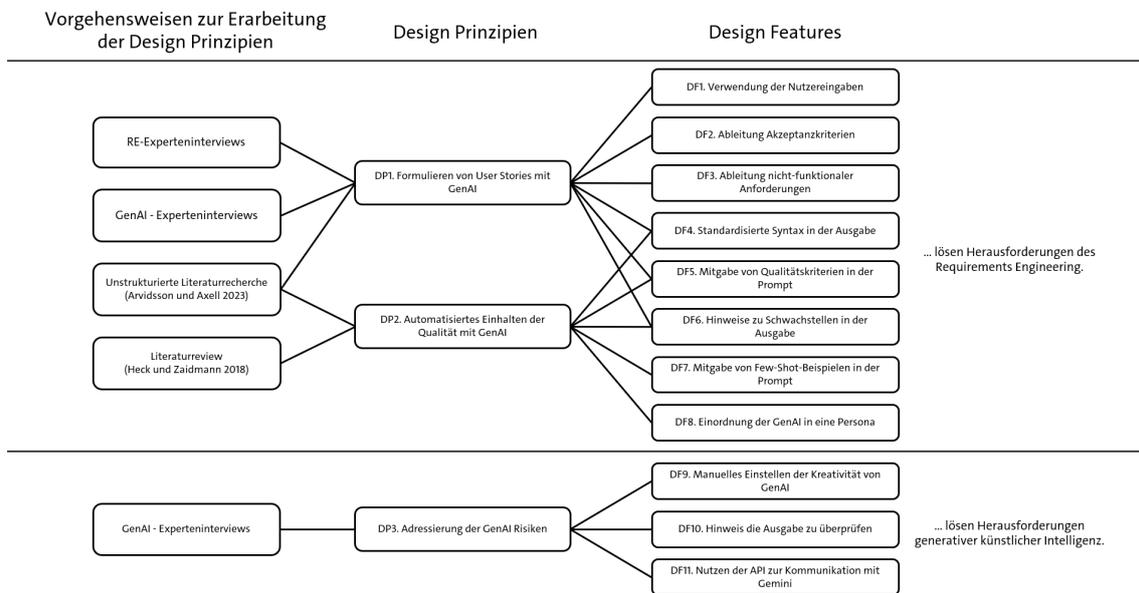


Abbildung 2: Design Prinzipien und Design Features des Ergebnisartefakts (Eigene Darstellung in Anlehnung an Meth et al. (2015))

DP1. Formulieren von User Stories mit GenAI basiert auf der Identifikation der RE-Herausforderungen, die in den Kapiteln 4.1.1 und 4.1.2 dargestellt werden. Hierbei wird kein Lösungsansatz für bestimmte RE-Herausforderungen betrachtet, sondern allgemein untersucht, wie GenAI zur Bewältigung dieser eingesetzt werden kann. Hierzu werden die Einschätzungen von RE- und GenAI-Experten herangezogen, welche in Kapitel 4.2.1 vorgestellt werden. Losgelöst von den RE-Herausforderungen werden innerhalb der RE-Experteninterviews die genutzten Anforderungsdokumente erörtert. Vier der fünf Befragten nutzen User Stories in ihrer Arbeit im RE (Ei2 - Z. 52, Ei3 - Z. 140 f., Ei4 - Z. 53, Ei5 - Z. 90)². Die verbreitete Nutzung wird, wie schon in Kapitel 2.1 erläutert, auch durch die Literatur bestätigt. Deshalb erfolgt eine Fokussierung auf das automatische Erstellen von User Stories. Außerdem werden Richtlinien zum Prompt Engineering nach Arvidsson und Axell (2023) berücksichtigt, die der unstrukturierten Literaturrecherche hervorgegangen sind.

² Dabei handelt es sich um eine Tendenz und kein Ergebnis einer quantitativen Auswertung, da es sich in diesem Falle um ein qualitatives Vorgehen handelt.

DP2. Automatisiertes Einhalten der Qualität entsteht aus den in Kapitel 4.1.3 definierten Qualitätskriterien und der Empfehlung von Heck und Zaidman (2018) für ein Tool zur Automatisierung der Einhaltung dieser Kriterien. Dies wird ebenfalls bei der Formulierung der Prompt berücksichtigt, indem Richtlinien nach Arvidsson und Axell (2023) einbezogen werden.

DP3. Adressierung der GenAI Risiken wird separat dargestellt, da es nicht die Anwendung von GenAI im RE behandelt, sondern allgemeine Herausforderungen des GenAI Einsatzes adressiert.

4.2.1 Bewältigung der RE-Herausforderungen mithilfe von GenAI

Die RE-Experten werden im Rahmen der Experteninterviews auch nach möglichen Bewältigungsstrategien für die genannten RE-Herausforderungen befragt. Die Antworten der RE-Experten konzentrieren sich zunächst nicht auf den Einsatz von GenAI. Einige der Herausforderungen lassen sich demnach durch verbesserte Kommunikation in Form von Workshops, Meetings, Retrospektiven und Nachbereitungen lösen (Ei1 - Z. 108 ff., Ei2 - Z. 170 ff., Ei4 - Z. 150 ff., Ei5 - Z. 210 f.). Auch wird die Bedeutung eines gemeinsamen Regelwerks zur Vereinheitlichung von Prozessen und Anforderungsdokumenten betont (Ei1 - Z. 112 ff.). Fehlendes Wissen wird beispielsweise durch eigene Recherchen ergänzt (Ei5 - Z. 211).

Zusätzlich werden die RE-Experten nach Ideen gefragt, wie GenAI ihrer Ansicht nach zur Bewältigung der diskutierten RE-Herausforderungen beitragen könnte. Die Antworten konzentrieren sich vor allem auf die Textgenerierung, beispielsweise beim automatisierten Formulieren von Anforderungen oder dem Erstellen von Anforderungsdokumenten (Ei1 - Z. 178, Ei3 - Z. 65 ff., Ei4 - Z. 22 ff., 172 ff.) dem Entwickeln von Akzeptanzkriterien (Ei2 - Z. 202 f., Ei4 - Z. 196 ff.) und Testszenarien (Ei4 - Z. 196 ff.). Dies bildet die Grundlage für das Design Prinzip **DP1** und daraus entstehende Design Features. Weitere genannte Ideen lassen sich mit den Lösungsansätzen der GenAI-Experten abgleichen, welche im Anschluss diskutiert werden. Dabei handelt es sich insbesondere um die Qualitätssicherung von Anforderungsdokumenten (Ei2 - Z. 213 f.), die auf *DP1* und *DP2* verweist, und die Berücksichtigung nicht-funktionaler Anforderungen (Ei4 - Z. 211 ff.) für Design Prinzip *DP1*. Abschließend ist zu erwähnen, dass die Integration von GenAI in die eigene Arbeitsweise bei allen befragten RE-Experten befürwortet wird (Ei1 - Z. 211 ff., Ei2 - Z. 224 ff., Ei3 - Z. 338 ff., Ei4 - Z. 253 ff., Ei5 - Z. 286 ff.).

Auswertung der GenAI-Experteninterviews

In den GenAI-Experteninterviews werden die identifizierten RE-Herausforderungen den Möglichkeiten von GenAI gegenübergestellt, um die Teilforschungsfrage (TF 2) zu adressieren. Für die Auswertung wird dabei ebenfalls ein Kategoriensystem nach Mayring (2010) aufgestellt, welches in Tabelle 11 dargestellt ist.

Die erste Kategorie K1, betitelt als *Expertise*, dient der Klassifikation der Interviewteilnehmer als Experten. An dieser Stelle wird die bisherige Erfahrung der Teilnehmer in den Bereichen KI und GenAI eingeordnet. Dabei liegt ein Fokus auf den Anwendungsbereichen sowie den Stärken und Schwächen von GenAI.

Kategorie K2 befasst sich mit dem Einfluss von *GenAI im RE*. Es wird untersucht, inwiefern GenAI eine unterstützende Rolle bei den RE-Herausforderungen einnehmen kann. Zusätzlich erfolgt eine Beurteilung potenzieller Risiken und Strategien für den Umgang mit diesen bei der Nutzung einer solchen Technologie.

Insgesamt werden 292 Segmente aus sechs durchgeführten GenAI-Interviews für die Kodierung herangezogen. Die Aufschlüsselung der Zuordnung der Textsegmente ist ebenfalls der angehängten Datei zu entnehmen.

Hauptkategorie	Subkategorie 1. Ebene	Subkategorie 2. Ebene
K1 Expertise - 0	K1.1 Erfahrung in KI - 6	-
	K1.2 Erfahrung in GenAI - 6	K1.2.1 Anwendungsbereiche - 17
		K1.2.2 Stärken - 15
		K1.2.3 Schwächen - 2
K2 GenAI im RE - 0	K2.1 Unterstützend - 126	-
	K2.2 Nicht unterstützend - 72	-
	K2.3 Evaluation des Ergebnisartefakt - 0	K2.3.1 Risiken - 25
		K2.3.2 Begegnung dieser Risiken - 23
Summe: 0	Summe: 210	Summe: 82

*Notationsweise: KX: Hauptkategorie KX.Y: Subkategorie KX.Y.Z: Subsubkategorie - #: Häufigkeit

Tabelle 11: Kategoriensystem und Häufigkeiten der Kodierungen der GenAI-Experteninterviews (Eigene Darstellung)

In der Auswertung der Interviews wird analysiert, welche RE-Herausforderungen durch den Einsatz von GenAI reduziert oder gelöst werden können. Die Ergebnisse dieses Abgleichs sind im Anhang in Tabelle B.3 festgehalten. Die Stichprobengröße der befragten Personen ist zu klein, um definitive Aussagen zu treffen. Es ist jedoch eine Tendenz ableitbar, welche RE-Herausforderungen durch GenAI potenziell gemindert oder gelöst werden könnten. Aufgrund zeitlicher Beschränkungen in den Interviews können jedoch nicht immer alle Lösungsansätze bei positiven Antworten weiterverfolgt werden.

Es wird festgestellt, dass nicht alle Herausforderungen vollständig lösbar sind. GenAI kann jedoch unterstützend wirken, indem es beim Erkennen dieser Herausforderungen hilft (Ei9 - Z. 40 ff.). Es wird auch betont, dass die Einbindung eines Menschen im Prozess sinnvoll ist, um die Ergebnisse zu überprüfen (Ei7 - Z. 182 f.). Bei weiteren Herausforderungen ist eine Unterstützung durch GenAI denkbar, wobei es jedoch wichtiger ist, die Menschen für die jeweilige Herausforderung zu sensibilisieren und zu schulen (Ei8 - Z. 159 f.).

In einigen Lösungsansätzen werden technische Implementierungen gefordert, wie die Einführung von Guardrails zur Bewältigung der *Unzulänglichkeit* (Ei6 - Z. 124 f.). Zudem werden Wissensstände innerhalb des Sprachmodells vorausgesetzt, etwa in Form von bekannten vergleichbaren Systemen, um der Herausforderung *fehlender Anforderungen* entgegenzuwirken (Ei9 - Z. 50 ff.). Es wird darauf hingewiesen, dass es der GenAI an Kontextdaten mangeln könnte (Ei7 - Z. 68 ff.) oder Training benötigt wird (Ei8 - Z. 213), insbesondere bei der Prioritätensetzung innerhalb von Anforderungsdokumenten. Zudem wird betont, dass die Datenqualität eines Sprachmodells bestimmte Standards erfüllen muss, um effektiv verschiedene Herausforderungen zu adressieren (Ei6 - Z. 95 f., Ei9 - Z. 117 ff.). Diese Aspekte sind jedoch nur begrenzt beeinflussbar, da sie grundlegende technische Eigenschaften der Sprachmodelle betreffen, die innerhalb dieser Arbeit nicht modifiziert werden.

Innerhalb des Prompt Engineering gibt es Möglichkeiten bestimmten Herausforderungen entgegenzuwirken und damit die Design Prinzipien *DP1* und *DP2* mit Design Features zu unterstützen. Beispielsweise kann innerhalb eines Prompts ein Standard vorgegeben werden, wie eine User Story inklusive Akzeptanzkriterien aussehen sollte (Ei11 - Z. 56 ff.), um die Herausforderung der *Unbestimmtheit* von User Stories anzugehen. Ebenfalls könnte innerhalb eines Prompts explizit definiert werden, nicht-funktionale Anforderungen zu berücksichtigen, um *deren Fehlen* entgegenzuwirken (Ei6 - Z. 152 ff.). Hinweise im GenAI Output, was innerhalb des Kontextes einer bestimmten User Story an Informationen fehlen könnte, wird bei der Herausforderung *fehlende Anforderungen* als Hilfestellung genannt (Ei11 - Z. 154 ff.). Durch Few-Shot-Beispiele könnte verdeutlicht werden, wie User Stories aussehen sollten, was laut der GenAI Experten Abhilfe bei *schlecht formulierten Anforderungen, mangelndem Wissen in Fachabteilungen über die Beschreibung von Use Cases* und *unreflektierter Prozessübernahme* schaffen kann (Ei10 - Z. 191 ff., 432 ff., 484 ff.). Dabei wird auch empfohlen, Qualitätsmetriken aufzustellen und diese in dem Prompt zu integrieren (Ei10 - Z. 483 ff.). *Schlecht formulierte Anforderungen* könnten zudem durch verbesserte Vorschläge einer GenAI überarbeitet werden (Ei11 - Z. 139 f.). Ein *weniger erfahrenes und qualifiziertes Team* könnte durch allgemeine Textgenerierung unterstützt werden (Ei9 - Z. 129 f.), sowie durch schrittweise Anleitungen in Form einer bestimmten Persona, die die GenAI einnehmen sollte (Ei10 - Z. 213 ff.).

Während der GenAI-Experteninterviews werden zudem verschiedene Vorteile von GenAI hervorgehoben. Zu den genannten Vorteilen zählen das *Brainstorming*, insbesondere im Kontext von HI (Ei10 - Z. 38 ff.). Des Weiteren wird die *Geschwindigkeit*, mit der GenAI Aufgaben erledigen kann, betont (Ei9 - Z. 20 ff., Ei11 - Z. 35). Die Experten unterstreichen auch die Fähigkeit zur *Informationsverarbeitung* (Ei6 - Z. 13 ff., Ei7 - Z. 16 ff., Ei8 - Z. 17 ff., Ei9 - Z. 17 f., Ei10 - Z. 17 ff., Ei11 - Z. 13 ff.) sowie zur *Inhaltsgenerierung* (Ei6 - Z. 28 ff., Ei8 - Z. 38 ff.). Außerdem wird die *Prozessautomatisierung* als ein weiterer Anwendungsbereich von GenAI genannt (Ei7 - Z. 9 ff., Ei10 - Z. 42 ff.). Die vollständige Auflistung dieser Vorteile ist im Anhang in Tabelle B.4 dokumentiert.

4.2.2 Herausforderungen von GenAI

Um mögliche Herausforderungen und Risiken bei der Nutzung von GenAI im RE vorzubeugen, die im Design Prinzip **DP3** des Ergebnisartefakts berücksichtigt werden, werden verschiedene Bedenken behandelt.

Die GenAI-Experten weisen auf die Herausforderung der *Halluzinationen* hin (Ei6 - Z. 351 f., Ei7 - Z. 286 ff., Ei8 - Z. 364 f., Ei9 - Z. 342 ff., Ei10 - Z. 498 ff., Ei11 - Z. 317). Dieses Risiko wird auch in Kapitel 2.2.5 diskutiert. Halluzinationen können auftreten, wenn die Kreativität des Modells zu hoch eingestellt ist (Ei6 - Z. 351 f.). Um diesem Risiko entgegenzuwirken, wird empfohlen, die Temperatur des Modells anpassbar zu machen (Ei6 - Z. 388, Ei9 - Z. 401 ff.), was im Design Feature *DF9* berücksichtigt wird. Jedoch wird gleichzeitig das *Fehlen von Kreativität* als mögliche Schwäche angeführt, wenn man versucht Halluzinationen zu vermeiden (Ei9 - Z. 346 ff.). Mit der Möglichkeit zur Einstellung der Temperatur wird ein Kompromiss geschaffen, der es dem Nutzer überlässt, hier Anpassungen vorzunehmen.

Eine weitere Herausforderung ist das *übermäßige Vertrauen von Nutzern* in die Ausgabe von GenAI (Ei6 - Z. 370 ff., Ei7 - Z. 293 ff., Ei8 - Z. 365 ff., Ei9 - Z. 323 ff.), was auch in den RE-Experteninterviews (Ei4 - Z. 206 ff.) und in Kapitel 2.2.5 thematisiert wird. Dies kann weitergehend zu einem möglichen Wissensverlust führen, wenn Nutzer beispielsweise Code von GenAI erstellen lassen, den sie am Ende nicht verstehen, aber dennoch implementieren (Ei8 - Z. 32 ff., Ei11 - Z. 303 ff.). Ein Hinweis die Ausgabe der GenAI zu überprüfen, in Form des Design Features *DF10*, soll hier Abhilfe schaffen. *Unzureichende Trainingsdaten* der Sprachmodelle stellen ebenfalls ein Risiko dar, da diese zu voreingenommenen Ergebnissen führen können (Ei8 - Z. 364 ff., Ei9 - Z. 335 ff.).

Zudem wird *Sicherheit* ebenfalls als Risiko identifiziert, welches durch lokale Implementierungen und das Nutzen von Schnittstellen begrenzt werden kann (Ei6 - Z. 363 f., Ei10 - Z. 504 ff.). Dies wird durch das Design Feature *DF11* implementiert.

Die vollständige Liste der von den GenAI-Experten genannten Herausforderungen und Risiken im Einsatz von GenAI wird im Anhang in Tabelle B.5 dargelegt.

4.3 Lösungsgestaltung

Das Ergebnisartefakt dieser Arbeit ist ein Prompt, der durch eine Webanwendung angesprochen wird, um RE-Prozesse zu unterstützen. Dieser Prompt ist unformatiert im Anhang B.2 und formatiert in einer separaten Datei beigefügt. Über ein Interface können Personen, die im RE tätig sind, Projektinformationen eingeben. Hierzu gehören unter anderem Requirements Engineers, Business Analysten und Product Owner. Basierend auf diesen Daten werden User Stories ausgegeben, die anschließend von den Nutzern manuell validiert werden müssen. Nach der Validierung soll es möglich sein, diese Dokumente als User Stories in den eigenen Projektverlauf zu integrieren. Die Gestaltung der Prompt wird in Kapitel 4.3.1 diskutiert. Der Aufbau dieser Prompt wird in Kapitel 4.3.2 beschrieben.

Der Fokus dieser Forschungsarbeit liegt auf der Entwicklung eines Prompts zur Erstellung von User Stories. Diese werden in der Literatur (Schön, Thomaschewski et al. 2017) und auch von interviewten RE-Experten als häufig verwendetes Artefakt zur Dokumentation von Anforderungen im RE herausgestellt.

Bei der Auswahl des verwendeten Sprachmodells fällt die Entscheidung auf Gemini Pro, ein von Google entwickeltes Sprachmodell (Gemini-Team et al. 2023). Y. Wang und Y. Zhao (2023) führen einen direkten Vergleich von Gemini Pro mit anderen Modellen in drei Kategorien des logischen Denkens durch. Ihre Ergebnisse zeigen, dass Gemini Pro eine vergleichbare Leistungsfähigkeit zu GPT-3.5 Turbo aufweist und weniger leistungsstark als GPT-4 Turbo ist (Y. Wang und Y. Zhao 2023). Ein Vorteil von Gemini Pro gegenüber den Modellen von OpenAI ist die Kostenstruktur. Während bei OpenAI bereits bei jeder Anfrage Kosten anfallen (*OpenAI Pricing* 2024), entstehen bei der Nutzung der Gemini API erst bei einer bestimmten Anzahl verwendeter Token Kosten (*Gemini API Pricing | Google AI for Developers* 2024).

Es wird kein Fokus auf eine Neuimplementierung von Schnittstellen zwischen der Webanwendung und dem Sprachmodell gelegt. Das Backend wird von einem internen Projekt des Partnerunternehmens adaptiert und in Abstimmung mit dem betreuenden Unternehmensmitarbeiter implementiert. Das Frontend entsteht in Kooperation mit einem Frontend-Entwickler und dient der Bereitstellung einer Benutzeroberfläche für die Durchführung der Evaluationen. Es ist jedoch wichtig zu betonen, dass der Code, sowohl für Backend als auch für das Frontend, aufgrund von Compliance-Richtlinien nicht dieser Arbeit angehängt ist.³ Diese Transparenz ermöglicht es, Screenshots des Frontends im weiteren Verlauf

³ Folglich soll das Backend und das Frontend nicht als Bewertungsgrundlage herangezogen werden.

dieses Kapitels einzubinden, um den Prozess innerhalb der Evaluation veranschaulichen zu können. Das Frontend wird in Kapitel 4.3.3 beschrieben. Der Hintergrund der technischen Implementierung wird in Kapitel 4.3.4 kurz dargelegt.

4.3.1 Entwicklung der Prompt

Im Folgenden wird die Entwicklung der Prompt vorgestellt. Durch die Anwendung von Prompt Engineering lassen sich gezielt Einflüsse auf die Ausgabe von GenAI steuern (White et al. 2023). Dies wird von Erkenntnissen aus der Literatur nach Arvidsson und Axell (2023) und Vorgaben von Unternehmen, die solche Sprachmodelle entwickeln (Google 2024; OpenAI 2024), unterstützt.

Der Prompt wird gemäß einer Auswahl an Richtlinien konzipiert, die speziell für den Einsatz in Text-zu-Text Aufgaben im RE-Bereich aufgestellt wurden (Arvidsson und Axell 2023). Dabei orientiert sich die Auswahl zunächst an den von Arvidsson und Axell (2023) vorgestellten Richtlinien, die für diese Aufgabenstellungen optimiert sind. Nicht alle vorgeschlagenen Richtlinien werden übernommen, sondern nur jene, die einen konkreten Mehrwert für die Anforderungen dieser Forschungsarbeit bieten. Die ausgewählten Richtlinien, die in dieser Arbeit Anwendung finden, sind in Tabelle 12 detailliert aufgeführt.

Wichtige Aspekte der Prompt-Konstruktion umfassen die Übergabe von Kontextinformationen, um präzisere Ergebnisse zu erzielen und das Risiko von Halluzinationen zu verringern (C1, C2, C3, C4) (Arvidsson und Axell 2023). Bei den Kontextinformationen handelt es sich um Informationen, die von Nutzern an die Webanwendung übergeben werden. Dies wird innerhalb von Design Feature *DF1* integriert. In C5 wird zudem empfohlen diese Kontextinformationen vor der Beschreibung der Aufgabenstellung an das Sprachmodell zu übergeben (Arvidsson und Axell 2023).

Die Einordnung des Sprachmodells in eine spezifische Persona dient der Steigerung der Konsistenz und Qualität im Output (P1) (Arvidsson und Axell 2023) und wird deshalb auch als Design Feature (*DF8*) implementiert. Zusätzlich trägt die Verwendung von Templates auch zur Verbesserung der Ergebnisse bei (T1) (Arvidsson und Axell 2023).

Die Auflösung von Mehrdeutigkeiten ist ein weiteres Ziel (Arvidsson und Axell 2023). Dabei soll sich das Modell an die bereitgestellten Informationen halten (D1) (Arvidsson und Axell 2023), was durch *DF1* berücksichtigt wird. Die Erkennung und Hervorhebung von Schwachstellen in den Anforderungsspezifikationen durch das Sprachmodell ist eine weitere empfohlene Richtlinie (D2) (Arvidsson und Axell 2023). Dies findet sich in Design Feature *DF6* wieder. Die Integration von Schlüsselwörtern kann die Ausgabe verbessern, indem relevante Kontextinformationen hervorgehoben werden (K2) (Arvidsson und Axell 2023).

Die Verwendung von Few-Shot-Beispielen wird spezifiziert, um das Modell besser an die Aufgabenstellung anzupassen (F1 - F4) (Arvidsson und Axell 2023). Dabei soll eine klare Unterscheidung zwischen Eingabe und Ausgabe berücksichtigt werden (F4) (Arvidsson und Axell 2023). Dies wird durch das Design Feature *DF7* implementiert.

Thema	Beschreibung
Kontext (C)	<ol style="list-style-type: none"> 1. Kontext zu Beispielen in Prompts hinzuzufügen, produziert effizientere und informativere Ergebnisse. 2. Fügen Sie allen Prompts Kontext hinzu, um Halluzinationen in der Ausgabe zu vermeiden. 3. Geben Sie den Kontext des Prompts an, um ein eng verwandtes Ergebnis zu gewährleisten. 5. Stellen Sie den Kontext zum Thema des Prompts dar, bevor Sie eine Aufgabe beschreiben.
Persona (P)	<ol style="list-style-type: none"> 1. Verbessert die Generierungsqualität, indem der Prompt mit einer Identität wie „Python-Programmierer“ oder „Mathematik-Nachhilfelehrer“ konditioniert wird.
Templates (T)	<ol style="list-style-type: none"> 1. Um die Logik und den gesunden Menschenverstand in der Ausgabe zu verbessern, folgen Sie einer Vorlage
Mehrdeutigkeitsauflösung (D)	<ol style="list-style-type: none"> 1. Stellen Sie sicher, dass alle Bereiche potenzieller Missverständnisse oder Unklarheiten erfasst werden, indem Sie einen detaillierten Umfang bereitstellen: <ul style="list-style-type: none"> - „Innerhalb dieses Umfangs“, - „Berücksichtigen Sie diese Anforderungen oder Spezifikationen“ 2. Um Schwachstellen in einer Anforderungsspezifikation zu finden, erwägen Sie Folgendes einzubeziehen: <ul style="list-style-type: none"> - „Weisen Sie auf Bereiche der Unklarheit oder potenziell unbeabsichtigte Ergebnisse hin“
Schlüsselwörter (K)	<ol style="list-style-type: none"> 2. Das Voranstellen von Schlüsselwörtern zu Aufforderungen verbessert die Leistung erheblich, indem dem Sprachmodell der geeignete Kontext bereitgestellt wird.
Few-Shot Prompts (F)	<ol style="list-style-type: none"> 1. Die Einbeziehung von „Frage:“ und „Antwort:“ verbessert die Antwort, gibt aber selten eine binäre Antwort. 2. Zur einfacheren Verständlichkeit sollten Beispiele in der Few-Shot-Aufforderung nummeriert werden. 3. Das Format von [EINGABE] und [AUSGABE] sollte sprachlich die Beziehung zwischen ihnen implizieren. 4. Spezifikationen können jedem [EINGABE] und [AUSGABE]-Paar hinzugefügt werden, um einen zusätzlichen Einblick in komplizierte Probleme zu geben.

Tabelle 12: Auswahl der genutzten Richtlinien für Prompt Engineering im RE nach Arvidsson und Axell (2023)

OpenAI und Google bieten spezifische Strategien und Empfehlungen für das Verfassen effektiver Prompts, die dazu beitragen, die Qualität der Ergebnisse von Sprachmodellen zu verbessern (Google 2024; OpenAI 2024). Diese Anweisungen sind teilweise den von Arvidsson und Axell (2023) formulierten Richtlinien für Prompt Engineering zuordenbar. Ein Aspekt der von OpenAI und Google vorgeschlagenen Strategien umfasst das Schreiben klarer Anweisungen. Es sollen detaillierte Informationen eingebunden werden, um relevantere Ausgaben zu generieren (Google 2024; OpenAI 2024). Die Vermeidung mehrdeutiger Begriffe ist eine weitere Möglichkeit die mitgegebenen Anweisungen eindeutig zu formulieren (Google 2024). Diese beiden Strategien korrespondieren direkt mit den Empfehlungen von Arvidsson und Axell (2023). Sie schlagen vor, in den Prompts umfassende Kontextinformationen bereitzustellen (Richtlinien C1, C2, C3, C5) (Arvidsson und Axell 2023). Zusätzlich empfehlen OpenAI und Google, komplexe Aufgaben in überschaubare Teilaufgaben zu zerlegen. Dies erleichtert die Verarbeitung durch das Sprachmodell und verbessert die Genauigkeit der Antworten (Google 2024; OpenAI 2024).

OpenAI hebt hervor, dass das Definieren einer Persona für das Sprachmodell (P1 (Arvidsson und Axell 2023)) hilfreich sein kann, um die Generierung von Inhalten zu spezifizieren (OpenAI 2024). Die Verwendung von Few-Shot-Beispielen, wie in den Richtlinien F1 bis F4 von Arvidsson und Axell (2023) beschrieben, unterstützt das Sprachmodell dabei, sich besser auf die spezifische Aufgabenstellung einzustellen und genauere Ergebnisse zu liefern (OpenAI 2024). Das Einfügen von Schlüsselwörtern, wie in Richtlinie K2 nach Arvidsson und Axell (2023) dargelegt, optimiert weiterhin die Ausgabe des Modells (Google 2024).

Die Prompt-Entwicklung innerhalb dieser Arbeit orientiert sich an den bewährten Richtlinien und Techniken aus der Literatur und den Strategien von OpenAI und Google. Die Einbeziehung dieser Strategien zur Informationsbereitstellung, Persona-Definition und Aufgabenzerlegung tragen zur Optimierung der Ausgaben von GenAI bei und fördern die Präzision sowie Relevanz der generierten Inhalte. Diese methodische Herangehensweise ist entscheidend für die Realisierung eines effektiven und zielgerichteten Prompts.

4.3.2 Aufbau der Prompt

Der Aufbau der Prompt, dargestellt in Abbildung 3, beginnt mit der Übergabe von Kontextinformationen. Diese umfassen die Aufgabenstellung an die GenAI und die Eingaben des Benutzers (*DF1*). Anschließend wird das Sprachmodell in eine spezifische Persona, in diesem Fall einen Requirements Engineer eines IT-Dienstleistungsunternehmens, hineinversetzt (*DF8*).

Durch drei optionale textuelle Eingabeparameter im Frontend der Webanwendung - „Zielgruppe“, „Funktionalität“ und „Nutzen“ - ergeben sich acht verschiedene Eingabemöglichkeiten. Diese Parameter sind auf das in 2.1 vorgestellten Schema von User Stories nach Lucassen, Dalpiaz, Werf et al. (2016b) zugeschnitten. In den Fällen, in denen ein oder mehrere dieser Felder nicht vom Nutzer ausgefüllt werden, soll die GenAI einen dem Projektkontext entsprechenden Parameter generieren und auf die User Story beziehen. Die Eingabefelder können in Abbildung 4 eingesehen werden. Alle acht Fallmöglichkeiten werden je nach Eingaben eines Nutzers innerhalb des Prompts behandelt. Zudem werden aus der Literatur identifizierte RE-Qualitätskriterien vorgegeben, die das Sprachmodell beim Erstellen der User Stories berücksichtigen soll, um den Design Prinzipien *DP1* und *DP2* mit dem Design Feature *DF5* gerecht zu werden.

Ein festgelegtes User Story Template strukturiert den Aufbau des User Story Satzes. Des Weiteren werden Akzeptanzkriterien nach Regeln, die Koelsch (2016, vgl. S. 314 f.) aus einem Artikel von Davu (2015) zusammengefasst, abgeleitet (*DF2*). Dabei werden auch explizit nicht-funktionale Anforderungen berücksichtigt (*DF3*). Testszenarien sind wiederum aus User Stories ableitbar und werden auch im Kontext der mitgegebenen

Informationen erstellt (Mollah und van den Bos 2023). Das Sprachmodell ist zudem darauf ausgerichtet, potenzielle Schwachstellen in der Dokumentation aufzuzeigen, um Denkanstöße für weitere User Stories zu liefern (DF6).

Zur Anpassung des Modells an die spezifische Aufgabe werden Few-Shot-Beispiele (DF7) genutzt (Arvidsson und Axell 2023). Die Ausgabe dieser User Story Beispiele ist einem realen Projekt des Partnerunternehmens entnommen. Diese User Stories sind erfolgreich innerhalb dieses Projektes umgesetzt worden. Zusätzlich geht dieser Teil der Prompt auf die Design Anforderungen ein. Abschließend wird ein einheitliches Ausgabeformat vorgegeben (DF4), um Konsistenz in den Ergebnissen zu gewährleisten.

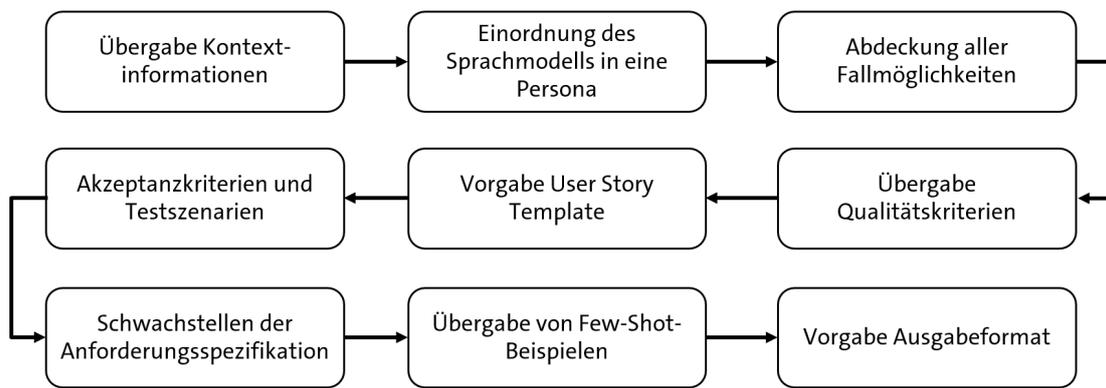


Abbildung 3: Aufbau der Prompt (Eigene Darstellung)

4.3.3 Webanwendung - Benutzeroberfläche

Die Benutzeroberfläche selbst ist kein Teil des Ergebnisartefakts. Sie wird jedoch vorgestellt, da die Benutzeroberfläche Grundlage des geplanten Umfangs der Evaluation darstellt. Sie dient der direkten Interaktion der Nutzer mit dem Sprachmodell und bietet einen Lösungsvorschlag mit der Integration von zwei Design Features (DF9, DF10). Die Benutzeroberfläche wird übersichtlich gestaltet, indem nur die für die Evaluation nötigen Elemente implementiert werden. Wie in Abbildung 4 ersichtlich, werden dafür Tooltips integriert. Diese Tooltips sollen den Evaluationsteilnehmern eine schnelle Einarbeitung innerhalb der begrenzten verfügbaren Zeit erleichtern. Im Zuge dessen wird die Richtlinie K2, also die Miteinbeziehung von Schlüsselwörtern, nach Arvidsson und Axell (2023) angewendet. Um dies zu unterstützen, wird der Benutzer in der Anwendung innerhalb der Tooltips auf die Verwendung von Schlüsselwörtern hingewiesen. Drei der Texteingabefelder sind auf das User Story Template abgestimmt und ein zusätzliches Textfeld ermöglicht es den Nutzern, den Projektkontext zu definieren.

Die Anpassung der Temperatur erlaubt eine Steuerung der Kreativität des Sprachmodells (Chen und Ding 2023). Die Einstellung der Kreativität ist dafür manuell anpassbar, um das Design Prinzip *DP3* zu berücksichtigen.

Moin! Gib hier ein paar Details ein...
... um dir ein Dokument generieren zu lassen.

Deine Angaben

LLM-Modell: Gemini
Sprache: Deutsch
Anforderungsdokument: User Story
Generator-Optionen: Akzeptanzkriterien, Testfälle

Projektkontext*

Gib hier eine gewünschte Funktionalität an innerhalb des mitgegebenen Projektkontextes ein.

Zielgruppe

Funktionalität

Nutzen

Kreativität:
Konsistent ————— Kreativ ————— Fantasiereich

Generieren

Abbildung 4: Ausgabe Webanwendung - Teil 1

Abbildung 5 präsentiert eine Nachstellung einer beispielhaften Anfrage aus der Evaluation. Die daraus resultierende User Story wird in Abbildung 6 gezeigt. Die User Story beginnt mit dem Namen dieser, gefolgt vom User Story Template. Anschließend werden Akzeptanzkriterien und Testszenarien dargestellt sowie mögliche Schwachstellen innerhalb der generierten User Story aufgezeigt. Zum Abschluss wird bei jeder generierten User Story ein Hinweis zur Überprüfung der generierten Ausgabe mitgegeben.

Moin! Gib hier ein paar Details ein...
... um dir ein Dokument generieren zu lassen.

Deine Angaben

LLM-Modell: Gemini
Sprache: Deutsch
Anforderungsdokument: User Story
Generator-Optionen: Akzeptanzkriterien, Testfälle

Projektkontext*
Wir entwickeln ein neues Gastronomie Service System, welches es ermöglicht, Bestellungen von Gästen über IoT Devices am Tisch aufzunehmen. Dazu gehört die Angabe der Bestellbereitschaft der Kunden. Das Ganze soll in real-time erfolgen, die Kunden sollen dadurch einen entspannteren Bestellprozess haben, das Personal soll entlastet werden und sich auf key tasks fokussieren können. Zusätzlich wird das Küchenpersonal rechtzeitig über Bestellungen und entsprechende Vorbereitungen informiert.

Zielgruppe
Gast

Funktionalität
Bestellbereitschaft über Tablet am Tisch signalisieren

Nutzen
Gestreampelnte Interaktion mit dem Servicepersonal, ohne dass Gäste gestört werden, bzw. sich Zeit nehmen können.

Kreativität:
Konsistent ————— Kreativ ————— Fantasereich

[Circular Arrow Button]

Abbildung 5: Ausgabe Webanwendung - Teil 2

4.3.4 Technische Implementierung

Der Fokus dieser Arbeit liegt nicht auf der technischen Implementierung. Dennoch ist es für das Verständnis der Funktionsweise der Webanwendung notwendig einige technische Aspekte zu erläutern.

Sobald Nutzer im Frontend auf den „Generieren“ Button drücken, wird der eingegebene Input vom Backend, einer REST-API, in Form eines Payloads entgegengenommen. Die einzelnen Parameter des Inputs werden extrahiert, in den entwickelten Prompt eingefügt und mittels eines HTTP POST-Requests an die Gemini-API gesendet. Die von Gemini Pro generierte Ausgabe wird vom Backend empfangen und anschließend im Frontend dargestellt.

Die Evaluationsphase der Anwendung wird durch den Einsatz von Google Cloud Services unterstützt, welche sowohl das Frontend als auch das Backend bereitstellen. Diese Hosting-Lösung gewährleistet einen störungsfreien Betrieb während der Evaluationsphase.

Zielgruppe
Gast

Funktionalität
Bestellbereitschaft über Tablet am Tisch signalisieren

Nutzen
Gestreamlinete Interaktion mit dem Servicepersonal, ohne dass Gäste gestört werden, bzw. sich Zeit nehmen können.

Kreativität:
Konsistent Kreativ Fantasiereich

Generieren

User Story Name: Bestellbereitschaft digital signalisieren

Als Gast **möchte ich** meine Bestellbereitschaft über das Tablet am Tisch signalisieren, **um** eine gestreamlinete Interaktion mit dem Servicepersonal zu erzielen, ohne dass ich gestört werde oder mir Zeit nehmen muss.

Akzeptanzkriterien:
Der Gast kann seine Bestellbereitschaft durch Antippen einer Schaltfläche oder eines Symbols auf dem Tablet signalisieren.
Das Signal wird in Echtzeit an das Servicepersonal gesendet.
Das Servicepersonal erhält einen Hinweis, dass der Gast bereit ist, seine Bestellung aufzugeben.
Die Bestellbereitschaft des Gastes wird auf dem Tablet visuell angezeigt.

Testszzenarien:
Szenario 1: Der Gast tippt auf die Schaltfläche "Bestellbereit" auf dem Tablet. Das Servicepersonal erhält eine Benachrichtigung.
Szenario 2: Der Gast tippt mehrfach auf die Schaltfläche "Bestellbereit". Das Servicepersonal erhält nur eine Benachrichtigung.
Szenario 3: Das Servicepersonal ignoriert die Benachrichtigung des Gastes für mehrere Minuten. Der Gast kann seine Bestellbereitschaft erneut signalisieren.

Schwachstellen in der Anforderungsspezifikation:
Es ist nicht klar, ob die Bestellbereitschaft des Gastes auf dem Tablet angezeigt wird, nachdem er das Signal gesendet hat.
Es ist nicht klar, ob das Servicepersonal eine Möglichkeit hat, die Bestellbereitschaft des Gastes zurückzusetzen.

Hinweis: Bitte überprüfe den Text, bevor du ihn als User Story in Deinem Projekt benutzt. Es handelt sich hierbei um einen Output einer generativen künstlichen Intelligenz, die nur die Parameter kennt, die Du ihr übergeben hast.

Kopieren

Abbildung 6: Ausgabe Webanwendung - Teil 3

4.3.5 Abgrenzung

Weitere Funktionen, wie etwa eine Chat-Funktionalität, die Goel et al. (2023) empfehlen, sind in der aktuellen Version der Webanwendung bewusst nicht integriert. Obwohl solche Funktionen nachweisbare Vorteile bieten können (Goel et al. 2023), hätte ihre Implementierung den Entwicklungsaufwand dieser Arbeit überschritten. Zumal der Fokus nicht auf der technischen Implementierung liegt. Die Erwägung weiterer Funktionalitäten wird in Kapitel 6.5 diskutiert.

5 Evaluationsergebnisse

Im Rahmen der Auswertung der Evaluationsergebnisse wird zunächst das nach Mayring (2010) entwickelte Kategoriensystem der Interviews vorgestellt, die nach der Thinking-Aloud-Simulation durchgeführt wurden. Diese Kategorisierung ist in Tabelle 13 dokumentiert.

Die erste Kategorie, K1, erfasst die Gesamtauswertung der *Evaluation*, insbesondere die Qualität der generierten User Stories und die Produktivität des Prozesses. Innerhalb dieser Kategorie wird die Teilforschungsfrage (TF 3) adressiert, indem die wahrgenommenen Vorteile und Herausforderungen im Einsatz mit der Webanwendung kategorisiert werden. Zusätzlich werden Ideen für fehlende Funktionalitäten aufgenommen.

Kategorie K2, *Expertise*, klassifiziert die Teilnehmer der Evaluation, die vorher nicht an den RE-Experteninterviews beteiligt waren. Hierzu werden deren beruflicher Hintergrund im RE und ihre Erfahrung mit GenAI erörtert.

Insgesamt basiert die Kodierung auf 175 Segmenten aus zehn Transkriptionen. Die Aufschlüsselung dieser Zuordnung der Textsegmente ist der angehängten Datei zu entnehmen.

Hauptkategorie	Subkategorie 1. Ebene	Subkategorie 2. Ebene
K1 Evaluation - 0	K1.1 Qualität (neutral) - 6	K1.1.1 Gute Qualität - 15
		K1.1.2 Schlechte Qualität - 7
	K1.2 Produktivität - 14	-
	K1.3 Praxisnutzung - 22	K1.3.1 Vorteile - 39
		K1.3.2 Herausforderungen - 26
K1.4 Fehlende Funktionalitäten - 29	-	
K2 Expertise - 0	K2.1 Beruflicher Hintergrund im RE - 10	-
	K2.2 Erfahrung in GenAI - 7	-
Summe: 0	Summe: 88	Summe: 87

*Notationsweise: KX: Hauptkategorie KX.Y: Subkategorie KX.Y.Z: Subsubkategorie - #: Häufigkeit

Tabelle 13: Kategoriensystem und Häufigkeiten der Kodierungen der Evaluationsinterviews (Eigene Darstellung)

Innerhalb der zehn Thinking-Aloud Simulationen werden in etwas mehr als vier Stunden 61 User Stories generiert, wie in Abbildung 7 dargestellt. 23 dieser werden von den Evaluationsteilnehmern verworfen. Von den verbleibenden 38 werden 17 User Stories unbearbeitet übernommen und 21 mit manuell durchgeführten Modifikationen der Evaluationsteilnehmer akzeptiert. Dazu gehören alle User Stories, bei denen mindestens ein Wort innerhalb der Ausgabe bearbeitet, hinzugefügt oder entfernt wird. Dies ergibt eine Summe von 38 übernommenen User Stories. Alle übernommenen User Stories sind in der separat angehängten Datei dokumentiert.

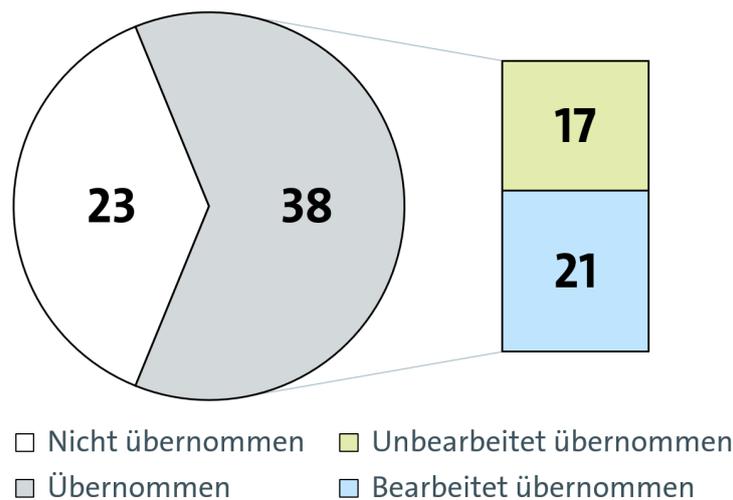


Abbildung 7: Generierte User Stories innerhalb der Evaluation (Eigene Darstellung)

5.1 Produktivität

Im Kontext der Produktivität beim Einsatz von GenAI im RE wird ein Teil der zentralen Hauptforschungsfrage adressiert. Die Nutzung von GenAI zur Erstellung von User Stories hat sich in den Evaluationen positiv auf die Produktivität der Teilnehmer ausgewirkt. Neun der zehn Teilnehmer berichten, dass sie durch die Webanwendung in der Lage waren, User Stories schneller zu erstellen (Et1 - Z. 13 ff., Et2 - Z. 11 ff., Et3 - Z. 29 ff., Et4 - Z. 33 ff., Et5 - Z. 29 ff., Et6 - Z. 61 ff., Et7 - Z. 74 ff., Et8 - Z. 45 ff., Et9 - Z. 45 ff.). Diese Beobachtung wird von einigen Teilnehmern erneut explizit als wahrgenommener Vorteil hervorgehoben, wie in Kapitel 5.3 dargestellt.

Ein Teilnehmer, bezeichnet als Et6, hebt jedoch Bedenken in diesem Zusammenhang hervor. Obwohl das grundlegende Erstellen von User Stories nach dessen Empfinden schneller erfolgt, könnte die schnelle Generierung vieler User Stories zu einem Mehraufwand führen. Dies gilt insbesondere für die Gewährleistung ihrer gegenseitigen Konsistenz (Et6 - Z. 62 ff.). Diese Problematik wird auch als Herausforderung in Kapitel 5.4 diskutiert.

Ein anderer Teilnehmer, identifiziert als Et10, bewertet den Zeitaufwand im Vergleich zur manuellen Methode als etwa gleichbleibend (Et10 - Z. 35 ff.). Dies wird damit begründet, dass jede generierte User Story einer gründlichen Überprüfung bedarf (Et10 - Z. 35 ff.).

Zusammenfassend lässt sich feststellen, dass der Einsatz von GenAI in dieser Stichprobe überwiegend zu einer Beschleunigung bei der Erstellung erster Entwürfe von User Stories führt, wenngleich diese auch eine sorgfältige Überprüfung erfordern.

5.2 Qualität

Im Rahmen der Evaluation der Qualitätsaspekte von GenAI im RE, die einen Teil der zentralen Hauptforschungsfrage behandelt, lässt sich die Qualität der generierten User Stories nicht unmittelbar bewerten. Einige Rückmeldungen der Evaluationsteilnehmer deuten darauf hin, dass die Qualität der Ergebnisse im Vergleich zur manuellen Methode in manchen Aspekten als überlegen und in anderen als unterlegen angesehen wird. Die dokumentierte Analyse der Teilnehmernaussagen zur Qualität, aufgeführt in Tabelle 13, umfasst 28 Aussagen:

- 15 Aussagen weisen auf eine gute oder verbesserte Qualität der generierten User Stories hin.
- Sieben Aussagen bezeichnen die Qualität als schlecht oder verschlechtert, wobei kritische Rückmeldungen häufig durch die Nichtübernahme der entsprechenden User Stories adressiert werden.
- Sechs neutrale Bewertungen reflektieren eine gemischte Wahrnehmung.

Ein Teilnehmer, bezeichnet als Et2, erläutert beispielsweise: *„Sowohl von mir selbst entwickelte User Stories haben ihre Schwachstellen, wo vielleicht die KI jetzt irgendwie schlaue Hinweise gegeben hat, die ich jetzt auch das allererste Mal nicht auf dem Schirm hatte. Gleichzeitig fehlen aber auch in der KI so ein paar Sachen. Also es ist jetzt nicht besser oder schlechter, es ist aber anders.“* (Et2 - Z. 4 ff.)

Diese Ergebnisse zeigen, dass die Verwendung von GenAI zur Generierung von User Stories unterschiedlich wahrgenommen wird, wobei die Qualität je nach spezifischem Aspekt variieren kann. Bemängelte Qualität führt dazu, dass die entsprechenden User Stories verworfen und neu generiert werden. Die Evaluationsteilnehmer werden angeregt, nur jene User Stories zu übernehmen und bei Bedarf anzupassen, die sie als qualitativ hochwertig einschätzen. Dazu tragen die 38 von 61 generierten User Stories zu einer durchschnittlich positiven Gesamtbewertung bei. Dies reflektiert die Fähigkeit der GenAI, effektiv auf die Anforderungen der Nutzer zu reagieren, insbesondere wenn diese ihre Eingaben daraufhin anpassen und mit der nächsten generierten Ausgabe zufrieden sind.

5.3 Wahrgenommene Vorteile

Ein Teil der Teilforschungsfrage (TF 3) untersucht die Wahrnehmung der Unterstützung durch GenAI im RE aus Sicht der befragten Experten. Die Evaluationsergebnisse in Bezug auf die wahrgenommenen Vorteile bei der Nutzung der Webanwendung sind in Tabelle 14 dokumentiert. Hervorzuheben sind jene Vorteile, die sich auf die identifizierten Herausforderungen im RE abbilden lassen, wie in Abbildung 8 dargestellt.

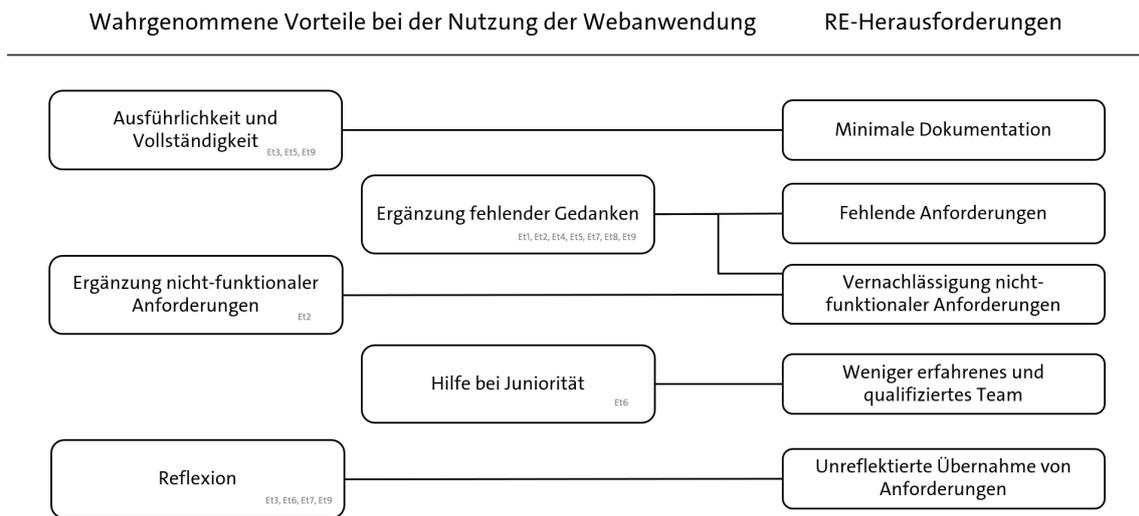


Abbildung 8: Zuordnung der wahrgenommenen Vorteile auf die identifizierten RE-Herausforderungen (Eigene Darstellung)

Ein Vorteil ist die Ausführlichkeit und Vollständigkeit der von Gemini Pro generierten Ausgabe. Einige Teilnehmer empfinden die bereitgestellten Inhalte teilweise als selbstverständlich und hätten diese selbst nicht dokumentiert (Et3 - Z. 61 ff., Et5 - Z. 55 ff., Et9 - Z. 94 ff.). Dies kann die Dokumentation innerhalb des RE-Prozesses vervollständigen und geht auf die Herausforderung der *minimalen Dokumentation* ein. Auch die Ergänzung nicht-funktionaler Anforderungen (Et2 - Z. 88 ff.) und fehlender Gedanken (Et1 - Z. 31 ff., Et2 - Z. 92 ff., Et4 - Z. 55 ff., Et5 - Z. 40 ff., Et7 - Z. 32 ff., Et8 - Z. 54 ff., Et9 - Z. 101 ff.) werden als Vorteile aufgenommen. Diese Ergänzungen tragen dazu bei *vernachlässigte nicht-funktionale Anforderungen* und *fehlende Anforderungen* zu adressieren. Die Funktionalität des Prompts, Schwachstellen innerhalb der Anforderungsspezifikation zu identifizieren, wird in diesem Zusammenhang als nützlich bewertet (Et4 -Z. 15 ff., 56 f. Et9 - Z. 37 f.). Diese Funktion unterstützt die Nutzer bei der Identifizierung und Verbesserung potenzieller Mängel in den Entwürfen der GenAI. Dadurch wird eine Reflexion und kritische Auseinandersetzung mit den generierten User Stories gefördert (Et3 - Z. 85 f., Et6 - Z. 77 ff., Et7 - Z. 77 ff., Et9 - Z. 98 ff.). Somit wird die RE-Herausforderung der *unreflektierten Übernahme von Anforderungen* adressiert.

Wahrgenommener Vorteil	Referenz
Ausführlichkeit und Vollständigkeit	Et3 - Z. 61 ff. Et5 - Z. 55 ff. Et9 - Z. 94 ff.
Ergänzung nicht-funktionaler Anforderungen	Et2 - Z. 88 ff.
Zeitersparnis	Et1 - Z. 31 Et2 - Z. 89 Et3 - Z. 47 ff. Et4 - Z. 53 ff. Et5 - Z. 50 ff. Et8 - Z. 54 ff. Et9 - Z. 93 ff.
Ergänzung fehlender Gedanken	Et1 - Z. 31 ff. Et2 - Z. 92 ff. Et4 - Z. 55 ff. Et5 - Z. 40 ff. Et7 - Z. 32 ff. Et8 - Z. 54 ff. Et9 - Z. 101 ff.
Hilfe bei Junorität	Et6 - Z. 77 ff.
Reflexion	Et3 - Z. 85 f. Et6 - Z. 77 ff. Et7 - Z. 77 ff. Et9 - Z. 98 ff.
Einnahme weiterer Perspektive	Et3 - Z. 85 ff. Et7 - Z. 74 ff. Et8 - Z. 53 ff.
Abdeckung weiterer Prozesse	Et7 - Z. 42 ff., 110 ff.
Einfachheit	Et9 - Z. 33 ff.
Einheitlichkeit	Et2 - Z. 95, 101 ff.

Tabelle 14: Wahrgenommene Vorteile in der Evaluation

Ein Vorteil ist die Ausführlichkeit und Vollständigkeit der von Gemini Pro generierten Ausgabe. Einige Teilnehmer empfinden die bereitgestellten Inhalte teilweise als selbstverständlich und hätten diese selbst nicht dokumentiert (Et3 - Z. 61 ff., Et5 - Z. 55 ff., Et9 - Z. 94 ff.). Dies kann die Dokumentation innerhalb des RE-Prozesses vervollständigen

und geht auf die Herausforderung der *minimalen Dokumentation* ein. Auch die Ergänzung nicht-funktionaler Anforderungen (Et2 - Z. 88 ff.) und fehlender Gedanken (Et1 - Z. 31 ff., Et2 - Z. 92 ff., Et4 - Z. 55 ff., Et5 - Z. 40 ff., Et7 - Z. 32 ff., Et8 - Z. 54 ff., Et9 - Z. 101 ff.) werden als Vorteile aufgenommen. Diese Ergänzungen tragen dazu bei *vernachlässigte nicht-funktionale Anforderungen* und *fehlende Anforderungen* zu adressieren. Die Funktionalität des Prompts, Schwachstellen innerhalb der Anforderungsspezifikation zu identifizieren, wird in diesem Zusammenhang als nützlich bewertet (Et4 - Z. 15 ff., 56 f. Et9 - Z. 37 f.). Diese Funktion unterstützt die Nutzer bei der Identifizierung und Verbesserung potenzieller Mängel in den Entwürfen der GenAI. Dadurch wird eine Reflexion und kritische Auseinandersetzung mit den generierten User Stories gefördert (Et3 - Z. 85 f., Et6 - Z. 77 ff., Et7 - Z. 77 ff., Et9 - Z. 98 ff.). Somit wird die RE-Herausforderung der *unreflektierten Übernahme von Anforderungen* adressiert.

Ein Evaluationsteilnehmer betont, dass die Webanwendung insbesondere für weniger erfahrene Teammitglieder von Vorteil sein könnte, indem sie ihnen hilft, von Beginn an professionelle User Stories zu verfassen (Et6 - Z. 77 ff.). Diese Beobachtung spricht die RE-Herausforderung eines *weniger erfahrenen und qualifizierten Teams* an und zeigt, dass GenAI sowohl als Werkzeug zur Inhaltentwicklung als auch als Lern- und Trainingsinstrument dienen kann.

Zusammenfassend lassen die Evaluationsergebnisse der wahrgenommenen Vorteile erkennen, dass GenAI einige der in der Literatur und den RE-Experteninterviews identifizierten RE-Herausforderungen lösen kann. Gleichzeitig kann sowohl die Effizienz im RE gesteigert werden als auch dazu beigetragen werden die Qualität der Anforderungsdokumentation zu verbessern. Dies kann beispielsweise durch Reflexion der generierten Anforderungen ermöglicht werden.

5.4 Wahrgenommene Herausforderungen

Die Beantwortung des zweiten Teils der Teilforschungsfrage (TF 3) beleuchtet die wahrgenommenen Herausforderungen der RE-Stakeholder beim Einsatz von GenAI im RE. Diese Herausforderungen werden in drei Kategorien gegliedert: identifizierte Herausforderungen, technisch lösbare Herausforderungen und Hypothesen.

Zunächst spiegeln die wahrgenommenen Herausforderungen, dokumentiert in Tabelle 15, jene wider, die auch die GenAI-Experten beschreiben (siehe Kapitel 4.2.2). Während der Evaluationen reflektieren die Teilnehmer über die von Gemini Pro generierten User Stories. Jedoch wird das *übermäßige Vertrauen von Nutzern* in die Ausgaben als ein potenziell auftretendes Problem aufgeführt (Et2 - Z. 52 ff., Et3 - Z. 71 ff., Et4 - Z. 50 ff., Et8 - Z. 58 ff., Et9 - Z. 64 ff., Et10 - Z. 73 ff.). Dies kann zu einer Vernachlässigung der Überprüfung der

Ausgaben führen. Insbesondere könnte dies in zeitkritischen Situationen auftreten, wodurch Fehler in den Entwicklungsprozess gelangen können. Zudem werden *Halluzinationen* als Herausforderung identifiziert (Et2 - Z. 44 ff., Et5 - Z. 15 ff., Et6 - Z. 82 ff., Et9 - Z. 61 ff.). Halluzinationen könnten beispielsweise in Verbindung mit einem übermäßigen Vertrauen zur Übernahme von Fehlern führen.

Wahrgenommene Herausforderung	Referenz
Übermäßiges Vertrauen von Nutzern	Et2 - Z. 52 ff. Et3 - Z. 71 ff. Et4 - Z. 50 ff. Et8 - Z. 58 ff. Et9 - Z. 64 ff. Et10 - Z. 73 ff.
Halluzinationen	Et2 - Z. 44 ff. Et5 - Z. 15 ff. Et6 - Z. 82 ff. Et9 - Z. 61 ff.

Tabelle 15: Wahrgenommene Herausforderungen in der Evaluation

Die technisch lösbaren Herausforderungen, aufgeführt in Tabelle 16, umfassen unter anderem das Fehlen von Daten im Sprachmodell. Durch Lernmechanismen könnte dies verbessert werden, sodass das Modell lernt, wie gute User Stories für die nutzende Person aussehen (Et1 - Z. 47 ff.). Die Integration von Best Practices, die Nutzer selbst einbringen könnten, wird ebenfalls in diesem Kontext vorgeschlagen (Et7 - Z. 103 ff.). Aktuell fehlende Multimodalität, wie der Upload von Bildern, und eine fehlende Interaktivität, wie etwa eine Chatfunktion (Et2 - Z. 106 ff.), könnten technisch ergänzt werden. Datenschutzbedenken auf Projektebene (Et4 - Z. 48 ff.) werden in diesem Fall durch die Nutzung der Gemini-API adressiert.

Technisch Lösbar / Neue Funktionalität	Referenz
Fehlende Daten innerhalb des Modells	Et1 - Z. 39 ff.
Fehlende Multimodalität	Et2 - Z. 104 ff.
Fehlende Interaktivität	Et2 - Z. 106 ff.
Datenschutzbedenken	Et4 - Z. 48 ff.

Tabelle 16: Technisch lösbare wahrgenommene Herausforderungen in der Evaluation

Im Rahmen der Herausforderungen erwähnte Aussagen der Evaluationsteilnehmer können als Hypothesen in Tabelle 17 interpretiert werden, die in weiteren DSR-Zyklen überprüfbar wären. So könnte beispielsweise die Effizienz von GenAI bei der Erstellung anderer Ausgaben im RE-Prozess, wie beispielsweise BPMN-Diagrammen, getestet werden (Et2 - Z. 39 ff.). Die Vermutung, dass eine derzeitige Unvertrautheit mit der Webanwendung zu schlechteren Ergebnissen führt (Et1 - Z. 35 ff., Et7 - Z. 20 ff.), könnte durch weitere Iterationen mit den gleichen Personen untersucht werden. Es wird auch vermutet, dass GenAI in komplexeren Projektumfeldern weniger nützlich sein könnte (Et2 - Z. 12 ff.). Abschließend weißt ein Evaluationsteilnehmer darauf hin, wie auch schon in Kapitel 5.1 aufgeführt, dass ein Mehraufwand entstehen könnte, um Konsistenz in den User Stories zu gewährleisten, wenn zu viele generiert werden (Et6 - Z. 62 ff.).

Hypothesen	Referenz
Schnelligkeit bei anderen Ausgaben könnte stagnieren	Et2 - Z. 39 ff.
Unvertrautheit mit der Webanwendung führt aktuell zu schlechterer Ausgabe	Et1 - Z. 35 ff. Et7 - Z. 20 ff.
Kein Nutzen bei höherer Komplexität des Projektkontextes	Et2 - Z. 12 ff.
Mehraufwand zur Erstellung Konsistenz	Et6 - Z. 62 ff.

Tabelle 17: Aufgestellte Hypothesen in der Evaluation

5.5 Beobachtungen während der Simulation

Die Simulationen mit der Thinking-Aloud Methode innerhalb der Evaluation ergeben weitere Beobachtungen zur Nutzung der GenAI-gestützten Webanwendung.

Trotz der Verfügbarkeit von drei optionalen textuellen Eingabefeldern füllen die Teilnehmer in fast allen Fällen alle Textfelder aus. Dies führt zu einer seltenen Nutzung spezifischer Funktionalitäten innerhalb des Prompts. Beispielsweise die Möglichkeit, dass zu einer Zielgruppe im Projektkontext eine Funktionalität und ein Nutzen generiert werden soll.

Einige Teilnehmer experimentieren mit der Kreativitätseinstellung. Es kann hierbei jedoch keine Präferenz herausgestellt werden, da Ausgaben mit geringerer Kreativität teilweise

qualitativ besser und teilweise schlechter wahrgenommen werden. Gleiches gilt bei erhöhter Kreativität. Dies könnte ebenfalls im Rahmen weiterer Forschungsarbeiten untersucht werden.

Die ausgegebenen User Stories enthalten in einigen Fällen die erwarteten Ergebnisse zu den definierten Testszenarien. Dies wird positiv aufgenommen. Es kann diskutiert werden, ob diese Ergebnisse der Testszenarien fester Bestandteil der generierten User Story sein sollten. Ebenfalls werden die von der GenAI hervorgehobenen Schwachstellen in den User Stories als nützlich betrachtet. Diese werden teilweise direkt in die Akzeptanzkriterien oder Testszenarien der zuletzt generierten User Story eingearbeitet oder zur Verbesserung der nächsten Eingabe verwendet.

Ein Evaluationsteilnehmer sendet wiederholt dieselben Eingaben ab, um Konsistenz innerhalb der Ausgaben zu überprüfen. Dabei wird angeregt, eine Funktion einzuführen, die es ermöglicht, mehrere User Stories zu einer Eingabe gleichzeitig zu generieren und diese vergleichend darzustellen. Zudem zeigt sich, dass mit der Eingabe umfangreicherer Kontextinformationen zielgerichtetere Ergebnisse erzielt werden können.

Wenn User Stories mehr als eine Funktionalität umfassen, werden sie als Epics betrachtet (Cohn 2004). Innerhalb der Evaluationen wurden User Stories teilweise als Epics interpretiert. Dies wurde neutral bewertet, da aus diesen Epics User Stories abgeleitet werden können.

Diese Beobachtungen deuten darauf hin, dass die Feinabstimmung des Ergebnisartefakts ein Aspekt für die Optimierung des Einsatzes von GenAI im RE darstellt.

6 Diskussion

In diesem Kapitel werden die Ergebnisse der Forschungsarbeit interpretiert. Außerdem wird die Methodik kritisch reflektiert sowie ein möglicher Ausblick für zukünftige Forschungsbestrebungen analysiert, die an diese Forschungsarbeit anschließen können.

6.1 Interpretation der Ergebnisse

In diesem Abschnitt der Diskussion werden die Ergebnisse aus der Interaktion der Evaluationsteilnehmer mit der entwickelten Webanwendung für das Erstellen von User Stories interpretiert. In der Analyse wird insbesondere untersucht, wie die Teilnehmer die Anwendung wahrnehmen und welche spezifischen Vorteile und Herausforderungen sie identifizieren.

Ein wesentliches Ergebnis der Evaluation ist, dass ein Großteil der Teilnehmer einen zeitlichen Vorteil beim Einsatz der Webanwendung im Vergleich zur manuellen Methode zur Erstellung von User Stories sieht. Diese Tendenz zeigt, dass das Erstellen von User Stories mithilfe von GenAI schneller möglich ist. Diese Produktivitätssteigerung wird als Vorteil wahrgenommen und hebt die praktische Relevanz der Anwendung hervor.

Die Evaluationsergebnisse verdeutlichen darüber hinaus, dass durch die Nutzung der Webanwendung im hybriden Ansatz User Stories von höherer Qualität erstellt werden können. Im Rahmen dieses Ansatzes ist zudem eine sorgfältige Prüfung und gegebenenfalls eine Anpassung der generierten User Stories ein integraler Bestandteil des Prozesses. GenAI ergänzt fehlende Gedanken der Evaluationsteilnehmer und liefert somit vollständigere Entwürfe im Vergleich zu einem manuellen Ansatz zur Erstellung von User Stories. Dieses Ergebnis unterstreicht die Fähigkeit der Anwendung, umfassende und detaillierte User Stories generieren zu können, die sonst möglicherweise wichtige Aspekte vernachlässigen würden.

Der Zuordnung der wahrgenommenen Vorteile auf bestehende RE-Herausforderungen kann entnommen werden, dass die Webanwendung zur Lösung oder Minderung mehrerer spezifischer Herausforderungen beiträgt. Dazu zählen unter anderem die *minimale Dokumentation* und *Fehlende Anforderungen*, wie in Abbildung 8 dargestellt.

Trotz der positiven Aspekte werden auch Herausforderungen im Einsatz der Webanwendung wahrgenommen. Diese liegen im *übermäßigen Vertrauen von Nutzern* in die Ausgaben von GenAI und in möglich auftretenden *Halluzinationen* (siehe Tabelle 15). Diese Punkte spiegeln die Notwendigkeit wider, die Nutzer hinsichtlich der Möglichkeiten und Grenzen der GenAI weiter zu sensibilisieren.

Abgesehen von den Ergebnissen der Evaluation wurde innerhalb der Forschungsarbeit die Entscheidung getroffen, auf die direkte Nutzung einer Chat-Oberfläche eines generativen Sprachmodells zu verzichten. Diese Entscheidung basiert auf mehreren Vorteilen. Zum einen nutzt die im Rahmen der Forschungsarbeit entwickelte Webanwendung einen Prompt, der auf Richtlinien der wissenschaftlichen Literatur basiert. Die Gestaltung der Prompt nach diesen Richtlinien gewährleistet Konsistenz und hohe Qualität im Input (Arvidsson und Axell 2023). Dies führt zu verlässlichen und konstanten Ausgaben, im Gegensatz zu oft variablen Ergebnissen einfacher Chat-Prompts (Ye et al. 2023). Zum anderen adressiert die Nutzung des Sprachmodells über eine API Datenschutzbedenken. Im Gegensatz zur direkten Nutzung von beispielsweise Gemini Pro werden hierbei keine sensiblen Daten unkontrolliert übermittelt. Dies bietet nicht nur technische, sondern auch datenschutzrechtliche Vorteile, die in verschiedenen Unternehmenskontexten von Bedeutung sind.

6.2 Beitrag zur Literatur

Die Entwicklung und Implementierung des Ergebnisartefakts im Rahmen dieser Forschungsarbeit stellt einen Beitrag zur wissenschaftlichen Literatur dar. Dieser Abschnitt thematisiert diesen Beitrag und ordnet ihn in den bestehenden Forschungskontext ein.

Zunächst bildet die systematische, wissenschaftliche Herangehensweise nach der DSR-Methodik nach Kuechler und Vaishnavi (2008) die Grundlage dieser Forschungsarbeit. Sie umfasst eine Literaturrecherche, die Durchführung von elf Experteninterviews, die Entwicklung des Ergebnisartefakts und zehn anschließende Evaluationen. Dieser holistische Ansatz ermöglicht es, Theorie und Praxis zu verbinden und eine interpretative Auswertung der Ergebnisse aus beiden Perspektiven vorzunehmen.

Das Ergebnisartefakt ist als *situative Implementierung* im DSR-Wissensbeitrags-Framework nach Gregor und Hevner (2013) einzuordnen. Es handelt sich um ein Softwareprodukt, das als „*Verbesserung*“ klassifiziert wird (Gregor und Hevner 2013). Dabei werden neue Lösungen für bereits bekannte Probleme entwickelt, was eine substanzielle Erweiterung der bestehenden Wissensbasis ermöglicht (Gregor und Hevner 2013). Diese neue Lösung bietet einen praktischen Lösungsansatz zur Minderung einiger der in der Literatur und Praxis identifizierten RE-Herausforderungen.

Die neu identifizierten RE-Herausforderungen, die aus den RE-Experteninterviews hervorgehen, können die bestehende Wissensbasis ergänzen. Um dies zu erreichen müssen diese RE-Herausforderungen jedoch noch durch weitere wissenschaftliche Methoden, wie beispielsweise einem quantitativen Ansatz, bestätigt werden.

Diese Forschungsarbeit ist in den aktuellen wissenschaftlichen Diskurs eingebettet und stellt eine Verbindung zu relevanten Studien im Bereich des Einsatzes von GenAI im RE

her. In diesem Kontext ist beispielsweise die Arbeit von Arvidsson und Axell (2023) zu erwähnen, die sich mit Richtlinien für das Prompt Engineering im RE beschäftigt. Die für das erstellte Ergebnisartefakt verwendeten Richtlinien sind somit aus der Literatur abgeleitet.

Ein weiterer Bezugspunkt ist die Arbeit von Goel et al. (2023), welche die Erstellung von Personas mittels GenAI untersucht. Diese Forschung zeigt auf, dass die mit GenAI generierten Personas in vielen Aspekten mit manuell erstellten Personas vergleichbar sind (Goel et al. 2023). Ähnliche Ergebnisse lassen sich aus der vorliegenden Arbeit ziehen. Die im Rahmen dieser Forschung generierten User Stories werden von den Evaluationsteilnehmern positiv bewertet. Dies macht sie vergleichbar mit manuell erstellten User Stories. Weiterhin unterstreicht dies nicht nur die Validität des entwickelten Prompts, sondern auch dessen praktische Relevanz im Kontext des RE. Darüber hinaus bietet die Arbeit von Goel et al. (2023) praktische Empfehlungen für einen effektiven und verantwortungsbewussten Einsatz dieser Technologie. Die für das in dieser Arbeit erstellte Ergebnisartefakt relevante Empfehlung wird in Kapitel 6.5 diskutiert.

Zusammenfassend leistet diese Forschungsarbeit durch die Entwicklung des Ergebnisartefakts einen Beitrag zur wissenschaftlichen Literatur. Sie bietet neue Perspektiven und Lösungsansätze für bestehende Probleme und eröffnet Möglichkeiten für weitere Forschungen im Bereich des Einsatzes von GenAI im RE.

6.3 Beitrag zur Praxis

Die Entwicklung und Bereitstellung der Webanwendung zur Erstellung von User Stories repräsentiert einen Fortschritt in der praktischen Anwendung von GenAI im RE. Das Ergebnisartefakt dieser Arbeit demonstriert, wie HI Ansätze die Produktivität und Qualität in der Erstellung von User Stories erhöhen können. Zum einen ermöglicht der Einsatz der GenAI Technologie eine Beschleunigung des Prozesses zur Erstellung von Anforderungsdokumenten, wie in diesem Fall von User Stories. Dies kann in agilen Softwareprojekten relevant sein, in welchen Zeit ein Erfolgsfaktor sein kann (Chow und D.-B. Cao 2008). Zum anderen unterstützt die Webanwendung verschiedene Stakeholder im Bereich des RE – darunter Requirements Engineers, Product Owner und Business Analysten – bei der effizienten Erstellung von qualitativ hochwertigen User Stories.

Die Einführung einer solchen GenAI gestützten Webanwendung trägt zudem zur Sensibilisierung für den Einsatz von HI im RE bei. Es wird durch die Bewertungen der Teilnehmer deutlich, dass ein solcher HI Einsatz nicht nur unterstützende Funktionen übernehmen kann, sondern auch aktiv zur Gestaltung und Verbesserung von Prozessen beiträgt. Gleichzeitig fordert der Einsatz einer solchen Technologie die Nutzer auf, sich kritisch mit den Grenzen

und Möglichkeiten von GenAI auseinanderzusetzen. Insbesondere die Herausforderungen wie das *übermäßige Vertrauen von Nutzern* in die generierten Ausgaben und das Phänomen der *Halluzinationen* werden adressiert.

Zum aktuellen Zeitpunkt besteht somit bereits Potenzial die Webanwendung in den Arbeitsalltag im RE zu integrieren, jedoch steht die vollständige Produktreife noch aus. Zukünftige Entwicklungen könnten nicht nur die Erstellung von User Stories, sondern auch weitere Anforderungsdokumente wie Lasten- und Pflichtenhefte sowie eine Unterstützung bei der Erstellung von Projektangeboten umfassen. Solche Erweiterungen würden es Unternehmen ermöglichen, schnell auf Marktveränderungen zu reagieren.

Die Entwicklung und Integration solcher Tools in die Praxis, basierend auf einem wissenschaftlichen Ansatz, kann entscheidend sein, um eine effektive Verbindung zwischen theoretischer und praktischer Anwendung herzustellen. Zusätzlich ist es möglich die Leistungsfähigkeit der RE-Prozesse in IT-Dienstleistungsunternehmen nachhaltig zu steigern.

6.4 Kritische Reflexion und Limitationen

In der kritischen Reflexion der durchgeführten Arbeit zeigen sich verschiedene Limitationen und Bereiche für potenzielle Verbesserungen.

Die Arbeit beschränkt sich auf einen einzigen Durchlauf im DSR-Zyklus, der eigentlich als iterativer Prozess konzipiert ist. Diese Einschränkung ist mit dem zeitlich definierten Zeitraum und den verfügbaren personellen Ressourcen zu begründen. Jedoch kann dies dazu geführt haben, dass nicht alle möglichen Verbesserungen und Anpassungen erfasst wurden, die durch weitere Iterationen hätten identifiziert werden können.

Die Auswahl der Interview- und Evaluationsteilnehmer erfolgt ausschließlich aus dem Partnerunternehmen. Diese stammen zwar aus verschiedenen Bereichen, was eine gewisse Perspektivdiversität bietet, jedoch könnten zusätzliche Sichtweisen aus anderen Unternehmen die Generalisierbarkeit und Relevanz der Ergebnisse erhöhen. Zusätzlich würde dies helfen, mögliche Biases zu verringern. Die Interviews und Evaluationen unterliegen innerhalb der Arbeit zudem zeitlichen Beschränkungen, die durch die verfügbaren Ressourcen der Teilnehmer getroffen wurden. Eine weitere Limitation ist die Evaluation der Webanwendung, die von zehn RE-Experten durchgeführt wurde. Eine Evaluation mit einer größeren Anzahl von Experten könnte die Validität der Ergebnisse zusätzlich stärken.

Eine mögliche Verzerrung in den Interviews durch das Auflisten der RE-Herausforderungen und -Qualitätskriterien ist zudem zu berücksichtigen. Das qualitative Vorgehen der Arbeit könnte durch weitere methodische Ansätze, wie beispielsweise einem quantitativen Vorgehen, ergänzt werden, um ein umfassenderes Problemverständnis zu erlangen.

Eine weitere Limitation liegt in der ausschließlichen Konzentration auf die Generierung von User Stories. Eine Ausweitung auf weitere Anforderungsdokumente könnte die Anwendbarkeit und den Nutzen der Webanwendung erhöhen. Zudem wurde die Evaluation nur in einem spezifischen fiktiven Projektkontext durchgeführt, was die Übertragbarkeit der Ergebnisse auf andere Kontexte einschränkt.

Der Fokus liegt auf dem Gemini 1.0 Pro Modell in deutscher Sprache, da sowohl die Forschungsarbeit als auch die Mehrheit der Mitarbeiter und Kunden des Partnerunternehmens deutschsprachig sind. Zukünftige Untersuchungen könnten davon profitieren, verschiedene Sprachmodelle und Sprachen zu testen. Dies könnte die Einsatzmöglichkeiten und Leistungsfähigkeit der Webanwendung weiter diversifizieren.

Abschließend liegt kein Fokus auf der technischen Umsetzung der Webanwendung oder einer ausgereiften Benutzeroberfläche. Der Fokus liegt auf der pragmatischen Umsetzung, wie GenAI bei der Erstellung von User Stories helfen kann. Dies kann zur Folge haben, dass das volle Potenzial der Webanwendung nicht demonstriert werden kann oder keine breite Akzeptanz unter den Endnutzern erreicht wird.

6.5 Ausblick

Aus den identifizierten Limitationen der vorliegenden Arbeit ergibt sich eine mögliche Richtung für zukünftige Forschungsarbeiten. Diese können das Potenzial für Verbesserungen und Erweiterungen der entwickelten Webanwendung aufzeigen.

Eine Empfehlung ist die Durchführung weiterer Iterationen im DSR-Zyklus. Dabei könnten die gesammelten Ergebnisse und das Feedback in die Weiterentwicklung der Webanwendung einbezogen werden. Insbesondere bietet weiteres Prompt Engineering Potenzial für Verbesserungen, indem der verwendete Prompt auf Basis des Nutzerfeedbacks angepasst wird. Außerdem könnten zukünftige Forschungsarbeiten darauf abzielen, die bereits zugeordneten Vorteile der Webanwendung auf die RE-Herausforderungen weiter zu analysieren, um effektive Lösungen zur Minimierung dieser Herausforderungen zu entwickeln. Zusätzlich könnten die wahrgenommenen Herausforderungen und die identifizierten Hypothesen der Evaluationsteilnehmer bei weiteren DSR-Iterationen in Betracht gezogen werden.

Ein weiterer Schritt ist die Ausweitung der Evaluation auf weitere Unternehmen. Dies könnte nicht nur die Validität der Ergebnisse stärken, sondern auch einen weiteren Einblick in den Praxiseinfluss der Webanwendung ermöglichen. Durch diese Ausweitung könnten weitere unterschiedliche Unternehmenskontexte und Anforderungen berücksichtigt werden.

Um die Validität der Ergebnisse weiter zu stärken, könnte eine Ergänzung des qualitativen Vorgehens um ein quantitatives Vorgehen hilfreich sein. Quantitative Methoden könnten

dabei unterstützen, die Wirksamkeit der Webanwendung statistisch zu belegen und Einblicke in verschiedene Nutzungskontexte ermöglichen.

Die Nutzung der Webanwendung in verschiedenen realen Kontexten könnte ebenfalls in Betracht gezogen werden. Dies könnte die Flexibilität und Skalierbarkeit der Webanwendung unter Beweis stellen. Zusätzlich könnte damit gezeigt werden, wie sich diese in unterschiedlichen Projektkontexten bewährt.

Darüber hinaus bietet die Analyse weiterer Sprachmodelle die Möglichkeit, die Leistungsfähigkeit der Prompt zu vergleichen und zu verbessern. Die Performance von Gemini Pro unterliegt beispielsweise der von GPT-4 Turbo in allgemeinen Fähigkeiten zur Entwicklung von Text und Code (C. Lu et al. 2024). Insbesondere der Einsatz unterschiedlicher Modelle könnte aufzeigen, welche Konfigurationen in verschiedenen Anwendungsgebieten die besten Ergebnisse liefern. Außerdem könnte auch der Ansatz des Blendings analysiert werden. Bei diesem Ansatz werden mehrere kleinere Sprachmodelle kombiniert. Diese Kombination ist leistungsfähiger als größere Modelle wie GPT-3.5 (X. Lu et al. 2024).

Eine Erweiterung der Webanwendung um zusätzliche Techniken innerhalb des Sprachmodells wie dem Fine-Tuning, Guardrails, Reranking und weiteren Few-Shot-Beispielen könnte ebenfalls analysiert werden. Eine mögliche Erweiterung für die Webanwendung könnte die Implementierung einer Funktion sein, die es erlaubt, die von einem Nutzer übernommenen und angepassten User Stories zu speichern. Die gesammelten Daten könnten dazu verwendet werden, das Modell weiter zu trainieren. Dies kann dazu führen, dass es mehr über den übergebenen Projektkontext lernt und besser versteht, wie qualitativ hochwertige User Stories für den Nutzer aussehen. Die Integration einer Chat-Funktionalität mit der generierten Ausgabe könnte zudem beispielsweise eine interaktive und nutzerzentrierte Komponente hinzufügen. Diese würde es in Iterationen ermöglichen, präzisere und kontextbezogenere Anforderungen zu generieren (Goel et al. 2023).

Ein weiterer Aspekt könnte die Forschung innerhalb der User Experience (UX) sein, um die Benutzerinteraktion mit der Webanwendung zu verbessern. Die Entwicklung einer intuitiven und benutzerfreundlichen Benutzeroberfläche könnte ein entscheidender Schritt sein, die Akzeptanz und das Engagement der Endnutzer zu erhöhen.

Schließlich könnte die Möglichkeit einer Erweiterung des Umfangs der Webanwendung auf weitere Anforderungsdokumente in Betracht gezogen werden. Dafür müsste der Prompt dahingehend angepasst werden. Dies könnte die Anwendbarkeit in weiteren Bereichen des RE erweitern und dabei helfen, weitere Anforderungsdokumente zu erstellen.

Durch die Umsetzung dieser Empfehlungen könnte die Webanwendung nicht nur in seiner Funktionalität und Benutzerfreundlichkeit verbessert werden, sondern auch einen Beitrag zur Forschung und Praxis im Bereich des RE leisten.

7 Fazit

Diese Forschungsarbeit untersucht die Anwendung von GenAI zur Steigerung der Qualität und Produktivität im RE im Kontext von HI. Ziel der Arbeit ist es, Potenziale und Herausforderungen innerhalb dieser Anwendung zu identifizieren und zu evaluieren. Die Durchführung erfolgt auf Basis der DSR-Methodik. Im Rahmen des DSR werden die Problemstellung präzisiert, das Ergebnisartefakt entwickelt und dessen Evaluation durchgeführt. Das entwickelte Ergebnisartefakt ist ein Prompt, der durch die Einbindung zusätzlicher Informationen von Nutzern die Generierung von User Stories unterstützt. Zudem wird eine Webanwendung bereitgestellt, die eine Schnittstelle zur Interaktion mit dem Sprachmodell bietet.

Anhand der zentralen Forschungsfrage wird untersucht, welchen Einfluss der Einsatz von GenAI im Kontext von HI auf die Produktivität und Qualität im RE-Prozess hat. Die Evaluationsergebnisse zeigen, dass der Einsatz von GenAI im RE-Prozess zu einer Steigerung der Produktivität führt. Diese Produktivitätssteigerung ist auf die beschleunigte Generierung von User Stories zurückzuführen. Darüber hinaus zeigt sich eine Verbesserung der Qualität der User Stories durch den HI-Ansatz. Der Einsatz von GenAI ermöglicht eine Beschleunigung im Vergleich zur manuellen Erstellung von User Stories. Der hybride Ansatz stellt sicher, dass die generierten Inhalte vollständig und qualitativ hochwertig sind. Die erreichte Qualität der User Stories wäre manuell, insbesondere unter Berücksichtigung der zeitlichen Beschränkungen in der Evaluation, schwer erreichbar. Die Hypothese, dass HI eine erhöhte Produktivität und Qualität bewirkt (siehe 1.2), wird durch diese Ergebnisse unterstützt. Allerdings werden auch User Stories generiert, die von den Evaluationsteilnehmern als unzureichend bewertet und daher verworfen werden.

Um die zentrale Forschungsfrage zu beantworten, werden verschiedene Teilforschungsfragen adressiert. Im Rahmen der Teilforschungsfrage (TF 1) wird die Problemstellung präzisiert. Dabei werden Herausforderungen im RE identifiziert. Die literaturbasierte Untersuchung ergibt 26 RE-Herausforderungen. Zusätzlich werden in den RE-Experteninterviews acht weitere Herausforderungen ermittelt.

Zur Beantwortung der Teilforschungsfrage (TF 2) wird untersucht, wie GenAI genutzt werden kann, um die identifizierten RE-Herausforderungen zu mindern oder zu lösen. Anhand der Aussagen der Experten werden Design Prinzipien aufgestellt. Diese Design Prinzipien legen fest, dass GenAI sowohl zur automatischen Erstellung von User Stories als auch zur Einhaltung von Qualitätsstandards eingesetzt werden soll. Darüber hinaus werden Herausforderungen und Limitationen von GenAI berücksichtigt, die über den Einsatz im RE hinausgehen. Dazu gehören Halluzinationen, die Sicherheit der bereitgestellten Daten der Nutzer und das übermäßige Vertrauen der Nutzer in die Ausgaben von GenAI.

Das Ergebnisartefakt wird in der Evaluation zur automatischen Erstellung von User Stories eingesetzt. Die Unterstützung durch GenAI wird von Experten im Rahmen der Beantwortung der Teilforschungsfrage (TF 3) überwiegend positiv bewertet. Sie nehmen eine Produktivitätssteigerung und eine mögliche Erhöhung der Qualität der generierten User Stories wahr. Eine Zuordnung der genannten Vorteile der Webanwendung durch die Experten zu den identifizierten Herausforderungen im RE zeigt, dass einige dieser Herausforderungen durch den Einsatz von GenAI adressiert werden können. Dennoch identifizieren die Experten zusätzliche Herausforderungen, die bei der Zusammenarbeit mit GenAI entstehen können. Dabei werden die Herausforderungen genannt, die bereits im Rahmen der GenAI-Experteninterviews aufgenommen wurden. Diese Ergebnisse unterstreichen die Notwendigkeit einer sorgfältigen Überprüfung und Validierung der von GenAI generierten Inhalte.

Die Ergebnisse dieser Forschungsarbeit deuten darauf hin, dass GenAI als Werkzeug eingesetzt werden kann, um bestehenden Herausforderungen im Bereich des RE zu begegnen. Diese Forschungsarbeit leistet einen Beitrag zur wissenschaftlichen Literatur, indem sie die Integration von GenAI im RE untersucht und Implikationen für die praktische Nutzung von GenAI aufzeigt. Darüber hinaus bietet sie Erkenntnisse für die Weiterentwicklung von GenAI-Anwendungen im RE und betont die Bedeutung einer ausgewogenen Nutzung von menschlicher und künstlicher Intelligenz.

Zukünftige Forschungsarbeiten könnten die Erweiterung der Webanwendung um zusätzliche Funktionen oder die Generierung weiterer Anforderungsdokumente untersuchen. Ebenso könnten langfristige Auswirkungen der Nutzung von GenAI im RE analysiert und Strategien zur Bewältigung weiterer identifizierter Herausforderungen entwickelt werden.

A Literaturverzeichnis

- Abualhaija, S., Sabetzadeh, M. und Trujillo, J. (Sep. 2023): „AIRE 2023: 10th International Workshop on Artificial Intelligence and Requirements Engineering“. In: 2023 IEEE 31st International Requirements Engineering Conference Workshops (REW). Hannover, Germany: IEEE, S. 42–43. ISBN: 9798350326918. DOI: 10.1109/REW57809.2023.00014. [Letzter Aufruf: 18. 12. 2023].
- Ajagbe, M. und Zhao, L. (Aug. 2022): „Retraining a BERT Model for Transfer Learning in Requirements Engineering: A Preliminary Study“. In: 2022 IEEE 30th International Requirements Engineering Conference (RE), S. 309–315. DOI: 10.1109/RE54965.2022.00046. [Letzter Aufruf: 11. 01. 2024].
- Akata, Z., Balliet, D., de Rijke, M., Dignum, F., Dignum, V., Eiben, G., Fokkens, A., Grossi, D., Hindriks, K., Hoos, H., Hung, H., Jonker, C., Monz, C., Neerincx, M., Oliehoek, F., Prakken, H., Schlobach, S., van der Gaag, L., van Harmelen, F., van Hoof, H., van Riemsdijk, B., van Wynsberghe, A., Verbrugge, R., Verheij, B., Vossen, P. und Welling, M. (Aug. 2020): A Research Agenda for Hybrid Intelligence: Augmenting Human Intellect With Collaborative, Adaptive, Responsible, and Explainable Artificial Intelligence. In: *Computer* 53(8), S. 18–28. ISSN: 1558-0814. DOI: 10.1109/MC.2020.2996587. [Letzter Aufruf: 04. 10. 2023].
- Alam, S., Bhatti, S., Asim, S. und Shah, S. (Jan. 2017): *Impact and Challenges of Requirement Engineering in Agile Methodologies: A Systematic Review*. Techn. Ber. DOI: 10.14569/IJACSA.2017.080455#sthash.grDwPWWr.dpuf.
- Alavi, M., Leidner, D. und Mousavi, R. (Jan. 2024): A Knowledge Management Perspective of Generative Artificial Intelligence. In: *Journal of the Association for Information Systems* 25(1), S. 1–12. ISSN: 1536-9323. DOI: 10.17705/1jais.00859.
- Amna, A. R. und Poels, G. (Mai 2022): Ambiguity in User Stories: A Systematic Literature Review. In: *Information and Software Technology* 145, S. 106824. ISSN: 0950-5849. DOI: 10.1016/j.infsof.2022.106824. [Letzter Aufruf: 27. 12. 2023].
- Arvidsson, S. und Axell, J. (Aug. 2023): *Prompt Engineering Guidelines for LLMs in Requirements Engineering*. [Letzter Aufruf: 20. 03. 2024].
- Bano, M. (Aug. 2015): „Addressing the Challenges of Requirements Ambiguity: A Review of Empirical Literature“. In: 2015 IEEE Fifth International Workshop on Empirical Requirements Engineering (EmpiRE), S. 21–24. DOI: 10.1109/EmpiRE.2015.7431303. [Letzter Aufruf: 02. 01. 2024].

- BCG (Sep. 2023): *Generative AI*. Online abrufbar unter: <https://www.bcg.com/capabilities/artificial-intelligence/generative-ai>.
- Benbya, H., Strich, F. und Tamm, T. (Jan. 2024): Navigating Generative Artificial Intelligence Promises and Perils for Knowledge and Creative Work. In: *Journal of the Association for Information Systems* 25(1), S. 23–36. ISSN: 1536-9323. DOI: 10.17705/1jais.00861.
- Bjarnason, E., Unterkalmsteiner, M., Borg, M. und Engström, E. (Sep. 2016): A Multi-Case Study of Agile Requirements Engineering and the Use of Test Cases as Requirements. In: *Information and Software Technology* 77, S. 61–79. ISSN: 0950-5849. DOI: 10.1016/j.infsof.2016.03.008. [Letzter Aufruf: 26. 04. 2024].
- Bommasani, R. et al. (Juli 2022): *On the Opportunities and Risks of Foundation Models*. DOI: 10.48550/arXiv.2108.07258. arXiv: 2108.07258 [cs]. [Letzter Aufruf: 17. 01. 2024].
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I. und Amodei, D. (Juli 2020): *Language Models Are Few-Shot Learners*. DOI: 10.48550/arXiv.2005.14165. arXiv: 2005.14165 [cs]. [Letzter Aufruf: 16. 01. 2024].
- Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., Bamber, G. J., Beltran, J. R., Boselie, P., Lee Cooke, F., Decker, S., DeNisi, A., Dey, P. K., Guest, D., Knoblich, A. J., Malik, A., Paauwe, J., Papagiannidis, S., Patel, C., Pereira, V., Ren, S., Rogelberg, S., Saunders, M. N. K., Tung, R. L. und Varma, A. (2023): Human Resource Management in the Age of Generative Artificial Intelligence: Perspectives and Research Directions on ChatGPT. In: *Human Resource Management Journal* 33(3), S. 606–659. ISSN: 1748-8583. DOI: 10.1111/1748-8583.12524. [Letzter Aufruf: 20. 01. 2024].
- Cao, L. und Ramesh, B. (Jan. 2008): Agile Requirements Engineering Practices: An Empirical Study. In: *IEEE Software* 25(1), S. 60–67. ISSN: 1937-4194. DOI: 10.1109/MS.2008.1. [Letzter Aufruf: 29. 12. 2023].
- Chandra Kruse, L., Seidel, S. und Gregor, S. (März 2015): Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions. In: *Proceedings of the Annual Hawaii International Conference on System Sciences* 2015, S. 4039–4048. DOI: 10.1109/HICSS.2015.485.
- Chen, H., Jiao, F., Li, X., Qin, C., Ravaut, M., Zhao, R., Xiong, C. und Joty, S. (Jan. 2024): *ChatGPT's One-year Anniversary: Are Open-Source Large Language Models*

- Catching Up?* DOI: 10.48550/arXiv.2311.16989. arXiv: 2311.16989 [cs]. [Letzter Aufruf: 28. 05. 2024].
- Chen, H. und Ding, N. (Okt. 2023): *Probing the Creativity of Large Language Models: Can Models Produce Divergent Semantic Association?* DOI: 10.48550/arXiv.2310.11158. arXiv: 2310.11158 [cs]. [Letzter Aufruf: 15. 05. 2024].
- Chow, T. und Cao, D.-B. (Juni 2008): A Survey Study of Critical Success Factors in Agile Software Projects. In: *Journal of Systems and Software*. Agile Product Line Engineering 81(6), S. 961–971. ISSN: 0164-1212. DOI: 10.1016/j.jss.2007.08.020. [Letzter Aufruf: 12. 05. 2024].
- Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A., Yee, L. und Zimmel, R. (2023): *Economic Potential of Generative AI | McKinsey*. Online abrufbar unter: <https://www.mckinsey.com/capabilitiesmckinsey-digi/tal/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier%5C#business-and-society> [letzter Aufruf: 22. 02. 2024].
- Chui, M. und Yee, L. (Juli 2023): AI Could Increase Corporate Profits by \$4.4 Trillion a Year, According to New Research. In: *McKinsey & Company*.
- Cohn, M. (Feb. 2004): *User Stories Applied: For Agile Software Development*. USA: Addison Wesley Longman Publishing Co., Inc. ISBN: 978-0-321-20568-1.
- Curcio, K., Navarro, T., Malucelli, A. und Reinehr, S. (Mai 2018): Requirements Engineering: A Systematic Mapping Study in Agile Software Development. In: *Journal of Systems and Software* 139, S. 32–50. ISSN: 0164-1212. DOI: 10.1016/j.jss.2018.01.036. [Letzter Aufruf: 28. 12. 2023].
- Davu, S. (Sep. 2015): *What Characteristics Make Good Agile Acceptance Criteria?* [Letzter Aufruf: 21. 03. 2024].
- Dellermann, D., Ebel, P., Söllner, M. und Leimeister, J. M. (Jan. 2019): Hybrid Intelligence. In: *Business & Information Systems Engineering* 61(5), S. 637–643. ISSN: 2363-7005. DOI: 10.1007/s12599-019-00595-2.
- Dencik, J., Goehring, B. und Marshall, A. (Jan. 2023): Managing the Emerging Role of Generative AI in Next-Generation Business. In: *Strategy & Leadership* 51(6), S. 30–36. ISSN: 1087-8572. DOI: 10.1108/SL-08-2023-0079. [Letzter Aufruf: 29. 03. 2024].

- Devlin, J., Chang, M.-W., Lee, K. und Toutanova, K. (Mai 2019): *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. DOI: 10.48550/arXiv.1810.04805. arXiv: 1810.04805 [cs]. [Letzter Aufruf: 19. 02. 2024].
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koochang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., Carter, L., Chowdhury, S., Crick, T., Cunningham, S. W., Davies, G. H., Davison, R. M., Dé, R., Dennehy, D., Duan, Y., Dubey, R., Dwivedi, R., Edwards, J. S., Flavián, C., Gauld, R., Grover, V., Hu, M.-C., Janssen, M., Jones, P., Junglas, I., Khorana, S., Kraus, S., Larsen, K. R., Latreille, P., Laumer, S., Malik, F. T., Mardani, A., Mariani, M., Mithas, S., Mogaji, E., Nord, J. H., O’Connor, S., Okumus, F., Pagani, M., Pandey, N., Papagiannidis, S., Pappas, I. O., Pathak, N., Pries-Heje, J., Raman, R., Rana, N. P., Rehm, S.-V., Ribeiro-Navarrete, S., Richter, A., Rowe, F., Sarker, S., Stahl, B. C., Tiwari, M. K., van der Aalst, W., Venkatesh, V., Viglia, G., Wade, M., Walton, P., Wirtz, J. und Wright, R. (Aug. 2023): Opinion Paper: “So What If ChatGPT Wrote It?” Multidisciplinary Perspectives on Opportunities, Challenges and Implications of Generative Conversational AI for Research, Practice and Policy. In: *International Journal of Information Management* 71, S. 102642. ISSN: 0268-4012. DOI: 10.1016/j.ijinfomgt.2023.102642. [Letzter Aufruf: 17. 01. 2024].
- Elghariani, K. und Kama, N. (Aug. 2016): „Review on Agile Requirements Engineering Challenges“. In: 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), S. 507–512. DOI: 10.1109/ICCOINS.2016.7783267. [Letzter Aufruf: 27. 12. 2023].
- Ezzini, S., Abualhaija, S., Arora, C. und Sabetzadeh, M. (Feb. 2023): *AI-based Question Answering Assistance for Analyzing Natural-language Requirements*. DOI: 10.48550/arXiv.2302.04793. arXiv: 2302.04793 [cs]. [Letzter Aufruf: 11. 01. 2024].
- Ezzini, S., Abualhaija, S. und Sabetzadeh, M. (Juni 2022): *WikiDoMiner: Wikipedia Domain-specific Miner*. DOI: 10.48550/arXiv.2206.10218. arXiv: 2206.10218 [cs]. [Letzter Aufruf: 18. 12. 2023].
- Ferrara, E. (Nov. 2023): Should ChatGPT Be Biased? Challenges and Risks of Bias in Large Language Models. In: *First Monday*. ISSN: 1396-0466. DOI: 10.5210/fm.v28i11.13346. arXiv: 2304.03738 [cs]. [Letzter Aufruf: 27. 04. 2024].
- Floridi, L. und Chiriatti, M. (Dez. 2020): GPT-3: Its Nature, Scope, Limits, and Consequences. In: *Minds and Machines* 30(4), S. 681–694. ISSN: 1572-8641. DOI: 10.1007/s11023-020-09548-1. [Letzter Aufruf: 17. 01. 2024].

- Gemini API Pricing* | *Google AI for Developers* (2024). Online abrufbar unter: <https://ai.google.dev/pricing> [letzter Aufruf: 18. 05. 2024].
- Gemini-Team et al. (Dez. 2023): *Gemini: A Family of Highly Capable Multimodal Models*. DOI: 10.48550/arXiv.2312.11805. arXiv: 2312.11805 [cs]. [Letzter Aufruf: 19. 02. 2024].
- Gläser, J. und Laudel, G. (Juli 2010): *Experteninterviews Und Qualitative Inhaltsanalyse*. VS Verlag für Sozialwissenschaften Wiesbaden. ISBN: 978-3-531-17238-5. [Letzter Aufruf: 03. 01. 2024].
- Glinz, M. (Juli 2022): Glossary of Requirements Engineering Terminology. In: 2.0.1, S. 27. [Letzter Aufruf: 26. 12. 2023].
- Goel, T., Shaer, O., Delcourt, C., Gu, Q. und Cooper, A. (Sep. 2023): „Preparing Future Designers for Human-AI Collaboration in Persona Creation“. In: Proceedings of the 2nd Annual Meeting of the Symposium on Human-Computer Interaction for Work (CHIWORK '23). New York, NY, USA: Association for Computing Machinery, S. 1–14. ISBN: 9798400708077. DOI: 10.1145/3596671.3598574. [Letzter Aufruf: 28. 02. 2024].
- Google (2024): *Tips to Write Prompts for Gemini - Google Workspace Learning Center*. Online abrufbar unter: <https://support.google.com/a/users/answer/14200040?hl=en> [letzter Aufruf: 31. 03. 2024].
- Görer, B. und Aydemir, F. B. (Sep. 2023): „Generating Requirements Elicitation Interview Scripts with Large Language Models“. In: 2023 IEEE 31st International Requirements Engineering Conference Workshops (REW). Hannover, Germany: IEEE, S. 44–51. ISBN: 9798350326918. DOI: 10.1109/REW57809.2023.00015. [Letzter Aufruf: 18. 12. 2023].
- Gregor, S. und Hevner, A. (Juni 2013): Positioning and Presenting Design Science Research for Maximum Impact. In: *MIS Quarterly* 37, S. 337–356. DOI: 10.25300/MISQ/2013/37. 2.01.
- Hallmann, D. (2020): ““I Don’t Understand!”: Toward a Model to Evaluate the Role of User Story Quality”. In: V. Stray, R. Hoda, M. Paasivaara, and P. Kruchten (eds.): *Agile Processes in Software Engineering and Extreme Programming (Lecture Notes in Business Information Processing)*. Cham: Springer International Publishing, pp. 103–112. ISBN: 978-3-030-49392-9. DOI: 10.1007/978-3-030-49392-9_7.
- Heck, P. und Zaidman, A. (Jan. 2018): A Systematic Literature Review on Quality Criteria for Agile Requirements Specifications. In: *Software Quality Journal* 26(1), S. 127–160. ISSN: 0963-9314. DOI: 10.1007/s11219-016-9336-4.

- Hecker, D., Döbel, I., Petersen, U., Schmitz, V., Voss, A. und Rauschert, A. (2017): Zukunftsmarkt Künstliche Intelligenz - Potenziale und Anwendungen. In:
- Helfferrich, C. (2014): *Leitfaden- und Experteninterviews*. In: N. Baur und J. Blasius (Hrsg.): *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer Fachmedien, S. 559–574. ISBN: 978-3-531-18939-0. DOI: 10.1007/978-3-531-18939-0_39. [Letzter Aufruf: 03.01.2024].
- Hevner, A., March, S. T., Sudha, R. und Park, J. (Jan. 2004): *Design Science in Information Systems Research*. Bd. 28.
- Hoy, Z. und Xu, M. (Jan. 2023): Agile Software Requirements Engineering Challenges—Solutions—A Conceptual Framework from Systematic Literature Review. In: *Information* 14(6), S. 1–19. ISSN: 2078-2489. DOI: 10.3390/info14060322.
- Inayat, I., Salim, S. S., Marczak, S., Daneva, M. und Shamshirband, S. (Jan. 2015): A Systematic Literature Review on Agile Requirements Engineering Practices and Challenges. In: *Computers in Human Behavior* 51, S. 915–929. ISSN: 0747-5632. DOI: 10.1016/j.chb.2014.10.046.
- Jiang, E., Olson, K., Toh, E., Molina, A., Donsbach, A., Terry, M. und Cai, C. J. (Apr. 2022): „PromptMaker: Prompt-based Prototyping with Large Language Models“. In: *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA '22)*. New York, NY, USA: Association for Computing Machinery, S. 1–8. ISBN: 978-1-4503-9156-6. DOI: 10.1145/3491101.3503564. [Letzter Aufruf: 16.01.2024].
- Kar, S., Roy, C., Das, M., Mullick, S. und Saha, R. (Sep. 2023): AI Horizons: Unveiling the Future of Generative Intelligence. In: *International Journal of Advanced Research in Science, Communication and Technology*, S. 387–391. DOI: 10.48175/IJARST-12969.
- Koelsch, G. (Jan. 2016): *Requirements Writing for System Engineering*. Berkeley, CA: Apress. ISBN: 978-1-4842-2098-6. DOI: 10.1007/978-1-4842-2099-3.
- Köller, O. (2009): *Evaluation pädagogisch psychologischer Maßnahmen*. In: E. Wild und J. Möller (Hrsg.): *Pädagogische Psychologie (Springer-Lehrbuch)*. Berlin, Heidelberg: Springer, S. 333–352. ISBN: 978-3-540-88573-3. DOI: 10.1007/978-3-540-88573-3_14. [Letzter Aufruf: 08.02.2024].
- Kuechler, W. und Vaishnavi, V. (Jan. 2008): On Theory Development in Design Science Research: Anatomy of a Research Project. In: *EJIS* 17, S. 489–504.

- Kuenzel, F. (Juni 2022): *Unterschied Querschnittsanalyse & Längsschnittanalyse*. Online abrufbar unter: <https://www.acad-write.com/ratgeber/quantitative-datenanalyse/querschnittsanalysen-laengsschnittanalysen/> [letzter Aufruf: 03.01.2024].
- Lester, B., Al-Rfou, R. und Constant, N. (Nov. 2021): „The Power of Scale for Parameter-Efficient Prompt Tuning“. In: M.-F. Moens, X. Huang, L. Specia und S. W.-t. Yih (Hrsg.): *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Online und Punta Cana, Dominican Republic: Association for Computational Linguistics, S. 3045–3059. DOI: 10.18653/v1/2021.emnlp-main.243. [Letzter Aufruf: 16.01.2024].
- Lewis, C. (Jan. 1982): *Using the "Thinking-Aloud" Method in Cognitive Interface Design*.
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. und Stoyanov, V. (Juli 2019): *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. DOI: 10.48550/arXiv.1907.11692. arXiv: 1907.11692 [cs]. [Letzter Aufruf: 19.02.2024].
- Lu, C., Qian, C., Zheng, G., Fan, H., Gao, H., Zhang, J., Shao, J., Deng, J., Fu, J., Huang, K., Li, K., Li, L., Wang, L., Sheng, L., Chen, M., Zhang, M., Ren, Q., Chen, S., Gui, T., Ouyang, W., Wang, Y., Teng, Y., Wang, Y., Wang, Y., He, Y., Wang, Y., Wang, Y., Zhang, Y., Qiao, Y., Shen, Y., Mou, Y., Chen, Y., Zhang, Z., Shi, Z., Yin, Z. und Wang, Z. (Jan. 2024): *From GPT-4 to Gemini and Beyond: Assessing the Landscape of MLLMs on Generalizability, Trustworthiness and Causality through Four Modalities*. DOI: 10.48550/arXiv.2401.15071. arXiv: 2401.15071 [cs]. [Letzter Aufruf: 19.02.2024].
- Lu, X., Liu, Z., Liusie, A., Raina, V., Mudupalli, V., Zhang, Y. und Beauchamp, W. (Jan. 2024): *Blending Is All You Need: Cheaper, Better Alternative to Trillion-Parameters LLM*. DOI: 10.48550/arXiv.2401.02994. arXiv: 2401.02994 [cs]. [Letzter Aufruf: 05.05.2024].
- Lucassen, G., Dalpiaz, F., van der Werf, J. M. E. M. und Brinkkemper, S. (Sep. 2016a): *Improving Agile Requirements: The Quality User Story Framework and Tool*. In: *Requirements Engineering* 21(3), S. 383–403. ISSN: 1432-010X. DOI: 10.1007/s00766-016-0250-x. [Letzter Aufruf: 18.12.2023].
- Lucassen, G., Dalpiaz, F., Van Der Werf, J. M. E., and Brinkkemper, S. (Aug. 2015): “Forging High-Quality User Stories: Towards a Discipline for Agile Requirements”. In: *2015 IEEE 23rd International Requirements Engineering Conference (RE)*. Ottawa, ON, Canada: IEEE, pp. 126–135. ISBN: 978-1-4673-6905-3. DOI: 10.1109/RE.2015.7320415. [Letzter Aufruf: 12/16/2023].

- Lucassen, G., Dalpiaz, F., Werf, J. M. E. M. van der und Brinkkemper, S. (2016b): „The Use and Effectiveness of User Stories in Practice“. In: M. Daneva und O. Pastor (Hrsg.): *Requirements Engineering: Foundation for Software Quality (Lecture Notes in Computer Science)*. Cham: Springer International Publishing, S. 205–222. ISBN: 978-3-319-30282-9. DOI: 10.1007/978-3-319-30282-9_14.
- Lund, B., Wang, T., Mannuru, N. R., Nie, B., Shimray, S. und Wang, Z. (Mai 2023): ChatGPT and a New Academic Reality: Artificial Intelligence-Written Research Papers and the Ethics of the Large Language Models in Scholarly Publishing. In: *Journal of the Association for Information Science and Technology* 74(5), S. 570–581. ISSN: 2330-1635, 2330-1643. DOI: 10.1002/asi.24750. arXiv: 2303.13367 [cs]. [Letzter Aufruf: 18. 12. 2023].
- Ma, X., Zhang, X., Pradeep, R. und Lin, J. (Mai 2023): *Zero-Shot Listwise Document Reranking with a Large Language Model*. DOI: 10.48550/arXiv.2305.02156. arXiv: 2305.02156 [cs]. [Letzter Aufruf: 28. 05. 2024].
- Maalej, W., Nayebi, M., Johann, T. und Ruhe, G. (Jan. 2016): Toward Data-Driven Requirements Engineering. In: *IEEE Software* 33(1), S. 48–54. ISSN: 1937-4194. DOI: 10.1109/MS.2015.153. [Letzter Aufruf: 24. 05. 2024].
- Marczak-Czajka, A. und Cleland-Huang, J. (Sep. 2023): „Using ChatGPT to Generate Human-Value User Stories as Inspirational Triggers“. In: 2023 IEEE 31st International Requirements Engineering Conference Workshops (REW), S. 52–61. DOI: 10.1109/REW57809.2023.00016. [Letzter Aufruf: 18. 12. 2023].
- Mayring, P. (1994): *Qualitative Inhaltsanalyse*. In: A. Boehm, A. Mengel und T. Muhr (Hrsg.): *Texte verstehen : Konzepte, Methoden, Werkzeuge (Schriften zur Informationswissenschaft)*. Bd. 14. Konstanz: UVK Univ.-Verl. Konstanz, S. 159–175. ISBN: 978-3-87940-503-9.
- Mayring, P. (Jan. 2010): *Qualitative Inhaltsanalyse*. In: G. Mey und K. Mruck (Hrsg.): *Handbuch Qualitative Forschung in Der Psychologie*. 1. Auflage, 2010. Wiesbaden: VS Verlag, S. 601–613. ISBN: 978-3-531-16726-8. DOI: 10.1007/978-3-531-92052-8_42.
- Mayring, P. und Fenzl, T. (2019): *Qualitative Inhaltsanalyse*. In: N. Baur und J. Blasius (Hrsg.): *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer Fachmedien, S. 633–648. ISBN: 978-3-658-21308-4. DOI: 10.1007/978-3-658-21308-4_42. [Letzter Aufruf: 01. 02. 2024].
- McCarthy, J., Minsky, M. L., Rochester, N. und Shannon, C. E. (1955): *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955*. Online

abrufbar unter: <https://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>
[letzter Aufruf: 17. 01. 2024].

McKinsey (Jan. 2022): *McKinsey & Company - The State of AI in 2022—and a Half Decade in Review* | *McKinsey*. Online abrufbar unter: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review%5C#review>.

McKinsey (Jan. 2023): *McKinsey & Company - The State of AI in 2023: Generative AI's Breakout Year* | *McKinsey*. Online abrufbar unter: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year>.

Memmert, L. und Tavanapour, N. (Mai 2023): „Towards Human-AI-Collaboration in Brainstorming: Empirical Insights into the Perception of Working with a Generative AI“. In.

Meth, H., Mueller, B. und Maedche, A. (Sep. 2015): Designing a Requirement Mining System. In: *Journal of the Association of Information Systems* 16, S. 799–837. DOI: 10.17705/1jais.00408.

Meuser, M. und Nagel, U. (2009): *Das Experteninterview — konzeptionelle Grundlagen und methodische Anlage*. In: S. Pickel, G. Pickel, H.-J. Lauth und D. Jahn (Hrsg.): *Methoden der vergleichenden Politik- und Sozialwissenschaft: Neue Entwicklungen und Anwendungen*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 465–479. ISBN: 978-3-531-91826-6. DOI: 10.1007/978-3-531-91826-6_23. [Letzter Aufruf: 03. 01. 2024].

Mollah, H. und van den Bos, P. (Apr. 2023): „From User Stories to End-to-end Web Testing“. In: 2023 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW), S. 140–148. DOI: 10.1109/ICSTW58534.2023.00036. [Letzter Aufruf: 21. 03. 2024].

Nguyen, T. H., Vo, B. Q., Lumpe, M. und Grundy, J. (Sep. 2012): „REInDetector: A Framework for Knowledge-Based Requirements Engineering“. In: *Proceedings of the 27th IEEE/ACM International Conference on Automated Software Engineering (ASE '12)*. New York, NY, USA: Association for Computing Machinery, S. 386–389. ISBN: 978-1-4503-1204-2. DOI: 10.1145/2351676.2351754. [Letzter Aufruf: 11. 01. 2024].

OpenAI (2024): *OpenAI Platform - Prompt Engineering*. Online abrufbar unter: <https://platform.openai.com/docs/guides/prompt-engineering> [letzter Aufruf: 18. 03. 2024].

- OpenAI et al. (Dez. 2023): *GPT-4 Technical Report*. DOI: 10.48550/arXiv.2303.08774. arXiv: 2303.08774 [cs]. [Letzter Aufruf: 22. 02. 2024].
- OpenAI Pricing* (2024). Online abrufbar unter: <https://openai.com/api/pricing/> [letzter Aufruf: 18. 05. 2024].
- Orrù, G., Piarulli, A., Conversano, C. und Gemignani, A. (Mai 2023): Human-like Problem-Solving Abilities in Large Language Models Using ChatGPT. In: *Frontiers in Artificial Intelligence* 6, S. 1–13. DOI: 10.3389/frai.2023.1199350. [Letzter Aufruf: 19. 12. 2023].
- Parikh, N. A. (Juni 2023): *Empowering Business Transformation: The Positive Impact and Ethical Considerations of Generative AI in Software Product Management – A Systematic Literature Review*. DOI: 10.48550/arXiv.2306.04605. arXiv: 2306.04605 [cs]. [Letzter Aufruf: 05. 10. 2023].
- Park, J. S., O’Brien, J. C., Cai, C. J., Morris, M. R., Liang, P. und Bernstein, M. S. (Aug. 2023): *Generative Agents: Interactive Simulacra of Human Behavior*. DOI: 10.48550/arXiv.2304.03442. arXiv: 2304.03442 [cs]. [Letzter Aufruf: 28. 05. 2024].
- Pohl, K. and Rupp, C. (2015): *Requirements Engineering Fundamentals: A Study Guide for the Certified Professional for Requirements Engineering Exam, Foundation Level, IREB Compliant*. Second edition. Santa Barbara, CA: Rocky Nook. ISBN: 978-1-937538-77-4.
- Rasheed, A., Zafar, B., Shehryar, T., Aslam, N. A., Sajid, M., Ali, N., Dar, S. H. und Khalid, S. (Jan. 2021): Requirement Engineering Challenges in Agile Software Development. In: *Mathematical Problems in Engineering* 2021, S. 1–18. ISSN: 1024-123X. DOI: 10.1155/2021/6696695.
- Rebedea, T., Dinu, R., Sreedhar, M., Parisien, C. und Cohen, J. (Okt. 2023): *NeMo Guardrails: A Toolkit for Controllable and Safe LLM Applications with Programmable Rails*. DOI: 10.48550/arXiv.2310.10501. arXiv: 2310.10501 [cs]. [Letzter Aufruf: 28. 05. 2024].
- Schön, E.-M., Thomaschewski, J. und Escalona, M. J. (Jan. 2017): Agile Requirements Engineering: A Systematic Literature Review. In: *Computer Standards & Interfaces* 49, S. 79–91. ISSN: 09205489. DOI: 10.1016/j.csi.2016.08.011.
- Schön, E.-M., Winter, D., Escalona, M. J., and Thomaschewski, J. (2017): “Key Challenges in Agile Requirements Engineering”. In: H. Baumeister, H. Lichter, and M. Riebisch (eds.): *Agile Processes in Software Engineering and Extreme Programming (Lecture Notes in Business Information Processing)*. Cham: Springer International Publishing, pp. 37–51. ISBN: 978-3-319-57633-6. DOI: 10.1007/978-3-319-57633-6_3.

- Scriven, M. (1972): *Die Methodologie Der Evaluation*. In: C. Wulf (Hrsg.): *Evaluation - Beschreibung Und Bewertung von Unterricht, Curricula Und Schulversuchen*. München: R. Piper & Co., S. 60–91.
- Seeber, I., Bittner, E., Briggs, R. O., Vreede, T., Vreede, G.-J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G. und Söllner, M. (Jan. 2020): *Machines as Teammates: A Research Agenda on AI in Team Collaboration*. In: *Information & Management* 57(2), S. 1–22. ISSN: 03787206. DOI: 10.1016/j.im.2019.103174.
- Shah, T. und Patel, S. (Aug. 2014): *A Review of Requirement Engineering Issues and Challenges in Various Software Development Methods*. In: *International Journal of Computer Applications* 99, S. 36–45. DOI: 10.5120/17451-8370.
- Soares, H. F., Alves, N. S., Mendes, T. S., Mendonça, M. und Spínola, R. O. (Apr. 2015): „Investigating the Link between User Stories and Documentation Debt on Software Projects“. In: 2015 12th International Conference on Information Technology - New Generations, S. 385–390. DOI: 10.1109/ITNG.2015.68. [Letzter Aufruf: 28. 12. 2023].
- Vasconcelos, H., Jörke, M., Grunde-McLaughlin, M., Gerstenberg, T., Bernstein, M. und Krishna, R. (Jan. 2023): *Explanations Can Reduce Overreliance on AI Systems During Decision-Making*. DOI: 10.48550/arXiv.2212.06823. arXiv: 2212.06823 [cs]. [Letzter Aufruf: 27. 04. 2024].
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. und Polosukhin, I. (Juli 2017): *Attention Is All You Need*. DOI: 10.48550/arXiv.1706.03762. arXiv: 1706.03762 [cs]. [Letzter Aufruf: 30. 03. 2024].
- Venable, J., Pries-Heje, J. und Baskerville, R. (Jan. 2016): *FEDS: A Framework for Evaluation in Design Science Research*. In: *European Journal of Information Systems* 25(1), S. 77–89. ISSN: 1476-9344. DOI: 10.1057/ejis.2014.36.
- Vieira, R., Ferreira, D., Borbinha, J. und Gaspar, G. (Jan. 2012): *A Requirements Engineering Analysis of MoReq*. In: *Records Management Journal* 22(3), S. 212–228. ISSN: 0956-5698. DOI: 10.1108/09565691211284407. [Letzter Aufruf: 16. 12. 2023].
- von Krogh, G. (Dez. 2018): *Artificial Intelligence in Organizations: New Opportunities for Phenomenon-Based Theorizing*. In: *Academy of Management Discoveries* 4(4), S. 404–409. ISSN: 2168-1007. DOI: 10.5465/amd.2018.0084. [Letzter Aufruf: 16. 01. 2024].
- Wagner, S., Fernández, D. M., Felderer, M., Vetrò, A., Kalinowski, M., Wieringa, R., Pfahl, D., Conte, T., Christiansson, M.-T., Greer, D., Lassenius, C., Männistö, T., Nayebi, M., Oivo, M., Penzenstadler, B., Prikladnicki, R., Ruhe, G., Schekelmann, A.,

- Sen, S., Spínola, R., Tuzcu, A., Vara, J. L. D. L. und Winkler, D. (Feb. 2019): Status Quo in Requirements Engineering: A Theory and a Global Family of Surveys. In: *ACM Transactions on Software Engineering and Methodology* 28(2), 9:1–9:48. ISSN: 1049-331X. DOI: 10.1145/3306607. [Letzter Aufruf: 11.01.2024].
- Wake, B. (Aug. 2003): INVEST in Good Stories, and SMART Tasks. In: *Bill Wake*.
- Wang, D., Churchill, E., Maes, P., Fan, X., Shneiderman, B., Shi, Y. und Wang, Q. (Apr. 2020): „From Human-Human Collaboration to Human-AI Collaboration: Designing AI Systems That Can Work Together with People“. In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. New York, NY, USA: Association for Computing Machinery, S. 1–6. ISBN: 978-1-4503-6819-3. DOI: 10.1145/3334480.3381069. [Letzter Aufruf: 28.02.2024].
- Wang, Y. und Zhao, Y. (Dez. 2023): *Gemini in Reasoning: Unveiling Commonsense in Multimodal Large Language Models*. DOI: 10.48550/arXiv.2312.17661. arXiv: 2312.17661 [cs]. [Letzter Aufruf: 18.05.2024].
- Webster, J. und Watson, R. T. (2002): Analyzing the Past to Prepare for the Future: Writing a Literature Review. In: *MIS Quarterly* 26(2), S. xiii–xxiii. ISSN: 0276-7783. JSTOR: 4132319. [Letzter Aufruf: 02.01.2024].
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J. und Schmidt, D. C. (Feb. 2023): *A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT*. DOI: 10.48550/arXiv.2302.11382. arXiv: 2302.11382 [cs]. [Letzter Aufruf: 20.03.2024].
- Wilde, T. und Hess, T. (Aug. 2007): Forschungsmethoden der Wirtschaftsinformatik. In: *WIRTSCHAFTSINFORMATIK* 49(4), S. 280–287. ISSN: 1861-8936. DOI: 10.1007/s11576-007-0064-z. [Letzter Aufruf: 03.01.2024].
- Wu, T., Terry, M. und Cai, C. J. (Apr. 2022): „AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts“. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. New York, NY, USA: Association for Computing Machinery, S. 1–22. ISBN: 978-1-4503-9157-3. DOI: 10.1145/3491102.3517582. [Letzter Aufruf: 16.01.2024].
- Ye, Q., Axmed, M., Pryzant, R. und Khani, F. (Nov. 2023): *Prompt Engineering a Prompt Engineer*. DOI: 10.48550/arXiv.2311.05661. arXiv: 2311.05661 [cs]. [Letzter Aufruf: 19.02.2024].

Zhou, X., Chen, Z., Jin, X. und Wang, W. Y. (Apr. 2021): „HULK: An Energy Efficiency Benchmark Platform for Responsible Natural Language Processing“. In: D. Gkatzia und D. Seddah (Hrsg.): Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations. Online: Association for Computational Linguistics, S. 329–336. DOI: 10.18653/v1/2021.eacl-demos.39. [Letzter Aufruf: 27.04.2024].

B Anhang

B.1 AIS eLibrary Suchstring Problemdefinition:

((abstract:(Requirements Engineer OR User Stor* OR Product Owner OR Business Analy*) OR title:(Requirements Engineer* OR User Stor* OR Product Owner OR Business Analy*) OR subject:(Requirements Engineer* OR User Stor* OR Product Owner OR Business Analy*)) AND ((abstract:Review OR title:Review OR subject:Review)) AND (abstract:(Challeng* OR Difficult* OR Issue* OR Problem* OR Complex* OR Quality Criteria) OR title:(Challeng* OR Difficult* OR Issue* OR Problem* OR Complex* OR Quality Criteria) OR subject:(Challeng* OR Difficult* OR Issue* OR Problem* OR Complex* OR Quality Criteria)) NOT (abstract:Estimat* OR title:Estimat*))*

B.2 Ergebnisartefakt - Prompt zur Erstellung von User Stories

Projektkontext: {{projectContext}}

Dokumentsprache: {{language}}

{% if requirementsDocument == "UserStory" %}

Erstelle eine User Story aus der Sicht eines Requirements Engineers eines IT-Dienstleister-Unternehmens.

Strukturhinweise:

- - Konzentriere dich so detailliert wie möglich auf eine spezifische Funktionalität.
- - Vermeide die Vermischung mehrerer Funktionen.
- - Bleibe unbedingt im gegebenen Projektkontext.

Parameter:

{% if not persona and not feature and not benefit %}

- Generiere Zielgruppe, Funktionalität und Nutzen für den folgenden Projektkontext “{{projectContext}}“

{% elif persona and feature and benefit %}

- Die Parameter sind wie folgt definiert: Zielgruppe = “{{persona}}“, Funktionalität = “{{feature}}“, Nutzen = “{{benefit}}“

{% elif persona and not feature and not benefit %}

- Generiere Funktionalität und Nutzen für die Zielgruppe “{{persona}}“ im folgenden Projektkontext “{{projectContext}}“

{% elif not persona and feature and not benefit %}

- Generiere Zielgruppe und Nutzen für die Funktionalität “{{feature}}” im folgenden Projektkontext “{{projectContext}}”

{% elif not persona and not feature and benefit %}

- Generiere Zielgruppe und Funktionalität für den Nutzen “{{benefit}}” im folgenden Projektkontext “{{projectContext}}”

{% elif persona and feature and not benefit %}

- Generiere Nutzen für die Zielgruppe “{{persona}}” und die Funktionalität “{{feature}}” im folgenden Projektkontext “{{projectContext}}”

{% elif persona and not feature and benefit %}

- Generiere Funktionalität für die Zielgruppe “{{persona}}” und den Nutzen “{{benefit}}” im folgenden Projektkontext “{{projectContext}}”

{% elif not persona and feature and benefit %}

- Generiere Zielgruppe für den Nutzen “{{benefit}}” und die Funktionalität “{{feature}}” im folgenden Projektkontext “{{projectContext}}”

{% endif %}

Vergib nun einen Namen für die User Story. Schreibe die User Story in folgendem Template: Als 'Zielgruppe' möchte ich 'Funktionalität', um 'Benefit' zu erzielen.

Arbeite die folgenden Qualitätskriterien impliziert ein, ohne sie direkt im Text aufzuführen:

- - Atomarität: Beschreibe mit der User Story nur eine Funktionalität.
- - Klarheit/Eindeutigkeit: Nutze eindeutige Fakten und vermeide Mehrdeutigkeit.
- - Prägnanz: Schreibe die User Story minimal, klar und präzise.
- - Konfliktfreiheit: Achte darauf, dass die User Story sich nicht mit anderen Funktionalitäten überschneidet.
- - Beachte die INVEST-Kriterien (I - Independent, N - Negotiable, V - Valuable, E - Estimable, S - Small, T - Testable).
- - Problemorientiert: Spezifiziere keine Lösungen.
- - Uniformität: Beachte ein einheitliches Schema.
- - Durchführbarkeit: Generiere nur User Stories, die sich auch wirklich durchführen lassen.

```
{% if acceptanceCriteria == "true" %}
```

- - Generiere Akzeptanzkriterien für die Funktionalität “{{ feature }}” nach den folgenden Regeln:
 - - Achte darauf, dass jedes Akzeptanzkriterium ein klares Ergebnis “bestanden/nicht bestanden” liefert und sich nicht auf Teilergebnisse stützt.
 - - Beziehe sowohl funktionale als auch nichtfunktionale Anforderungen in die Akzeptanzkriterien ein.
 - - Nutze die Akzeptanzkriterien, um die Grenzen einer User Story zu definieren
 - - Definiere mit den Akzeptanzkriterien die Parameter einer User Story.
 - - Formuliere die Akzeptanzkriterien klar und unmissverständlich.
 - - Halte die Akzeptanzkriterien implementierungsunabhängig.
 - - Formuliere die Akzeptanzkriterien spezifisch, messbar und neutral.

```
{% endif %} {% if testCases == "true" %}
```

- - Generiere Testszenarien für die Funktionalität “feature” nach den folgenden Regeln.
 - - Formuliere die Testszenarien so, dass sie für alle Stakeholder verständlich sind.
 - - Verwende dabei eine gemeinsame, domänenspezifische Sprache.
 - - Stelle sicher, dass die Testszenarien sowohl die Funktionalität als auch die Leistung überprüfen.

```
{% endif %} {% endif %}
```

Um Schwachstellen in einer Anforderungsspezifikation zu finden, solltest du folgendes tun: Weise auf unklare Bereiche oder potenziell unbeabsichtigte Ergebnisse am Ende des Outputs hin, aus denen dann neue User Stories entstehen können.

Orientiere dich bei der Entwicklung neuer User Stories an der Struktur der mitgelieferten Beispiele, die mit ““““ Beispiele ““““ eingegrenzt wurden. Dabei handelt es sich um vier anschauliche User Stories. Diese sind zu Anleitungszwecken jeweils in Input- und Output Abschnitte unterteilt. Input: Dabei handelt es sich um die gleichen Parameter, die auch dieser Prompt übergeben werden können. Output: Der Output soll beispielhaft zeigen, wie User Stories im gegebenen Kontext formuliert werden. Auch wenn “Schwachstellen in der Anforderungsspezifikation“ in diesen Beispielen nicht erwähnt werden, ist die Identifizierung und Behebung solcher Schwachstellen bei der Formulierung neuer User Stories von entscheidender Bedeutung. Achte dennoch darauf, dass der Abschnitt “Schwachstellen in der Anforderungsspezifikation“ nicht doppelt aufgelistet wird.

““““

Beispiel 1:

Input: {“language“: “deutsch“, “creativity“: 2, “projectContext“: “Es geht darum einen Prüfstand zu entwickeln, in den Softwareprojekte deployed werden. Automatisierte Ökobilanzierung verteilter Softwaresysteme. Anschließend soll der Prüfstand den Energieverbrauch, den CO2e-Fußabdruck und weitere Größen ausgeben, Versionen vergleichen und Handlungsempfehlungen für eine möglichst nachhaltige Umsetzung geben. Die beiden Kern-Differenzierungsmerkmale gegenüber existierenden Lösungen sind (1.) Messung verteilter Systeme inkl. Netzwerkinfrastruktur (2.) Vollständige Ökobilanz inklusive aller Ökologischen Wirkdimensionen. Die systementwickelnde Person ist dabei eine der Hauptnutzergruppen des Prüfstandes. Der Digital Twin ist dabei eine virtuelle Abbildung des Realsystems. Die Komponenten des Realsystems werden in der Infrastrukturdefinition übergeben.“, “persona“: “systementwickelnde Person“, “feature“: “nach Bereitstellung meiner Infrastrukturdefinition müssen automatisch alle benötigten VM’s für den Digital Twin erzeugt werden.“, “acceptanceCriteria“: “true“, “testCases“: “false“}

Output: User Story Name: Erzeugung des Digital Twins anhand der Infrastrukturdefinition
Als systementwickelnde Person möchte ich einen Prozess, der nach Bereitstellung meiner Infrastrukturdefinition automatisch die Erzeugung der benötigten VM’s einleitet, damit bei der Messung alle erforderlichen Komponenten im Digital Twin berücksichtigt werden.

Akzeptanzkriterien

- Sobald die Infrastrukturdefinition bereitgestellt wird, soll sie zunächst in eine virtuelle Infrastrukturdefinition (abzüglich irrelevanter Parameter für die VM) übersetzt werden und danach das System automatisch mit der Erzeugung der entsprechenden VMs beginnen.
- Für jeden in der Infrastrukturdefinition spezifizierten Knoten oder jede Entität, auf der Workloads ausgeführt werden sollen, muss eine entsprechende VM erzeugt werden.
- Die erzeugten VMs müssen die in der Infrastrukturdefinition festgelegten Spezifikationen und Anforderungen bestmöglich widerspiegeln.
- Die erzeugten VMs müssen in der Lage sein, gemäß der in der Infrastrukturdefinition festgelegten Anforderungen, untereinander zu kommunizieren.

Beispiel 2:

Input: {“language“: “deutsch“, “creativity“: 2, “projectContext“: “Es geht darum einen Prüfstand zu entwickeln, in den Softwareprojekte deployed werden. Automatisierte Ökobilanzierung verteilter Softwaresysteme. Anschließend soll der Prüfstand den Energieverbrauch, den CO2e-Fußabdruck und weitere Größen ausgeben, Versionen vergleichen und

Handlungsempfehlungen für eine möglichst nachhaltige Umsetzung geben. Die beiden Kern-Differenzierungsmerkmale gegenüber existierenden Lösungen sind (1.) Messung verteilter Systeme inkl. Netzwerkinfrastruktur (2.) Vollständige Ökobilanz inklusive aller Ökologischen Wirkdimensionen. Innerhalb dieser User Story soll die bereits bestehende Linux Messumgebung auf irrelevante Hintergrundprozesse geprüft werden.“, “persona“: “systementwickelnde Person“, “feature“: “Vertestung meiner Anwendung in einer sterilen Linux Messumgebung“, “acceptanceCriteria“: “true“, “testCases“: “true“ }

Output: User Story Name: Verifizierung Linux Messumgebung

Als systementwickelnde Person möchte ich sicherstellen, dass meine Anwendung in einer sterilen Messumgebung getestet wird, um die präzisesten und zuverlässigsten Daten über ihren Energieverbrauch zu erhalten.

Akzeptanzkriterien:

- Eine gründliche Analyse der aktiven Prozesse in der Linux-Messumgebung ist durchgeführt, um sicherzustellen, dass keine irrelevanten Hintergrundprozesse laufen.
- Nicht für das System under Test (SUT) oder die Last relevante Prozesse sind nach Möglichkeit identifiziert und eliminiert.
- Es soll sichergestellt werden, dass alle Prozesse auf dem Host, die zur Messumgebung gehören, erfasst werden.
- Regelmäßige Überprüfungen der Messumgebung sind etabliert, um ihre Reinheit und Eignung für genaue Messungen zu gewährleisten.
 - - Hierfür wird eine Checkliste im Confluence (an die gleiche Stelle, wie die Schritt-für-Schritt Anleitung zum Aufsetzen der Messumgebung) hinterlegt.

Testszenerien:

- Szenario 1: Analyse der aktiven Prozesse: Führe eine Analyse der aktiven Prozesse in der Linux-Messumgebung durch, um sicherzustellen, dass keine irrelevanten Hintergrundprozesse laufen, die die Messergebnisse beeinflussen könnten.
- Szenario 2: Regelmäßige Überprüfungen der Messumgebung: Führe einen Test durch, bei dem die Schritte der im Confluence hinterlegten Checkliste befolgt werden, um die Messumgebung aufzusetzen und zu verifizieren. Dies dient dazu, die Praktikabilität und Vollständigkeit der Anleitung zu bewerten.

Beispiel 3:

Input: { “language“ : “deutsch“, “creativity“: 3, “projectContext“: “Es geht darum einen Prüfstand zu entwickeln, in den Softwareprojekte deployed werden. Automatisierte Ökobilanzierung verteilter Softwaresysteme. Anschließend soll der Prüfstand den Energiever-

brauch, den CO₂e-Fußabdruck und weitere Größen ausgeben, Versionen vergleichen und Handlungsempfehlungen für eine möglichst nachhaltige Umsetzung geben. Die beiden Kern-Differenzierungsmerkmale gegenüber existierenden Lösungen sind (1.) Messung verteilter Systeme inkl. Netzwerkinfrastruktur (2.) Vollständige Ökobilanz inklusive aller Ökologischen Wirkdimensionen.“, “persona“ : “Teil des Entwicklerteams“, “feature“ : “automatisierte Git-Pipeline“, “benefit“ : “neueste Versionen unserer Python- und Bash-Skripte auf dem Testserver. Und Vereinfachung des Prozesses für das Hinzufügen von Tests.“, “acceptanceCriteria“ : “true“, “testCases“ : “true“ }

Output: User Story Name: Automatisierte Aktualisierung von Skripten auf dem Testserver
Als Teil des Entwicklerteams möchte ich eine automatisierte Pipeline einrichten, damit ich stets die neuesten Versionen unserer Python- und Bash-Skripte auf dem Testserver habe (und den Prozess für spätere Erweiterungen, wie das Hinzufügen von Tests, oder prüfen auf Code-Formatierung vereinfache).

Akzeptanzkriterien

- Die Pipeline soll automatisch auslösen, sobald ins Git-Repository gepusht wird.
- Ob dabei alle, oder nur der main-Branch beobachtet werden sollen muss im Laufe des Tickets entschieden werden
- Ggf. sollte evaluiert werden, ob die Pipeline für die anderen Branches manuell ausgeführt wird.
- Die neueste Version der Python- und Bash-Skripte muss nach jedem Push ins Git-Repository automatisch auf dem Testserver verfügbar sein.
- Die Pipeline soll mit Bitbucket oder Jenkins umgesetzt werden
- Die Pipeline soll so konfiguriert sein, dass sie problemlos um zusätzliche Funktionen, wie automatisierte Tests, erweitert werden kann.

Testszenarios

- Szenario 1: Trigger der Pipeline durch Push ins Git-Repository. Push in den main-Branch eines Git-Repositories ausführen und prüfen, ob die Pipeline automatisch ausgelöst wird
- Szenario 2: Automatische Aktualisierung auf dem Testserver: Nach einem erfolgreichen Push ins Git-Repository die Verfügbarkeit der neuesten Version der Python- und Bash-Skripte auf dem Testserver überprüfen.
- Szenario 3: Verwendung von Bitbucket: Konfiguration der Pipeline in Bitbucket überprüfen, um sicherzustellen, dass sie gemäß den Anforderungen eingerichtet wurde.

Beispiel 4:

Input: { "language" : "deutsch", "creativity": 1, "projectContext": "Es geht darum einen Prüfstand zu entwickeln, in den Softwareprojekte deployed werden. Automatisierte Ökobilanzierung verteilter Softwaresysteme. Anschließend soll der Prüfstand den Energieverbrauch, den CO2e-Fußabdruck und weitere Größen ausgeben, Versionen vergleichen und Handlungsempfehlungen für eine möglichst nachhaltige Umsetzung geben. Die beiden Kern-Differenzierungsmerkmale gegenüber existierenden Lösungen sind (1.) Messung verteilter Systeme inkl. Netzwerkinfrastruktur (2.) Vollständige Ökobilanz inklusive aller Ökologischen Wirkdimensionen. Die systementwickelnde Person ist dabei eine der Hauptnutzergruppen des Prüfstandes. Sie will in dieser User Story, dass die Messergebnisse abgespeichert werden können.", "persona" : "systementwickelnde Person", "feature" : "Abspeicherung der erfassten Messdaten", "acceptanceCriteria" : "true", "testCases" : "false" }

Output: User Story Name: Metriken sichern

Als systementwickelnde Person möchte ich, dass die erfassten Daten der Messung gespeichert werden, damit ich Versionen und Produkte vergleichen und historischen Trends analysieren kann.

Akzeptanzkriterien:

- Die erfassten Messdaten können in einer sicheren und strukturierten Datenbank oder einem nachvollziehbaren Dateiformat gespeichert werden, das für die Speicherung vieler Messwerte optimiert ist.
- Es gibt eine Möglichkeit, ältere Messreihen abzurufen.

Die gesamte Form der User Story sollte also wie folgt aussehen:

User Story Name:

“User Story Template“

{% if acceptanceCriteria == "true" %}

Akzeptanzkriterien:

{% endif %}

{% if testCases == "true" %}

Testszzenarien:

{% endif %}

Schwachstellen in der Anforderungsspezifikation:

B.3 Ergebnisse der Experteninterviews

Tabelle B.1: Teilartefakt: Abgleich der RE-Herausforderungen der Literatur mit RE-Experten (Eigene Darstellung in Anlehnung an Webster und Watson (2002))

RE-Herausforderung	Experteninterviews					
	Ei1	Ei2	Ei3	Ei4	Ei5	Summe
Unbestimmtheit	X	X		X	X	4
Inkonsistenz	X	X	X	X	X	5
Unzulänglichkeit	X	X	X	X		4
Duplikate	X			X		2
Mehrdeutigkeit	X	X			X	3
Minimale Dokumentation	X	X		X	X	4
Vernachlässigung nicht-funktionaler Anforderungen	X	X	X	X		4
Bewältigung von Anforderungsänderungen (und ihrer Bewertung)	X	X	X	X	X	5
Ungeeignete Methode der Prioritätensetzung	X	X		X		3
Schlecht formulierte Anforderungen	X	X		X		3
Weniger erfahrenes und qualifiziertes Team	X	X	X			3
Fehlende Anforderungen	X	X	X	X	X	5
Kundenverfügbarkeit	X	X	X	X	X	5
Schätzung von Budget und Zeitplan	X	X	X			3
Ungeeignete Architektur	X	X	X	X		4

Fortsetzung auf der nächsten Seite

Tabelle B.1: Teilartefakt: Abgleich der RE-Herausforderungen der Literatur mit RE-Experten (Eigene Darstellung in Anlehnung an Webster und Watson (2002)) - Fortsetzung

RE-Herausforderung	Experteninterviews					
	Ei1	Ei2	Ei3	Ei4	Ei5	Summe
Unfähigkeit und Zustimmung des Kunden	X	X	X	X		4
Vertragliche Beschränkungen und Volatilität der Anforderungen	X	X	X	X		4
Fehlendes (Kunden-)wissen	X	X		X	X	4
Kommunikation (-smethoden)	X	X	X	X	X	5
Prototyping Problematik	X	X	X		X	4
Weniger Vorplanung, Konzentration und keine anfängliche Beteiligung des Teams	X	X	X	X		4
Implizites Wissen	X	X		X	X	4
Funktionale und technische Abhängigkeiten zwischen Teams	X	X	X	X	X	5
Stakeholderakzeptanz, dass Teams eigene Entscheidungen treffen	X	X				2
Wahrung Gesamtüberblick	X	X		X	X	4
Zusammenarbeit mit den Endnutzern		X	X	X		3

RE-Qualitätskriterium	Experteninterviews					
	Ei1	Ei2	Ei3	Ei4	Ei5	Summe
Syntax	X	X	X	X	X	5
Atomarität	X	X	X	X		4
Prägnanz	X	X	X	X	X	5
Semantik	X	X	X		X	5
Konsistenz und Korrektheit	X	X	X	X	X	5
Klarheit	X	X		X	X	4
Pragmatik	X	X	X	X	X	5
Vollständigkeit	X	X	X	X	X	5
Einheitlichkeit	X	X		(X)	X	4
Einzigartigkeit	X	X		X	X	4
Lexikalität			X		X	2
Formalität	X	X	X	X	X	5
Sättigung		X				1
INVEST	X	X		X		3
Card, Conversation, Confirmation	(X)					1
Cohn's Guideliness						0
Durchführbarkeit	X	X	X			3
Abstraktheit	X		X	X	X	4
Überprüfbarkeit	X	X	X	X	X	5

Tabelle B.2: Teilartefakt: Abgleich der RE-Qualitätskriterien der Literatur in den RE-Experteninterviews (Eigene Darstellung in Anlehnung an Webster und Watson (2002))

Tabelle B.3: Teilartefakt: Einschätzung bezüglich des Einsatzes von GenAI zum Lösen der RE-Herausforderungen in den GenAI-Experteninterviews (Eigene Darstellung in Anlehnung an Webster und Watson (2002))

RE-Herausforderung	Experteninterviews						
	Ei6	Ei7	Ei8	Ei9	Ei10	Ei11	Summe
Unbestimmtheit	X	X	X	X	X	X	6
Inkonsistenz	X	X		X	X		4
Unzulänglichkeit	X	(X)	(X)	(X)	X	X	6
Duplikate	X	X	X	X	X	X	6
Minimale Dokumentation	X	X	X	X	X	X	6
Vernachlässigung nicht-funktionaler Anforderungen	X			X	X		3
Bewältigung von Anforderungsänderungen (und ihrer Bewertung)	X	X	X	X	X	X	6
(Ungeeignete Methode der) Prioritätensetzung	X	X	(X)	X	X	X	6
schlecht formulierte Anforderungen	(X)	X	(X)	X	X	X	6
Weniger erfahrenes und qualifiziertes Team	X	X	(X)	X	X		5
fehlende Anforderungen	X	(X)		X	X	X	5
Kundenverfügbarkeit					X		1
Schätzung von Budget und Zeitplan	X				X		2
Ungeeignete Architektur					X		1
Unfähigkeit und Zustimmung des Kunden	(X)						1
Vertragliche Beschränkungen und Volatilität der Anforderungen		X		X			2
Fehlendes (Kunden-)wissen	X	(X)		X	X	X	5

Fortsetzung auf der nächsten Seite

Tabelle B.3: Teilartefakt: Einschätzung bezüglich des Einsatzes von GenAI zum Lösen der RE-Herausforderungen in den GenAI-Experteninterviews (Eigene Darstellung in Anlehnung an Webster und Watson (2002))
- Fortsetzung

RE-Herausforderung	Experteninterviews						Summe
	Ei6	Ei7	Ei8	Ei9	Ei10	Ei11	
Kommunikation (-smethoden)	X		X	X	X	X	5
Prototyping Problematik				X			1
weniger Vorplanung, Konzentration und keine anfängliche Beteiligung des Teams			X				1
Implizites Wissen	(X)			X	X		3
Funktionale und technische Abhängigkeiten zwischen Teams		X			X		2
Stakeholderakzeptanz, dass Teams eigene Entscheidungen treffen			X	(X)			2
Wahrung Gesamtüberblick	X	X			X	X	4
Zusammenarbeit mit den Endnutzern	(X)	X	X		X	X	5
Kunde kennt seine Anforderungen nicht	X		X	X	X		4
Kunde denkt eher in Lösungen, als in Anforderungen	X	X	X		X		4
Fachabteilungen können Use Cases nicht formulieren	X	X	X	X	X	X	6
Fachabteilungen können die Prozesse nicht beschreiben	X		X	X	X		4
Fehlende Nachvollziehbarkeit von früher umgesetzten Anforderungen	X		X	X	X		4
zu späte Einbindung Softwareentwicklerteams		X	X	X	X		4
Anforderungen werden oftmals aus der falschen Rolle formuliert	X	X	X		X		4

Fortsetzung auf der nächsten Seite

Tabelle B.3: Teilartefakt: Einschätzung bezüglich des Einsatzes von GenAI zum Lösen der RE-Herausforderungen in den GenAI-Experteninterviews (Eigene Darstellung in Anlehnung an Webster und Watson (2002))
- Fortsetzung

RE-Herausforderung	Experteninterviews						Summe
	Ei6	Ei7	Ei8	Ei9	Ei10	Ei11	
Unreflektierte Übernahme von Anforderungen			X	X	X		3

Mit (X) gekennzeichnet, wenn viele Vorinformationen und Daten zur Verfügung stehen müssen (Ei9 - Z. 50 ff.), Training des LLMs nötig ist (Ei8 - Z. 213) oder GenAI nur indirekt unterstützen kann (Ei6 - Z. 295-297). Die Trennung der Tabelle zeigt in der oberen Hälfte die literaturbasierten RE-Herausforderungen und die untere Hälfte zeigt die RE-Herausforderungen aus den RE-Experteninterviews.

GenAI-Vorteil	Referenz
Verständnis natürlicher Sprache	Ei6 - Z. 46 ff.
Prozessautomatisierung	Ei7 - Z. 9 ff. Ei10 - Z. 42 ff.
Informationsverarbeitung	Ei6 - Z. 13 ff. Ei7 - Z. 16 ff. Ei8 - Z. 17 ff. Ei9 - Z. 17 f. Ei10 - Z. 17 ff. Ei11 - Z. 13 ff.
Inhaltsgenerierung	Ei6 - Z. 30 ff. Ei8 - Z. 16 ff.
Brainstorming	Ei6 - Z. 28 ff. Ei10 - Z. 38 ff.
Kosteneinsparungen	Ei11 - Z. 27 f.
Innovativität	Ei11 - Z. 28
Geschwindigkeit	Ei9 - Z. 20 ff. Ei11 - Z. 35
Überelegenheit gegenüber klassischer KI	Ei11 - Z. 28 ff.
Hilfestellungen bei der Programmierung	Ei6 - Z. 33 ff. Ei10 - Z. 24 ff.

Tabelle B.4: Teilartefakt: GenAI-Vorteile aus den GenAI-Experteninterviews

GenAI-Herausforderung	Referenz
Übermäßiges Vertrauen von Nutzern	Ei6 - Z. 370 ff. Ei7 - Z. 293 ff. Ei8 - Z. 365 ff. Ei9 - Z. 323 ff.
Halluzinationen	Ei6 - Z. 351 f. Ei7 - Z. 286 ff. Ei8 - Z. 364 f. Ei9 - Z. 342 ff. Ei10 - Z. 498 ff. Ei11 - Z. 317
Sicherheit	Ei6 - Z. 363 f. Ei10 - Z. 504 ff.
Prompt Informationen reichen nicht aus	Ei6 - Z. 352 f., 356 f.
User hat keinen Einblick in den Prompt	Ei6 - Z. 361 ff.
Langfristige Perspektive fehlt	Ei7 - Z. 297 f.
Unzureichende Trainingsdaten	Ei8 - Z. 364 ff. Ei9 - Z. 335 ff.
Geringe Kreativität	Ei9 - Z. 346 ff.
Kosten	Ei10 - Z. 517
Hohe Performancekriterien	Ei10 - Z. 523 ff.
Missbrauchspotenzial	Ei10 - Z. 525 ff.
Wissensverlust im Unternehmen	Ei8 - Z. 32 ff. Ei11 - Z. 303 ff.

Tabelle B.5: Teilartefakt: GenAI-Herausforderungen aus den GenAI-Experteninterviews

B.4 Leitfaden RE-Experteninterviews

Leitfaden & Protokoll für ein semistrukturiertes Interview - RE-Herausforderungen und
-Qualitätskriterien

Interviewpartner:in	
Daten zu Interviewpartner:in	
Datum	
Start- und Endzeit	

Einleitung

- Kurze Vorstellung des Forschungskontexts und des Ziels des Interviews
- Zusicherung der Vertraulichkeit und Pseudonymität der Antworten
- Erläuterung des Interviewablaufs und der geschätzten Dauer
- Einverständnis zur Aufzeichnung des Gespräches einholen

Forschungskontext

Das zentrale Anliegen meiner Masterarbeit ist die Untersuchung des potenziellen Mehrwerts und Einflusses generativer künstlicher Intelligenz (GenAI) innerhalb des Requirements Engineering (RE). Das übergeordnete Vorhaben besteht darin, Potenziale zur Steigerung der Produktivität der Anforderungserhebung und zur Erhöhung der Qualität der daraus resultierenden Anforderungen zu identifizieren.

Ziel des Interviews

Mit diesem Interview möchte ich die gesammelten Erkenntnisse aus meiner Literaturrecherche zu den RE-Herausforderungen und -Qualitätskriterien sowie zum möglichen Einsatz von GenAI in diesem Zusammenhang ergänzen. Durch das Sammeln von praktischer Erfahrung und Einblicken soll eine Brücke zwischen Theorie und Praxis geschlagen werden, um die Forschungsfrage aus beiden Perspektiven umfassend zu beleuchten.

Vorerfahrung im RE und GenAI
Wie lange bist Du schon im Bereich des RE tätig?
Welche Rollen / Funktionen hattest Du bereits in Projekten im Kontext des RE?
Welche Erfahrung hast du bereits mit GenAI gemacht? (Beruflich oder privat?)

Tabelle B.6: Interviewleitfaden RE - Fragen des ersten Abschnitts

Vorgehen im RE
Welche Art von Anforderungsdokumenten nutzt du?
Welche Methoden/Tools verwendest du aktuell im Rahmen deiner Arbeit als Requirements Engineer, um diese Anforderungsdokumente zu erstellen/dokumentieren?

Tabelle B.7: Interviewleitfaden RE - Fragen des zweiten Abschnitts

RE-Qualitätskriterien und -Herausforderungen
Welche Qualitätskriterien sind dir in Bezug auf RE bekannt? Und nach welchen Kriterien gehst du vor, wenn du Anforderungsdokumente erstellst?
Basierend auf den in der Literatur identifizierten RE-Qualitätskriterien (werden an dieser Stelle erwähnt): Welche davon hast du bereits für die Erstellung von Anforderungsdokumenten genutzt?
Im Kontext des RE: Welche Herausforderungen hast du in deinen Projekten in den bereits erwähnten Rollen erlebt?
Basierend auf den in der Literatur identifizierten RE-Herausforderungen (werden an dieser Stelle erwähnt): Bitte gib an, wenn Du auf diese Herausforderung im Laufe Deiner Projekterfahrung gestoßen bist.
Welche spezifischen Methoden oder Strategien setzt du ein, um den bekannten und genannten Herausforderungen, wie sie in der Praxis auftreten, zu begegnen?

Tabelle B.8: Interviewleitfaden RE - Fragen des dritten Abschnitts

Einsatz von GenAI im RE (Tiefe in dieses Thema abhängig vom Erfahrungsgrad der interviewten Person)
In Bezug auf die zuvor diskutierten Herausforderungen im RE: Welches Potenzial siehst du für den Einsatz von GenAI-Systemen, um diese Herausforderungen zu bewältigen? Kannst du dir vorstellen, dass bestimmte Herausforderungen durch den Einsatz solcher Systeme reduziert oder eliminiert werden und falls ja, wie?
Könntest du spezifische Anwendungsbereiche innerhalb des RE identifizieren, in denen der Einsatz von GenAI besonders vorteilhaft erscheint? Gibt es bestimmte Aspekte des RE-Prozesses, die von solchen Technologien profitieren könnten?
<i>Angenommen, es gäbe eine Webanwendung, die auf GenAI basiert und RE-Prozesse unterstützt: Beispielsweise mit der Übergabe einiger wichtiger Parameter (wie bspw. Projektkontext, Zielgruppe, definierten allgemeinen Anforderungen) würde ein Prompt an eine GenAI geschickt werden, die dann Anforderungsdokumente, wie z.B. User Stories, generiert.</i>
Würdest du ein solches System in deine Arbeitsweise integrieren wollen? Welche Gründe könnten dafür oder dagegen sprechen?
Welche zusätzlichen Funktionen sollte ein solches System haben, um das Tool effektiv in deinen Arbeitsalltag zu integrieren?

Tabelle B.9: Interviewleitfaden RE - Fragen des vierten Abschnitts

B.5 Leitfaden GenAI-Experteninterviews

Leitfaden & Protokoll für ein semistrukturiertes Interview – Bewältigung
RE-Herausforderungen mit GenAI

Interviewpartner:in	
Daten zu Interviewpartner:in	
Datum	
Start- und Endzeit	

Einleitung

- Kurze Vorstellung des Forschungskontexts und des Ziels des Interviews
- Zusicherung der Vertraulichkeit und Pseudonymität der Antworten
- Erläuterung des Interviewablaufs und der geschätzten Dauer
- Einverständnis zur Aufzeichnung des Gespräches einholen

Forschungskontext

Das zentrale Anliegen meiner Masterarbeit ist die Untersuchung des potenziellen Mehrwerts und Einflusses von generativer künstlicher Intelligenz (GenAI) innerhalb des Requirements Engineering (RE). Das übergeordnete Vorhaben besteht darin, Potenziale zur Steigerung der Produktivität der Anforderungserhebung und zur Erhöhung der Qualität der daraus resultierenden Anforderungen zu identifizieren.

Ziel des Interviews

Die aus der Literatur und Expertengesprächen identifizierten Herausforderungen im RE den Möglichkeiten von GenAI gegenüberzustellen. Dabei soll untersucht werden, ob GenAI-Systeme diese Herausforderungen adressieren können und mit welchen Risiken ein solches Vorgehen behaftet sein kann.

Allgemein GenAI/KI
Wie lange bist du bereits im Kontext der künstlichen Intelligenz tätig? Wie lange davon im Bereich der generativen KI?
Welche grundlegenden Anwendungsbereiche von GenAI hältst du für besonders relevant?
Wo siehst du die größten Stärken von GenAI aus deiner Erfahrung heraus?

Tabelle B.10: Interviewleitfaden GenAI - Fragen des ersten Abschnitts

Einsatz von GenAI im RE
Basierend auf einer Liste von Herausforderungen, welche im RE auftreten, wenn nur Menschen im Anforderungsprozess beteiligt sind, die ich durch Literaturrecherche und Gesprächen mit RE-Experten zusammengestellt habe: Welche dieser Herausforderungen lassen sich deiner Meinung nach mit GenAI lösen bzw. reduzieren? Falls ja, wie?
Könntest du spezifische Anwendungsbereiche innerhalb des RE's identifizieren, in denen der Einsatz von GenAI besonders vorteilhaft erscheint? Gibt es bestimmte Aspekte des RE-Prozesses, die von solchen Technologien profitieren könnten?
<i>Angenommen, es gäbe eine Webanwendung, die auf GenAI basiert und RE-Prozesse unterstützt: Beispielsweise mit der Übergabe einiger wichtiger Parameter (wie bspw. Projektkontext, Zielgruppe, definierten allgemeinen Anforderungen) würde ein Prompt an eine GenAI geschickt werden, die dann Anforderungsdokumente, wie z.B. User Stories, generiert.</i>
Welche spezifischen Risiken, Gefahren und Schwachpunkte könnten mit der Nutzung einer solchen Anwendung einhergehen?
Falls noch nicht erwähnt: Wie ordnest du Halluzinationen, Prompt Injection und sonstige Sicherheitsrisiken im Kontext dieses Tools ein?
Wie kann man diesen Risiken (vor allem technisch) vorbeugen?

Tabelle B.11: Interviewleitfaden GenAI - Fragen des zweiten Abschnitts

B.6 Evaluation

B.6.1 Leitfaden Nachevaluation

Leitfaden & Protokoll für ein semistrukturiertes Interview - Evaluation

Interviewpartner:in	
Daten zu Interviewpartner:in	
Datum	
Start- und Endzeit	

Einleitung

- Kurze Vorstellung des Forschungskontexts und des Ziels der Evaluation
- Zusicherung der Vertraulichkeit und Pseudonymität der Antworten
- Erläuterung des Evaluationsablaufs und der geschätzten Dauer
- Einverständnis zur Aufzeichnung des Gespräches einholen

Forschungskontext

Das zentrale Anliegen meiner Masterarbeit ist die Untersuchung des potenziellen Mehrwerts und Einflusses generativer künstlicher Intelligenz (GenAI) innerhalb des Requirements Engineering (RE). Das übergeordnete Vorhaben besteht darin, Potenziale zur Steigerung der Produktivität der Anforderungserhebung und zur Erhöhung der Qualität der daraus resultierenden Anforderungen zu identifizieren.

Ziel des Interviews

Mit dieser Evaluation möchte ich die gewonnenen Erkenntnisse aus der Literaturrecherche und den RE- und GenAI-Experteninterviews überprüfen. Dafür stelle ich eine Webanwendung zur Verfügung. Diese soll den RE-Prozess mittels Generierung von User Stories vereinfachen. Dieser Prozess, der hybriden Intelligenz soll hinsichtlich Qualität, Produktivität sowie möglicher Vorteile und Herausforderungen evaluiert werden.

Vorerfahrung im RE und GenAI (Bei noch nicht interviewten Personen kurz die Erfahrung erfragen)
Wie lange bist Du schon im Bereich des RE tätig?
Welche Rollen / Funktionen hattest Du bereits in Projekten im Kontext des RE?
Welche Erfahrung hast du bereits mit GenAI gemacht? (Beruflich oder privat?)

Tabelle B.12: Interviewleitfaden Evaluation - Fragen des ersten Abschnitts

Nachevaluation
Wie schätzt du die Qualität der mit dieser Anwendung generierten User Stories im Vergleich zu den traditionell entwickelten User Stories in deinem Arbeitsumfeld ein?
Wie bewertest du den Zeitaufwand für die Nutzung dieser Anwendung im Vergleich zur herkömmlichen Methode der Erstellung von User Stories in deinem Arbeitskontext?
Würdest Du ein solches Tool in deinen Alltag als Requirements Engineer integrieren? Welche Vorteile und Herausforderungen identifizierst du?
Welche spezifischen Funktionen hast du vermisst, die deiner Meinung nach notwendig wären, um die Ergebnisqualität zu verbessern?

Tabelle B.13: Interviewleitfaden Evaluation - Fragen des zweiten Abschnitts

B.6.2 Agenda und Aufgabenstellung

Agenda

1. **10 Minuten** - Durchlesen der Aufgabenstellung und Vorstellung des Projektkontextes.
2. **25 Minuten** - Thinking-Aloud: Aktives Entwickeln von User Stories.
3. **15 Minuten** - Abschließendes Interview, um Einblicke in die Gedanken und Entscheidungen während des Prozesses zu gewinnen.
4. **10 Minuten** - Flexibler Zeitpuffer für unvorhergesehene Verzögerungen.

Aufgabenstellung

Die Evaluation wird mithilfe der *Thinking-Aloud* (Lewis 1982) Methode durchgeführt. Dabei sollst Du Deine Gedanken kontinuierlich und offen aussprechen, während du die

Aufgabe bearbeitest.

- Du hast zunächst 5 Minuten Zeit, um die gesamte Aufgabenstellung durchzulesen, falls Du dies noch nicht im Vorhinein getan hast.
- Anschließend hast du 25 Minuten Zeit zur Bearbeitung der Aufgabe.
- Als Product Owner des „SmartServiceSystems“ ist es Deine Aufgabe, auf Basis des Projektkontextes (welcher Dir im Termin vorgelesen wird) und mithilfe der zur Verfügung gestellten Webanwendung User Stories zu entwickeln.
- Es gibt keine Vorgabe, wie viele User Stories Du erstellen sollst.
- Die User Stories sollen auf einer allgemeinen Ebene gehalten werden, um einen Überblick über die Funktionalitäten des „SmartServiceSystems“ zu geben und als Grundlage für Diskussionen mit dem Product Management dienen.
- Es ist möglich und auch gewollt die vom Tool erstellten User Stories von dir händisch anzupassen, falls du es für notwendig hältst. Die User Stories sollen in einer Qualität erstellt werden, wie du sie selbst in einem realen Szenario verfassen würdest.
- Konzentriere Dich außerdem dabei auf die Funktionalitäten des Tools, ohne die Benutzerfreundlichkeit in den Vordergrund zu stellen.
- Je genauer Du Deine Anfragen an das Tool stellst, desto zielgerichtete Ergebnisse sind zu erwarten. Um Deine Anfragen an das Tool zu konkretisieren kannst du hierfür die optionalen Parameter „Zielgruppe“, „Funktionalität“ und „Nutzen“ verwenden.
- Übertrage alle User Stories, die Du für qualitativ hältst oder anpassen möchtest, in ein Dokument Deiner Wahl (z.B. Word oder ein Textdokument) ohne eine bestimmte Form einzuhalten.
- Du bekommst gleich den Projektkontext, mögliche Funktionalitäten des Systems und Stakeholder vorgelesen. Versetze Dich in die Lage eines realen Projektes und mach Notizen zu den gegebenen Informationen.
- Verständnisfragen sind jederzeit möglich, jedoch keine Fragen zu speziellen Funktionalitäten. Treffe, falls nötig, begründete Annahmen.
- Sage nun der Person bescheid, die die Evaluation durchführt.
- Dir wird der Link für die Webanwendung im Termin bereitgestellt. Teile Deinen Bildschirm, sobald Du bereit bist. Ab diesem Zeitpunkt startet die Bearbeitungszeit von 25 Minuten.

B.6.3 Beispiel-Projektkontext

Projektname: SmartServiceSystem

Projektkontext: Im Rahmen der stetigen Entwicklung der Gastronomiebranche streben Restaurants danach, das Kundenerlebnis zu optimieren, ohne dabei den persönlichen Touch und Servicequalität zu verlieren. Das SmartServiceSystem zielt darauf ab, ein intelligentes Supportsystem zu etablieren, das nicht nur die Kommunikation zwischen Gästen und Personal verbessert, sondern auch die Gesamteffizienz des Serviceprozesses steigert. Einschließlich der Optimierung der Kommunikation zwischen Service- und Küchenpersonal. Dieser Aspekt ist wesentlich, um die Bestell- und Zubereitungsprozesse zu synchronisieren, was wiederum zu einer beschleunigten und fehlerfreien Speisenausgabe führt. Indem die unmittelbare Interaktion zwischen Servicepersonal und Gästen für die Bestellaufnahme erhalten bleibt, gewährleistet die Einführung des Systems weiterhin eine persönliche Betreuung.

Funktionsweise:

- **IoT-Display am Tisch:**
 - Gäste können ihre Bestellbereitschaft signalisieren
 - Gäste können in Echtzeit den Status der aufgegebenen Bestellung verfolgen
 - Anzeige Speisekarte und aktuelle Angebote des Restaurants
- **Bestell- und Bezahlssystem:**
 - Servicepersonal wird über Bestellbereitschaft der Tische auf einem eigenen Gerät informiert
 - Bestellaufnahme über dieses Gerät möglich
 - Die Bestellungen werden priorisiert, sortiert und anschließend in die Küche gesendet
 - Zahlung über dieses Gerät möglich
- **Analyse und Trenderfassung:**
 - Automatische Datenerfassung über die Dauer des Gesamtprozesses
 - Visualisierung der gesammelten Daten und Identifizierung von Trends

Stakeholder:

- Gäste: Nutzen das IoT-Display am Tisch.
- Servicepersonal: Erhält Benachrichtigungen über Gästeanfragen und ist für die

persönliche Betreuung und Bestellaufnahme zuständig.

- Küchenpersonal: Erhält Bestellungen des Systems direkt in die Küche
- Restaurantmanagement: Überwacht die Implementierung und Nutzung des Systems zur Optimierung des Serviceablaufs und zur Steigerung der Kundenzufriedenheit. Erhalten wertvolle Daten zur Gesamteffizienz des Prozesses.