

Akzeptanz von Leitfäden für Mitarbeitendengespräche

*Eine Untersuchung zur Akzeptanz von KI-generierten und
psychologisch entwickelten Leitfäden für
Mitarbeitendengespräche im Anwendungskontext der
Personalentwicklung.*

BACHELOR THESIS

Juni 2025

Autorin
Rania Ghazal

Betreuungsperson
Dr. Yasemin Tahris

Praxispartnerin
FLOWIT AG
Kontaktperson Bilâl Tahris

Abstract

Traditionelle Mitarbeitendengespräche gelten als zentrales Instrument der Personalentwicklung, sind in der Praxis jedoch oft zeitaufwändig und wenig standardisiert. Vor dem Hintergrund zunehmender Digitalisierung und begrenzter HR-Ressourcen gewinnen KI-gestützte Systeme an Bedeutung. Die vorliegende Arbeit untersucht, inwiefern KI-generierte Gesprächsleitfäden (ChatGPT) eine vergleichbare Akzeptanz erfahren wie Leitfäden, die von psychologisch ausgebildeten Fachpersonen erstellt wurden. Ausgangspunkt der Untersuchung ist die FLOWIT-Plattform, ein digitales HR-Tool, das Unternehmen bei der strukturierten Gesprächsführung unterstützt. Ziel der Arbeit ist es, die technologische Akzeptanz solcher KI-Leitfäden im Anwendungskontext zu analysieren und Implikationen für den praktischen Einsatz abzuleiten. Die Methodik umfasst ein quantitatives Studiendesign zur Erfassung von zentralen Konstrukten der Technologieakzeptanz. Als theoretische Grundlage dienen u.a. das Technology Acceptance Model (TAM), die Unified Theory of Acceptance and Use of Technology (UTAUT) sowie die Computers Are Social Actors Theory (CASA). Die Ergebnisse deuten auf interessante Unterschiede und Gemeinsamkeiten in der Wahrnehmung beider Leitfadenarten hin. Ergänzende qualitative Rückmeldungen geben zudem Hinweise auf potenzielle Gestaltungs- und Entwicklungsbedarfe von FLOWIT. Die Arbeit liefert praxisnahe Erkenntnisse zur Rolle von KI in der Personalentwicklung und bietet eine fundierte Grundlage für weiterführende Forschung zu hybriden Formen der Gesprächsführung im digitalen HR-Kontext.

Schlagerwörter: Human Resources, Personalentwicklung, Personalentwicklungsmethoden, Mitarbeitendengespräche, Gesprächsvorbereitung, Leitfadenerstellung, Künstliche Intelligenz, ChatGPT, Akzeptanz, technologische Akzeptanz

Anzahl Zeichen: 124'732 (inkl. Leerzeichen und aller Bestandteile des Berichts, exkl. Anhang)

Inhaltsverzeichnis

1	Einleitung.....	1
2	FLOWIT	4
3	Theoretische Fundierung.....	6
3.1	Begriffsbestimmung Personalentwicklung	6
3.2	Etablierte Methoden der Personalentwicklung.....	6
3.2.1	360°-Feedback	6
3.2.2	Feedforward-Methode	7
3.2.3	Lösungsorientierte Gesprächsführung.....	8
3.3	Mitarbeitendengespräche	8
3.3.1	Zielsetzung von Mitarbeitendengesprächen	9
3.3.2	Zentraler Nutzen von Mitarbeitendengesprächen	9
3.4	Verschiedene Arten von Mitarbeitendengesprächen.....	11
3.4.1	Jährliches Mitarbeitendengespräch.....	12
3.4.2	Probezeitgespräch	12
3.4.3	Austrittsgespräch.....	13
3.4.4	Kritisches Gespräch.....	14
3.5	Gesprächsvorbereitung als zentraler Erfolgsfaktor	14
3.6	Künstliche Intelligenz	15
3.6.1	Large-Language-Models.....	16
3.6.2	Funktionsweise von Large-Language-Models.....	16
3.6.3	Chancen und Grenzen von Large-Language-Models	17
3.7	ChatGPT und GPT4o.....	17
3.7.1	Funktionsweise von ChatGPT4o	18
3.7.2	Chancen und Grenzen von ChatGPT4o.....	18
3.8	Custom GPTs.....	18
3.8.1	Funktionsweise von Custom GPTs	19
3.8.2	Erstellung von Custom GPTs	19
3.8.3	Chancen und Grenzen von Custom GPTs.....	19
3.9	Mensch-Maschine-Interaktion	20
3.10	Akzeptanztheorien.....	22
3.10.1	Technology Acceptance Model	22
3.10.2	Unified Theory of Acceptance and Use of Technology	24
3.10.3	Computers Are Social Actors Theory	25
4	Konstruktion des Custom GPT	27
4.1	Erstellung der KI-generierten Leitfäden.....	28
5	Konstruktion der psychologischen Leitfäden.....	29
6	Methodik	30
6.1	Untersuchungsdesign	31
6.2	Operationalisierung	32
6.3	Konstruktion des Fragebogens	32

6.3.1	Qualitativer Pretest.....	33
6.3.2	Finaler Fragebogen	34
6.3.3	Finale Umfragestruktur.....	36
6.3.4	Technischer Pretest	38
6.4	Stichprobengrösse und Sampling	38
6.5	Datenerhebung.....	39
6.6	Datenauswertung	40
6.6.1	Deskriptive Statistik	40
6.6.2	Prüfung der Daten	40
6.6.3	Inferenzstatistische Analysen	41
6.6.4	Qualitative Zusatzanalysen	41
7	Ergebnisse.....	43
7.1	Gesamthafte Ergebnisse	43
7.2	Ergebnisse der Einzelvergleiche der Items	44
7.3	Ergebnisse der qualitativen Zusatzanalysen	50
8	Zusammenfassung und Interpretation der Ergebnisse.....	53
8.1	Theoretische Einordnung der Befunde	55
9	Limitationen.....	56
10	Implikationen, kritische Würdigung und Ausblick	57
11	Literaturverzeichnis	60
12	Tabellenverzeichnis	66
13	Abbildungsverzeichnis	67

1 Einleitung

Die vorliegende Arbeit erfolgt in Zusammenarbeit mit der FLOWIT AG. FLOWIT ist eine Onlineplattform, die als KI-unterstützter digitaler Coach Unternehmen bei der Mitarbeitendenbindung und -entwicklung unterstützt (FLOWIT, 2024). Den Ausgangspunkt für die Entwicklung von FLOWIT stellt die Erkenntnis dar, dass traditionelle Methoden zur jährlichen Leistungsbeurteilung, insbesondere Mitarbeitendengespräche, nicht mehr der modernen Arbeitswelt entsprechen. Starre Prozesse und wenig individualisierte Herangehensweisen führen häufig zu einem geringeren Wirkungsgrad. Es wurde deutlich, dass eine innovative und zukunftsorientierte Lösung benötigt wird. FLOWIT adressiert diese Problematik durch digital gestaltete Entwicklungs- und Feedbackprozesse kombiniert mit psychologischen, technologischen und HR-bezogenen Perspektiven. Die Plattform bietet Funktionen wie Pulsbefragungen, 360°-Feedback, regelmässige Zielevaluationen, E-Learnings und Mitarbeitendengespräche. Über 60 Unternehmen mit insgesamt mehr als 10'000 Mitarbeitenden nutzen FLOWIT. Dabei stammen die Unternehmen aus den unterschiedlichsten Branchen wie z.B. Gesundheits- und Bildungswesen, Gast- und Baugewerbe oder Industrie- und Dienstleistungssektor (FLOWIT, 2024).

Die Lösungen von FLOWIT werden durch ein interdisziplinäres Team entwickelt, das sich aus Fachpersonen aus den Bereichen Datenwissenschaft und KI, Personalwesen, IT, Psychologie sowie Unternehmensführung zusammensetzt (FLOWIT, 2024). Eine koordinierende Funktion bei der Einführung und Nutzung der Plattform spielt das interne Customer-Success-Team (CS-Team), welches Unternehmen von der Implementierung bis zur laufenden Anwendung begleitet. Die genaue Ausgestaltung hängt dabei von den lizenzierten Modulen ab. In der vorliegenden Arbeit liegt der Fokus auf einer der meistgenutzten Funktionen: den *Reviews*. Unter dem Oberbegriff *Reviews* werden im Anwendungskontext von FLOWIT verschiedene Gesprächsarten zusammengefasst, darunter jährliche Mitarbeitendengespräche (MAG), Probezeitgespräche, Austrittsgespräche sowie kritische Gespräche. Die Funktion ermöglicht es, strukturierte Gesprächsleitfäden zu definieren, digital auszufüllen und durch Künstliche Intelligenz (KI) auswerten zu lassen. Die KI-generierte Zusammenfassung dient Führungskräften und Mitarbeitenden als Vorbereitung auf das anschliessende persönliche Gespräch. In der

Praxis variiert die Ausgangslage hinsichtlich der vorhandenen Gesprächsleitfäden stark. Drei typische Ausgangslagen lassen sich dabei unterscheiden:

1. Unternehmen verfügen bereits über etablierte Vorlagen, die direkt in die Plattform integriert werden.
2. Unternehmen bringen vorhandene Vorlagen mit, wünschen aber eine Überarbeitung oder Optimierung durch das CS-Team.
3. Unternehmen besitzen keine Leitfäden, sodass neue entwickelt werden müssen.

In den Ausgangslagen 2 und 3 übernimmt eine psychologisch ausgebildete Fachperson des CS-Teams die Ausgestaltung der Leitfäden. Ziel dabei ist es, Leitfäden bereitzustellen, die sowohl psychologisch fundiert als auch an die spezifischen Anforderungen des jeweiligen Unternehmens angepasst sind. In den Entwicklungsprozess fliessen dabei unterschiedliche Einflussfaktoren, darunter Gesprächsart und -ziel, Branche, Unternehmensgrösse, Beschäftigtengruppe (Desktop- oder Non-Desktop-Mitarbeitende) sowie aktuelle kulturelle Themen oder organisationale Herausforderungen ein (FLOWIT, 2024).

In der Praxis zeigt sich trotz dieser durchdachten Vorgehensweise, dass standardisierte Vorlagen die spezifischen Anforderungen einzelner Unternehmen oftmals nicht vollständig abdecken. Die individuelle Erstellung oder Anpassung von Leitfäden ist folglich mit einem erheblichen personellen und zeitlichen Aufwand verbunden. Gerade vor dem Hintergrund begrenzter Ressourcen im operativen HR-Alltag gewinnt der Einsatz KI-gestützter Systeme an Bedeutung (Garnitz & Schaller, 2023; Luger & Sellen, 2016). Während KI in Bereichen wie Rekrutierung oder Performance-Monitoring bereits punktuell Anwendung findet, besteht im Feld der inhaltlichen Gesprächsvorbereitung noch grosses ungenutztes Potenzial (Garnitz & Schaller, 2023; Hermeier, Heupel & Fichtner-Rosada, 2019).

Das Ziel dieser Arbeit besteht daher darin, den Prozess der Leitfadenerstellung effizienter zu gestalten und gleichzeitig die Qualität der Ergebnisse zu sichern. Konkret wird dabei geprüft, inwieweit sich Leitfäden, die mit ChatGPT erstellt wurden, hinsichtlich ihrer Akzeptanz mit jenen vergleichen lassen, die von psychologisch ausgebildeten Fachpersonen erstellt wurden.

Dabei zielt die vorliegende Arbeit darauf ab, folgende Fragestellung zu beantworten:

Wie ausgeprägt ist die technologische Akzeptanz von KI-generierten Leitfäden für typische Mitarbeitendengespräche der FLOWIT?

Die dazugehörige Hypothese lautet:

Leitfäden, die mit Künstlicher Intelligenz (ChatGPT) erstellt wurden, weisen eine äquivalente Akzeptanz auf wie Leitfäden, die von psychologisch ausgebildeten Fachpersonen erstellt wurden.

Zur Beantwortung der Forschungsfrage wird ein quantitatives Untersuchungsdesign gewählt. Basierend auf den erstellten Leitfäden wird ein strukturierter Fragebogen zur systematischen Erfassung der Akzeptanz entwickelt. Die erhobenen Daten werden statistisch ausgewertet, um Rückschlüsse auf die Forschungsfrage und Hypothese zu ermöglichen. Ergänzend dienen qualitative Rückmeldungen dazu, die Ergebnisse differenzierter zu interpretieren und potenzielle Optimierungsbedarfe der Plattform zu identifizieren.

Die Arbeit gliedert sich in einen theoretischen und einen empirischen Teil. Nach einer Einführung und der Vorstellung des Praxispartners FLOWIT werden im theoretischen Teil zentrale Konzepte der Personalentwicklung, Gesprächsführung, KI sowie Technologieakzeptanz erläutert. Darauf aufbauend wird die Konstruktion der KI-generierten sowie der psychologisch entwickelten Gesprächsleitfäden beschrieben. Der empirische Teil umfasst das methodische Vorgehen, die Durchführung der Studie sowie die Auswertung der quantitativen und qualitativen Daten. Abschliessend folgen eine Interpretation der Ergebnisse, Limitationen, Implikationen sowie ein Ausblick auf zukünftige Forschungsansätze.

2 FLOWIT

FLOWIT ist ein Schweizer Unternehmen im Bereich HR-Technologie, welches eine KI-gestützte Plattform für Personalentwicklung entwickelt hat (FLOWIT, 2024). FLOWIT verfolgt das Ziel, Personalentwicklungsprozesse zeitgemäss zu gestalten und Mitarbeitende aktiv in die Entwicklung einzubinden, da diese in klassischen Personalentwicklungsprozessen oft zu wenig berücksichtigt werden.



Abbildung 1. FLOWIT Logo (FLOWIT, 2024)

Die Idee zu FLOWIT entstand im Jahr 2017 als erkannt wurde, dass traditionelle Personalentwicklungsmethoden in einer zunehmend moderneren Arbeitswelt nicht mehr ausreichen. Eine innovative All-In-One-Lösung wurde gesucht. Im Dezember 2020 gegründet, ging die Plattform 2021 live. Sie wird branchenübergreifend eingesetzt, etwa im Gesundheits-, Bau- oder Bildungswesen sowie in Industrie- und Dienstleistungsunternehmen. FLOWIT wurde mehrfach ausgezeichnet, u.a. mit dem Swiss HR Award (2022), dem Swiss Economic Award (2024) und dem HR Award (2025). In einer Finanzierungsrunde im August 2024 wurden vier Millionen Schweizer Franken Kapital eingeworben (FLOWIT, 2024).

FLOWIT stellt eine All-in-One-Plattform zur Unterstützung der Personalentwicklung in Unternehmen dar. Ziel ist es, Unternehmen in der strukturierten Umsetzung von Entwicklungsprozessen zu unterstützen. Die Plattform ermöglicht eine ortsunabhängige Zusammenarbeit, das Setzen individueller Entwicklungsziele sowie die Nachverfolgung entsprechender Massnahmen. Zudem soll sie beitragen, bislang wenig sichtbare Kompetenzen zu erfassen und unabhängig von Funktion oder Sprache Entwicklungsmöglichkeiten aufzuzeigen. Die Gestaltung der Feedbackprozesse erfolgt

mehrdimensional und soll die Kommunikation zwischen Führungskräften und Mitarbeitenden fördern (FLOWIT, 2024). Die Plattform ist modular aufgebaut und lässt sich anhand verschiedener Module an spezifische Bedürfnisse anpassen:

Engagement

- Regelmässige Befragungen zur Erfassung der Teamstimmung
- Identifikation von Herausforderungen und Stimmungen im Unternehmen
- Förderung einer offenen, vertrauensvollen Feedbackkultur

Performance

- Digitale Gesprächsformate für Mitarbeitendengespräche und Zielvereinbarungen
- KI-gestützte Auswertungen zur besseren Einschätzung von Entwicklungsthemen
- Förderung strukturierter und regelmässiger Zielverfolgung

Development

- Systematisches 360°-Feedback aus unterschiedlichen Perspektiven
- Erkennen und Fördern individueller Stärken
- Dokumentation von Entwicklungspfaden und gezielte Förderung von Talenten

Growth (Add-On)

- Integriertes digitales Lernangebot für individuelle Weiterbildung
- Automatische Vorschläge zu passenden Lerninhalten
- Verbindung von Lernen und Arbeiten im Alltag

FLOWIT ist nicht ausschliesslich als technisches Hilfsmittel zur Abbildung von HR-Prozessen konzipiert, sondern wird als strategisches Instrument eingesetzt. Es unterstützt zeitgemässe Führung, fördert die Zusammenarbeit und ermöglicht eine messbare, kontinuierliche Weiterentwicklung von Mitarbeitenden. Ziel des Einsatzes ist es, digitale Strukturen im Personalwesen zu etablieren und vorhandenes Entwicklungspotenzial systematisch zu nutzen (FLOWIT, 2024).

In der vorliegenden Arbeit dient FLOWIT als praxisnaher Anwendungskontext, um die Akzeptanz von KI-generierten Gesprächsleitfäden im Rahmen von MAGs zu untersuchen.

3 Theoretische Fundierung

In diesem Kapitel werden zentrale Begriffe, Theorien und Konzepte im Themenfeld Personalentwicklung und KI systematisch aufgearbeitet. Ziel ist es, eine theoretische Grundlage bereitzustellen, die für die Bearbeitung der Forschungsfrage von Bedeutung ist.

3.1 Begriffsbestimmung Personalentwicklung

Die Personalentwicklung bezeichnet alle spezifische Massnahmen eines Unternehmens, die darauf abzielen, die fachlichen, methodischen und persönlichen Kompetenzen der Mitarbeitenden zu erhalten und weiterzuentwickeln (Hermeier, Heupel & Fichtner-Rosada, 2019; Schuler, 2014). Personalentwicklung wird in der Literatur als kontinuierlicher Prozess verstanden, der eng mit der strategischen Ausrichtung und dem langfristigen Unternehmenserfolg verknüpft ist (Achouri, 2015). Dabei verfolgt die Personalentwicklung das Ziel, bestehende oder absehbare Qualifikationslücken zu schliessen, die Leistungsfähigkeit der Mitarbeitenden zu sichern und die Mitarbeitenden auf zukünftige Anforderungen vorzubereiten (Achouri, 2015; Rowold, 2013). Sowohl Mitarbeitende als auch Führungskräfte tragen gemeinsam Verantwortung für eine wirksame Umsetzung dieser Ziele (Schuler, 2014).

3.2 Etablierte Methoden der Personalentwicklung

Die folgenden Abschnitte geben einen Überblick über zentrale Methoden der Personalentwicklung im Kontext von MAGs und erläutern deren jeweilige Zielsetzungen und Anwendungslogiken. Im Fokus stehen dabei das 360°-Feedback, die Feedforward-Methode sowie die lösungsorientierte Gesprächsführung, die in den folgenden Unterkapiteln beschrieben werden.

3.2.1 360°-Feedback

Das 360°-Feedback stellt ein etabliertes Verfahren zur systematischen und multiperspektivischen Evaluation von Mitarbeitenden und Führungskräften dar (Achouri, 2015). Im Mittelpunkt steht die beurteilte Person, die Feedback von verschiedenen

Stakeholdern innerhalb ihres Arbeitsumfelds erhält (Jöns & Bungard, 2018). Zu den Stakeholdern gehören bspw. Kunden, Führungskräfte, Mitarbeitende oder andere relevante Fokuspersonen (Hossiep, Zens & Berndt, 2020).

Ziel des 360°-Feedbacks ist es, durch die Kombination von Selbst- und Fremdeinschätzungen, ein umfassendes Bild der Leistung und Entwicklungspotenziale einer Person zu gewinnen, um daraus gezielte Entwicklungsmaßnahmen ableiten zu können (Scherer & Sarges, 2019). Darüber hinaus stellt das 360°-Feedback eine wertvolle Grundlage für MAGs dar, indem es eine umfassende Beurteilung von Verhaltensweisen und Leistungen der mitarbeitenden Person liefert (Schirrmacher, 2023; Werther, 2015). Die im Rahmen des 360°-Feedback gewonnenen Erkenntnisse können im MAG aufgegriffen und besprochen werden, um Entwicklungspotenzial zu identifizieren und gezielte Verbesserungsmaßnahmen zu definieren (Jöns & Bungard, 2018).

3.2.2 Feedforward-Methode

Die Feedforward-Methode von Marshall Goldsmith beschreibt ein zukunftsorientiertes Feedback im organisationalen Kontext (Thorn, McLeod & Goldsmith, 2007). Im Gegensatz zum klassischen, retrospektiven Feedback, bei dem Feedback zur bisherigen Leistung oder dem bisherigen Verhalten gegeben wird, wird ein zukunftsorientiertes Feedback gegeben (Hossiep et al., 2020). Die Feedforward-Methode konzentriert sich also auf konkrete Handlungsmöglichkeit für die Zukunft, mit dem Ziel, das Feedback positiv zu gestalten. Studien deuten darauf hin, dass diese Form des Feedbacks als weniger bedrohlich, sondern effizienter und positiver erlebt wird (Thorn et al., 2007). Personen, die Feedback nach dieser Methode erhalten, empfinden das Feedback im Vergleich zum normalen Feedback nicht als Angriff oder Kränkung, sondern können in der Praxis konstruktiver damit umgehen (Thorn et al., 2007).

In MAGs kann die Anwendung der Feedforward-Methode helfen, den Fokus weg von rückblickender Beurteilung hin zu zukunftsbezogenen Entwicklungszielen zu lenken. Dies ermöglicht eine Fokussierung auf die zukünftige Entwicklung und die damit verbundenen nötigen Entwicklungsmaßnahmen, welche ein Hauptziel der MAGs darstellen (Hossiep et al., 2020; Thorn et al., 2007).

3.2.3 Lösungsorientierte Gesprächsführung

Die Lösungsorientierte Gesprächsführung stellt einen methodischen Ansatz dar, der den Fokus auf Ressourcen, Ziele und konkrete Handlungsmöglichkeiten der mitarbeitenden Person richtet (Greif, Möller & Scholl, 2018). Anstatt Probleme zu analysieren, wird gezielt nach bereits funktionierenden Ansätzen und erwünschten Entwicklungen gefragt (Schmitz & Billen, 2008). Diese zukunftsgerichtete Haltung kann eine positive Gesprächsatmosphäre fördern und die Eigenverantwortung stärken.

Im Kontext von MAGs dient der Ansatz dazu, Entwicklungspotenziale zu erkennen und gemeinsam realistische Ziele zu formulieren (Greif et al., 2018; Schmitz & Billen, 2016). Führungskräfte können durch klare, strukturierte Fragen Kompetenzen sichtbar machen und die Perspektive auf Lösungen lenken. Eine respektvolle und anerkennende Haltung seitens Führungskraft wird dabei als zentrale Voraussetzung für eine Lösungsorientierte Gesprächsführung beschrieben (Greif et al., 2018). Bei der Erstellung von Gesprächsleitfäden bietet die Lösungsorientierte Gesprächsführung wichtige Anwendungsansätze. Leitfragen sollten gezielt dazu anregen, über Stärken, bisherige Erfolge und mögliche nächste Schritte nachzudenken. Statt vergangenheitsbezogenem Feedback steht die Förderung von Handlungsspielräumen im Vordergrund, die Mitarbeitende zur aktiven Mitgestaltung motivieren kann. Die Lösungsorientierte Gesprächsführung eignet sich laut Fachliteratur besonders gut für MAGs, da sie Motivation fördern und einen strukturierten Rahmen für Veränderungsprozesse bieten kann. Zudem wird angenommen, dass auf Stärken fokussierte Vereinbarungen tendenziell konkreter, realistischer und nachhaltiger ausfallen (Greif et al., 2018; Schmitz & Billen, 2016).

3.3 Mitarbeitendengespräche

Das MAG stellt ein zentrales Instrument in der Personalentwicklung dar und zählt zu den etabliertesten Personalführungsinstrumenten (Brenner, 2020). Der Begriff *Mitarbeitendengespräch* ist nicht einheitlich definiert und kann je nach Kontext unterschiedliche Bedeutungen annehmen (Mentzel, 2020). Laut Schuler (2014) stellt das MAG einen strukturierten Dialog zwischen einer mitarbeitenden Person und einer Führungskraft dar, wobei je nach Gesprächsart eine bestimmte Zielsetzung besteht. Das Gespräch findet ausserhalb des operationalen Alltags statt, ist planungsintensiv und

benötigt ein gewisses Mass an Vorbereitung. Die Gesprächsdurchführung verläuft anhand vordefinierter Themen, die je nach Gesprächsanlass variieren können (Hossiep et al., 2020).

3.3.1 Zielsetzung von Mitarbeitendengesprächen

Es gibt verschiedene Beweggründe, MAGs durchzuführen, wobei die jeweilige Zielsetzung je nach Art des Gesprächs variiert (Schuler, 2014). Typischerweise werden folgende Ziele verfolgt:

- Definition und Vereinbarung klarer Ziele und Entwicklungsmassnahmen; Abstimmung individueller Aufgaben mit Unternehmenszielen (Hossiep et al., 2020).
- Geben und Empfangen von Feedback; Reflexion des Führungsstils und Verbesserung der Zusammenarbeit (Neuberger, 2015).
- Klärung gegenseitiger Erwartungen, Erfolgskriterien, Verantwortlichkeiten und Entscheidungsbefugnissen (Hossiep et al., 2020; Neuberger, 2015).
- Förderung der gemeinsamen Zusammenarbeit; Stärkung des Beitrags zur Unternehmensstrategie (Schuler, 2014).

Darüber hinaus können MAGs auch genutzt werden, um die Angemessenheit von Unternehmensstrukturen, Arbeitsprozessen und -mitteln zu bewerten und ggf. Optimierungen vorzunehmen (Mentzel, 2020). Damit tragen MAGs dazu bei, individuelle Entwicklungspläne mit strategischen Unternehmenszielen in Einklang zu bringen und eine konstruktive Zusammenarbeit gezielt weiterzuentwickeln (Kaudela-Baum, Meldau & Brassler, 2022).

3.3.2 Zentraler Nutzen von Mitarbeitendengesprächen

Als vielseitiges Führungsinstrument bringen MAGs aus verschiedenen Perspektiven potenzielle Nutzen mit sich (Schuler, 2014). Die jeweiligen Schwerpunkte und Nutzenaspekte variieren dabei je nach Perspektive. (Brenner, 2020; Hossiep, Zens & Berndt, 2020).

Nutzen für Mitarbeitende

Für Mitarbeitende stellt das MAG eine wertvolle Gelegenheit dar, die Selbstwahrnehmung mit der Fremdwahrnehmung ihrer Führungskraft abzugleichen. Dies kann zur persönlichen und beruflichen Entwicklung beitragen und eine gezielte Steuerung der eigenen Karriere ermöglichen (Mentzel, 2020). Zu den zentralen Nutzenaspekten gehören:

- Erhalt von Feedback zu erbrachten Leistungen zur realistischen Einschätzung der eigenen Stärken und Entwicklungspotenziale (Brenner, 2020).
- Klärung zukünftiger Ziele betreffend Arbeit und Weiterentwicklung (Hossiep, Zens & Berndt, 2020).
- Vereinbarung von Verbesserungs- und Entwicklungsmaßnahmen, Aufzeigen von langfristigen Perspektiven (Neuberger, 2015).

Nutzen für Führungskräfte

Mitarbeitendengespräche geben Führungskräften Einblicke in die Perspektiven ihrer Mitarbeitenden (Brenner, 2020). Diese Rückmeldungen ermöglichen eine gezielte Weiterentwicklung der eigenen Führungsarbeit und helfen, die Zusammenarbeit im Team zu optimieren. Konkret profitieren Führungskräfte durch:

- Einblicke in die Sichtweise der Mitarbeitenden hinsichtlich Arbeitsprozesse und -organisation zur Identifikation von Verbesserungspotenzialen
- Feedback zur eigenen Führungsarbeit, um Stärken und mögliche Verbesserungsansätze zu identifizieren (Neuberger, 2015).
- Verständnis für die Zielvorstellungen der Mitarbeitenden, um individuelle Bedürfnisse und Karrierewünsche einzubeziehen (Mentzel, 2020).
- Erkennen de- und motivierender Faktoren, um eine gezieltere Aufgabenzuteilung und nachhaltige Förderung der Arbeitsmotivation zu ermöglichen (Brenner, 2020; Hossiep et al., 2020).

Nutzen für Unternehmen

Auf organisationaler Ebene wird MAGs eine strategische Relevanz zugeschrieben. Sie gelten als potenzielles Instrument, um betriebliche Abläufe zu reflektieren, Arbeitsprozesse effizienter zu gestalten und die Zufriedenheit, Motivation und Leistungsfähigkeit der Mitarbeitenden zu fördern (Schuler, 2014).

- Optimierung von Arbeitsprozessen: Gewonnene Erkenntnisse zeigen Defizite und Verbesserungspotenziale auf und führen zu effizienteren Abläufen und höherer Arbeitsqualität.
- Förderung der Zusammenarbeit: Eine offene Kommunikationskultur stärkt Teamgefühl, verbessert die Kooperation und führt zu einem produktiveren Arbeitsumfeld
- Erhöhung von Arbeitszufriedenheit, Motivation und Commitment. Das Bundesministerium für Arbeit und Soziales (2014) konnte nachweisen, dass eine gesteigerte Arbeitszufriedenheit und Motivation mit einer geringeren Fluktuation einhergehen. Auch Hossiep et al. (2020) fanden signifikante Zusammenhänge zwischen dem Führen von MAGs und höherem Engagement sowie stärkerem Commitment der Mitarbeitenden.

Zusammenfassend stellen MAGs ein essenzielles Führungsinstrument dar. Sie können auf mehreren Ebenen Impulse für individuelle Entwicklung und organisationale Weiterentwicklung liefern. Der konkrete Nutzen hängt jedoch massgeblich von der Umsetzung, der Gesprächsqualität und dem organisationalen Kontext ab (Brenner, 2020; Hossiep et al., 2020; Mentzel, 2020).

3.4 Verschiedene Arten von Mitarbeitendengesprächen

In der Praxis gibt es eine Vielzahl von Gesprächsarten, welche je nach Anlass und Zielsetzung zum Einsatz kommen (Minar, 2024). Im Anwendungskontext von FLOWIT zählen das jährliche MAG, das Probezeitgespräch, das Austrittsgespräch sowie das kritische Gespräch zu den häufigsten Gesprächsarten. Jede Art erfüllt eine spezifische Funktion und trägt dazu bei, die Zusammenarbeit im Unternehmen zu stärken und eine positive Arbeitsumgebung zu fördern (Hossiep et al., 2020). Die folgenden Unterkapitel erläutern die Gesprächsarten exemplarisch in ihrer theoretischen und praktischen Bedeutung.

3.4.1 Jährliches Mitarbeitendengespräch

Das jährliche MAG gilt als zentraler Bestandteil der Personalentwicklung (Schuler, 2014). Es bietet einen strukturierten und offenen Austausch zwischen Mitarbeitenden und Führungskräften und dient dazu, die vergangene Arbeitsperiode zu reflektieren, Erwartungen zu klären und gemeinsame Ziele für die Zukunft zu setzen (Neuberger, 2015). Als wichtiges Führungsinstrument unterstützt das MAG die strategische Personalentwicklung und kann für Motivation, Zufriedenheit und Engagement entscheidend sein (Achouri, 2015). Durch eine offene und wertschätzende Gestaltung wird es zu einem wirkungsvollen Instrument, das nicht nur individuelle Karrieren fördert, sondern auch die Unternehmenskultur nachhaltig stärkt (Stangier, 2019). Voraussetzung für ein erfolgreiches MAG ist eine sorgfältige und strukturierte Vorbereitung. Folglich kann ein vorbereitetes MAG Klarheit über bisherige Leistungen und Entwicklungspotenziale schaffen und den Grundstein für eine erfolgreiche Zusammenarbeit legen (Minar, 2024). Neben der Retrospektive der erzielten Ergebnisse stehen die Arbeitsbedingungen, die Qualität der Zusammenarbeit und individuelle Entwicklungswünsche im Mittelpunkt (Brenner, 2020; Neuberger, 2015; Stangier, 2019). Die Fachliteratur betont zudem, dass durch eine aktive Beteiligung beider Seiten ein wertvoller Austausch entstehen kann, welcher ein gemeinsames Verständnis für Erwartungen und Herausforderungen fördert und zur Verbesserung der Zusammenarbeit beiträgt (Brenner, 2020; Mentzel, 2020).

3.4.2 Probezeitgespräch

Das Probezeitgespräch stellt ein strukturiertes Gespräch zwischen Führungskraft und Mitarbeitenden dar, welches typischerweise ab der dritten Woche der Probezeit stattfindet (Schmitz & Billen, 2016). Die Probezeit selbst dient gegenseitig als Testphase, ob das Arbeitsverhältnis auf Dauer fortgeführt werden soll. Das Gespräch kann diesen Prozess unterstützen, indem es eine erste Zwischenbilanz ermöglicht und Klarheit über Erwartungen, Leistungen und mögliche Entwicklungspotenziale schafft (Brenner, 2020). In der Fachliteratur wird das Probezeitgespräch als spezielle Form allgemeiner Beurteilungsgespräche angesehen. Ziel ist, Potenziale der mitarbeitenden Person zu erkennen und zu fördern sowie Rückmeldungen zu den bisherigen Arbeitsergebnissen zu geben (Schmitz & Billen, 2016). Ergänzend geht es dabei auch um die Einschätzung des

Entwicklungspotenzials, die Vereinbarung neuer Ziele sowie die Förderung der Zusammenarbeit und Verbesserung von Rahmenbedingungen (Brenner, 2020).

Eine sorgfältige und gezielte Vorbereitung auf das Probezeitgespräch gilt als wichtige Voraussetzung, um einen offenen und konstruktiven Austausch zu ermöglichen (Steiger & Lippmann, 2013). Dabei sollte sich die mitarbeitende Person im Vorfeld gezielt mit den eigenen Stärken, Schwächen sowie den bisherigen Arbeitsergebnissen auseinandersetzen. Diese Selbstreflexion trägt wesentlich dazu bei, unterschiedliche Wahrnehmungen zwischen Führungskraft und mitarbeitender Person nachvollziehen und besprechen zu können. Nur so lässt sich ein produktiver Abgleich zwischen Selbst- und Fremdbild herstellen, der als Grundlage für gemeinsame Lösungen und weitere Entwicklungsschritte dient (Hossiep et al., 2020).

3.4.3 Austrittsgespräch

Bei der Beendigung eines Beschäftigungsverhältnisses wird, unabhängig von wem die Trennung ausgeht, i.d.R. ein Austrittsgespräch geführt (Hossiep et al., 2020). Aus Sicht des Unternehmens ist es von Interesse, die Gründe, die zu einer Kündigung geführt haben, zu erfahren. Das Austrittsgespräch ermöglicht darüber hinaus das Identifizieren bestehender Probleme innerhalb des Unternehmens und deren Auswirkung auf die Mitarbeitenden. Daraus lassen sich gezielte Massnahmen entwickeln, um Arbeitsbedingungen verbessern zu können (Schuler, 2014). Wird das Arbeitsverhältnis durch das Unternehmen beendet, zielt das Austrittsgespräch darauf ab, die Beweggründe zu beleuchten und der betroffenen Person die Möglichkeit für Feedback zu geben. Gleichzeitig kann das Gespräch dabei helfen, Hinweise für künftige berufliche Perspektiven zu sammeln (Hossiep et al., 2020). Ein Austrittsgespräch dient darüber hinaus auch dazu, das Arbeitsverhältnis in einem wertschätzenden Rahmen zu beenden, indem man die Leistungen der mitarbeitenden Person anerkennt und die Trennung respektvoll gestaltet; ein wichtiger Aspekt sowohl für die persönliche Wertschätzung als auch für das Image des Unternehmens (Schuler, 2014).

3.4.4 Kritisches Gespräch

Ein kritisches Gespräch ist ein strukturiert geführter Austausch zwischen Führungskraft und mitarbeitender Person, der auf einen konkreten Anlass zurückgeht und situativ geführt wird, wenn eine bestimmte Situation oder ein Verhalten eine Klärung oder Veränderung erforderlich machen (Hossiep et al., 2020). Ziel ist, Missstände oder Unzufriedenheit offen anzusprechen und gleichzeitig konstruktive Perspektiven auf Lösungen und positive Weiterentwicklungen zu richten. Kritische Gespräche gelten als besonders herausfordernd. Umso wichtiger ist es, dass sie in einem respektvollen und geschützten Rahmen stattfinden, vorzugsweise im persönlichen Austausch. Dies schafft Raum für Offenheit und ermöglicht es beiden Seiten, ihre Sichtweisen einzubringen (Schmitz & Billen, 2016).

Eine sorgfältige Vorbereitung ist unerlässlich, da kritische Gespräche beiderseits oft Unsicherheit, Befürchtungen oder sogar Ängste auslösen (Steiger & Lippmann, 2013). In einer guten Gesprächsführung ist Kritik lösungsorientiert, sachlich und im Sinne einer Weiterentwicklung gestaltet. Hilfreich ist, die Inhalte des Gesprächs klar zu strukturieren, konkrete Beispiele für beobachtetes Verhalten zu benennen und die Gesprächsatmosphäre so zu gestalten, dass ein respektvoller Dialog möglich ist.

3.5 Gesprächsvorbereitung als zentraler Erfolgsfaktor

Der Prozess eines MAG umfasst mehrere aufeinander abgestimmte Schritte, wobei die Gesprächsvorbereitung als einer der zentralsten Erfolgsfaktoren gilt (Hossiep et al., 2020). Eine gründliche Vorbereitung legt den Grundstein für ein strukturiertes und zielführendes Gespräch und trägt wesentlich zur Gesprächsqualität und zum Nutzen für alle Beteiligten bei (Brenner, 2020). Die Vorbereitung beginnt mit der frühzeitigen Terminvereinbarung und schliesst die inhaltliche Planung des Gesprächs ein (Steiger & Lippmann, 2013). Laut Brenner (2020) trägt insbesondere eine transparente Kommunikation im Vorfeld dazu bei, Unsicherheiten zu reduzieren und ein vertrauensvolles Gesprächsklima zu schaffen. Ein wesentliches Element der Vorbereitung ist die Erstellung eines Gesprächsleitfadens (Kaudela-Baum et al., 2022). Dieser dient als Orientierungshilfe und strukturiert das Gespräch entlang definierter Themenbereiche (Steiger & Lippmann, 2013). Ein durchdachter Leitfaden berücksichtigt die Perspektiven beider Gesprächspartner.

Die genaue Ausgestaltung der Fragen hängt dabei von Gesprächsart und Zielsetzungen ab (Schmitz & Billen, 2016). Typische Themenbereiche sind Zusammenarbeit, persönliche und berufliche Entwicklung, Zielvereinbarungen, Feedbackkultur und allgemeine Arbeitszufriedenheit (Kaudela-Baum et al., 2022). Schuler und Kanning (2014) betonen, dass MAGs nicht nur der Leistungsbewertung dienen, sondern auch als strategisches Instrument der Personalentwicklung und Motivation verstanden werden müssen.

Ein strukturierter Leitfaden unterstützt nicht nur einen reibungslosen Gesprächsverlauf, sondern fördert auch Transparenz, Vergleichbarkeit und eine positive Gesprächsatmosphäre (Hossiep et al., 2020). Zusätzlich hilft er dabei, das Gespräch lösungsorientiert zu gestalten, klare Vereinbarungen zu treffen und eine zielgerichtete Nachbereitung zu ermöglichen (Brenner, 2020). Vor diesem theoretischen Hintergrund hat FLOWIT eine Funktion entwickelt, bei der KI-gestützte Zusammenfassungen als Ausgangspunkt für das anschliessende persönliche Gespräch genutzt werden. Dieses Beispiel verdeutlicht, wie theoretische Konzepte aus der Personalentwicklung in digitale Anwendungen überführt werden können. Zugleich wirft es die Frage auf, wie solche Lösungen im operationalen Alltag tatsächlich wahrgenommen und akzeptiert werden.

3.6 Künstliche Intelligenz

Künstliche Intelligenz bezeichnet die Fähigkeit von Computersystemen, Aufgaben zu lösen, die typischerweise menschliche Intelligenz erfordern, wie etwa das Verstehen von Sprache, das Treffen von Entscheidungen oder das Erkennen von Mustern (Taulli, 2022). In der heutigen Arbeitswelt gewinnt KI insbesondere in der Personalentwicklung zunehmend an Bedeutung (Kauffeld, 2014). Die Wurzeln der KI reichen bis in die 1950er Jahre zurück (Ertel, 2021). Seither entwickelt sich das Feld in Innovationsschüben weiter, insbesondere durch den zunehmenden Fortschritt seit den 2010er Jahren durch Machine Learning, Deep Learning und die Verfügbarkeit grosser Datenmengen (Ertel, 2021; Kreutzer, 2023). In der Praxis kommt fast ausschliesslich sogenannte schwache KI zum Einsatz. Dabei handelt es sich um Systeme, die auf klar umrissene Aufgabenbereiche spezialisiert sind (Kauffeld, 2014).

Im Kontext der vorliegenden Arbeit ist insbesondere die Anwendung von KI im Bereich der Sprachverarbeitung relevant. Dabei steht der Einsatz sogenannter Large-

Language-Models (LLMs) im Mittelpunkt, die in den folgenden Unterkapiteln näher erläutert werden.

3.6.1 Large-Language-Models

Large-Language-Models sind eine spezielle Form von KI-Modellen, die auf die Verarbeitung und Generierung natürlicher Sprache spezialisiert sind (Ertel, 2021). Einen der bekanntesten Vertreter dieser Modellklasse stellt GPT (Generative Pre-trained Transformer) dar (Paaß & Hecker, 2020). LLMs basieren auf neuronalen Netzwerken mit Milliarden von Parametern und werden auf umfangreichen Textkorpora trainiert (Ertel, 2021; Taulli, 2022). Ziel ist es, durch das Erkennen sprachlicher Muster sinnvolle, kohärente und kontextbezogene Texte zu generieren (Kreutzer, 2023; Taulli, 2022). Einsatzgebiete umfassen u.a. Chatbots, automatisierte Textzusammenfassungen sowie Sprachassistenzsysteme. Zudem finden sie Anwendung in der Personalentwicklung, etwa bei der Erstellung von Lerninhalten oder der Analyse von Mitarbeitendenfeedback (Bünnagel, 2024).

3.6.2 Funktionsweise von Large-Language-Models

Die Funktionsweise von LLMs beruht auf sogenannten Transformer-Architekturen, die erstmals 2017 vorgestellt wurden (Paaß & Hecker, 2020). Diese ermöglichen es dem Modell, komplexe sprachliche Muster zu erkennen und langfristige Abhängigkeiten im Text zu erfassen (Bünnagel, 2024; Ertel, 2021). In der Vortrainingsphase lernt das Modell anhand grosser Mengen unstrukturierter Texte, Sprachstrukturen, Grammatik und semantische Zusammenhänge zu erkennen. Anschliessend kann das Modell durch sogenanntes *Fine-Tuning* auf spezifische Aufgaben oder Fachgebiete angepasst bzw. trainiert werden (Bünnagel, 2024). Während der Anwendung analysiert das Modell den Kontext eines Texteingangs und erzeugt auf dieser Basis die wahrscheinlichste Fortsetzung oder Antwort (Paaß & Hecker, 2020). Die Fähigkeit, Sprache auf menschlich wirkende Weise zu generieren, ist dabei ein zentrales Merkmal von LLMs (Ertel, 2021; Taulli, 2022).

3.6.3 Chancen und Grenzen von Large-Language-Models

Large-Language-Models bieten vielfältige Möglichkeiten zur Unterstützung in der Personalentwicklung (Hasenbein, 2023). Sie können zur Erstellung strukturierter Inhalte, automatisierter Beantwortung von Anfragen oder Analyse von Sprachdaten eingesetzt werden (Hasenbein, 2023). Eine besondere Stärke liegt in der Fähigkeit, Inhalte flexibel an verschiedene Zielgruppen, Sprachniveaus oder Gesprächsziele anzupassen (Kreutzer, 2023). Den Potenzialen stehen jedoch auch Grenzen gegenüber (Paaß & Hecker, 2020). Die Qualität der Ergebnisse hängt stark von den Trainingsdaten ab (Bünnagel, 2024; Ertel, 2021). Ein zentrales Problem der LLMs stellt das Phänomen der sogenannten Halluzination dar, bei dem die Modelle plausibel klingende jedoch faktisch falsche Inhalte erzeugen. Datenschutz, Transparenz und ethische Fragen, etwa im Umgang mit sensiblen Mitarbeiterdaten, stellen weitere zentrale Herausforderungen bei der Nutzung von LLMs dar (Bünnagel, 2024; Ertel, 2021; Taulli, 2022). Ein bewusster und reflektierter Einsatz ist daher unerlässlich.

3.7 ChatGPT und GPT4o

ChatGPT ist ein von OpenAI entwickeltes KI-basiertes Sprachmodell, das auf der GPT-Architektur (Generative Pre-trained Transformer) basiert (Bünnagel, 2024). Es wurde entwickelt, um menschenähnliche Sprache zu verstehen und zu generieren und wird mittlerweile in zahlreichen Anwendungsfeldern eingesetzt, zunehmend auch im HR-Kontext (OpenAI, 2024). Mit GPT4o (omni) präsentierte OpenAI im Mai 2024 erstmals ein multimodales Modell, das Text-, Bild- und Audioeingaben innerhalb einer einzigen Architektur verarbeiten kann. Im Vergleich zu früheren Versionen bietet GPT4o eine höhere Verarbeitungsgeschwindigkeit, eine natürliche Dialogführung in Echtzeit sowie eine verbesserte Fähigkeit zur kontextsensitiven Reaktion. Diese Eigenschaften machen das Modell besonders für den praktischen Einsatz in Bereichen, in denen strukturierte Kommunikation und Anpassungsfähigkeit erforderlich sind, relevant. Ein möglicher Anwendungsbereich dafür ist die automatisierte Erstellung von Gesprächsleitfäden in der Personalentwicklung (OpenAI, 2024).

3.7.1 Funktionsweise von ChatGPT4o

GPT4o basiert auf der Transformer-Architektur und wurde mit einer grossen Bandbreite an Textdaten trainiert (OpenAI, 2024). Die aktuelle Version kann Sprache, Bilder und Audioinformationen verstehen und darauf reagieren. Die multimodale Verarbeitung erlaubt es, verschiedene Eingabeformen miteinander zu verknüpfen und entsprechend angepasste Ausgaben zu generieren (OpenAI, 2024). Im Anwendungskontext der Leitfadenerstellung bedeutet dies, dass Inhalte kontextbezogen strukturiert und sprachlich auf unterschiedliche Anforderungen angepasst werden können. Diese Anpassungsfähigkeit macht es zu einem wertvollen Werkzeug in der digitalen Unterstützung personalisierter Kommunikationsformate (OpenAI, 2024).

3.7.2 Chancen und Grenzen von ChatGPT4o

Aus theoretischer Perspektive lassen sich beim Einsatz von ChatGPT4o in der Personalentwicklung mehrere Potenziale identifizieren (Scheuer, 2020). So ermöglichen LLMs eine effiziente, sprachlich konsistente und individualisierbare Leitfadenerstellung. In der Praxis kann dies Fachpersonen entlasten, Prozesse beschleunigen und ermöglicht eine flexible Anpassung an unterschiedliche Gesprächsanlässe. Gleichzeitig sind bestimmte Herausforderungen zu beachten (OpenAI, 2024). Die Qualität der generierten Inhalte ist stark abhängig von der Qualität und Breite der zugrunde liegenden Trainingsdaten. Zudem besteht das Risiko sogenannter Halluzinationen, bei denen inhaltlich unzutreffende, aber plausibel wirkende Aussagen erzeugt werden. Auch Fragen des Datenschutzes und der Transparenz spielen insbesondere im HR-Kontext eine wichtige Rolle. Der Einsatz von GPT4o sollte daher bewusst gesteuert und reflektiert erfolgen, um die Potenziale gezielt zu nutzen und mögliche Risiken zu minimieren (OpenAI, 2024).

3.8 Custom GPTs

Mit der Einführung von sogenannten Custom GPTs durch OpenAI im November 2023, besteht die Möglichkeit, massgeschneiderte Versionen von ChatGPT zu entwickeln. Durch gezielte Konfiguration lassen sich diese Modelle auf bestimmte Aufgaben, Sprachstile oder Wissensdomänen ausrichten, was ihren Einsatz in spezialisierten Kontexten

erleichtert. Theoretisch ermöglichen Custom GPTs folglich, LLMs kontextbezogen anzupassen und gezielt in bestehende Prozesse zu integrieren (OpenAI, 2024).

3.8.1 Funktionsweise von Custom GPTs

Custom GPTs basieren auf der GPT-4-Architektur und erlauben Nutzenden, das Verhalten des Modells durch benutzerdefinierte Anweisungen und eigene Inhalte zu steuern. Die Konfiguration erfolgt über eine benutzerfreundliche Oberfläche und setzt keine Programmierkenntnisse voraus. Zudem kann festgelegt werden, welche Funktionen das Modell unterstützen soll wie etwa Webzugang, Dateiupload oder Bildverarbeitung. Durch diese konfigurierbare Flexibilität entstehen vielfältige Anwendungsmöglichkeiten wie etwa in der Kundenkommunikation, im Projektmanagement oder in der Wissensvermittlung. Die Kombination aus individueller Steuerung und leistungsfähiger Sprachverarbeitung macht Custom GPTs zu einer vielseitig einsetzbaren Technologie (OpenAI, 2024).

3.8.2 Erstellung von Custom GPTs

Die Erstellung eines Custom GPT erfolgt direkt über die bereitgestellte Konfigurationsumgebung auf der Plattform von OpenAI. Nutzende können dabei in wenigen Schritten festlegen, welche Aufgaben das Modell übernehmen soll, welche Informationen es verwenden darf und wie es sprachlich auftreten soll. Der Prozess ist niederschwellig gestaltet und für fachfremde Personen zugänglich. So lassen sich innerhalb kurzer Zeit angepasste Modelle entwickeln, die auf spezifische Ziele ausgerichtet sind. Die konkrete Ausgestaltung des im Rahmen dieser Arbeit erstellten Custom GPTs wird in Kapitel 4 detailliert erläutert.

3.8.3 Chancen und Grenzen von Custom GPTs

Custom GPTs bieten vielfältige Möglichkeiten zur Individualisierung von Verhalten, Sprachstil und Informationsbasis angepasst an spezifische Anwendungsanforderungen. Dies ermöglicht die Entwicklung massgeschneiderter Lösungen, Prozessoptimierungen und das Automatisieren manueller Aufgaben. Die benutzerfreundliche Oberfläche erlaubt

Konfiguration auch ohne Programmierkenntnisse. Gleichzeitig ergeben sich gewisse Herausforderungen, welche berücksichtigt werden müssen. Die Offenheit der Systeme birgt Risiken etwa durch fehlerhafte Eingaben oder potenzielle Manipulation. Zudem ist beim Umgang mit sensiblen Daten besondere Sorgfalt gefragt. Auch die Abhängigkeit von OpenAI hinsichtlich Infrastruktur und Nutzungsrichtlinien sollte reflektiert betrachtet werden.

Zusammenfassend bieten Custom GPTs ein hohes Potenzial zur Unterstützung organisatorischer Prozesse. Ihr effektiver und sicherer Einsatz setzt jedoch ein reflektiertes Verständnis technischer und ethischer Rahmenbedingungen voraus.

3.9 Mensch-Maschine-Interaktion

Nach der Betrachtung technologischer Grundlagen richtet sich der Blick nun auf die Schnittstelle zwischen Mensch und Maschine (Schirmacher, 2023). Die Frage, wie Mensch und KI effektiv zusammenarbeiten können, ist zentral für den erfolgreichen Einsatz dieser Technologien, insbesondere in komplexen Bereichen wie der Personalentwicklung (Epley, Waytz & Cacioppo, 2007; Kauffeld, 2014). Erkenntnisse aus Forschung und Praxis zeigen dabei deutlich, dass die grössten Potenziale durch die Kombination maschineller Verarbeitungsstärke mit menschlicher Urteilskraft liegen (Ertel, 2021; Nass & Moon, 2000). KI-gestützte Sprachmodelle wie ChatGPT oder individuell konfigurierte Custom GPTs sind in der Lage, grosse Datenmengen in kurzer Zeit zu analysieren, Muster zu erkennen und strukturierte Ausgaben zu generieren (Ertel, 2021; Taulli, 2022). Ihre Stärken liegen in der Geschwindigkeit, Skalierbarkeit und Reproduzierbarkeit (Ertel, 2021; Taulli, 2022). Der Mensch wiederum bringt unverzichtbare Fähigkeiten wie Kontextverständnis, Empathie, ethische Reflexion und strategisches Denken ein, also Kompetenzen, die Maschinen bislang nicht leisten können (Luger & Sellen, 2016; Nass & Moon, 2000). Im Zusammenspiel entsteht das Prinzip der *Hybrid Intelligence*, bei dem die KI unterstützend wirkt, Entscheidungen jedoch weiterhin durch Menschen getroffen und verantwortet werden (Epley et al., 2007). In der Personalentwicklung bedeutet dies bspw., dass KI-basierte Systeme Inhalte vorschlagen oder auswerten, während Zielsetzungen, Kommunikation und Feinevaluation in der Hand der Fachkräfte bleiben (Hasenbein, 2023).

In der praktischen Umsetzung kann zur Leitfadenerstellung KI als Assistenzsystem eingesetzt und anschliessend durch menschliche Expertise überarbeitet und individualisiert

werden (Garnitz & Schaller, 2023). Ziel ist dabei nicht die Substitution menschlicher Arbeit, sondern ihre gezielte Ergänzung (Bauer, Ganz, Hämmerle & Renner, 2019). Über den Bereich der Personalentwicklung hinaus zeigen sich vielfältige Einsatzmöglichkeiten für KI im HR-Bereich (Schirmmacher, 2023). Vom Verfassen von Stellenanzeigen über die Vorauswahl von Bewerbenden bis hin zur Unterstützung bei Vorstellungsgesprächen, administrativen Prozessen und der strategischen Personalplanung (Bauer et al., 2019; Hermeier et al., 2019). In einer zunehmend wettbewerbsintensiven Arbeitswelt gewinnt der Einsatz von KI für HR-Bereiche vornehmlich an Bedeutung, um qualifizierte Talente zu gewinnen, zu fördern und langfristig zu binden (Hasenbein, 2023).

Trotz dieser vielversprechenden Perspektiven zeigt sich aktuell noch ein zurückhaltendes Bild in der Umsetzung (Garnitz & Schaller, 2023). Es bestehen stets Bedenken hinsichtlich Datenschutzes oder Verzerrungen, welche maßgeblich die Akzeptanz von KI-Systemen beeinflussen. Derzeit setzen lediglich rund 5% der Unternehmen KI aktiv im HR-Bereich ein, etwa ein Viertel plant dies künftig. Als besonders relevant wird dabei das Automatisierungspotenzial zur Effizienzsteigerung von Personalprozessen eingeschätzt (Garnitz & Schaller, 2023; Hermeier et al., 2019). Dennoch ist zu erwarten, dass der Einfluss KI-gestützter Systeme im HR-Bereich in den kommenden Jahren weiter zunimmt (Garnitz & Schaller, 2023). Vor dem Hintergrund zunehmender Leistungsfähigkeit von LLMs wie ChatGPT erscheint es denkbar, dass KI künftig nicht nur unterstützende Aufgaben übernimmt, sondern aktiv an komplexeren HR-Prozessen mitwirkt, wo strukturierte Daten, wiederkehrende Abläufe und zeitkritische Entscheidungen eine Rolle spielen (Hermeier et al., 2019).

Die Mensch-Maschine-Interaktion steht damit im Zentrum eines zukunftsorientierten verantwortungsvollen KI-Einsatzes (Bauer et al., 2019). Nur wenn die jeweiligen Stärken von Mensch und KI gezielt kombiniert werden, können nachhaltige Lösungen entstehen (Garnitz & Schaller, 2023; Hasenbein, 2023; Nass & Moon, 2000). Dieses Zusammenspiel bildet auch die Grundlage für die weiteren Überlegungen und Perspektiven dieser Arbeit insbesondere in Hinblick auf die Akzeptanz KI-generierter Leitfäden im Kontext von MAG.

3.10 Akzeptanztheorien

Vor dem Hintergrund der zunehmenden Verbreitung von KI-Anwendungen wie ChatGPT stellt sich verstärkt die Frage, unter welchen Bedingungen solche Technologien von Nutzenden wahrgenommen und akzeptiert werden. In der vorliegenden Arbeit steht die Akzeptanz von KI-generierten Leitfäden für MAGs im Zentrum des Interesses. Zur theoretischen Fundierung dieser Fragestellung werden in diesem Kapitel zentrale Ansätze zur Technologieakzeptanz vorgestellt. Als zentrale theoretische Grundlagen dienen das Technology Acceptance Model (TAM) und dessen Weiterentwicklung, die Unified Theory of Acceptance and Use of Technology (UTAUT). Darüber hinaus wird die Computers Are Social Actors (CASA)-Theorie betrachtet, die beschreibt, wie Menschen in der Interaktion mit Maschinen soziale Verhaltensweisen anwenden.

3.10.1 Technology Acceptance Model

Das TAM stellt ein etabliertes theoretisches Modell zur Erklärung und Vorhersage der Akzeptanz neuer Technologien dar (Davis & Granić, 2024). Es wurde von Fred Davis im Jahr 1987 entwickelt und basiert auf der Annahme, dass die Nutzung einer Technologie primär durch zwei Hauptkonstrukte determiniert wird:

- Wahrgenommene Nützlichkeit (PU)
- Wahrgenommene Benutzerfreundlichkeit (PEOU)

Die wahrgenommene Nützlichkeit (PU) beschreibt die subjektive Einschätzung einer Person hinsichtlich des Nutzens einer Technologie für die eigene Arbeitsleistung (Scheuer, 2020). Die wahrgenommene Benutzerfreundlichkeit (PEOU) bezieht sich auf die erwartete Leichtigkeit der Nutzung. Dabei beeinflusst die Benutzerfreundlichkeit indirekt die Nützlichkeitswahrnehmung, da eine geringe kognitive Belastung die wahrgenommene Effizienz der Technologie erhöhen kann (Davis & Granić, 2024).

Ergänzt wird das TAM durch weitere Konstrukte:

- Einstellung zur Nutzung (ATU)
- Verhaltensabsicht zur Nutzung (BI)

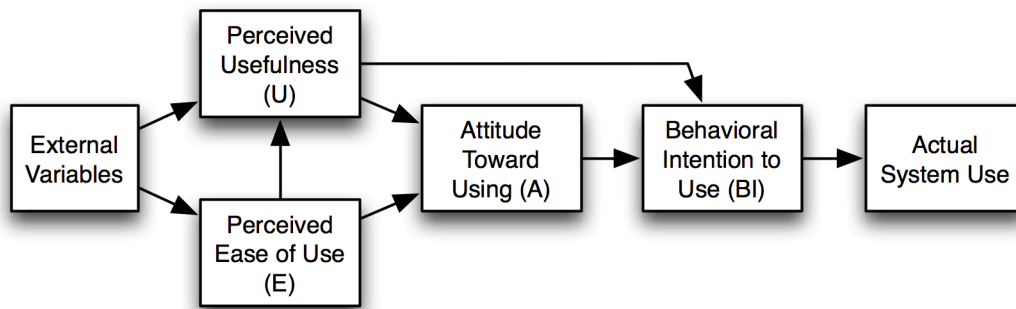


Abbildung 2. Technology-Acceptance-Model (Venkatesh & Davis, 2000)

Die Einstellung zur Nutzung (ATU) beschreibt die individuelle Bewertung der Technologie. Sie wird sowohl durch die wahrgenommene Benutzerfreundlichkeit als auch durch die wahrgenommene Nützlichkeit beeinflusst. Die Verhaltensabsicht zur Nutzung (BI) wiederum gibt an, in welchem Masse eine Person beabsichtigt, die Technologie künftig zu verwenden. Die Verhaltensabsicht (BI) wird massgeblich durch die Einstellung zur Nutzung (ATU) sowie direkt durch die wahrgenommene Nützlichkeit (PU) beeinflusst (Davis, 1987).

Aufgrund der empirisch nachgewiesenen Relevanz des Modells wurde es in späteren Studien durch Venkatesh und Davis (2000) zum TAM2 weiterentwickelt. Diese Version integriert soziale Einflussfaktoren wie subjektive Normen und Image sowie kognitive Instrumentalprozesse, darunter Job-Relevanz, Output-Qualität und Ergebnismachbarkeit (Venkatesh & Davis, 2000). Das TAM hat sich in einer Vielzahl empirischer Studien als robustes und prädiktives Modell etabliert (Scheuer, 2020). Aufgrund seiner hohen Erklärungskraft wird es sowohl in der Wissenschaft als auch in der Praxis zur Technologieakzeptanzforschung und Optimierung von Technologien verwendet. Es ermöglicht eine fundierte Analyse möglicher Akzeptanzbarrieren und kann damit zur Gestaltung effektiver Implementierungsstrategien in Unternehmen beitragen (Davis & Granić, 2024; Venkatesh & Davis, 2000).

Die Wahl des TAM als theoretische Grundlage ergibt sich aus seiner weitreichenden Anwendung in der Erforschung der Technologieakzeptanz und seiner hohen Erklärungskraft für die Nutzung neuer Systeme. Obwohl TAM2 zusätzliche Einflussfaktoren berücksichtigt, wurde in dieser Arbeit bewusst TAM verwendet. Der Fokus liegt auf der grundlegenden Nutzungsbereitschaft, die durch die Konstrukte Wahrgenommene Nützlichkeit (PU) und Benutzerfreundlichkeit (PEOU) ausreichend erfasst wird.

3.10.2 Unified Theory of Acceptance and Use of Technology

Die Unified Theory of Acceptance and Use of Technology (UTAUT) stellt ein integratives Modell dar, das die Determinanten der Akzeptanz und Nutzung von Technologien systematisch erfasst und analysiert (Scheuer, 2020). Entwickelt von Venkatesh, Morris, Davis und Davis (2003) basiert das Modell auf der Synthese mehrerer etablierter Ansätze zur Technologieakzeptanz, einschliesslich dem TAM, der Theorie des geplanten Verhaltens (TPB) sowie der Innovationsdiffusionstheorie (IDT). Durch diese Konsolidierung soll die UTAUT eine umfassendere Erklärung der Nutzendenakzeptanz technologischer Systeme in organisationalen Kontexten ermöglichen (Venkatesh et al., 2003). Kernbestandteile des UTAUT-Modells sind vier zentrale Konstrukte:

- Leistungserwartung (PE)
- Aufwandserwartung (EE)
- Sozialer Einfluss (SI)
- Unterstützende Rahmenbedingungen (FC)

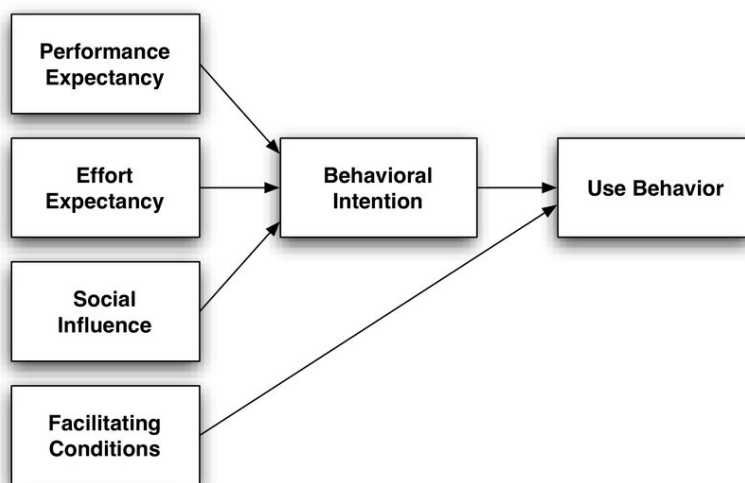


Abbildung 3. Unified Theory of Acceptance and Use of Technology (Venkatesh et al. 2003)

Die Leistungserwartung (PE) beschreibt die subjektive Einschätzung der Nutzenden hinsichtlich der potenziellen Leistungssteigerung durch den Einsatz der Technologie. Die Aufwandserwartung (EE) reflektiert die antizipierte kognitive und physische Anstrengung, die mit der Nutzung der Technologie verbunden ist. Der soziale Einfluss (SI) bezieht sich auf die wahrgenommene Relevanz der Meinungen signifikanter sozialer Akteure im

Entscheidungsprozess (Scheuer, 2020). Unterstützende Rahmenbedingungen (FC) umfassen externe, infrastrukturelle und organisatorische Faktoren, die die Implementierung und Nutzung der Technologie erleichtern können (Venkatesh et al., 2003). Die Leistungserwartung, Aufwandserwartung und der soziale Einfluss wirken direkt auf die Verhaltensabsicht zur Nutzung, während unterstützende Rahmenbedingungen in erster Linie das tatsächliche Nutzungsverhalten beeinflussen. Zusätzlich werden diese Beziehungen durch Moderatoren wie Alter, Geschlecht, Erfahrung und Freiwilligkeit der Nutzung beeinflusst, wodurch sich die Wirkung der einzelnen Konstrukte je nach Anwendungskontext unterscheiden kann (Venkatesh et al., 2003).

Empirische Studien belegen die hohe Erklärungskraft des UTAUT für die Technologieakzeptanz in verschiedenen organisationalen und nicht-organisationalen Kontexten (Scheuer, 2020). In der Praxis ermöglicht das Modell eine evidenzbasierte Ableitung von Strategien zur Förderung der Implementierung neuer Technologien, indem es mögliche Akzeptanzbarrieren identifiziert und adressiert (Venkatesh et al., 2003).

Im Rahmen der vorliegenden Arbeit wird die UTAUT als ergänzendes theoretisches Modell herangezogen, da sie verschiedene Technologieakzeptanzmodelle zusammenführt und eine hohe Erklärungskraft für die Nutzung technologischer Systeme besitzt. Die im Modell vorgesehenen moderierenden Variablen werden in dieser Untersuchung jedoch nicht berücksichtigt, da sie für die spezifische Zielsetzung der Arbeit nicht im Fokus stehen.

3.10.3 Computers Are Social Actors Theory

Die Computers Are Social Actors (CASA)-Theorie besagt, dass Menschen in der Interaktion mit technischen Systemen unbewusst soziale Verhaltensmuster aktivieren, welche eigentlich für den Austausch mit anderen Menschen gedacht sind (Gambino, Fox & Ratan, 2020). Laut Nass und Moon (2000) werden technologische Systeme, insbesondere solche mit interaktiven oder kommunikativen Eigenschaften, nicht nur als neutrale Werkzeuge wahrgenommen, sondern als soziale Akteure, mit denen Nutzende nach vertrauten Mustern interagieren. Dieses psychologische Phänomen beruht auf Denkgewohnheiten, die Menschen helfen, in ihrer Umwelt schnell und effizient zu handeln. Die Zuschreibung sozialer Eigenschaften an Maschinen geschieht dabei unbewusst, weil unser Gehirn soziale Signale erkennt und entsprechend reagiert (Gambino et al., 2020).

Damit liefert CASA einen Erklärungsrahmen für die Tatsache, dass Nutzende auch bei Maschinen soziale Erwartungen entwickeln, selbst wenn sie deren Künstlichkeit erkennen (Nass & Moon, 2000). Mit der zunehmenden Verbreitung von KI hat die CASA-Theorie erneut an Relevanz gewonnen. Moderne KI-Technologien wie z.B. ChatGPT imitieren zunehmend menschliche Kommunikationsstile. Systeme wie virtuelle Assistenten, Chatbots oder interaktive Agenten können mittlerweile Mimik, Sprachintonationen und kontextabhängige Reaktionen simulieren, was zu einer stärkeren sozialen Bindung der Nutzenden an diese Technologien führt. Gleichzeitig zeigen aktuelle Studien, dass Menschen mit der Zeit spezifische mentale Modelle für den Umgang mit intelligenten Systemen entwickeln (Nass & Moon, 2000). Während klassische CASA-Annahmen von einer generellen und stabilen Übertragung sozialer Regeln auf Maschinen ausgehen, legen neuere Befunde nahe, dass Nutzende differenziertere Erwartungen an KI-Technologien haben und zwischen menschlichen und künstlichen Interaktionspartnern zunehmend bewusster unterscheiden (Epley, Waytz & Cacioppo, 2007).

Die CASA-Theorie stellt ein zentrales Modell in der Erforschung der Mensch-Computer-Interaktion dar und liefert wertvolle Erkenntnisse für die Gestaltung und Akzeptanz digitaler Systeme (Gambino et al., 2020). Sie erklärt, warum menschenähnliche Eigenschaften in technischen Systemen Vertrauen und soziale Nähe fördern können. Gleichzeitig zeigen neuere Entwicklungen, dass CASA als Modell weiterentwickelt werden muss, um dynamische Anpassungen an technologische Fortschritte zu berücksichtigen (Gambino et al., 2020). Während anthropomorphe Gestaltungselemente weiterhin die Interaktion mit Maschinen erleichtern, erfordert der wachsende Einfluss von KI eine differenzierte Betrachtung der Nutzerwahrnehmung (Epley et al., 2007). Die Forschung im Bereich der Human-Computer-Interaktion wird sich daher weiterhin mit der Frage beschäftigen, wie sich die sozialen Mechanismen der Mensch-Maschine-Interaktion im Zuge technologischer Innovationen verändern.

Die (CASA)-Theorie dient in dieser Untersuchung als theoretischer Rahmen. Im Kontext der technologischen Akzeptanz von KI-generierten Leitfäden für MAGs liefert CASA eine relevante Perspektive darauf, inwiefern Nutzende KI-gestützte Systeme als soziale Akteure wahrnehmen und entsprechend mit ihnen interagieren.

4 Konstruktion des Custom GPT

Im Rahmen dieser Arbeit wurde ein eigenes Custom GPT mithilfe der Plattform ChatGPT von OpenAI erstellt. Ziel war es, ein Sprachmodell zu entwickeln, das auf strukturierte Weise Leitfäden für unterschiedliche Arten von MAGs generieren kann. Dabei sollte das GPT unter Berücksichtigung psychologischer, methodischer und sprachlicher Aspekte inhaltlich fundierte und angemessene Leitfäden entwickeln. Die Entwicklung des Custom GPT erfolgte schrittweise. Zunächst wurde über die ChatGPT-Oberfläche ein neues Custom GPT angelegt. Anschliessend stellte ChatGPT automatisch Fragen zur Konfiguration des Modells, etwa zur Festlegung eines Titels, zur Generierung eines Profilbilds sowie zur Sammlung wichtiger Informationen über den Zweck und die Funktionsweise des GPT. Die eigentliche Konfiguration erfolgte iterativ. Hierbei wurde das Modell gefragt, welche Aspekte bei der Konfiguration eines Custom GPT beachtet werden müssen. Auf Basis dieser Hinweise wurden die Konfigurationsbereiche systematisch abgearbeitet. Dazu gehörten:

- Zweck des Custom GPT
- Zielgruppe
- Eingrenzung der Gesprächsarten
- Relevante Daten, auf die sich das Custom GPT stützen soll

Zur inhaltlichen Anreicherung des GPT wurden mehrere relevante Dokumente hochgeladen, darunter:

- Exposé mit Fragestellung, Hypothese und Zielsetzung der Arbeit
- Dokumente mit theoretischem Grundlagewissen zu den einzelnen Gesprächsarten
- Dokumente mit psychologischem Grundlagewissen der Personalentwicklung
- Leitfaden zur gendergerechten Sprache und Grundlagen zur Fragenkonstruktion

Diese Dateien wurden in bearbeitbaren Formaten (.docx) bereitgestellt, um eine optimale Verarbeitung zu ermöglichen. Dabei wurde auf die Einhaltung von Dateigrösse (max. 512 MB) und Token-Limit (8192 Tokens pro Anfrage) geachtet. Ausserdem wurden personenbezogene oder vertrauliche Daten entfernt, um den Anforderungen des Datenschutzes gerecht zu werden.

Ein wichtiger Bestandteil des Entwicklungsprozesses war das wiederholte Testen des GPTs mit realistischen Eingaben. Auf Grundlage der Rückmeldungen wurden Anfragen präzisiert und Inhalte nachjustiert. Dieses iterative Feintuning stellte sicher, dass die

Leitfäden sowohl inhaltlich korrekt als auch sprachlich angemessen waren. Im Entwicklungsverlauf traten jedoch mehrere Herausforderungen auf, die den Prozess beeinflussten. Eine zentrale Schwierigkeit bestand darin, das Custom GPT wiederholt explizit anzuweisen, alle Antworten konsequent auf Deutsch zu formulieren. Trotz entsprechender Voreinstellungen reagierte das Modell gelegentlich mit englischen Abschnitten, was eine zusätzliche manuelle Nachbearbeitung erforderte. Darüber hinaus stellte die technisch bedingte Maximalbegrenzung der Dateigrösse beim Upload von Hintergrundmaterialien eine Einschränkung dar, da umfassende theoretische Dokumente teilweise aufgeteilt werden mussten. Diese Rahmenbedingungen erforderten eine präzise Strukturierung der Eingaben und teilweise kreative Lösungen im Umgang mit den Systemgrenzen. Die Erstellung des Custom GPT war somit nicht nur ein technischer, sondern auch ein konzeptioneller und gestalterischer Prozess. Durch eine strukturierte Planung, iterative Verbesserung und gezielte Anpassung konnte ein leistungsfähiges, zielgerichtetes GPT-Modell entwickelt werden, das als Grundlage für die Erstellung KI-generierter Leitfäden dient.

4.1 Erstellung der KI-generierten Leitfäden

Die Erstellung der KI-generierten Leitfäden für MAGs erfolgte mithilfe des zuvor entwickelten Custom GPT. Um eine klare und einheitliche Struktur der Leitfäden sicherzustellen, wurde dem Custom GPT zur Erstellung der Leitfäden eine detaillierte Anleitung gegeben. Ziel war es, dem GPT möglichst präzise Vorgaben zu geben, sodass über alle generierten Leitfäden hinweg der Aufbau sehr ähnlich aussieht. Die erstellten Anleitungen sind im Anhang A zu finden.

Durch die vorab hochgeladenen theoretischen Grundlagen zu verschiedenen Gesprächsarten konnte das GPT auf eine umfangreiche Wissensbasis zurückgreifen. In Kombination mit dem hochgeladenen psychologischen Wissen war das Modell in der Lage, massgeschneiderte, kontextualisierte Leitfäden zu erstellen, die auf die jeweilige Gesprächsform zugeschnitten waren. Diese wurden anschliessend auf Verständlichkeit, Praxistauglichkeit und inhaltliche Stimmigkeit geprüft. Gezieltes und wiederholtes Prompting zur Anpassung der Leitfäden wurde zur Vergleichbarkeit der Leitfäden nicht eingesetzt. Sprachmodelle wie ChatGPT verbessern ihre Ausgabe durch gezielte und

mehrfache Prompteingaben schrittweise, was zu inhaltlichen Verfeinerungen aber auch zu potenziellen Verschiebungen führen kann. Da ein solches Vorgehen die Leitfäden stark individualisiert und inhaltlich verändert hätte, wurde bewusst darauf verzichtet. Ziel war es, eine standardisierte Grundlage zu schaffen, bei der die Qualität der generierten Inhalte möglichst direkt auf der bereitgestellten Wissensbasis beruht. Die KI-generierten Leitfäden sind im Anhang C zu finden.

5 Konstruktion der psychologischen Leitfäden

Zur Ergänzung der KI-generierten Leitfäden wurden Leitfäden durch psychologisch ausgebildete Fachpersonen erstellt. Insgesamt konnten acht Fachpersonen gewonnen werden, welche auf Basis ihres psychologischen Fachwissens Leitfäden für unterschiedliche MAGs entwickelten.

Zur Sicherstellung vergleichbarer Ausgangsbedingungen erhielten die beteiligten psychologisch ausgebildeten Fachpersonen die gleichen strukturierten Anweisungen, wodurch die von ihnen erstellten und die KI-generierten Leitfäden auf einer gemeinsamen methodischen Grundlage basieren. Dies ermöglicht einen systematischen Vergleich beider Herangehensweisen im weiteren Verlauf der Untersuchung. Sowohl die erstellten Anleitungen als auch die von psychologisch ausgebildeten Fachpersonen erstellten Leitfäden sind im Anhang B und D zu finden.

6 Methodik

Die vorliegende Hauptstudie folgt einem quantitativen Forschungsansatz. Es handelt sich um eine empirische Untersuchung, die als systematische Replikationsstudie konzipiert ist (Döring & Bortz, 2016). Das Design orientiert sich an bestehenden Originalstudien in Scheuer (2020) zur Akzeptanz von KI, wurde jedoch gezielt an die Forschungsfrage und die dazugehörige Hypothese angepasst. Zusätzlich zur quantitativen Hauptstudie wurde in Anlehnung an Friese (2012) und Pentzold et al. (2018) eine ergänzende qualitative Erhebung durchgeführt, um vertiefte Einblicke in individuelle Wahrnehmungen und Einschätzungen der Teilnehmenden zu erhalten.

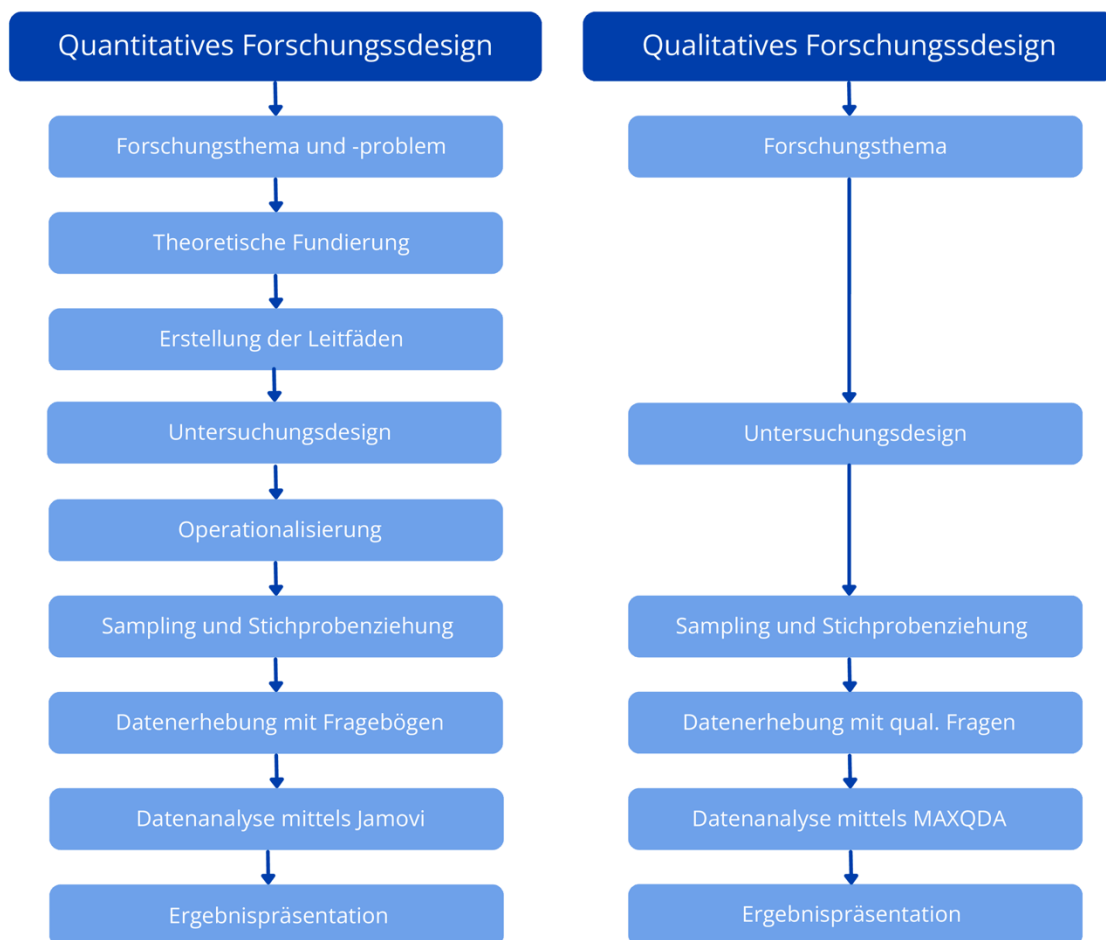


Abbildung 4. Untersuchungsdesign der vorliegenden Arbeit (in Anlehnung an Döring und Bortz, 2016)

In den darauffolgenden Kapiteln wird anhand der *Abbildung* die methodische Vorgehensweise der vorliegenden Arbeit detailliert beschrieben.

6.1 Untersuchungsdesign

Ziel dieser Arbeit ist es, den Erstellungsprozess von Gesprächsleitfäden für MAGs zu optimieren. Im Zentrum stand die Untersuchung der Akzeptanz von Leitfäden, die entweder mithilfe KI (ChatGPT) oder durch psychologisch ausgebildete Fachpersonen entwickelt wurden. Es sollte überprüft werden, inwieweit die Akzeptanz der beiden Leitfadenarten vergleichbar ist.

Die Hauptstudie folgte einem quantitativen, hypothesenprüfenden Forschungsansatz und war als systematische Replikationsstudie mit vergleichendem Design angelegt (Moosbrugger & Kelava, 2020). Das Untersuchungsdesign war quasi-experimentell aufgebaut und basierte auf einer zwischen-subjektlichen Gruppenzuordnung: Die Teilnehmenden wurden randomisiert entweder einem KI-generierten oder einem fachlich entwickelten Leitfaden zugewiesen und beurteilten diesen anschliessend anhand eines standardisierten Fragebogens. Die Erhebung erfolgte einmalig und webbasiert, womit ein Querschnittsdesign vorlag (Moosbrugger & Kelava, 2020). Ergänzend zur quantitativen Hauptauswertung wurden im Fragebogen drei offene Fragen integriert, die sich spezifisch auf das FLOWIT-Tool bezogen. Ziel dieser qualitativen Ergänzung war es, der Praxispartnerin die Möglichkeit zu geben, offenes und individuelles Feedback der Teilnehmenden zu erhalten und praxisrelevante Hinweise zur Weiterentwicklung des Tools zu gewinnen. Diese methodische Struktur ermöglichte es, sowohl Unterschiede in der Wahrnehmung der beiden Leitfadenarten quantitativ zu analysieren als auch vertiefte qualitative Einsichten zu gewinnen. Alle detaillierten methodischen Aspekte der Studie werden in den folgenden Kapiteln erläutert.

6.2 Operationalisierung

Zur Beantwortung der Forschungsfrage ist es erforderlich, die zentralen Konzepte der Studie in messbare Variablen zu überführen (Moosbrugger & Kelava, 2020). Die Operationalisierung legt fest, wie die theoretischen Konstrukte im Rahmen der Erhebung empirisch erfasst wurden (Döring & Bortz, 2016). Die unabhängige Variable der Studie stellte die Art der Erstellung des Gesprächsleitfadens dar und umfasste zwei Ausprägungen:

- KI-generierter Leitfaden (ChatGPT)
- von psychologisch ausgebildeten Fachpersonen erstellter Leitfaden

Die abhängige Variable bildete die Akzeptanz des jeweiligen Leitfadens und umfasste ebenfalls zwei Ausprägungen:

- Akzeptanz der ChatGPT-Gruppe
- Akzeptanz der Psychologie-Gruppe

Zur Erfassung dieser Akzeptanz wurde ein standardisierter Fragebogen verwendet, der verschiedene theoretisch fundierte Konstrukte umfasste. Die Operationalisierung erfolgte über mehrere Items, die auf einer fünfstufigen Likert-Skala beantwortet wurden. Die genaue Zusammensetzung der Items sowie ihre theoretische Zuordnung sind in Kapitel 6.3.2 dargestellt.

6.3 Konstruktion des Fragebogens

Um eine umfassende Erfassung der Akzeptanz von KI-generierten Leitfäden sicherzustellen, basierte die Konstruktion des Fragebogens auf den relevantesten Konstrukten der Technologieakzeptanzmodelle. Aus dem TAM wurden folgende Konstrukte integriert:

- Wahrgenommene Nützlichkeit (PU)
- Wahrgenommene Benutzerfreundlichkeit (PEOU)
- Einstellung zur Nutzung (ATU)
- Verhaltensabsicht zur Nutzung (BI)

Die Konstrukte stellen zentrale Prädiktoren für die Technologieakzeptanz dar und beeinflussen direkt, ob der Leitfaden als hilfreich und einfach nutzbar empfunden wird (Davis, 1987). Zudem ist es von Relevanz, zu erfassen, wie die Einstellung gegenüber

FLOWIT ist und ob eine Verhaltensabsicht zur Nutzung besteht. Aus der UTAUT wurden folgende Konstrukte genommen:

- Leistungserwartung (PE)
- Aufwandserwartung (EE)

Die benutzten Konstrukte der UTAUT ergänzen die Konstrukte der TAM, indem sie zusätzlich die subjektiven Erwartungen an die Wirksamkeit und den Aufwand der Nutzung erfassen (Venkatesh et al., 2003). Das Konstrukt des sozialen Einflusses (SI) wurde zur Konstruktion des Fragebogens nicht verwendet, da es nicht zur Beantwortung der Fragestellung beiträgt. Auch das Konstrukt der unterstützenden Bedingungen (FC) war in der vorliegenden Arbeit nicht von Interesse.

Die CASA-Theorie ergänzt die technologischen Akzeptanzmodelle um folgende Konstrukte:

- Wahrgenommene Menschlichkeit (PHL)
- Emotionale Beteiligung (EE)
- Wechselseitigkeit und Soziale Präsenz (RSP)

Diese Konstrukte sind entscheidend, da Menschen Medieninhalte sozial interpretieren und dies die Akzeptanz von KI-generierten Leitfäden beeinflussen kann (Epley, Waytz & Cacioppo, 2007; Luger & Sellen, 2016). Das Konstrukt Vertrauen in den Leitfaden (TR) wurde nicht zur Konstruktion des Fragebogens verwendet, da dieses Konstrukt in den Konstrukten Wahrgenommene Nützlichkeit (PU) der TAM und Leistungserwartung (PE) der UTAUT integriert ist.

6.3.1 Qualitativer Pretest

Im Rahmen der Fragebogenentwicklung wurde, um potenzielle Probleme frühzeitig zu identifizieren und anhand derer Anpassung vornehmen zu können, ein qualitativer Pretest durchgeführt. Basierend auf der Empfehlung von Döring und Bortz (2016) nahmen am qualitativen Pretest drei Fachexperten teil, die über fundierte methodische und inhaltliche Kenntnisse im Forschungsbereich verfügen. Diese Vorgehensweise ermöglichte eine detaillierte Reflexion über die Verständlichkeit, Struktur und Relevanz der Fragen. Zunächst wurde der Fragebogen den Fachexperten zum eigenständigen Durchlesen vorgelegt. Anschliessend wurden qualitative Interviews durchgeführt, in denen die

Fachexperten ihre Eindrücke schilderten. Dabei äusserten sie sich zu möglichen Unklarheiten oder Schwierigkeiten beim Verständnis einzelner Items und gaben gezielte Verbesserungsvorschläge. Zusätzlich kam eine kognitive Pretesttechnik zum Einsatz, um die Verständlichkeit und Eindeutigkeit der Fragen weiter zu überprüfen. Beim sogenannte Probing wurden gezielte Nachfragen zu den gegebenen Antworten gestellt, um herauszufinden, ob die Fragen wie beabsichtigt interpretiert wurden (Döring & Bortz, 2016). Basierend auf den Ergebnissen wurden gezielte Anpassungen des Fragebogens vorgenommen.

6.3.2 Finaler Fragebogen

Auf Basis des qualitativen Pretests wurde der Fragebogen angepasst und finalisiert. Einige Items wurden redundant wahrgenommen, andere enthielten nicht gendergerechte Formulierungen oder nicht eindeutige Begriffe. Diese Rückmeldungen führten zu sprachlichen sowie inhaltlichen Anpassungen. Um eine vergleichbare Aussagekraft der einzelnen Konstrukte sicherzustellen, wurde eine möglichst gleichmässige Anzahl an Items pro Konstrukt gewählt. Dadurch sollten Verzerrungen vermieden und eine höhere Reliabilität der Skalen erreicht werden. Jedes Konstrukt wurde durch jeweils zwei bis drei Items abgebildet, orientiert an der Struktur der Originalstudien. Jedes Item konnte durch eine fünfstufige Likert-Skala bewertet werden, welche direkt aus den Originalstudien übernommen wurde. Diese reichte von *1= Trifft überhaupt nicht zu* bis *5= Trifft voll und ganz zu*. Zusätzlich wurden am Ende des Fragebogens offene Fragen integriert. Die offenen Fragen dienten dazu, ein tieferes Verständnis über die Wahrnehmung, Nutzungserfahrungen und potenziellen Herausforderungen von FLOWIT zu gewinnen. Der finale Fragebogen sah wie folgt aus:

Technologieakzeptanzmodell (TAM)

Wahrgenommene Nützlichkeit (PU)

- Der Leitfaden ist nützlich für meine Gesprächsvorbereitung.
- Der Leitfaden hilft mir, mein Gespräch strukturierter zu führen.
- Der Leitfaden hilft mir, meine Vorbereitung effizienter zu gestalten.

Wahrgenommene Benutzerfreundlichkeit (PEOU)

- Der Leitfaden ist einfach zu verstehen und anzuwenden.
- Die Nutzung des Leitfadens erfordert wenig Aufwand.
- Die Struktur des Leitfadens wirkt verständlich und logisch aufgebaut.

Einstellung zur Nutzung (ATU)

- Ich halte den Leitfaden für eine sinnvolle Unterstützung bei Gesprächen.
- Die Nutzung des Leitfadens verbessert meine Vorbereitung auf Gespräche.
- Ich stehe dem Einsatz des Leitfadens positiv gegenüber.

Verhaltensabsicht zur Nutzung (BI)

- Ich werde den Leitfaden für zukünftige Gespräche nutzen.
- Ich würde den Leitfaden weiterempfehlen.

Unified Theory of Acceptance and Use of Technology (UTAUT)

Leistungserwartung (PE)

- Ich habe den Eindruck, dass der Leitfaden Gespräche erfolgreicher machen kann.
- Durch die Nutzung des Leitfadens fühle ich mich besser vorbereitet.
- Ich glaube, dass der Leitfaden meine Effizienz in der Gesprächsführung steigert.

Aufwandserwartung (EE)

- Der Leitfaden ist leicht in meinen Vorbereitungsprozess zu integrieren.
- Die Nutzung des Leitfadens erfordert keinen hohen Aufwand.
- Ich finde es unkompliziert, den Leitfaden für meine Gesprächsvorbereitung zu nutzen.

Computers Are Social Actors (CASA) Theorie

Wahrgenommene Menschlichkeit (PHL)

- Der Leitfaden vermittelt mir ein personalisiertes und anpassbares Nutzungserlebnis.
- Ich habe das Gefühl, dass der Leitfaden individuell auf meine Bedürfnisse eingeht.
- Der Leitfaden fühlt sich persönlich und nicht generisch an.

Emotionale Beteiligung (EE)

- Die Nutzung des Leitfadens fühlt sich angenehm und hilfreich an.
- Die Nutzung des Leitfadens gibt mir ein Gefühl von Sicherheit in der Gesprächsführung.
- Ich empfinde den Leitfaden als eine wertvolle emotionale Unterstützung bei meiner Vorbereitung.

Wechselseitigkeit und Soziale Präsenz (RSP)

- Ich habe das Gefühl, dass der Leitfaden mich aktiv durch meine Gesprächsvorbereitung begleitet.
- Der Leitfaden schafft eine Art Dialog, indem er mich gezielt in meiner Vorbereitung unterstützt.

Offene Fragen

- Welchen allgemeinen Eindruck hast du von FLOWIT gewonnen?
- Welche Aspekte der Nutzung von FLOWIT fandest du besonders positiv?
- Welche Aspekte der Nutzung von FLOWIT siehst du als verbesserungswürdig?

6.3.3 Finale Umfragestruktur

Der finalisierte Fragebogen wurde durch Instruktionstexte ergänzt, die den Teilnehmenden den Ablauf der Umfrage erläuterten und sie durch den Bearbeitungsprozess begleiteten. Ziel war es, eine klare Orientierung zu bieten, den Zweck der Erhebung transparent zu machen und ein einheitliches Verständnis der Aufgabenstellung sicherzustellen. Die Instruktionen enthielten sowohl praktische Hinweise zur Durchführung als auch Informationen zu Dauer, Datenschutz und Freiwilligkeit. Es wurde deutlich gemacht, dass die Teilnahme anonym erfolgt, keine Rückschlüsse auf die Identität möglich sind und die erhobenen Daten ausschliesslich zu wissenschaftlichen

Zwecken verwendet werden. Vor dem technischen Pretest wurden die Instruktionen gezielt auf Verständlichkeit und sprachliche Klarheit geprüft. Diese Überprüfung erfolgte durch Testpersonen, die stellvertretend für potenzielle Teilnehmende den Ablauf durchliefen. Basierend auf ihrem Feedback wurden einzelne Formulierungen überarbeitet, sprachlich vereinfacht und die Struktur der Instruktionstexte optimiert. Die überarbeiteten Instruktionen wurden anschliessend in die FLOWIT-Plattform integriert und den Teilnehmenden sowohl vor Beginn der Umfrage als auch an relevanten Stellen während des Ausfüllprozesses angezeigt. So wurde sichergestellt, dass sie jederzeit über den Ablauf und die Bedeutung einzelner Abschnitte informiert waren.

Im Rahmen der Instruktionen wurde bewusst auf eine explizite Kennzeichnung verzichtet, dass es sich um KI-generierte Leitfäden handelt. Diese Entscheidung basiert auf Erkenntnissen aus zahlreichen Literaturen, welche besagen, dass Menschen weniger skeptisch gegenüber KI sind, wenn diese sich menschlich verhält oder kaum von menschlichem Verhalten zu unterscheiden ist (Epley et al., 2007; Nass & Moon, 2000). Auch Studien zeigen, dass Menschen dazu tendieren, KI-generierte Inhalte unkritischer zu bewerten, wenn diese empathisch wirken oder vertraute Kommunikationsmuster nutzen (Luger & Sellen, 2016). Diese sogenannte *Mindlessness* beschreibt die unreflektierte Akzeptanz menschenähnlicher Technologien, insbesondere wenn deren künstlicher Ursprung nicht offensichtlich ist (Luger & Sellen, 2016). Diese Phänomene lassen sich unter dem Begriff *Anthropomorphismus* einordnen (Scheuer, 2020). Dabei wird nicht nur die äussere Form eines Menschen übertragen, sondern auch Verhalten oder menschliche Intelligenz (Scheuer, 2020). Besonders relevant sind die Konzepte der menschlichen Verkörperung und des menschlichen Kognitionsniveaus, welche im Kontext KI-gestützter Systeme gezielt eingesetzt werden können, um Akzeptanz zu fördern (Scheuer, 2020). Um eine möglichst neutrale Ausgangsposition für alle Teilnehmenden zu schaffen und deren Urteilsbildung nicht zu beeinflussen, wurde daher auf den Hinweis zum KI-Ursprung bewusst verzichtet. Ziel war es, Verzerrungen (Bias) im Antwortverhalten zu minimieren und die Wirkung der Leitfäden unabhängig von ihrer Herkunft zu erfassen.

Darüber hinaus wurde nicht die vollständige Anzahl der ursprünglich erstellten Leitfadenfragen in die Umfrage übernommen. Zum einen zeigte sich in der praktischen Anwendung von FLOWIT, dass eine reduzierte Anzahl gezielter und offener Fragen häufig zu

einem höheren Mehrwert führt (FLOWIT, 2024). Zum anderen weisen auch beschriebene Studien von Moosbrugger und Kelava (2020) darauf hin, dass kürzere Leitfäden nicht nur effizienter, sondern auch anwenderfreundlicher wahrgenommen werden. Da die Umfrage bereits eine gewisse Länge aufwies, wurde zudem im Sinne der Teilnehmenden eine bewusste Reduktion vorgenommen, um Ermüdungstendenzen zu vermeiden. Auch dieses Vorgehen stützt sich auf wissenschaftliche Empfehlungen von Döring und Bortz (2016). Die finale Umfragestruktur, einschliesslich der verwendeten Fragen für beide Leitfadenvarianten, ist im Anhang E und F dokumentiert.

6.3.4 Technischer Pretest

Nach der finalen Konstruktion des Fragebogens wurde dieser auf FLOWIT implementiert. Im Rahmen des technischen Pretests wurde überprüft, ob die technische Umsetzung fehlerfrei funktioniert und alle Fragen wie vorgesehen dargestellt werden. Dabei wurden insbesondere die Benutzerführung, die Skalierung der Antwortmöglichkeiten sowie die Speicherung der Eingaben getestet (Döring & Bortz, 2016).

6.4 Stichprobengrösse und Sampling

Die Festlegung der Stichprobengrösse erfolgte zur Sicherstellung einer ausreichend statistischen Teststärke auf Grundlage einer A-priori-Power-Analyse (Moosbrugger & Kelava, 2020). Da die Studie die Akzeptanz zweier unterschiedlicher Leitfäden untersucht, wurde ein zwischen-subjektliches Design mit zwei unabhängigen Gruppen gewählt. Für die Berechnung der Stichprobengrösse wurden folgende Parameter herangezogen:

- Signifikanzniveau (α): 0.05
- Teststärke ($1-\beta$): 0.95
- Effektgrösse (Cohens d): 0.5 (mittelgrosser Effekt)

Die Power-Analyse ergab eine Mindeststichprobengrösse von 210 Teilnehmenden mit 105 Personen pro Gruppe. Die Rekrutierung der Teilnehmenden erfolgte mittels einer nicht-probabilistischen Stichprobenziehung, konkret in Form eines Convenience Samplings (Gelegenheitsstichprobe) (Döring & Bortz, 2016). Dabei wurden über Online-Plattformen und soziale Netzwerke wie LinkedIn, Instagram, WhatsApp usw. Teilnehmende für die

Umfrage angesprochen. Diese Form der Stichprobenziehung zeichnet sich durch eine Auswahl von Teilnehmenden anhand deren Verfügbarkeit aus und ist mit einem vergleichsweise geringen Aufwand verbunden (Döring & Bortz, 2016). Als Einschlusskriterium galt, dass Teilnehmende entweder bereits Erfahrung mit MAGs gesammelt hatten, sei es als durchführende oder empfangende Person. Auf diese Weise sollte sichergestellt werden, dass die befragten Personen über relevante praktische Erfahrung zur fundierten Bewertung der vorgelegten Gesprächsleitfäden verfügten. Die Einhaltung dieses Kriteriums wurde im Rahmen des Fragebogens durch eine Screening-Frage überprüft: *Konntest du bereits Erfahrungen bezüglich Mitarbeitendengespräche machen? Beschreibung: Hattest du bereits selbst ein Mitarbeitendengespräch oder hast du bereits Erfahrung mit der Durchführung von Mitarbeitendengesprächen? (Probezeitgespräch, Austrittsgespräch, jährliches Mitarbeitendengespräch o.Ä.) und Antwortoptionen: Ja oder Nein.* Nur Teilnehmende, die diese Frage mit *Ja* beantworteten, wurden in die Auswertung einbezogen.

6.5 Datenerhebung

Die Datenerhebung erfolgte innerhalb eines Zeitraums von drei Wochen. Da im Rahmen der Studie acht unterschiedliche Leitfäden entwickelt wurden, ergaben sich daraus auch acht separate Umfragen. Um eine möglichst gleichmässige Verteilung der Teilnehmenden auf die verschiedenen Umfragen zu erzielen, wurde ein Sammel-Link über die Plattform nimblelinks.com erstellt. Dieser Sammel-Link beinhaltete die einzelnen URLs zu den acht Umfragen und ordnete die Teilnehmenden zufällig zu. Dadurch konnte eine annähernd ausgewogene Verteilung der Teilnehmenden auf die verschiedenen Leitfäden erzielt werden.

6.6 Datenauswertung

Nach Abschluss der Datenerhebung wurden die Rohdaten bereinigt und aufbereitet. In Fällen mit fehlenden Werten wurde eine Mittelwertimputation vorgenommen, anschliessend Extremwerte und potenzielle Ausreisser identifiziert. Vor Beginn der Analyse wurden die Datensätze in ein einheitliches Format (Excel) überführt und auf Konsistenz und Vollständigkeit geprüft, um mögliche Formatierungs- oder Eingabefehler auszuschliessen. Die Datenanalyse erfolgte mittels des Statistiktools Jamovi.

6.6.1 Deskriptive Statistik

In der vorliegenden Studie wurde auf die Erhebung klassischer demografischer Merkmale wie Alter, Geschlecht oder Berufserfahrung bewusst verzichtet. Dieser methodische Entscheid basiert auf mehreren Gründen. Zum einen lag der Fokus der Untersuchung auf der inhaltlichen Bewertung der Leitfäden und nicht auf personenspezifischen Unterschieden. Da demografische Daten weder hypothesenrelevant waren noch Subgruppenvergleiche vorgesehen wurden, erschien deren Erhebung entbehrlich (Döring & Bortz, 2016). Zum anderen sollte die Belastung der Teilnehmenden gering gehalten und die Antwortmotivation durch einen schlanken Fragebogen gefördert werden (Porst, 2014). Hinzu kommt, dass alle Teilnehmenden mit dem Format von MAGs vertraut sind, was eine weitgehend homogene Stichprobe hinsichtlich der Vorerfahrung im Kontext von MAGs nahelegt (Döring & Bortz, 2016). Um Rückschlüsse auf Einzelpersonen zu vermeiden, wurde aus datenschutzethischen Überlegungen auf sensible Angaben verzichtet (Porst, 2014).

6.6.2 Prüfung der Daten

Zur Überprüfung der Voraussetzungen für inferenzstatistische Analysen wurde ein Shapiro-Wilk-Test durchgeführt. Dieser fiel mit $p < .001$ signifikant aus und deutete auf eine Verletzung der Normalverteilungsannahme hin. Daher wurde für die weiteren Analysen der nichtparametrische Mann-Whitney-U-Test herangezogen. Eine Prüfung der Varianzhomogenität entfiel, da diese Voraussetzung bei nichtparametrischen Tests nicht erforderlich ist.

6.6.3 Inferenzstatistische Analysen

Zur Überprüfung der Hypothese wurde ein geeigneter Test auf Basis der Prüfergebnisse ausgewählt. Aufgrund einer signifikanten Verletzung der Normalverteilungsannahme im Shapiro-Wilk-Test ($p < .001$) wurde zur Prüfung der Hypothesen der nichtparametrische Mann-Whitney-U-Test eingesetzt. Die getesteten Hypothesen lauteten:

H_0 : Die durchschnittliche Akzeptanz der Leitfäden unterscheidet sich nicht zwischen der ChatGPT-Gruppe und der Psychologie-Gruppe.

H_1 : Die durchschnittliche Akzeptanz der Leitfäden unterscheidet sich zwischen der ChatGPT-Gruppe und der Psychologie-Gruppe.

6.6.4 Qualitative Zusatzanalysen

Ergänzend zur quantitativen Hauptauswertung wurden im Fragebogen drei offene Fragen integriert, um dem Praxispartner praxisrelevante Hinweise zur Weiterentwicklung des Tools zu ermöglichen. Die offenen Fragen lauteten wie folgt:

- Welchen allgemeinen Eindruck hast du von FLOWIT gewonnen?
- Welche Aspekte der Nutzung von FLOWIT fandest du besonders positiv?
- Welche Aspekte der Nutzung von FLOWIT siehst du als verbesserungswürdig?

Zur Auswertung der offenen Antworten wurde die NCT-Methode (Noticing-Collecting-Thinking) nach Friese (2012) herangezogen. Diese Methodik ermöglicht eine strukturierte aber zugleich flexible Annäherung an qualitative Daten (Pentzold, Bischof & Heise, 2018). Sie orientiert sich an den Prinzipien der Grounded Theory und begreift die Analyse als einen erkundenden Prozess, bei dem das Datenmaterial systematisch erschlossen wird (Friese, 2012; Pentzold et al., 2018). In der ersten Phase *Noticing* wurden das Datenmaterial gelesen und Memos festgehalten. Im Anschluss daran erfolgte in der Phase des *Collecting* die Entwicklung von Codes, welche auf inhaltlicher Nähe basierten und iterativ zu Kategorien zusammengefasst wurden. Die *Tabelle* zeigt das erarbeitete Kategoriensystem.

Tabelle 1

Kategoriensystem mit Haupt- und Subkategorien (eigene Darstellung)

Hauptkategorie	Subkategorie	Beispielhafte Aussagen
Gesamteindruck von FLOWIT		
	Positiver Gesamteindruck	<i>Modern, hilfreich</i>
	Neutraler Eindruck	<i>Kein starker Eindruck</i>
	Kritischer Gesamteindruck	<i>Oberflächlich, nicht hilfreich</i>
Stärken der Nutzung		
	Bedienung und Verständlichkeit	<i>Intuitiv, selbsterklärend</i>
	Design	<i>Visuell gut, ästhetisch</i>
	Digitale Strukturierung und Unterstützung	<i>Hilft, Gedanken zu ordnen</i>
	Interaktivität	<i>Wie ein Gespräch</i>
Optimierungspotenzial		
	Individualisierbarkeit	<i>Nicht personalisierbar</i>
	Technische Einschränkungen	<i>Skala unklar, nervige Registrierung</i>

Abschliessend wurden in der Phase des *Thinking* die codierten Inhalte miteinander in Beziehung gesetzt und hinsichtlich ihrer Bedeutung, Relevanz und Wiederkehr interpretiert (Friese, 2012). Zur Unterstützung des Analyseprozesses wurde die qualitative Datenanalysesoftware MAXQDA eingesetzt, die strukturiertes Codieren, die Organisation des Datenmaterials sowie die Visualisierung von Zusammenhängen erleichtert.

Die qualitative Auswertung diente ausschliesslich explorativen Zwecken und wurde nicht inferenzstatistisch interpretiert. Sie ergänzt die quantitativen Ergebnisse um subjektive Einschätzungen und unterstützt den Praxispartner bei der zielgerichteten Optimierung des Tools auf Grundlage realer Rückmeldungen aus der Zielgruppe.

7 Ergebnisse

In diesem Kapitel werden die Ergebnisse der quantitativen Datenanalyse systematisch dargestellt. Die Auswertung der Daten erfolgte anhand der im Kapitel 6 beschriebenen Methoden.

7.1 Gesamthafte Ergebnisse

Im Folgenden werden die aggregierten Ergebnisse (N= 165) über alle Leitfäden hinweg präsentiert. Dabei erfolgt ein Vergleich der durchschnittlichen Einschätzungen zwischen der ChatGPT-Gruppe und der Gruppe der psychologisch ausgebildeten Fachpersonen.

Aufgrund einer signifikanten Verletzung der Normalverteilungsannahme im Shapiro-Wilk-Test ($p < .001$) wurde ein nichtparametrischer Mann-Whitney-U-Test durchgeführt. Dieser hat ergeben, dass die durchschnittliche Gesamteinschätzung der ChatGPT-Gruppe ($M = 3.82$, $SD = 0.99$, 95%-KI [3.78, 3.86], $N = 94$) signifikant höher war als die der Psychologie-Gruppe ($M = 3.68$, $SD = 1.03$, 95%-KI [3.63, 3.72], $N = 71$), ($U = 1'910'000$, $p < .001$, $r = .08$). Dies entspricht einem kleinen Effekt. Auch die nicht überlappenden 95%-Konfidenzintervalle deutet darauf hin, dass ein signifikanter Unterschied zwischen den Gruppen besteht.

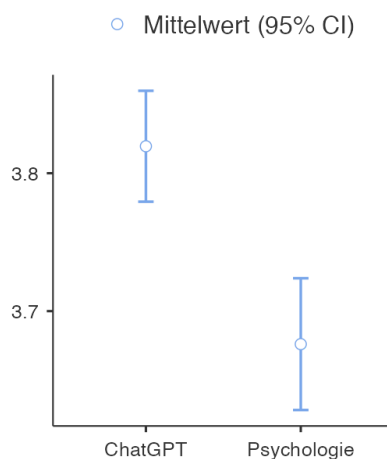


Abbildung 5. 95%- Konfidenzintervall der gesamten Ergebnisse (eigene Darstellung)

Die Reliabilitätsanalyse ergab ein Cronbachs Alpha von 0.0232, was auf eine sehr geringe interne Konsistenz hindeutet. Dieses Ergebnis ist jedoch im Kontext der konzeptionellen Struktur zu interpretieren. Die Messung basierte auf drei theoretischen Modellen und umfasste verschiedene Akzeptanzkonstrukte. Aufgrund dieser inhaltlichen Heterogenität ist ein niedriges Cronbachs Alpha erwartbar und stellt kein Indiz für eine mangelhafte Messqualität dar, sondern reflektiert die intendierte Vielfalt der erfassten Aspekte. Ergänzend wurde eine Post-Hoc-Analyse zur Bestimmung der Teststärke ($1-\beta$) durchgeführt. Die erreichte Teststärke lag bei 16 %, womit die Wahrscheinlichkeit, einen tatsächlich vorhandenen Unterschied korrekt zu entdecken, sehr gering war.

7.2 Ergebnisse der Einzelvergleiche der Items

Im Folgenden werden die Ergebnisse der Einzelvergleiche zwischen den Items der ChatGPT und der Psychologie-Gruppe vorgestellt, um Unterschiede in der Wahrnehmung der verschiedenen Aspekte des Leitfadens zu analysieren. Die Darstellung erfolgt in der Reihenfolge der Items, wie sie im Fragebogen vorzufinden waren.

Technologieakzeptanzmodell (TAM)

Wahrgenommene Nützlichkeit (PU)

Item 1: *Der Leitfaden ist nützlich für meine Gesprächsvorbereitung.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der Nützlichkeit für die ChatGPT-Gruppe ($M = 3.87$, $SD = 0.99$) und die Psychologie-Gruppe ($M = 3.87$, $SD = 0.86$) nicht signifikant unterschiedlich war ($U = 3251$, $p = .766$, $r = .03$). Dies entspricht einem kleinen Effekt.

Item 2: *Der Leitfaden hilft mir, mein Gespräch strukturierter zu führen.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der Gesprächsstrukturierung für die ChatGPT-Gruppe ($M = 3.93$, $SD = 1.01$) und die Psychologie-Gruppe ($M = 3.89$, $SD = 0.95$) nicht signifikant unterschiedlich war ($U = 3204$, $p = .644$, $r = .04$). Dies entspricht einem kleinen Effekt.

Item 3: *Der Leitfaden hilft mir, meine Vorbereitung effizienter zu gestalten.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der Effizienzsteigerung für die ChatGPT-Gruppe ($M = 3.86$, $SD = 0.93$) und die Psychologie-Gruppe ($M = 3.85$, $SD = 0.89$) nicht signifikant unterschiedlich war ($U = 3294$, $p = .880$, $r = .13$). Dies entspricht einem kleinen Effekt.

Wahrgenommene Benutzerfreundlichkeit (PEOU)

Item 4: *Der Leitfaden ist einfach zu verstehen und anzuwenden.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der Verständlichkeit und Anwendbarkeit für die ChatGPT-Gruppe ($M = 4.07$, $SD = 0.98$) und die Psychologie-Gruppe ($M = 4.13$, $SD = 1.01$) nicht signifikant unterschiedlich war ($U = 3179$, $p = .577$, $r = .05$). Dies entspricht einem kleinen Effekt.

Item 5: *Die Nutzung des Leitfadens erfordert wenig Aufwand.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung des benötigten Aufwands für die ChatGPT-Gruppe ($M = 4.18$, $SD = 0.82$) und die Psychologie-Gruppe ($M = 4.03$, $SD = 0.88$) nicht signifikant unterschiedlich war ($U = 3026$, $p = .275$, $r = .09$). Dies entspricht einem kleinen Effekt.

Item 6: *Die Struktur des Leitfadens wirkt verständlich und logisch aufgebaut.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der Strukturverständlichkeit für die ChatGPT-Gruppe ($M = 3.90$, $SD = 0.92$) und die Psychologie-Gruppe ($M = 3.93$, $SD = 0.96$) nicht signifikant unterschiedlich war ($U = 3265$, $p = .803$, $r = .02$). Dies entspricht einem kleinen Effekt.

Einstellung zur Nutzung (ATU)

Item 7: *Ich halte den Leitfaden für eine sinnvolle Unterstützung bei Gesprächen.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der Sinnhaftigkeit für die ChatGPT-Gruppe ($M = 3.96$, $SD = 0.92$) und die Psychologie-Gruppe ($M = 3.92$, $SD = 0.92$) nicht signifikant unterschiedlich war ($U = 3225$, $p = .699$, $r = .03$). Dies entspricht einem kleinen Effekt.

Item 8: *Die Nutzung des Leitfadens verbessert meine Vorbereitung auf Gespräche.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der Verbesserung der Vorbereitung für die ChatGPT-Gruppe (M = 4.00, SD = 0.94) und die Psychologie-Gruppe (M = 3.90, SD = 0.91) nicht signifikant unterschiedlich war (U = 3109, p = .430, r = .07). Dies entspricht einem kleinen Effekt.

Item 9: *Ich stehe dem Einsatz des Leitfadens positiv gegenüber.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einstellung gegenüber dem Leitfaden für die ChatGPT-Gruppe (M = 4.11, SD = 0.94) signifikant positiver ausfiel als für die Psychologie-Gruppe (M = 3.69, SD = 1.04), (U = 2529, p = .005, r = .24). Dies entspricht einem kleinen bis mittleren Effekt.

Verhaltensabsicht zur Nutzung (BI)

Item 10: *Ich werde den Leitfaden für zukünftige Gespräche nutzen.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Absicht zur zukünftigen Nutzung für die ChatGPT-Gruppe (M = 3.61, SD = 1.07) und die Psychologie-Gruppe (M = 3.34, SD = 1.18) nicht signifikant unterschiedlich war (U = 2938, p = .174, r = .12). Dies entspricht einem kleinen Effekt.

Item 11: *Ich würde den Leitfaden weiterempfehlen.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Bereitschaft zur Weiterempfehlung für die ChatGPT-Gruppe (M = 3.72, SD = 1.03) und die Psychologie-Gruppe (M = 3.56, SD = 1.08) nicht signifikant unterschiedlich war (U = 3065, p = .352, r = .08). Dies entspricht einem kleinen Effekt.

Unified Theory of Acceptance and Use of Technology (UTAUT)

Leistungserwartung (PE)

Item 12: *Ich habe den Eindruck, dass der Leitfaden Gespräche erfolgreicher machen kann.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der Gesprächserfolgshilfe für die ChatGPT-Gruppe (M = 3.94, SD = 0.84) signifikant höher ausfiel als für die Psychologie-Gruppe (M = 3.49, SD = 0.98), (U = 2452, p = .002, r = .27). Dies entspricht einem mittleren Effekt.

Item 13: *Durch die Nutzung des Leitfadens fühle ich mich besser vorbereitet.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der Vorbereitung durch den Leitfaden für die ChatGPT-Gruppe (M = 3.87, SD = 1.01) und die Psychologie-Gruppe (M = 3.70, SD = 0.99) nicht signifikant unterschiedlich war (U = 2964, p = .199, r = .11). Dies entspricht einem kleinen Effekt.

Item 14: *Ich glaube, dass der Leitfaden meine Effizienz in der Gesprächsführung steigert.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der Effizienzsteigerung für beide Gruppen (ChatGPT: M = 4.00, SD = 0.97; Psychologie: M = 4.00, SD = 1.01) nicht signifikant unterschiedlich war (U = 3014, p = .264, r = .10). Dies entspricht einem kleinen Effekt.

Aufwandsersparung (EE)

Item 15: *Der Leitfaden ist leicht in meinen Vorbereitungsprozess zu integrieren.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der Integration in den Vorbereitungsprozess für die ChatGPT-Gruppe (M = 4.07, SD = 0.92) und die Psychologie-Gruppe (M = 4.03, SD = 0.74) nicht signifikant unterschiedlich war (U = 3104, p = .410, r = .07). Dies entspricht einem kleinen Effekt.

Item 16: *Die Nutzung des Leitfadens erfordert keinen hohen Aufwand.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung des empfundenen Aufwands für die ChatGPT-Gruppe (M = 4.18, SD = 0.83) und die Psychologie-

Gruppe (M = 3.89, SD = 1.04) nicht signifikant unterschiedlich war (U = 2854, p = .092, r = .15). Dies entspricht einem kleinen bis mittleren Effekt.

Item 17: *Ich finde es unkompliziert, den Leitfaden für meine Gesprächsvorbereitung zu nutzen.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der Unkompliziertheit der Nutzung für die ChatGPT-Gruppe (M = 4.12, SD = 0.80) und die Psychologie-Gruppe (M = 4.04, SD = 0.85) nicht signifikant unterschiedlich war (U = 3200, p = .628, r = .41). Dies entspricht einem grossen Effekt.

Computers Are Social Actors (CASA) Theorie

Wahrgenommene Menschlichkeit (PHL)

Item 18: *Der Leitfaden vermittelt mir ein personalisiertes und anpassbares Nutzungserlebnis.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung des personalisierten Nutzungserlebnisses für die ChatGPT-Gruppe (M = 3.46, SD = 0.99) und die Psychologie-Gruppe (M = 3.35, SD = 1.00) nicht signifikant unterschiedlich war (U = 3123, p = .463, r = .06). Dies entspricht einem kleinen Effekt.

Item 19: *Ich habe das Gefühl, dass der Leitfaden individuell auf meine Bedürfnisse eingeht.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der individuellen Anpassung für die ChatGPT-Gruppe (M = 3.26, SD = 1.06) und die Psychologie-Gruppe (M = 3.25, SD = 1.09) nicht signifikant unterschiedlich war (U = 3317, p = .945, r = .01). Dies entspricht einem sehr kleinen Effekt.

Item 20: *Der Leitfaden fühlt sich persönlich und nicht generisch an.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der persönlichen Wahrnehmung des Leitfadens für die ChatGPT-Gruppe (M = 3.34, SD = 1.12) und die Psychologie-Gruppe (M = 3.15, SD = 1.17) nicht signifikant unterschiedlich war (U = 3049, p = .327, r = .10). Dies entspricht einem kleinen Effekt.

Emotionale Beteiligung (EE)

Item 21: *Die Nutzung des Leitfadens fühlte sich angenehm und hilfreich an.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung des angenehmen und hilfreichen Nutzungserlebnisses für die ChatGPT-Gruppe ($M = 3.91$, $SD = 0.81$) signifikant höher ausfiel als für die Psychologie-Gruppe ($M = 3.48$, $SD = 0.98$), ($U = 2459$, $p = .002$, $r = .26$). Dies entspricht einem mittleren Effekt.

Item 22: *Die Nutzung des Leitfadens gab mir ein Gefühl von Sicherheit in der Gesprächsführung.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der gefühlten Sicherheit für die ChatGPT-Gruppe ($M = 3.73$, $SD = 0.95$) und die Psychologie-Gruppe ($M = 3.58$, $SD = 1.04$) nicht signifikant unterschiedlich war ($U = 3054$, $p = .330$, $r = .09$). Dies entspricht einem kleinen Effekt.

Item 23: *Ich empfand den Leitfaden als eine wertvolle emotionale Unterstützung bei meiner Vorbereitung.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der emotionalen Unterstützung für die ChatGPT-Gruppe ($M = 3.59$, $SD = 1.13$) und die Psychologie-Gruppe ($M = 3.34$, $SD = 1.13$) nicht signifikant unterschiedlich war ($U = 2879$, $p = .118$, $r = .14$). Dies entspricht einem kleinen Effekt.

Wechselseitigkeit und Soziale Präsenz (RSP)

Item 24: *Ich hatte das Gefühl, dass der Leitfaden mich aktiv durch meine Gesprächsvorbereitung begleitete.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der aktiven Begleitung für die ChatGPT-Gruppe ($M = 3.55$, $SD = 1.07$) und die Psychologie-Gruppe ($M = 3.52$, $SD = 1.07$) nicht signifikant unterschiedlich war ($U = 3299$, $p = .897$, $r = .01$). Dies entspricht einem sehr kleinen Effekt.

Item 25: *Der Leitfaden schuf eine Art Dialog, indem er mich gezielt in meiner Vorbereitung unterstützte.*

Der Mann-Whitney-U-Test hat ergeben, dass die durchschnittliche Einschätzung der dialogischen Unterstützung für die ChatGPT-Gruppe ($M = 3.56$, $SD = 1.00$) und die Psychologie-Gruppe ($M = 3.45$, $SD = 0.97$) nicht signifikant unterschiedlich war ($U = 3139$, $p = .495$, $r = .06$). Dies entspricht einem kleinen Effekt.

7.3 Ergebnisse der qualitativen Zusatzanalysen

Im Rahmen einer qualitativen Inhaltsanalyse wurden die offenen Rückmeldungen der Teilnehmenden systematisch ausgewertet. Ziel war es, vertiefende Einsichten in die Wahrnehmung und Akzeptanz des Tools FLOWIT zu gewinnen. Die Aussagen wurden in thematische Kategorien codiert und hinsichtlich ihrer Häufigkeit und inhaltlichen Verknüpfungen analysiert.

Wahrnehmung des Gesamteindrucks

Die qualitative Inhaltsanalyse zeigt einen überwiegend positiven Gesamteindruck von FLOWIT. Besonders häufig wurden die moderne Gestaltung, die professionelle Wirkung sowie die Integration in zeitgemässe Arbeitskontexte genannt. Ebenso hoben viele Teilnehmende die einfache und intuitive Bedienbarkeit hervor. Die klare Struktur und die selbsterklärende Nutzung wurden als besonders vorteilhaft empfunden und trugen massgeblich zur Anwendungsfreundlichkeit und Akzeptanz des Tools bei. Darüber hinaus wurde FLOWIT als hilfreiches Mittel zur Gesprächsvorbereitung und Strukturierung eingeschätzt. Es unterstützte dabei, Gedanken zu ordnen und sich gezielt auf Gespräche vorzubereiten - ein Effekt, der häufig mit der wahrgenommenen Nützlichkeit digitaler Tools assoziiert wird (vgl. TAM). Auch das interaktive Format wurde positiv bewertet, da es eine natürliche und angenehme Nutzungsumgebung schaffe. FLOWIT konnte somit sowohl funktional als auch ästhetisch überzeugen.

Die *Abbildung* zeigt die Häufigkeit und Gewichtung der genannten Aspekte in Form einer Codewolke, wobei der positive Gesamteindruck, die intuitive Bedienbarkeit sowie die strukturierende Wirkung besonders häufig genannt wurden.



Abbildung 6. Codewolke (MAXQDA, 2024)

Ergänzend dazu zeigten sich einzelne neutrale oder unentschlossene Rückmeldungen, vor allem dann, wenn die Testphase als zu kurz empfunden wurde, um eine fundierte Bewertung abzugeben. Wenige, aber wiederkehrende kritische Hinweise betrafen eine als oberflächlich wahrgenommene Gestaltung sowie die begrenzte Individualisierbarkeit. Diese Aspekte lassen sich als Hinweise auf eingeschränkte Passung zu spezifischen Nutzungskontexten interpretieren, welches sowohl eine Anforderung im UTAUT als auch im TAM darstellt.

Optimierungspotenziale

Trotz der grundsätzlich positiven Grundhaltung wurden auch kritische Aspekte benannt und Optimierungspotenziale identifiziert. Eine zugehörige *Abbildung* visualisiert die Zusammenhänge zwischen den im Rahmen der Inhaltsanalyse kodierten Kategorien sowie die thematischen Verbindungen zwischen den Rückmeldungen. Deutlich wird, dass Aspekte wie der positive Gesamteindruck, die digitale Strukturierung und die Bedienbarkeit häufig gemeinsam genannt wurden. Gleichzeitig zeigen sich Verbindungen zu kritischen Rückmeldungen wie dem Wunsch nach höherer Individualisierbarkeit und Hinweisen auf technische Optimierungspotenziale.

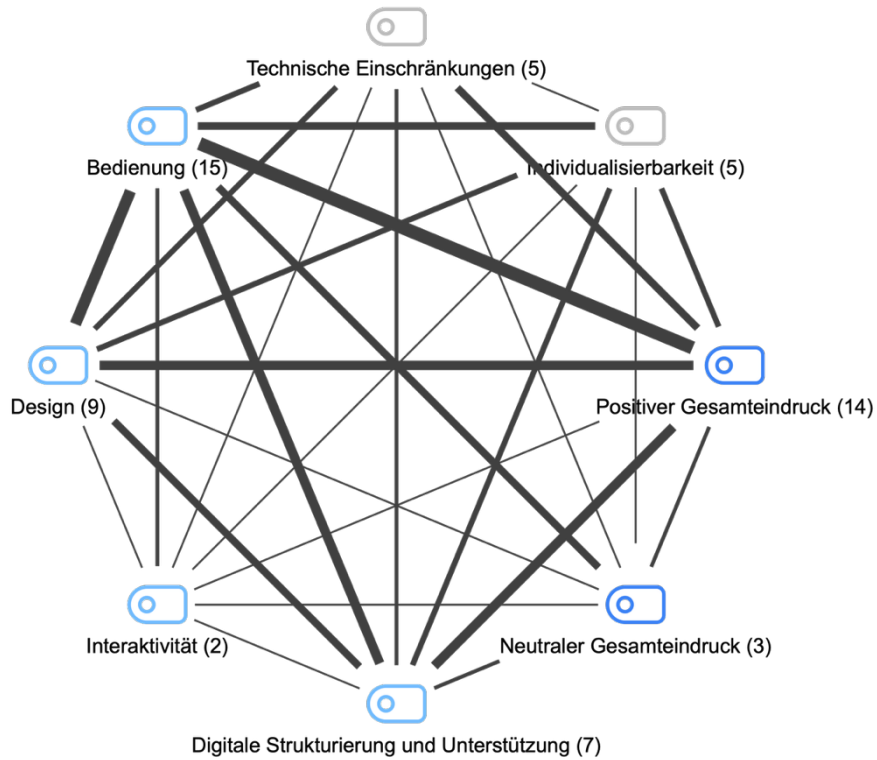


Abbildung 7. Code-Relations-Modell nach Vorkommen der Codes (MAXQDA, 2024)

Zusammenfassend lassen die Rückmeldungen erkennen, dass FLOWIT als grundsätzlich unterstützendes und positiv bewertetes Tool wahrgenommen wird. Gleichzeitig zeigen sich in den Bereichen Flexibilität, technischer Umsetzung und inhaltlicher Differenzierung konkrete Optimierungspotenziale.

8 Zusammenfassung und Interpretation der Ergebnisse

Im vorliegenden Kapitel werden die Ergebnisse der durchgeführten Studie zusammengefasst und interpretiert. Im Rahmen dieser Untersuchung wurde die folgende Hypothese untersucht:

Leitfäden, die mit Künstlicher Intelligenz (ChatGPT) erstellt wurden, weisen eine äquivalente Akzeptanz auf wie Leitfäden, die von psychologisch ausgebildeten Fachpersonen erstellt wurden.

Die dazugehörigen statistischen Hypothesen lauteten:

H_0 : Die durchschnittliche Akzeptanz der Leitfäden unterscheidet sich nicht zwischen der ChatGPT-Gruppe und der Psychologie-Gruppe.

H_1 : Die durchschnittliche Akzeptanz der Leitfäden unterscheidet sich zwischen der ChatGPT-Gruppe und der Psychologie-Gruppe.

Ein Mann-Whitney-U-Test ergab, dass die durchschnittliche Gesamteinschätzung der ChatGPT-Gruppe ($M = 3.82$, $SD = 0.99$) signifikant höher war als jene der Psychologie-Gruppe ($M = 3.68$, $SD = 1.03$), ($U = 1'910'000$, $p < .001$, $r = .08$). Aufgrund des signifikanten Testergebnisses wurde die Nullhypothese (H_0) verworfen und die Alternativhypothese (H_1) angenommen. Es bestand somit ein statistisch signifikanter Unterschied in der durchschnittlichen Akzeptanz zwischen den beiden Gruppen. Die Effektstärke ($r = .08$) deutet jedoch auf einen kleinen Effekt hin. Die Ergebnisse der Einzelvergleiche ergänzen das signifikante Gesamtergebnis, indem sie zeigen, dass bei einzelnen Items signifikante Unterschiede zwischen der ChatGPT- und der Psychologie-Gruppe festgestellt wurden. Besonders hervorzuheben sind die Items zur Gesprächserfolgförderung (Item 12) und zum angenehmen Nutzungserleben (Item 21), die jeweils zugunsten des KI-basierten Leitfadens ausfielen. Weitere Items zeigten p-Werte nahe an der Signifikanzgrenze. Bei den übrigen Items wurden keine signifikanten Unterschiede festgestellt, was das Ergebnis einer insgesamt vergleichbaren Akzeptanz zusätzlich unterstützt. Zusammengefasst bestätigen die Einzelvergleiche somit die Tendenz, dass KI-basierte Leitfäden in wesentlichen Konstrukten ähnlich bewertet wurden wie von psychologisch ausgebildeten Fachpersonen

entwickelte Leitfäden, mit punktuellen Vorteilen zugunsten der KI-gestützten Variante. Die qualitativen Zusatzanalysen ergänzen die quantitativen Befunde um vertiefende Einsichten in die Nutzungsperspektive. Eine inhaltsanalytisch orientierte Auswertung der offenen Rückmeldungen ergab drei zentrale Themenbereiche:

Positive Aspekte

Besonders häufig wurden die moderne Gestaltung, die einfache Bedienbarkeit sowie die strukturierte Gesprächsvorbereitung positiv hervorgehoben. Diese Dimensionen wurden mehrfach gemeinsam genannt und deuten auf eine hohe wahrgenommene Funktionalität hin. Diese Wahrnehmungen korrespondieren mit den quantitativen Einschätzungen zur Benutzerfreundlichkeit und Nützlichkeit des KI-Leitfadens.

Kritische Hinweise

Mehrere Teilnehmende äusserten Kritik an der begrenzten Individualisierbarkeit der Inhalte. Diese Hinweise verdeutlichen, dass trotz insgesamt positiver Bewertung insbesondere die Flexibilität der Lösung für unterschiedliche Gesprächskontexte verbesserungswürdig erscheint.

Vorschläge zur Weiterentwicklung

Konkrete Anregungen bezogen sich auf erweiterbare Templates, konfigurierbare Textbausteine oder branchenspezifische Anpassungen. Diese Vorschläge lassen sich als Indikatoren für eine grundsätzlich vorhandene Akzeptanz und Offenheit gegenüber KI-Tools interpretieren, allerdings unter der Voraussetzung, dass sie kontextsensitiv gestaltbar sind.

Die gewonnenen Erkenntnisse bieten somit eine fundierte Grundlage zur Beantwortung der zentralen Forschungsfrage: *Wie ausgeprägt ist die technologische Akzeptanz von KI-generierten Leitfäden für typische Mitarbeitendengespräche der FLOWIT?*

Die Gesamtergebnisse deuten darauf hin, dass KI-generierte Leitfäden in ihrer Akzeptanz mit jenen von psychologisch ausgebildeten Fachpersonen vergleichbar sind und punktuell sogar leicht besser bewertet werden. Die signifikant höhere Gesamteinschätzung des KI-Leitfadens (ChatGPT $M = 3.82$, Psychologie $M = 3.68$) bei gleichzeitig geringer Effektstärke ($r = .08$) spricht für eine grundsätzlich positive, aber differenzierte

Nutzerwahrnehmung. Diese Einschätzung wird durch qualitative Rückmeldungen gestützt, in denen insbesondere Benutzerfreundlichkeit, klare Struktur und moderne Gestaltung hervorgehoben wurden. Gleichzeitig liefern die offenen Antworten Hinweise auf potenzielle Verbesserungspunkte etwa in Bezug auf Individualisierbarkeit und technische Ausgestaltung. Die Verknüpfung beider Perspektiven verdeutlicht, dass KI-gestützte Lösungen grundsätzlich akzeptiert werden, deren Wirksamkeit jedoch stark von kontextsensitiver Anpassung und kontinuierlicher Weiterentwicklung abhängt.

Die Ergebnisse stützen damit die zugrundeliegende Hypothese weitgehend, dass KI-generierte Leitfäden eine äquivalente Akzeptanz aufweisen mit einzelnen Vorteilen gegenüber den psychologisch fundierten Leitfäden.

8.1 Theoretische Einordnung der Befunde

Im theoretischen Kontext lassen sich die Ergebnisse dieser Untersuchung in Bezug auf die verwendeten Akzeptanzmodelle interpretieren. Das TAM postuliert, dass die Akzeptanz technologischer Systeme primär von der wahrgenommenen Nützlichkeit (PU) und der wahrgenommenen Benutzerfreundlichkeit (PEOU) beeinflusst wird. Die Ergebnisse zeigen, dass KI-generierte Leitfäden hinsichtlich dieser beiden Konstrukte ähnlich positiv bewertet wurden wie von psychologisch ausgebildeten Fachpersonen erstellte Leitfäden, was mit den theoretischen Annahmen des TAM übereinstimmt. Auch im Rahmen der UTAUT, welche die Leistungserwartung (PE), Aufwandserwartung (EE) und weitere Einflussfaktoren auf die Technologieakzeptanz betont, zeigen sich ähnliche Tendenzen. Die Leistungserwartung wurde für beide Leitfadenversionen vergleichbar hoch eingeschätzt, während hinsichtlich der Aufwandserwartung keine negativen Unterschiede zugunsten der KI-basierten Variante festgestellt wurden. Ergänzend bietet die CASA-Theorie einen Erklärungsansatz für die positive Wahrnehmung des FLOWIT-Tools. Die Ergebnisse deuten darauf hin, dass Nutzende das KI-basierte System nicht nur als funktionales Hilfsmittel, sondern auch als sozial unterstützendes Instrument wahrnehmen.

Insgesamt legen die Ergebnisse nahe, dass die theoretischen Modelle TAM, UTAUT und CASA geeignete Rahmenwerke darstellen, um die Akzeptanz KI-generierter Gesprächsleitfäden zu erklären. Gleichzeitig zeigen sie auf, dass weitere empirische Untersuchungen erforderlich sind, um diese ersten Befunde zu bestätigen und zu vertiefen.

9 Limitationen

Bei der Interpretation der vorliegenden Ergebnisse sind verschiedene Limitationen zu berücksichtigen.

Die Stichprobengrösse von (N= 165) ist als vergleichsweise begrenzt einzustufen. Zwar ermöglichte diese Stichprobengrösse grundlegende inferenzstatistische Analysen, sie schränkt jedoch insbesondere die Verallgemeinerbarkeit der Ergebnisse ein. Auch die Post-Hoc-Analyse der Teststärke weist auf eine methodische Limitation hin. Die erreichte Teststärke ($1-\beta$) betrug lediglich 16 %, deutlich unterhalb des empfohlenen Wertes von 80 % (Cohen, 1988). Daraus ergibt sich, dass die Wahrscheinlichkeit, tatsächlich existierende Unterschiede zwischen den Gruppen zu entdecken, relativ gering war. Die gefundenen Effekte sollten daher vorsichtig interpretiert werden. Zusätzlich ergab die Reliabilitätsanalyse ein sehr niedriges Cronbachs Alpha von 0.0232. Eine mögliche Erklärung für diesen Befund könnte in der Heterogenität der erfassten Items liegen, da verschiedene Konstrukte der Akzeptanz abgefragt wurden. Dennoch bleibt die Messgenauigkeit der Skala eingeschränkt. Die qualitative Analyse der offenen Antworten stellt eine weitere Limitation dar. Da sie explorativ angelegt war und auf einer begrenzten Anzahl an Rückmeldungen basiert, sind die daraus gewonnenen Erkenntnisse nicht generalisierbar und sollten lediglich als erste Hinweise verstanden werden. Eine weitere Einschränkung ergibt sich aus der Konzentration auf ein einziges KI-gestütztes System (ChatGPT). Die Ergebnisse der Untersuchung beziehen sich ausschliesslich auf die Wahrnehmung dieses spezifischen Tools und lassen sich nicht ohne Weiteres auf andere KI-Systeme übertragen. Schliesslich wurde im Rahmen der Untersuchung der soziale Einfluss, ein zentraler Bestandteil des UTAUT-Modells, nicht explizit erfasst. Damit bleibt unklar, inwiefern Gruppennormen oder soziale Erwartungen die Akzeptanz der Leitfäden beeinflusst haben könnten.

Trotz dieser Limitationen liefert die vorliegende Arbeit relevante erste Erkenntnisse zur Akzeptanz KI-gestützter Gesprächsleitfäden und bildet eine wertvolle Grundlage für weiterführende Forschung.

10 Implikationen, kritische Würdigung und Ausblick

Die Ergebnisse dieser Untersuchung bieten erste Hinweise darauf, wie KI-generierte Gesprächsleitfäden im Vergleich zu solchen von psychologisch ausgebildeten Fachpersonen hinsichtlich ihrer Akzeptanz wahrgenommen werden. Die erzielte Signifikanz in der Gesamteinschätzung des KI-Leitfadens stellt trotz geringer Effektstärke eine relevante Grundlage für die weiterführende Auseinandersetzung mit KI-basierten Leitfäden im Kontext der Gesprächsvorbereitung dar. Die Ergebnisse legen nahe, dass KI-generierte Leitfäden grundsätzlich eine vergleichbare, teils sogar höhere Akzeptanz erfahren können als klassisch entwickelte Formate. Daraus ergeben sich praxisrelevante Anhaltspunkte für eine gezielte Integration solcher Leitfäden, insbesondere in Unternehmen, in denen die Notwendigkeit für standardisierte Gesprächsabläufe besteht und begrenzt Ressourcen verfügbar sind. Unter diesen Bedingungen könnten KI-basierte Leitfäden die manuelle Erstellung entlasten und gleichzeitig zur Konsistenz und Effizienz beitragen. Voraussetzung dafür ist jedoch eine sorgfältige Anpassung an den jeweiligen Anwendungsrahmen. Um die Akzeptanz zu sichern, sollte insbesondere auf Individualisierbarkeit, technische Stabilität und sprachliche Qualität geachtet werden. So können KI-basierte Lösungen im Vergleich zum IST-Zustand eine zweckmäßige und zugleich skalierbare Ergänzung in der Gesprächsvorbereitung darstellen. Zugleich zeigen die Befunde aber auch auf, dass der Einsatz von KI im HR-Bereich mit gewissen Grenzen verbunden ist. So stellt sich etwa die Frage, inwieweit Verantwortung für sensible, personenbezogene Inhalte an KI-Systeme abgegeben werden kann oder darf. Damit verbunden sind zentrale ethische Herausforderungen insbesondere im Hinblick auf Transparenz der Systemlogik, Datenschutz und potenzielle Verzerrungen durch die zugrunde liegenden Trainingsdaten. Solche Aspekte sollten kritisch reflektiert und kontextsensitiv adressiert werden.

Darüber hinaus lieferten die separat erhobenen qualitativen Rückmeldungen zu FLOWIT praxisrelevante Hinweise auf dessen Optimierungspotenziale, insbesondere im Hinblick auf Individualisierbarkeit und technische Gestaltung. Diese Erkenntnisse beziehen sich nicht auf die Hauptuntersuchung zur Akzeptanz der Leitfäden selbst, sondern ergänzen diese um weiterführende Perspektiven auf die konkrete Anwendung des Tools in der Praxis. Somit erweitern die qualitativen Befunde die statistischen Analysen um wertvolle kontextbezogene Perspektiven auf die tatsächliche Nutzung.

Überraschend war insbesondere, dass der KI-Leitfaden nicht nur als gleichwertig, sondern punktuell sogar als überlegen wahrgenommen wurde z.B. in Bezug auf Gesprächsstrukturierung und Nutzungserlebnis. Angesichts verbreiteter Skepsis gegenüber KI im HR-Bereich war dies keineswegs selbstverständlich. Diese Beobachtung wirft neue Fragen auf: Welche spezifischen Merkmale der KI-generierten Inhalte tragen zu dieser positiven Bewertung bei? Inwiefern spielen dabei standardisierte, neutrale Formulierungen eine Rolle? Und verändert sich die Akzeptanz, wenn der Einsatzkontext emotional aufgeladener oder konfliktbehafteter wird? Diese Fragen markieren relevante Anknüpfungspunkte für zukünftige Forschung.

Auch aus wissenschaftlicher Perspektive ergeben sich weiterführende Überlegungen. Die Befunde stützen zentrale Annahmen des TAM sowie der UTAUT, insbesondere in Bezug auf die Relevanz der Konstrukte Benutzerfreundlichkeit (PEOU), Leistungserwartung (PE) und Wahrgenommene Nützlichkeit (PU) als Einflussfaktoren für die Akzeptanz. Das Konstrukt des sozialen Einflusses, welches ein zentrales Element des UTAUT-Modells darstellt, wurde in der vorliegenden Studie nicht berücksichtigt, könnte jedoch in künftigen Untersuchungen eine wichtige ergänzende Rolle spielen. Darüber hinaus macht das sehr niedrige Cronbachs Alpha deutlich, dass bei der Skalenkonstruktion stärker auf Itemhomogenität geachtet werden sollte. Denkbar wäre auch, theoriebasierte Subskalen gezielter voneinander abzugrenzen und gesondert auszuwerten, um die Reliabilität gezielt zu steigern. Die Aussagekraft der vorliegenden Studie ist zudem durch eine begrenzte Stichprobengröße sowie eine geringe Teststärke eingeschränkt. Die Ergebnisse deuten dennoch auf relevante Tendenzen hin, die in weiterführenden Arbeiten überprüft und vertieft werden sollten. Insbesondere könnte eine grössere Stichprobe dazu beitragen, die Aussagekraft der Befunde zu erhöhen und mögliche höhere Effekte mit höherer statistischer Sicherheit nachzuweisen. Auch der Einsatz der Leitfäden in realen Gesprächen könnte zukünftige Studien bereichern und zur Validierung der Ergebnisse beitragen. Ergänzend wäre es denkbar, unterschiedliche KI-Systeme miteinander zu vergleichen, um die Akzeptanz verschiedener technologischer Lösungen differenziert zu untersuchen und mögliche systembedingte Unterschiede in der Wahrnehmung zu identifizieren. Darüber hinaus stellt sich die Frage, inwieweit durch gezieltes Prompting die

Qualität KI-generierter Leitfäden weiter gesteigert werden kann, was sich potenziell positiv auf deren Akzeptanz auswirken könnte.

Insgesamt liefern die Ergebnisse der vorliegenden Arbeit erste empirische Anhaltspunkte. Im Unterschied zu bisherigen Studien, die den Einsatz von KI überwiegend in allgemeinen oder hypothetischen Szenarien untersuchten, richtet sich der Fokus dieser Arbeit auf reale Anwendungskontexte im Bereich der Personalentwicklung. Statt hypothetischer Szenarien wurden reale Leitfäden eingesetzt. Sie verdeutlichen sowohl das Potenzial als auch die Notwendigkeit für weiterführende Forschung und Entwicklung in diesem Themenfeld.

11 Literaturverzeichnis

Achouri, C. (2015). *Human Resources Management: Eine praxisbasierte Einführung (2.Aufl.)*.

Wiesbaden: Gabler Verlag. <https://doi.org/10.1007/978-3-8349-4740-6>

Bauer, W., Ganz, W., Hämmerle, M. & Renner, T. (2019). KÜNSTLICHE INTELLIGENZ IN DER

UNTERNEHMENSPRAXIS. *Fraunhofer IAO*. Verfügbar unter: [https://publica-](https://publica-rest.fraunhofer.de/server/api/core/bitstreams/0dd40c4f-9c68-4f8e-ba56-e4743834805d/content)

[rest.fraunhofer.de/server/api/core/bitstreams/0dd40c4f-9c68-4f8e-ba56-](https://publica-rest.fraunhofer.de/server/api/core/bitstreams/0dd40c4f-9c68-4f8e-ba56-e4743834805d/content)

[e4743834805d/content](https://publica-rest.fraunhofer.de/server/api/core/bitstreams/0dd40c4f-9c68-4f8e-ba56-e4743834805d/content)

Brenner, D. (2020). *Mitarbeitergespräche souverän führen: Eine praxisorientiertes Manual*

für Führungskräfte. Wiesbaden: Springer Fachmedien Wiesbaden.

<https://doi.org/10.1007/978-3-658-31358-6>

Bünnagel, W. (2024). *Künstliche Intelligenz und Unternehmenswissen: Betriebliches*

Wissensmanagement auf morgen ausrichten. Berlin, Heidelberg: Springer Berlin

Heidelberg. <https://doi.org/10.1007/978-3-662-68779-6>

Davis, F. (1987). *USER ACCEPTANCE OF INFORMATION SYSTEMS: THE TECHNOLOGY*

ACCEPTANCE MODEL (TAM). Verfügbar unter:

<https://quod.lib.umich.edu/b/busadwp/images/b/1/4/b1409190.0001.001.pdf>

Davis, F. & Granić, A. (2024). *The Technology Acceptance Model: 30 Years of TAM*. Cham:

Springer Nature Switzerland. <https://doi.org/10.1007/978-3-030-45274-2>

Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und*

Humanwissenschaften. Berlin, Heidelberg: Springer Berlin Heidelberg.

<https://doi.org/10.1007/978-3-642-41089-5>

Epley, N., Waytz, A. & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of

anthropomorphism. *Psychological Review*, 114(4), 864–886.

<https://doi.org/10.1037/0033-295X.114.4.864>

- Ertel, W. (2021). *Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung*.
Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-32075-1>
- FLOWIT (2024). *Produkt*. Verfügbar unter: <https://www.flowit.ai/de/product/overview>
- FLOWIT (2024). *Unternehmen*. Verfügbar unter: <https://www.flowit.ai/de/about-us>
- Friese, S. (2012). *Getting to know atlas.ti*. Verfügbar unter:
<https://methods.sagepub.com/book/mono/qualitative-data-analysis-with-atlas/chpt/one-getting-know-atlasti>
- Gambino, A., Fox, J. & Ratan, R. (2020). Building a Stronger CASA: Extending the Computers Are Social Actors Paradigm. *Human-Machine Communication, 1*, 71–86.
<https://doi.org/10.30658/hmc.1.5>
- Garnitz, J. & Schaller, D. (2023). ChatGPT, Chatbots und mehr – wie wird Künstliche Intelligenz in den HR-Abteilungen von Unternehmen genutzt? *Leibniz-Institut für Wirtschaftsforschung, 76(09)*, 65-68. <https://hdl.handle.net/10419/279737>
- Greif, S. & Möller, H. (2018). *Handbuch Schlüsselkonzepte im Coaching*. In W. Scholl (Hrsg.). Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-49483-7>
- Hasenbein, M. (2023). *Mensch und KI in Organisationen: Einfluss und Umsetzung Künstlicher Intelligenz in wirtschaftspsychologischen Anwendungsfeldern*. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-66375-2>
- Hermeier, B. & Heupel, T. (2019). Wie die Digitalisierung unsere Arbeitsplätze und Arbeitsweisen verändert. In S. Fichtner-Rosada (Hrsg.), *Arbeitswelten der Zukunft*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-23397-6>

- Hossiep, R., Zens, J. E. & Berndt, W. (2020). *Mitarbeitergespräche* (2. Aufl.). Göttingen: Hogrefe Verlag. <https://doi.org/10.1026/03002-000>
- Jöns, I. (2018). Grundlagen, Gestaltungshinweise, Erfahrungsberichte. In W. Bungard (Hrsg.), *Feedbackinstrumente im Unternehmen*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-20759-5>
- Kaudela-Baum, S. & Meldau, S. (2022). Führung und Kollaboration in Zeiten der Digitalisierung und Transformation. In M. Brassler (Hrsg.), *Leadership und People Management*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-35521-0>
- Kauffeld, S. (Hrsg.). (2014). *Arbeits-, Organisations- und Personalpsychologie für Bachelor*. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-42065-8>
- Kreutzer, R. T. (2023). *Künstliche Intelligenz verstehen: Grundlagen – Use-Cases – unternehmenseigene KI-Journey*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-42598-2>
- Luger, E. & Sellen, A. (2016). „Like Having a Really Bad PA“: The Gulf between User Expectation and Experience of Conversational Agents. <https://doi.org/10.1145/2858036.2858288>
- Mentzel, W. (2020). *Mitarbeitergespräche*. Krefeld: Reemers Publishing Services GmbH. https://www.wiso-net.de/document/HAUF__da21729a4f501b05a6c8b2de5cb8b055cd4417d3
- Minar, E. (2024). *Praxistipps für Nachwuchs-Führungskräfte: Ratschläge eines Top-Managers*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-44030-5>

- Moosbrugger, H. (2020). *Testtheorie und Fragebogenkonstruktion*. In A. Kelava (Hrsg.). Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-61532-4>
- Nass, C. & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues*, 56(1), 81–103. <https://doi.org/10.1111/0022-4537.00153>
- Neuberger, O. (2015). *Das Mitarbeitergespräch: Praktische Grundlagen für erfolgreiche Führungsarbeit*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-07940-6>
- OpenAI (2024). *Research*. Verfügbar unter: <https://openai.com/research/index/>
- OpenAI (2024). *ChatGPT*. Verfügbar unter: <https://openai.com/chatgpt/overview/>
- Paaß, G. & Hecker, D. (2020). *Künstliche Intelligenz: Was steckt hinter der Technologie der Zukunft?* Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-30211-5>
- Pentzold, C. & Bischof, A. (2018). Theoriegenerierendes empirisches Forschen in medienbezogenen Lebenswelten. Ein Lehr- und Arbeitsbuch. In N. Heise (Hrsg.), *Praxis Grounded Theory*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-15999-3>
- Porst, R. (2014). *Fragebogen: Ein Arbeitsbuch*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-02118-4>
- Scherm, M. & Sarges, W. (2019). *360°-Feedback* (2. Aufl.). Bern: Hogrefe Verlag. <https://doi.org/10.1026/03000-000>
- Scheuer, D. (2020). *Akzeptanz von Künstlicher Intelligenz: Grundlagen intelligenter KI-Assistenten und deren vertrauensvolle Nutzung*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-29526-4>

- Schirmacher, U. (2023). *Kompetenzorientierte Personalentwicklung: Wie Sie in 9 Schritten ein individuelles Lernprogramm erstellen*. Wiesbaden: Springer Fachmedien
Wiesbaden. <https://doi.org/10.1007/978-3-658-41487-0>
- Schmitz, L. & Billen, B. (2016). *Lösungsorientierte Mitarbeitergespräche: zielorientiert planen, klar formulieren, erfolgreich Vereinbarungen treffen* (5. Aufl.). München: Redline Verlag. <https://katalog.ub.uni-leipzig.de/Record/0-655772707>
- Schuler, H. (2014). *Lehrbuch der Personalpsychologie* (3.Aufl.). In U. P. Kanning (Hrsg.). Göttingen Bern Wien Paris Oxford Prag Toronto Boston, Mass Amsterdam Kopenhagen Stockholm Florenz Helsinki: Hogrefe Verlag.
<https://doi.org/10.1026/978-3-840-92363-0>
- Stangier, V. (2019). *Über Mitarbeitergespräche Erfolge generieren: Dialoge des Vertrauens*. Tübingen: expert verlag. <https://doi.org/10.2357/978-3-823-37673-6>
- Steiger, T. (2013). Führungskompetenz und Führungswissen. In E. Lippmann (Hrsg.). *Handbuch Angewandte Psychologie für Führungskräfte*. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-34357-5>
- Taulli, T. (2022). *Grundlagen der Künstlichen Intelligenz: Eine nichttechnische Einführung*. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-66283-0>
- Thorn, A., McLeod, M. & Goldsmith, M. (2007). *Quick Reference for Peer Coaching*. Verfügbar unter: <https://marshallgoldsmith.com/wp-content/uploads/2015/10/Peer-Coaching-Overview.pdf>
- Venkatesh, V., Morris, M., Davis, G. & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425-478.
<https://doi.org/10.2307/30036540>

Venkatesh, V. & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance

Model: Four Longitudinal Field Studies. *Management Science*, 46(2), 186–204.

<https://doi.org/10.1287/mnsc.46.2.186.11926>

VERBI Software. (2024). *MAXQDA 2024 [Computer-Software]*. <https://www.maxqda.com>

Werther, S. (2015). *Einführung in Feedbackinstrumente in Organisationen: Vom 360°-*

Feedback bis hin zur Mitarbeiterbefragung. Wiesbaden: Springer Fachmedien

Wiesbaden. <https://doi.org/10.1007/978-3-658-10497-9>

12 Tabellenverzeichnis

Tabelle 1 Kategoriensystem mit Haupt- und Subkategorien	42
---	----

13 Abbildungsverzeichnis

Abbildung 1. FLOWIT Logo (FLOWIT, 2024)	4
Abbildung 2. Technology-Acceptance-Model (Venkatesh & Davis, 2000).....	23
Abbildung 3. Unified Theory of Acceptance and Use of Technology (Venkatesh et al. 2003)	24
Abbildung 4. Untersuchungsdesign der vorliegenden Arbeit (in Anlehnung an Döring und Bortz, 2016).....	30
Abbildung 5. 95%- Konfidenzintervall der gesamten Ergebnisse (eigene Darstellung)	43
Abbildung 6. Codewolke (MAXQDA, 2024).....	51
Abbildung 7. Code-Relations-Modell nach Vorkommen der Codes (MAXQDA, 2024)	52

Hilfsmittelverzeichnis mit Verwendungszweck

In der vorliegenden Arbeit wurden folgende KI-Assistenzsysteme eingesetzt:

KI-Assistenzsystem	Teile/Stelle(n) in der Arbeit	Einsatz
DeepL	-	Übersetzung von englischer Literatur
OpenAI	Kapitel: ChatGPT und GPT-40	Literaturrecherche zu ChatGPT
OpenAI	Kapitel: Custom GPTs	Literaturrecherche zu Custom GPTs
ChatGPT	Kapitel: Custom GPTs	Literaturrecherche zu Custom GPTs
ChatGPT	Gesamte Arbeit	Findung von Synonymen
ChatGPT	Abstract	Kürzung des Abstracts auf die vorgegebene Anzahl Wörter
ChatGPT	Gesamte Arbeit	Kürzen von Sätzen
ChatGPT	Gesamte Arbeit	Prüfung auf Rechtschreibung