

MAS HCID MASTERARBEIT 2026

Konzeption und Evaluation eines KI-unterstützten Analyseprozesses im Microsoft-365-Ökosystem

Autor

Dominik James Schendl

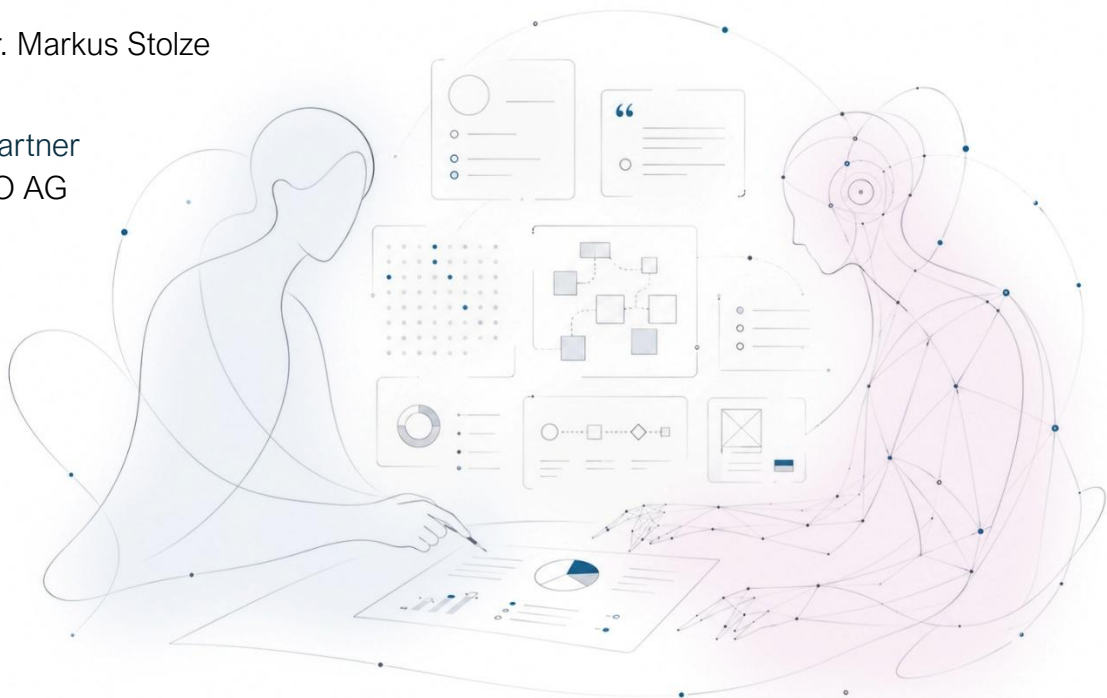
Betreuende

Dr. Susanne Schmidt-Rauch

Prof. Dr. Markus Stolze

Praxispartner

GARAIO AG



Eigenständigkeitserklärung

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbst und ohne fremde Hilfe durchgeführt habe, ausser derjenigen, welche explizit beschrieben sind, dass ich sämtliche verwendeten Quellen erwähnt und gemäss gängigen wissenschaftlichen Regeln korrekt zitiert habe, und dass ich keine durch Copyright geschützten Materialien (z.B. Bilder) in dieser Arbeit in unerlaubter Weise genutzt habe, und dass ich in dieser Arbeit keine Adressen, Telefonnummern und andere persönliche Daten von Personen, die nicht zum Kernteam gehören, publiziere.



Dominik James Schendl

Abstract

Diese Arbeit befasst sich mit der Konzeption, prototypischen Ausgestaltung und Evaluation eines KI-unterstützten Analyseprozesses im Microsoft-365-Ökosystem. Ausgangspunkt ist das im Employee Experience Bereich (nachfolgend EX-Bereich genannt) der GARAIO verantwortete M365-Evolution-Assessment, dessen fachlich tragfähige, aber aufwendige manuelle Auswertung den Einsatz auf grössere Projekte beschränkt. Dieser EX-Bereich kümmert sich um Microsoft 365 Systemlandschaften (nachfolgend M365 genannt). Daran knüpft das Light Assessment an, für das untersucht wird, wie ein KI-unterstützter Analyseprozess so gestaltet werden kann, dass der Analyseaufwand sinkt, die Ergebnisverlässlichkeit gewahrt bleibt und die fachliche Verantwortung beim Fachpersonal verbleibt.

Methodisch verbindet die Arbeit die menschenzentrierte Gestaltung nach DIN EN ISO 9241-210 mit Participatory Design und Garretts Five Elements of UX und ergänzt sie um qualitative Interviews, die Inhaltsanalyse nach Mayring (2022) sowie das eigens definierte Vibe Research. Entlang eines angepassten HCD-Kreislaufs werden zuerst der Nutzungskontext erschlossen und aus drei Interviews fünf Kernbedürfnisse samt Anforderungskatalog abgeleitet, anschliessend der Soll-Prozess partizipativ mit fünf M365-Fachpersonen erarbeitet und zu einem siebenschrittigen Light-Assessment-Prozess ausgestaltet. Dieser wechselt durchgängig zwischen Mensch und KI, wobei die KI die strukturierende Fleissarbeit gemäss dem extra definierten Vibe Research übernimmt und Wort-treuekontrolle, Priorisierung und Präsentation beim Fachpersonal bleiben. Parallel finden technische Gestaltungen für Proof of Concept der Lösung statt.

Die Evaluation mit drei Fachpersonen an realen Kundendaten, in der die KI im Wizard-of-Oz-Aufbau simuliert wird, zeigt eine hohe Nutzbarkeit und einen Effizienzgewinn von etwa 50 bis 60 auf etwa 20 Arbeitsstunden. Eine technische fünffache Wiederholung des Analyseprozesses zu Reliabilitätsprüfung, zeigt auf der entscheidungsrelevanten Ebene substanzielle bis nahezu perfekte Lauf-zu-Lauf-Stabilität. Die Zuordnung einzelner Belegstellen schwankt dagegen stärker. Aus dem Projekt wird zudem ein übertragbares, siebenschrittiges Vorgehensmodell zur KI-unterstützten Analyse in soziotechnischen Systemen abgeleitet.

Als zentrale Erkenntnis zeigt sich, dass die Verlässlichkeit eines solchen Prozesses weniger vom Sprachmodell als von einer prüfbar Prozessarchitektur, einer bewussten Mensch-KI-Aufgabenteilung und klar gesetzten Eingriffspunkten abhängt. Die menschliche Kontrolle ist dabei nicht als nachträgliche Prüfung, sondern als fest eingebauter Bestandteil des Prozesses zu verstehen. Als Grenze bleibt, dass die Ergebnisverlässlichkeit bislang nur an einem einzelnen Assessmentfall geprüft ist.

Inhaltsverzeichnis

1. Projektkontext.....	1
1.1. Ausgangslage.....	1
1.2. Problemstellung.....	3
1.3. Vorgaben	3
1.4. Abgrenzung	4
1.5. M365-Ökosystem und Copilot	6
1.6. Rahmenbedingungen: Datenschutz, Vertraulichkeit, Anonymisierung, Zugänge.....	7
2. Fragestellung und Zielsetzung.....	8
2.1. Zielsetzung und Forschungsfrage.....	8
2.2. Lernziele.....	9
3. Literaturvertiefung	9
3.1. Vorgehen bei der Quellenrecherche und Quellenkritik.....	9
3.2. KI- und LLM-gestützte Analyse: Stand der Technik und Forschung	10
3.3. Praxisrecherche und State of Practice	13
3.4. Begriffliche Einordnung: Vibe Research	14
3.5. Lernreflexion - Literaturvertiefung	17
4. Vorgehen und Methoden	19
4.1. Vorgehensmodell	19
4.2. Menschzentrierte Entwicklung als Vorgehensmodell	19
4.3. Qualitative Interviews.....	23
4.4. Vibe Research als reguliertes Vorgehen.....	24
4.5. Usability Testing.....	26
4.6. Stakeholder Analyse	27
4.7. Risikoanalyse.....	28
4.8. Qualitative Inhaltsanalyse nach Mayring als methodische Ergänzung zu Vibe Research.....	28
4.9. Lernreflexion – Vorgehen und Methoden.....	29
5. Planung des Vorgehens	32
5.1. Projektplanung und Umgang mit Planabweichungen.....	32
5.2. Stakeholder Analyse	33
5.3. Risikoanalyse.....	34

5.4. Lernreflexion – Planungsphase	36
6. Phase 1: Nutzungskontext verstehen	37
6.1. Stakeholder Interviews	37
6.2. Inhaltsanalyse mit Vibe Research.....	39
6.3. Herleitung der Kernbedürfnisse	43
6.4. Kernbedürfnisse.....	44
6.5. Ergebnisse	46
6.6. Lernreflexion – Phase 1	48
7. Phase 2: Konzeption	49
7.1. Participatory Design	49
7.2. Konkrete Durchführung	49
7.3. Der partizipative Prozess.....	51
7.4. Exemplarische Vertiefung: Transkript Bereinigung	54
7.5. Reflexionsworkshop	57
7.6. Ergebnisse	58
7.7. Lernreflexion – Phase 2	59
8. Phase 3: Gestaltung	60
8.1. Five Elements of UX.....	60
8.2. Konkrete Durchführung	60
8.3. Ergebnis: Der Light Assessmentprozess	63
8.4. Lernreflexion – Phase 3.....	65
9. Phase 4: Evaluation	68
9.1. Konkrete Durchführung	68
9.2. Ergebnisse	70
9.3. Exemplarische Verbesserungsvorschläge	73
9.4. Lernreflexion – Phase 4.....	76
10. Technischer Aspekt des KI-Einsatzes.....	78
10.1. Abgrenzung	78
10.2. Awareness für Harness Engineering	78
10.3. Kodierleitfaden für LLM-basierte Systeme.....	78
10.4. Modellwahl nach Wirtschaftlichkeit und Datensicherheit.....	81
10.5. Ergebnisverlässlichkeit: Lauf-zu-Lauf-Stabilität.....	81

10.6. Ergebnisse	82
11. Ergebnisse	84
11.1. Beantwortung der Forschungsfragen.....	84
11.2. Erreichen der Zielsetzung	85
11.3. Lernziele	85
12. Vorgehensmodell zur Erstellung eines Assessments für soziotechnische Systeme	87
12.1. Geltungsbereich und Voraussetzungen	87
12.2. Aufbau des Modells	87
12.3. Schritt 1: Das System analysieren	88
12.4. Schritt 2: Codebook und Empfehlungskatalog erstellen.....	88
12.5. Schritt 3: Standard-Erhebung planen	89
12.6. Schritt 4: Aufbereitung der Daten und Testen	89
12.7. Schritt 5: Aufbereitung der Daten definieren	90
12.8. Schritt 6: Technischer Test der Inhaltsanalyse.....	90
12.9. Schritt 7: Zielartefakte definieren.....	91
12.10. Einordnung	91
13. Leistungsreflexion	91
13.1. Eigene Reflexion	91
13.2. Beiträge und Mehrwert	92
13.3. Weitere Einschätzungen	93
14. Persönliche Reflexion	94
15. Literatur- und Quellenverzeichnis.....	96
16. Nutzung von KI-Tools in der Arbeit	100
16.1. Deklaration der Nutzung von KI-Tools.....	100
16.2. Reflexion zur Nutzung von KI-Tools.....	100
17. Anhang	102

1. Projektkontext

1.1. Ausgangslage

Die vorliegende Arbeit ist im EX-Bereich der GARAIO AG verortet. Sie knüpft an das dort verantwortete M365-Evolution-Modell und das darauf aufbauende Evolution Assessment an. Im Folgenden werden das organisationale Umfeld, das Modell, das ursprüngliche Assessment und die verdichtete Weiterentwicklung zum Light Assessment beschrieben, an die diese Arbeit unmittelbar anschliesst.

1.1.1. Der Auftraggeber

Die GARAIO AG

Ist ein 1994 gegründetes Schweizer Software- und IT-Dienstleistungsunternehmen, mit Hauptsitz in Bern und einer weiteren Niederlassung in Zürich, welches rund 90 Mitarbeitende beschäftigt. Als Microsoft Solutions Partner für digitale Transformation entwickelt sie massgeschneiderte Softwarelösungen und digitale Services, die Unternehmen dabei unterstützen, digitale Vorhaben praxisnah umzusetzen und sowohl für Mitarbeitende als auch für Kundinnen und Kunden konsistente digitale Erlebnisse bereitzustellen. Die fachlichen Schwerpunkte des Unternehmens liegen in den Bereichen Modern Workplace, Customer Experience, individueller Softwareentwicklung und E-Commerce. GARAIO verbindet dabei technologische Lösungsentwicklung mit einer ausgeprägten Orientierung an Nutzerbedürfnissen und einer partnerschaftlichen Arbeitsweise, was eine inhaltliche Anschlussfähigkeit zur menschenzentrierten Gestaltung herstellt. Für diese Arbeit bildet GARAIO den organisationalen Rahmen. Das Vorhaben ist im EX-Bereich des Unternehmens angesiedelt, der den Untersuchungsgegenstand operativ verantwortet und das Evolution Assessment als zentrale Dienstleistung führt.

1.1.2. Der Fachkontext

Der EX Bereich

Der Bereich Employee Experience, intern als EX-Bereich bezeichnet, verfolgt das Ziel, die digitale Arbeitsumgebung und die Mitarbeitenden-Erfahrung in Unternehmen nachhaltig zu verbessern. Das EX-Team begleitet Organisationen bei der Einführung und Optimierung von Microsoft 365 sowie bei den dazugehörigen Adoption- und Change-Management-Massnahmen. Damit wird sichergestellt, dass neue digitale Werkzeuge wie Microsoft Teams oder Viva nicht nur technisch verfügbar sind, sondern auch tatsächlich in die Arbeitsprozesse integriert und von den Mitarbeitenden produktiv genutzt werden. Eine zentrale Dienstleistung des EX-Bereichs ist das Employee-Experience-Evolution-Modell mit dem zugehörigen Assessment. Beides ist im EX-Bereich entstanden und wird dort gepflegt und weiterentwickelt. Für die vorliegende Arbeit ist der EX-Bereich damit zugleich Praxispartner, Auftraggeber und fachlicher Bezugsrahmen. Die im EX-Bereich verorteten Rollen bilden zudem den Kern der Stakeholder, die weiter ausdifferenziert werden (siehe [Abschnitt 5.2](#)).

Das M365 Evolution Modell

Das M365-Evolution-Modell ist ein dreistufiges Reifegradmodell mit den Stufen Work, Boost und Perform. Es ordnet den Reifegrad der Nutzung und Verankerung von Microsoft 365 in Unternehmen entlang einer aufsteigenden Logik ein, in der grundlegende Arbeitsfähigkeit, Effizienzgewinn und schliesslich produktivitätsfördernde Nutzung unterschieden werden. Hintergrund des Modells ist die Beobachtung, dass Schwierigkeiten nach der Einführung von M365-Systemen im Unternehmensalltag häufig nicht primär technischer Natur sind, sondern organisatorische, menschliche und kulturelle Ursachen haben. Das Modell bildet damit den fachlichen Bezugsrahmen, innerhalb dessen das Evolution Assessment ansetzt.

1.1.3. Die Produkte

Das M365 Evolution Assessment

Im nachfolgend Evolution Assessment genannt, ist ein strukturiertes Analyseverfahren zur Bestimmung des digitalen Reifegrads einer Organisation im Modern-Work-Umfeld. Es erfasst den Stand der digitalen Zusammenarbeit, macht Handlungsfelder sichtbar und leitet daraus konkrete Entwicklungsschritte für den weiteren Ausbau des digitalen Arbeitsplatzes ab. Methodische Grundlage bildet das Evolution-Modell mit seinen drei Reifegradstufen, das technische, organisatorische und kulturelle Voraussetzungen gleichermaßen einbezieht. Entstanden ist das Assessment in Zusammenarbeit mit einer internen UX-Fachperson und orientiert sich an einem Mixed-Methods-Vorgehen aus dem HCID-Umfeld. Methodisch folgt es einer menschenzentrierten Logik. Im Zentrum steht nicht allein, welche M365-Funktionen technisch verfügbar sind, sondern wie Mitarbeitende die digitale Arbeitsumgebung tatsächlich nutzen, welche Aufgaben sie darin bewältigen und welche organisationalen Bedingungen diese Nutzung prägen. In der Anwendung überzeugt das Assessment durch die fachliche Tragfähigkeit seiner Ergebnisse und durch die ganzheitliche Sicht, die es M365-Fachpersonen auf den jeweiligen Kundenkontext eröffnet. Zugleich zeigen sich strukturelle Hürden. Insbesondere der Auswertungsaufwand wird als hoch wahrgenommen, was den breiten Einsatz, vor allem bei kleineren Projekten, erschwert. An dieser Konstellation knüpft das Light Assessment an.

Das Light Assessment

Das Light Assessment ist eine fokussierte Weiterentwicklung des Evolution Assessments. Es verdichtet das bestehende Verfahren auf klare eingegrenzte Themenfelder und reduziert den explorativen Anteil gegenüber der klassischen Variante. Ziel ist es, den Erhebungs- und insbesondere den Auswertungsaufwand so weit zu reduzieren, dass das Angebot auch in kleineren Kundenprojekten wirtschaftlich tragfähig wird. Das klassische Evolution Assessment bleibt in Situationen relevant, in denen eine breitere und offenere Untersuchung erforderlich ist.

Die Idee zum Light Assessment entstand aus der praktischen Erfahrung mit dem Evolution Assessment. Im operativen Einsatz zeigte sich wiederholt, dass der hohe Analyseaufwand den Einsatz auf grössere Projekte beschränkt, obwohl die Methode auch in kleineren Kontexten Mehrwert stiften könnte. KI-Unterstützung und Automatisierung wurden in diesem Zusammenhang als mögliche Lösungsansätze identifiziert. Die Stossrichtung des Light Assessments verbindet zwei Perspektiven. Aus Sicht der M365-Fachpersonen reagiert sie auf das Bedürfnis, im Analyseprozess gezielt entlastet zu werden, ohne die fachliche Tiefe der Beurteilung zu verlieren. Aus Geschäftssicht reagiert sie auf das Bedürfnis, das Angebot in einem wirtschaftlich tragfähigen Rahmen anzubieten. Diese Verbindung von Nutzerbedürfnis und Geschäftsperspektive begründet den Anwendungskontext der vorliegenden Arbeit.

Im weiteren Verlauf der Arbeit wird das Evolution Assessment in seiner angepassten, im Light Assessment eingebetteten, Form als fachlicher Bezugsrahmen behandelt. Der eigentliche Untersuchungsgegenstand ist der innerhalb des Light Assessments angesiedelte Analyseprozess, weshalb sich die folgenden Kapitel primär auf diesen Analyseprozess beziehen.

1.2. Problemstellung

Aus der Ausgangslage ergibt sich ein Spannungsfeld zwischen einer methodisch tragfähigen Auswertung und einem für die Praxis vertretbaren Aufwand. Die Problemstellung wird im Folgenden entlang von vier Perspektiven entwickelt, die im weiteren Verlauf der Arbeit aufgenommen werden, nämlich die Belastung im bestehenden Evolution Assessment, die spezifischen Anforderungen des geplanten Light Assessments, die geschäftlichen Rahmenbedingungen sowie die daraus abgeleitete Aufgabenstellung.

1.2.1. Evolution Assessment

Die Ergebnisse des Evolution Assessments überzeugen fachlich. Der hohe Aufwand im Analyseprozess belastet jedoch die ohnehin stark ausgelasteten M365-Fachpersonen. Die Belastung ergibt sich aus zwei unterschiedlichen Kompetenzprofilen, die im Auswertungsprozess zusammengeführt werden müssen. M365-Fachpersonen verfügen über Erfahrung und Fachwissen, um Bedürfnisse und Anforderungen mit den Funktionsumfängen von M365 abzustimmen. Für die Analyse qualitativer und quantitativer empirischer Daten braucht es jedoch eine andere methodische Expertise, weshalb UX-Fachpersonen unterstützend einbezogen werden müssen.

Die eigentliche Hauptarbeit der M365-Fachpersonen besteht nicht in der manuellen Datenanalyse, sondern in der Konzeption und Umsetzung von Lösungen im M365-Umfeld. Die Analysephase bindet daher Ressourcen ausserhalb ihres fachlichen Kerns und reduziert die Skalierbarkeit des Angebots. Inhaltlich überzeugt das Assessment in der Anwendung, der Kostenpunkt stellt jedoch auf Kundenseite häufig eine Hemmschwelle dar. Auch wenn der manuelle Analyseprozess bereits beschleunigt wurde, beansprucht er zu viel Zeit. In der Folge wird das Assessment derzeit überwiegend in grösseren Kundenprojekten eingesetzt, obwohl die M365-Fachpersonen ein Interesse daran haben, es auch in kleineren Projekten anzubieten, da es dort frühzeitig zielgerichtete Handlungsfelder identifizieren würde.

1.2.2. Light Assessment

Das Light Assessment setzt an dieser Problemlage an. Es zielt darauf ab, den manuellen Analyseprozess sinnvoll und menschenzentriert durch KI zu unterstützen, ohne die fachliche Verantwortung der M365-Fachpersonen einzuschränken. Die Analyse empirischer Daten soll dabei zuverlässig automatisiert und in einer für Fachpersonen kontrollierbaren Form bereitgestellt werden, um Effizienzgewinn und Ergebnisverlässlichkeit zugleich zu adressieren. Aus Sicht der Nutzenden und der Geschäftsseite ist gleichzeitig akzeptiert, dass das Light Assessment nicht den vollen Detaillierungsgrad und Umfang des klassischen Evolution Assessments erreichen muss. Dieser Trade-off zwischen Tiefe und Effizienz bildet eine zentrale Gestaltungsannahme der Arbeit und wird in der Konzeption sowie in der Evaluation entsprechend berücksichtigt.

1.3. Vorgaben

1.3.1. Geschäftsvorgaben

Auf Geschäftsseite bestehen zwei strukturierende Vorgaben, die den Gestaltungsraum der Arbeit eingrenzen. Der gesamte Prozess muss inklusive der KI-Komponenten innerhalb der Microsoft-365-Systemwelt umgesetzt werden, um die Anschlussfähigkeit an bestehende Lizenzen, Datenflüsse und Datenschutzregelungen sicherzustellen. Der Gesamtaufwand des Assessments muss so gehalten werden, dass das Angebot in einem bestimmten Preisrahmen angeboten werden kann. Dieser Preisrahmen ergibt sich aus der angestrebten Einsatzbreite, insbesondere bei kleineren Kundenprojekten, und definiert eine wirtschaftliche Obergrenze für den Aufwand der KI-unterstützten Auswertung. Das gesamte Assessment muss in approximativ 40 Arbeitsstunden durchgeführt werden können, um diesen Preisrahmen zu erreichen, für den Analyseprozess selbst bedeutet dies etwa 20 Stunden.

1.3.2. Aufgabenstellung

Aus den vorangehenden Punkten ergibt sich die Aufgabenstellung dieser Arbeit. Untersucht wird, wie sich der manuelle Analyseprozess im Light Assessment durch einen KI-unterstützten Soll-Prozess so gestalten lässt, dass er den Aufwand für die M365-Fachpersonen reduziert, die Ergebnisverlässlichkeit absichert und gleichzeitig die geschäftlichen und technischen Rahmenbedingungen des M365-Ökosystems einhält. Aufgabe der Arbeit ist die Konzeption, prototypische Ausgestaltung und Evaluation dieses Analyseprozesses sowie die Ableitung eines Vorgehensmodells zur KI-unterstützten empirischen Datenanalyse im M365-Ökosystem.

1.4. Abgrenzung

Die Arbeit adressiert einen klar umrissenen Ausschnitt des M365-Evolution-Assessments. Im Folgenden wird explizit geklärt, was Gegenstand der Untersuchung ist und was bewusst nicht. Die Abgrenzung erfolgt entlang dreier Dimensionen, nämlich dem Scope des Assessment-Prozesses, dem Scope des begleitenden Pilot-Projekts und dem Scope des Prototypings. Damit werden zugleich die fachlichen und rechtlichen Randbedingungen sichtbar, insbesondere die geltenden NDAs, die Anforderungen an die Pseudonymisierung sowie die Doppelrolle als Autor und faktischer Auftraggeber dieser Arbeit, was in der Stakeholder Analyse erörtert wird.

1.4.1. Scope Assessment-Prozess

Der Gesamtprozess eines Assessments beinhaltet diverse Meetings, technische Abklärungen, die Erhebung der Daten und andere Tätigkeiten. Der Fokus dieser Arbeit liegt jedoch im Kernstück, dem Analyseprozess der empirischen Daten.

Im Scope

Innerhalb des Gesamtprozesses eines Assessments konzentriert sich diese Arbeit auf jene Prozessschritte, die der eigentlichen Erhebung empirischer Daten nachgelagert sind. Dazu zählen die Nachbearbeitung und Aufbereitung der qualitativen und quantitativen Daten, deren Analyse sowie die Erstellung des finalen Summaries, das wiederum als Lieferobjekt oder als Grundlage für die abschliessende Präsentation durch die Fachperson dient. Dieser Ausschnitt wird im Folgenden als Analyseprozess bezeichnet.

Nicht im Scope

Die administrativen und projektleitenden Tätigkeiten im Zusammenhang mit dem Assessment sowie die Datenerhebung selbst, sind nicht Teil der Masterarbeit selbst. Auch der Aufbau und die Durchführung der Interviews und der Umfrage sind nicht Gegenstand der Arbeit, ebenso wenig die abschliessende Präsentation der Ergebnisse und die zugehörigen Unterlagen.

Eine Übersicht des Prozessablaufs, mit out of Scope-Markierung ist nachfolgend dargestellt.

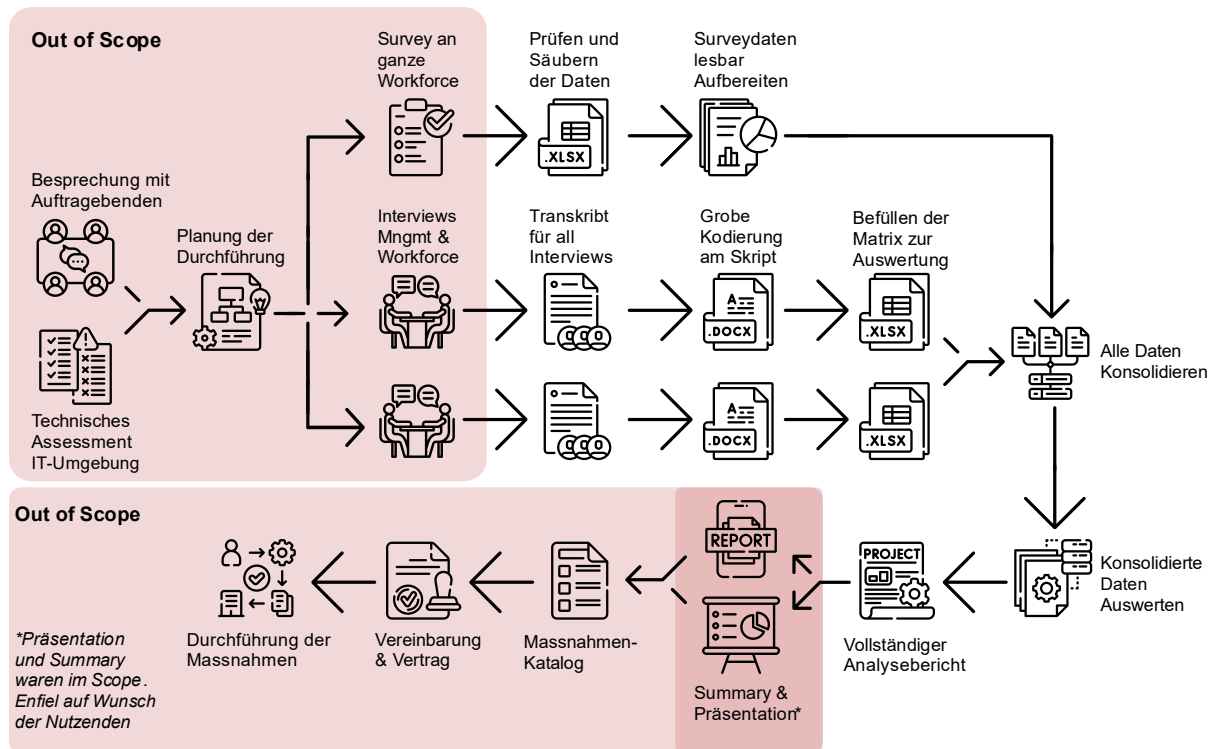


Abbildung 1: Übersicht des Evolution Assessments mit Markierungen zum Scope (Eigene Darstellung)

1.4.2. Scope Pilot-Projekt

Parallel zur Konzeption des Light Assessments wurde ein Pilot-Projekt für eine pseudonymisierte Kundenorganisation durchgeführt. Das Projekt umfasste eine Gruppe von fünf M365-Consultants sowie den Autor dieser Arbeit. Ziel des Pilot-Projekts war die Neukonzeption des Analyseprozesses im laufenden Projekt, weshalb der Zeitplan bewusst flexibel gehalten wurde.

Im Scope

Teil der Arbeit, ist die methodische Begleitung dieses Teams als Nutzergruppe. Die Mitglieder des Pilot-Projekts werden im Sinne des Participatory Design einbezogen. Die menschenzentrierten Methoden, die zur partizipativen Gestaltung des Analyseprozesses eingesetzt werden, sind Teil der Arbeit und werden entsprechend dargestellt. Die methodische Verortung des Participatory Design wird vertieft (siehe [Kapitel 4](#)).

Nicht im Scope

Die eigentlichen Projektarbeiten und die Rohdaten des Pilotprojekts sind nicht Teil der Masterarbeit. Dazu gehören Projektplanung, Korrespondenzen sowie die vollständigen erhobenen und ausgewerteten Kundendaten. Soweit Artefakte aus dem Pilotprojekt eine Interaktionsfunktion im Analyseprozess haben, werden ausgewählte exemplarische Ausschnitte mit Bezug zur Masterarbeit dokumentiert und im Text referenziert, sodass die Arbeit ohne NDA-Einschränkungen veröffentlicht werden kann.

1.4.3. Scope Prototyping

Das Produkt dieser Arbeit ist ein Prozess mit einer vorgegebenen Chat-Interaktionsschnittstelle. Das Ergebnis ist somit nicht ein eigenständiges KI-Werkzeug, sondern ein konzeptionell ausgearbeiteter Prozessablauf mit definierter Mensch-KI-Interaktion. Der zugehörige Prototyp ist primär ein konzeptioneller Prozessprototyp, der die Abfolge der Arbeitsschritte, die Verantwortlichkeiten und die Übergaben zwischen Fachperson und KI sichtbar macht. Ein einfacher Low-Fidelity-Prototyp zur isolierten Erprobung einzelner Prozessschritte wird als methodische Option offengehalten und je nach Bedarf in der Evaluation eingesetzt.

Im Scope

Der Fokus des Prototyps ist rein konzeptionell. Einzelne Teilschritte des Prozesses werden technisch umgesetzt, um die fachliche Funktionsweise nachvollziehbar zu machen. Diese technischen Umsetzungen haben den Charakter eines Proof of Concept und dienen der Veranschaulichung sowie der Erprobung kritischer Schritte, vergleichbar mit dem Bau eines Wire Flows oder Mock Ups.

Nicht im Scope

Technische Aspekte werden in der Arbeit nur erwähnt, falls sie für das Kontextverständnis notwendig sind. Nicht im Scope sind die eigenständige technische Entwicklung der KI-Komponenten, die Beurteilung des allgemeinen technologischen Reifegrads sowie die produktive Implementierung in bestehende Unternehmensprozesse. Der Prototyp erhebt entsprechend keinen Anspruch auf technische Reife oder vollumfängliche Funktionalität.

1.5. M365-Ökosystem und Copilot

In diesem Abschnitt werden grundlegende Aspekte der vorgegebenen Systemlandschaft um M365 erklärt. Diese Informationen dienen vorwiegend dazu, im Verlauf der Arbeit konzeptionelle Entscheidungen nachvollziehen zu können, welche hinsichtlich dieser Aspekte getroffen werden. Zentrale Fachbegriffe zu M365, KI und LLM sind im Fachglossar erläutert (siehe Anhang A, S. 103).

1.5.1. Microsoft 365

Microsoft 365 oder auch M365 genannt, bildet in dieser Arbeit die technische Grundlage für die spätere Umsetzung. Daher muss das Konzept auf die M365 Umgebung ausgerichtet werden. Bei M365 handelt es sich um eine cloudbasierte Plattform, die verschiedene Anwendungen für Zusammenarbeit und Produktivität bündelt, wie z. B. E-Mail, Dokumentenverwaltung, Chat und Kalender und weitere Programme, zur Steigerung der eigenen Produktivität. Alle Daten und Dienste werden innerhalb eines organisations-spezifischen Bereichs, dem sogenannten Tenant, eine organisatorische Einheit innerhalb von Microsoft 365, die eine eigene Datenhaltung, Benutzerverwaltung und Zugriffslogik umfasst, verwaltet. Dieser dient als abgegrenzte Umgebung, in der Benutzer, Inhalte und Berechtigungen zentral gesteuert werden. Der Zugriff auf Inhalte erfolgt über gemeinsame Identitäts- und Berechtigungsdienste, sodass Benutzer je nach Rolle auf relevante Daten zugreifen können.

1.5.2. Microsoft 365 Copilot

Es ist keine einzelne KI-Anwendung, sondern eine Orchestrierungsplattform, die drei Komponenten zusammenführt. Ein vortrainiertes Sprachmodell, meist ein GPT-basiertes Foundation Model aus dem Azure-OpenAI-Dienst, übernimmt die sprachliche und inferenzielle Verarbeitung. Der Microsoft Graph liefert die organisationsspezifischen Inhalte und Berechtigungen des Tenants, etwa E-Mails, Chats, Dokumente und Kalendereinträge. Die Microsoft-365-Anwendungen bilden die Umgebung, in der Copilot kontextbezogen eingebettet ist. Die Orchestrierung verknüpft diese Komponenten: Eine Anfrage wird mit Kontext aus dem Graph angereichert, vom Sprachmodell verarbeitet und im Anwendungskontext zurückgespielt.

Für die vorliegende Arbeit ist diese Architektur relevant, weil sie die Trennung von generativen und deterministischen Komponenten nahelegt. Generative Schritte eignen sich, wo Bedeutung interpretiert, Inhalte verdichtet oder Texte formuliert werden. Deterministische Komponenten, etwa Graph-Zugriffe, Programmlogik oder Tool-Aufrufe, sichern Verlässlichkeit bei exakten oder wiederholbaren Operationen. Diese Trennung wird in Konzeption und Prototyp übernommen.

1.5.3. Copilot Agents

Die Agents bilden ein zentrales Gestaltungsmittel innerhalb dieser Umgebung. Ein Agent ist nicht ein autonomes KI-System, sondern eine konfigurierte Dialogschnittstelle, die einem definierten Workflow folgt. Insbesondere die deklarative Variante erlaubt es, einen Prozess auf hoher Abstraktionsebene zu beschreiben, ohne jede einzelne Aktion vorzuprogrammieren. Der Orchestrator und das Sprachmodell übernehmen die Umsetzung in konkrete Schritte. Aus Gestaltungssicht ist entscheidend, dass ein Agent innerhalb klar gesetzter Grenzen arbeitet und in der Regel an definierten Punkten Rückmeldung oder Freigabe der nutzenden Person einholt. Diese Eigenschaft ist für die Mensch-KI-Aufgabenteilung im Analyseprozess unmittelbar tragend.

1.5.4. Datenschutz und Governance

Die Datenflüsse im Copilot-Ökosystem bleiben an den Tenant gebunden, eine Anfrage verlässt die geschützte M365-Umgebung nie. Der Graph vermittelt semantisch relevante Inhalte und respektiert bestehende Berechtigungen, eine Person sieht über Copilot nur, worauf sie auch sonst Zugriff hat. Ein- und Ausgaben fließen weder in die Modellweiterentwicklung ein, noch sind sie für andere Tenants sichtbar. Für eine Arbeit mit realen, unter NDA stehenden Kundendaten ist diese Datenbindung eine wichtige Voraussetzung.

Der Governance-Rahmen erweitert dieses Prinzip. Copilot unterliegt denselben Compliance-Zusagen wie die übrigen M365-Dienste: EU Data Boundary, vertraglich zugesicherte DSGVO-Konformität, Audit-Logs über Microsoft Purview sowie bestehende Sicherheitsrichtlinien wie Conditional Access, mehrstufige Authentifizierung und Data Loss Prevention. Inhaltsfilter prüfen ein- und ausgehende Prompts. Anforderungen an Vertraulichkeit, Nachvollziehbarkeit und Verantwortungsfähigkeit werden so teilweise durch die Plattform selbst adressiert.

Daraus folgen zwei Konsequenzen. Erstens entlastet die Plattform einen Teil der Datenschutz-Vorkehrungen und macht die prototypische Arbeit mit realen Kundendaten erst tragfähig. Zweitens bleiben die Grenzen der generativen Komponente bestehen: Outputs sind probabilistisch, nicht vollständig reproduzierbar und halluzinationsanfällig. Das begrenzte Kontextfenster erzwingt eine implizite Vorauswahl durch den Orchestrator, Format- und Konventionsanforderungen müssen über Prompts oder deterministische Nachverarbeitung gesichert werden und stille Modell-Updates können Verhalten und Ergebnisqualität unbemerkt verändern. Die Konzeption muss dies durch Aufgabenteilung zwischen generativen und deterministischen Schritten, prüfbare Zwischenergebnisse und fachliche Kontrollpunkte abfedern.

1.6. Rahmenbedingungen: Datenschutz, Vertraulichkeit, Anonymisierung, Zugänge

Der Autor steht unter den geschäftsweiten NDAs von GARAIO und hat dadurch vollen Zugriff auf reale Daten von Kunden. Die tatsächliche Nutzung und Verarbeitung dieser Daten im gesamten Projektkontext, betrifft jedoch fast ausschliesslich von der eigentlichen Masterarbeit ausgeschlossene Bereiche. Diese Rohdaten werden daher für die Nachvollziehbarkeit und Wertung nicht offengelegt. Für punktuellen Einsatz, wie dem Demonstrieren von arbeitsrelevanten Ergebnissen oder um Entscheidungen zu begründen, werden die notwendigen exemplarischen Auszüge der Daten anonymisiert, um Rückschlüsse auf Firmen und Personen ausschliessen zu können.

Die Firma GARAIO behält sich das Recht vor, in Absprache mit dem Autor gewisse Details zu firmeninternen Abläufen für die Veröffentlichung entweder zu pauschalisieren oder in der Arbeit nicht näher zu erörtern, um die Arbeit an sich uneingeschränkt verfügbar zu machen. Solche Anpassungen betreffen jedoch keine HCD-fachspezifischen Inhalte.

2. Fragestellung und Zielsetzung

2.1. Zielsetzung und Forschungsfrage

Aus der Aufgabenstellung leiten sich Forschungsfrage, Projektziel und Teilziele sowie persönliche Lernziele ab. Die folgenden Abschnitte überführen die zuvor entwickelte Problemstellung in eine klar formulierte Leitfrage und in nachvollziehbar geordnete Teilfragen. Sie definieren das angestrebte Projektergebnis und benennen die fachlichen Lerneffekte, die mit der Arbeit verbunden sind.

2.1.1. Forschungsfrage

Die übergreifende Leitfrage der Arbeit lautet:

Wie kann ein KI-unterstützter Analyseprozess zur systematischen Auswertung qualitativer und quantitativer empirischer Daten im Rahmen des M365-Evolution-Assessments im Microsoft-365-Ökosystem konzipiert und evaluiert werden, sodass der Analyseaufwand reduziert, die Ergebnisverlässlichkeit gewahrt und die fachliche Verantwortung beim Fachpersonal belassen wird?

Diese Leitfrage wird durch sechs HCD-relevante Teilfragen geschärft, die jeweils einen spezifischen Beitrag der Arbeit adressieren:

- Welche Schwachstellen, Anforderungen und Automatisierungspotenziale weist der bestehende Analyseprozess auf?
- Welche Aufgaben eignen sich im Analyseprozess für eine KI-Unterstützung und welche Schritte sollten aus Gründen der Qualität, Nachvollziehbarkeit und Verantwortung beim Fachpersonal verbleiben?
- Wie ist ein KI-unterstützter Soll-Prozess für die Analyse empirischer Daten im M365-Evolution-Assessment zu gestalten?
- Inwieweit verbessert ein prototypisch ausgestalteter KI-unterstützter Analyseprozess die Nutzbarkeit und den Effizienzgrad für M365-Fachpersonen?
- Inwieweit liefert der prototypisch ausgestaltete Analyseprozess über unterschiedliche Datensätze hinweg hinreichend verlässliche Ergebnisse?
- Welche Gestaltungsprinzipien und welches Vorgehensmodell lassen sich aus Konzeption und Evaluation ableiten?

2.1.2. Projektziel und Teilziele

Das übergeordnete Projektziel besteht in der Entwicklung und empirischen Evaluation eines KI-unterstützten Analyseprozesses für die Auswertung von M365-Evolution-Assessments im M365-Ökosystem. Untersucht wird, inwiefern ein solcher Prozess den für die Analyse erforderlichen Arbeitsaufwand reduzieren und gleichzeitig eine hohe Ergebnisverlässlichkeit sicherstellen kann. Ergebnisverlässlichkeit bezeichnet dabei die Genauigkeit der Aufbereitung und Analyse empirischer Daten sowie die Stabilität der Resultate bei wiederholter Analyse identischer Datensätze. Auf dieser Grundlage wird ein Vorgehensmodell zur KI-unterstützten empirischen Datenanalyse im M365-Ökosystem abgeleitet.

Aus dem Projektziel ergeben sich die folgenden Teilziele:

- Durchführung einer IST- und Bedarfsanalyse sowie Modellierung des bestehenden Analyseprozesses, um relevante Prozessschritte, Anforderungen und Einsatzpotenziale für KI-Unterstützung systematisch zu verorten
- Iterative Konzeption und prototypische Ausgestaltung eines KI-unterstützten Analyseprozesses unter Nutzung von M365-Werkzeugen und Copilot-Funktionalitäten
- Evaluation der Nutzbarkeit des prototypischen Analyseprozesses mit M365-Fachpersonen im Rahmen von User-Testings
- Evaluation der Effizienz anhand des Vergleichs bisher dokumentierter Arbeitsstunden aus vergangenen Assessments mit dem im Testing erhobenen Arbeitsaufwand sowie unter Bezug auf die unternehmensseitige Zielgrösse von maximal rund zwanzig Arbeitsstunden
- Evaluation der Ergebnisverlässlichkeit der automatisierten Analyse mit geeigneten Konsistenzmassen, um die Stabilität der Resultate bei wiederholter Analyse identischer Datensätze zu beurteilen
- Ableitung eines reflektierten Vorgehensmodells zur KI-unterstützten empirischen Datenanalyse im M365-Ökosystem

2.2. Lernziele

Über das Projektziel hinaus verfolgt die Arbeit drei fachliche Lern- und Erkenntnisziele:

- Vertiefung des Verständnisses für die Analyse empirischer Daten unter Einbezug KI-gestützter Auswertungsschritte und deren systematischer Überprüfung
- Vertiefung des Verständnisses für den menschenzentrierten Einsatz von KI-Systemen in der Nutzerforschung, insbesondere im Hinblick auf Potenziale, Grenzen und Risiken
- Aufbau praktischer Erfahrung in der Konzeption und prototypischen Ausgestaltung eines Analyseprozesses auf Basis von LLM-Schnittstellen sowie in der Planung, Durchführung und Auswertung entsprechender User-Testings

3. Literaturvertiefung

3.1. Vorgehen bei der Quellenrecherche und Quellenkritik

3.1.1. Quellenarten

Die Auswahl der herangezogenen Literatur folgt einer bewussten Schichtung entlang zweier Kriterien, nämlich der thematischen Nähe zum Untersuchungsgegenstand und der Aktualität des Forschungsstandes. Für den unmittelbaren Untersuchungsgegenstand, dem KI-unterstützten Analyseprozess im M365-Ökosystem, liegt bislang keine etablierte Forschungstradition mit kanonischen Standardwerken vor. Der Gegenstand verbindet mehrere Teilbereiche, die jeweils unterschiedlich weit entwickelt sind. Menschzentrierte Gestaltung und qualitative Analyseverfahren verfügen über etablierte Grundlagen. Der spezifische Einsatz generativer KI bzw. LLM-basierter Systemen und von M365-Copilot-Funktionalitäten in empirischen Analyseprozessen ist dagegen ein junges und dynamisches Feld. Daraus ergibt sich für die Arbeit eine Mischung aus etablierten Standardwerken, peer-reviewten Konferenz- und Journalbeiträgen sowie ausgewählten Preprints, Praxisberichten und Plattformquellen.

Standardwerke

Für konzeptionelle und methodische Grundlagen der menschenzentrierten Gestaltung wird auf etablierte Standardwerke und normative Grundlagen zurückgegriffen. Deren Aussagen sind weitgehend unabhängig von kurzfristigen technologischen Entwicklungszyklen gültig und stützen den Rahmen, in dem der KI-unterstützte Analyseprozess konzipiert wird.

Zeitaktuelle Arbeiten

Für die KI-bezogene Argumentation werden hingegen aktuelle Beiträge stärker gewichtet. Der Forschungsstand zu LLM-basierten Systemen entwickelt sich in kurzen Zyklen und das Verhalten solcher Systeme kann sich auch innerhalb derselben Modellfamilie über kurze Zeiträume verändern. Chen et al. (2023) zeigen dies anhand zeitlich getrennter Snapshots von GPT-3.5 und GPT-4, bei denen identische Aufgaben zu substantiell unterschiedlichen Ergebnissen führten. Für die Quellenarbeit folgt daraus eine erhöhte Sensibilität gegenüber dem Publikationsdatum und gegenüber Aussagen zu konkretem Modellverhalten. Auch die Arbeit mit LLM-basierten Systemen im Research-Bereich, namentlich Vibe Research, ist zu neu, um Standardwerke zu liefern, doch sind bereits genug vorhanden, um für ein neues Feld eine ausreichend belastbare Quellenlage zu bieten.

3.1.2. Quellenkritik

Die Literaturarbeit stützt sich aufgrund dieser Konstellation auf einen zusammengesetzten Bezugsrahmen. Etablierte Grundlagen werden dort herangezogen, wo sie methodische oder gestalterische Prinzipien absichern. Für direkte Aussagen zum KI- und M365-spezifischen Kontext werden wie erwähnt (siehe [Abschnitt 3.1.1](#)) aktuelle Beiträge berücksichtigt, sofern sie einen klaren thematischen oder methodischen Bezug zur Arbeit aufweisen.

Methodisch lehnt sich dieses Vorgehen an die Empfehlungen von Webster und Watson (2002) zu einer konzeptzentrierten Literaturarbeit sowie an vom Brocke et al. (2009) zur transparenten Dokumentation des Suchprozesses an. Die Auswahllogik wird damit nicht als Vollständigkeitsanspruch verstanden, sondern als begründete Fokussierung auf jene Beiträge, die für die spezifische Forschungsfrage den stärksten Belegwert liefern. Die vollständige Quellenkritik findet sich (siehe Anhang B, S. 108).

3.2. KI- und LLM-gestützte Analyse: Stand der Technik und Forschung

3.2.1. LLM und Foundation Models

Der Forschungsstand zur KI-gestützten Analyse qualitativer und quantitativer Daten ist eng mit der Entwicklung LLM-basierter Systeme verbunden. Bommasani et al. (2021) beschreiben diese Entwicklung mit dem Begriff der Foundation Models und meinen damit breit vortrainierte Modelle, die mit vergleichsweise wenig Anpassung für unterschiedliche Aufgaben eingesetzt werden können. Dazu gehört auch die Verarbeitung von Interviewtranskripten, offenen Umfrageantworten und weiteren textbasierten Forschungsdaten. Für die vorliegende Arbeit sind solche Modelle nicht als Ersatz etablierter empirischer Analyseverfahren zu verstehen. Klassische Verfahren, insbesondere die qualitative Inhaltsanalyse nach Mayring (2022), liefern einen methodischen Bezugspunkt. An ihnen müssen sich KI-gestützte Verarbeitungsschritte messen lassen. Die ausführliche methodische Verortung erfolgt in den späteren Kapiteln. Der vorliegende Abschnitt fasst vor allem zusammen, welche Chancen, Grenzen und Risiken die aktuelle Forschung für KI-unterstützte Analyseprozesse aufzeigt.

3.2.2. Chancen

Die Chancen werden vor allem dort sichtbar, wo sich wiederkehrende Kodier- und Strukturierungsschritte teilweise entlasten lassen. KI kann grössere Datenmengen schneller erschliessen, Muster sichtbar machen und erste Verdichtungen vorbereiten. Stige et al. (2024) zeigen in einer systematischen Literaturübersicht

zu KI im UX-Design, dass KI-gestützte Verfahren in einzelnen Schritten der Nutzerforschung Effizienz und Innovationsspielraum erhöhen können.

Lu et al. (2024a) präzisieren diesen Befund aus einer Human-Centered-AI-Perspektive. UX-Forschende nutzen KI demnach vor allem für klar begrenzte, unterstützende Aufgaben. Die übergreifende analytische Einordnung, die empathische Interpretation und die Verantwortung für die Ergebnisse bleiben dagegen beim Fachpersonal. Praxisnahe Erfahrungsberichte zeigen zudem ein wiederkehrendes Muster. Erzielten Zeitgewinnen im eigentlichen Kodiervorgang steht ein erheblicher Aufwand für Prompt-Entwicklung und manuelle Validierung gegenüber, was die Nettoeffizienz relativiert. Konkrete Workflow-Beispiele werden vertieft (siehe [Abschnitt 3.3](#)).

3.2.3. Grenzen und Risiken

Diesen Potenzialen stehen Grenzen gegenüber, die für die Konzeption eines KI-unterstützten Analyseprozesses zentral sind. Bender et al. (2021) bezeichnen LLM-basierte Systeme als «stochastic parrots» und beziehen sich damit auf die Wahrscheinlichkeitslogik, auf welcher solche Systeme basieren. Damit kritisieren sie, dass Modelle sprachliche Muster erzeugen können, ohne zwingend über ein tragfähiges Verständnis der Bedeutung zu verfügen. Für Analyseprozesse entsteht daraus ein Risiko, weil Modellausgaben plausibel wirken können, obwohl sie verzerrt, unvollständig oder nicht ausreichend belegt sind. Zudem kann unklar werden, wer für die fachliche Qualität eines Ergebnisses verantwortlich ist.

Die Halluzinationsforschung konkretisiert dieses Risiko. Huang et al. (2024) unterscheiden zwischen intrinsischen Widersprüchen zum Prompt und extrinsischen Widersprüchen zur Faktenlage. Sie diskutieren Verfahren zur Erkennung und Begrenzung solcher Fehler, weisen aber zugleich auf deren strukturelle Grenzen hin. Kalai et al. (2025) erklären Halluzinationen als Spezialfall binärer Klassifikationsfehler. Ihre Analyse zeigt, dass heutige Bewertungssysteme das Raten häufig stärker belohnen, als das explizite Ausdrücken von Unsicherheit. Für die Prozessgestaltung folgt daraus, dass Falschaussagen nicht als Ausnahmefehler behandelt werden dürfen, sondern als grundsätzlich möglich eingeplant und durch geeignete Sicherungen adressiert werden müssen. Eng damit verbunden ist die Frage der Reproduzierbarkeit. Chen et al. (2023) zeigen empirisch, dass GPT-3.5 und GPT-4 zwischen März und Juni 2023 bei identischen Aufgaben deutlich unterschiedliche Genauigkeiten erreichten. In einem Beispiel zur Erkennung von Primzahlen sank die Erfolgsquote von GPT-4 von 84 auf 51 Prozent, obwohl die Modellbezeichnung gleich blieb. Cui und Alexander (2026) bestätigen diese Problematik für datenanalytische Aufgaben. Über 480 Versuche mit sechs Modellen, vier Temperatureinstellungen und zwei Prompt-Strategien hinweg zeigen sie, dass auch bei identischer Konfiguration und Temperatur null unterschiedliche Resultate auftreten können.

Für die methodische Gestaltung bedeutet dies, dass eine einzelne LLM-gestützte Analyse nur begrenzt belastbar ist. Ergebnisse müssen wiederholbar geprüft, miteinander verglichen und durch fachliche Sichtung abgesichert werden. Mehrere unabhängige Ausführungen, systematische Vergleichsanalysen und klare Validierungsschritte sind deshalb keine optionalen Ergänzungen, sondern ein notwendiger Teil eines verlässlichen KI-gestützten Analyseprozesses.

3.2.4. Mensch-KI-Aufgabenteilung

Diese Befunde stehen in engem Zusammenhang mit der Diskussion zur Mensch-KI-Aufgabenteilung. Amershi et al. (2019) formulieren in den Guidelines for Human-AI Interaction achtzehn Leitlinien für die Gestaltung interaktiver KI-Systeme. Für die vorliegende Arbeit besonders relevant sind Anforderungen an Transparenz, Korrekturmöglichkeiten und den Umgang mit Unsicherheit. KI-Systeme sollen ihre Fähigkeiten und Grenzen nachvollziehbar machen. Nutzende müssen Fehler effizient korrigieren und Änderungen im Systemverhalten einordnen können.

Eine konzeptionelle Grundlage für diese Aufgabenteilung liefern zwei weitere Beiträge. Mosqueira-Rey et al. (2023) ordnen Human-in-the-Loop in einer systematischen Übersicht und unterscheiden Active Learning, Interactive Machine Learning und Machine Teaching nach dem Kriterium, wer den Lernprozess steuert. Shneiderman (2020) ergänzt diese Linie um eine programmatische Perspektive der Human-Centered AI und formuliert fünfzehn Empfehlungen für Reliable, Safe und Trustworthy AI auf den Ebenen Team, Organisation und Industrie. Beide Beiträge stützen die in dieser Arbeit verfolgte Logik, hohe menschliche Kontrolle mit hoher Automatisierung zu verbinden, ohne dass beides einander ausschließt. Matter et al. (2024) verschärfen diese Perspektive. Nutzende geraten in der Interaktion mit GenAI-Systemen teilweise in eine Assistenzrolle. Dabei bleibt unklar, wie viel Verantwortung sie gegenüber einem Output tatsächlich übernehmen können, wenn weder sie selbst noch das System ein vertieftes Verständnis des verarbeiteten Inhalts besitzen. Die Autoren beschreiben diese Konstellation als «Double Chinese Room». Für den Analyseprozess folgt daraus, dass die Rolle des Fachpersonals nicht nur formal vorgesehen sein darf. Sie muss so gestaltet werden, dass echte Kontrolle, Korrektur und fachliche Verantwortung möglich bleiben.

Für den vorliegenden Untersuchungsgegenstand ergibt sich aus diesen Befunden ein klarer Schluss. KI kann die Strukturierung, das Kodieren und die Verdichtung von Daten unterstützen, sie kann diese Aufgaben aber nicht eigenständig verantworten. Die Ergebnisverlässlichkeit hängt deshalb von einer bewusst gestalteten Mensch-KI-Aufgabenteilung ab. Zusätzlich braucht es methodische Vorkehrungen gegen Halluzinationen, gegen instabile Resultate und gegen eine unklare Verantwortungszuordnung.

3.2.5. Context Engineering

In dieser Arbeit wird Context Engineering als die bewusste, aufgabenbezogene Zusammenstellung, Strukturierung und situative Bereitstellung jenes Bezugsmaterials verstanden, mit dem ein LLM-basiertes System eine konkrete Aufgabe bearbeiten soll (Mei et al. 2025, Anthropic 2025, LangChain o. J.). Der Ansatz geht über Prompt Engineering hinaus: Während Prompt Engineering primär auf die Formulierung wirksamer Instruktionen zielt, fokussiert Context Engineering darauf, welche Informationen, Beispiele, Regeln, Werkzeuge und Ausgangsdokumente einem Modell in welcher Form bereitgestellt werden müssen, damit eine Aufgabe nachvollziehbar und konsistent bearbeitet werden kann. Damit wird nicht nur die sprachliche Anweisung gestaltet, sondern auch das Material, auf das sie sich bezieht, als eigenständige methodische Grundlage behandelt.

Für das methodische Vorgehen lassen sich daraus drei Grundprinzipien ableiten. Erstens wird pro Aufgabe ein kuratierter Kontext-Pool gepflegt, der nur relevante, überprüfbare und fachlich tragfähige Inhalte enthält. Zweitens werden Anweisungen aufgabengerecht angereichert, indem passende Ausschnitte daraus bereitgestellt werden, etwa Definitionen, Ankerbeispiele, methodische Vorgaben oder Ausgangsdokumente. Drittens wird der Kontext situativ bereitgestellt, weil er einem LLM-basierten System beim einzelnen Aufruf nicht automatisch vollständig zur Verfügung steht, sondern gezielt mitgegeben, abgerufen oder verdichtet werden muss.

Hinweis zur Quellenlage

Context Engineering und Harness Engineering sind junge Begriffe, deren wissenschaftliche Konsolidierung noch nicht abgeschlossen ist. Die herangezogenen Quellen umfassen einen wissenschaftlichen Survey im Preprint-Status (Mei et al. 2025), einen technischen Beitrag des LLM-Anbieters Anthropic (2025) sowie einen wissenschaftlichen Preprint zu AI Harness Engineering (Zhong und Zhu 2026). Diese Quellenlage reflektiert den aktuellen Diskurs, der derzeit wesentlich durch industrielle Praxis und erste akademische Systematisierungsversuche geprägt ist.

3.3. Praxisrecherche und State of Practice

Die Praxisrecherche ist als kompakte Konkurrenzanalyse aufgebaut. Sie umfasst keine vollständige Marktübersicht, sondern stellt bewusst zwei Lernfälle in den Vordergrund. Beide Fälle stammen aus dem Schweizer Hochschulkontext und behandeln den praktischen Einsatz von KI in Forschungs- oder Analysearbeiten. Sie eignen sich deshalb besonders als Vergleichsobjekte für das vorliegende Vorhaben. Im Fokus stehen jene Parallelen, aus denen sich für die Konzeption und die prototypische Umsetzung konkrete Übernahmen begründen lassen.

3.3.1. Lernfall 1: Fedosov, Kern und Barth (2026)

Das Projekt

Fedosov et al. (2026) dokumentieren im Rahmen eines Innosuisse-geförderten Projekts zu Active and Assisted Living einen dreistufigen LLM-gestützten Workflow für die qualitative Analyse von Interviewdaten in Schweizerdeutsch. Der Workflow gliedert sich in eine Audio-zu-Text-Transkription mit dem auf Schweizerdeutsch spezialisierten Werkzeug stt4sg, eine KI-gestützte Sprechersegmentierung mit explizit nicht-generativem Prompt und ein KI-gestütztes qualitatives Coding mit strukturiertem Output. Die Autoren begründen die phasenweise Trennung der Schritte mit der Reduktion von Fehlerfortpflanzung und mit der besseren Möglichkeit gezielter Prompt-Gestaltung.

Das Learning

Aus diesem Lernfall lassen sich für die vorliegende Arbeit vier konkrete Übernahmen ableiten. Erstens bestätigt der Fall die Phasentrennung als Designprinzip. Sie erhöht die Lokalisierbarkeit von Fehlern und erlaubt die gezielte Optimierung einzelner Prompts. Zweitens stützt der Fall den Codebook-basierten Zugang. Fedosov et al. (2026) empfehlen einen hybriden Ansatz mit vordefiniertem Codebook und kontrollierter Erweiterung. Die vorliegende Arbeit übernimmt diese Logik in einer strengeren, vollständig deduktiven Variante, weil der Anschluss an den verbindlichen Kodierungsrahmen des bestehenden M365-Evolution-Assessments erforderlich ist. Drittens wird die Trennung zwischen semantischen und formalen Aufgaben übernommen. Fedosov et al. (2026) setzen den Segmentierungsschritt strikt nicht-generativ ein. Die vorliegende Arbeit überträgt dieses Prinzip auf die Architekturebene, indem deterministische Skripte alle formalen Transformationen verantworten und Sprachmodellläufe ausschliesslich semantische Aufgaben übernehmen. Viertens wird die Validierungsverantwortung beim Fachpersonal festgehalten. Der Lernfall zeigt, dass Forschende trotz weitgehender Automatisierung für Oversight, Prompt-Design, Validierung und Interpretation verantwortlich bleiben.

3.3.2. Lernfall 2: Matter, Wiedemann, Leikert-Böhm und Pleger (2024)

Das Projekt

Matter et al. (2024) verdichten Beobachtungen aus fünf Innosuisse-geförderten Anwendungsprojekten zum praktischen Einsatz von GenAI in unterschiedlichen Domänen. Konkret untersuchen sie GenAI als UX-Researcher bei der Erstellung einer Customer Journey für eine Schweizer Bank, GenAI als IT-Security-Berater im Rahmen eines Ransomware-Readiness-Assessments, GenAI als Enabler für Prozessdigitalisierung in KMU, GenAI in Knowledge Management bei einem Immobilienunternehmen und GenAI für

Texterzeugung in einer Hochschulübung. In allen Settings waren die Outputs der GenAI-gestützten Konfigurationen objektiv mindestens so gut wie jene der traditionellen Vergleichsfälle, teilweise mit deutlichem Effizienzgewinn. Die Autoren beobachten zugleich, dass Nutzende subjektiv häufig in eine Subordinationsrolle gegenüber dem System geraten, mit wachsender Frustration und Rückzug aus der aktiven Kontrolle.

Das Learning

Aus diesem Lernfall lassen sich drei Übernahmen ableiten. Erstens bestätigt der Fall die Notwendigkeit einer bewussten Mensch-KI-Aufgabenteilung. Die fachliche Entscheidungshoheit verbleibt in der vorliegenden Arbeit explizit beim M365-Fachpersonal und wird über die Bewertungstabellen als sichtbarer Übergabepunkt institutionalisiert. Zweitens stützt der Fall die Unterscheidung zwischen prozeduralem und konzeptuellem Wissen. Matter et al. (2024) beobachten, dass GenAI-Unterstützung bei prozeduralem Wissen hohen subjektiven Nutzen stiftet, während konzeptuelles Wissen ambivalent bleibt. Die vorliegende Arbeit setzt diese Trennung in der Architektur um, indem deterministische Skripte prozedurale Aufgaben übernehmen und KI-Schritte ausschliesslich semantische Verdichtung leisten. Drittens stützt der Fall den methodisch differenzierten Umgang mit dem Ziel der Effizienzsteigerung. Effizienzgewinne sind real, werden aber durch Subjektivempfinden und Verantwortungsfragen relativiert. Die vorliegende Arbeit nimmt diese Relativierung methodisch auf, indem Effizienz, Nutzbarkeit und Ergebnisverlässlichkeit als gleichrangige Evaluationsdimensionen behandelt werden.

Eine methodisch begründete Lücke gegenüber dem Lernfall bildet die Breite der Domänen. Matter et al. (2024) werten fünf unterschiedliche Anwendungsdomänen aus, während die vorliegende Arbeit auf eine einzige, klar umrissene Domäne fokussiert. Diese Verengung ermöglicht die vertiefte Auseinandersetzung mit der konkreten Anwendung im M365-Evolution-Assessment.

3.3.3. Übernahmen in die eigene Arbeit

Beide Lernfälle weisen mehrere Überschneidungen auf, welche für die vorliegende Arbeit zusätzlich relevant sind. Beide betonen die Notwendigkeit, die Validierung von Modelloutputs beim Fachpersonal zu belassen. Beide diskutieren den Aufwand für Prompt-Engineering als methodisch zentralen Faktor, der die Nettoeffizienz relativiert. Beide weisen auf die strukturelle Limitierung von Reproduzierbarkeit und Konsistenz bei LLM-Outputs hin. Beide behandeln die Versionsabhängigkeit von Sprachmodellen als methodisches Risiko und beide reflektieren das Spannungsfeld zwischen Effizienzgewinn und Verlässlichkeitsrisiko, das die Zielsetzung der vorliegenden Arbeit prägt.

Daraus ergeben sich für die folgenden Kapitel konkrete Verankerungen. In der Konzeption wird die Phasentrennung als Strukturprinzip festgelegt und durch klare Übergaben zwischen den Phasen institutionalisiert. Der deduktive Kodierungsrahmen bildet die Grundlage des Codings, und die Trennung zwischen semantischen und formalen Aufgaben wird als architektonische Entscheidung umgesetzt. In der prototypischen Umsetzung führen diese Entscheidungen zu einer sequenziellen Verarbeitung mit prüfbaren Zwischenergebnissen. Sprachmodellläufe übernehmen ausschliesslich semantische Verdichtung, deterministische Skripte realisieren alle formalen Transformationen. Die fachliche Entscheidungshoheit wird über Bewertungstabellen institutionalisiert. In der Evaluation wird die in beiden Lernfällen dokumentierte Konsistenz- und Reproduzierbarkeitsproblematik durch eine Prüfung mehrfacher Durchläufe adressiert.

3.4. Begriffliche Einordnung: Vibe Research

Der Begriff ist neu und bislang nicht als standardisierte wissenschaftliche Methode etabliert. Auf Grund der Erkenntnisse aus der Literaturforschung und der Natur des Projektrahmens, kann Vibe Research als ein gültiges Vorgehen akzeptiert werden. Doch dafür muss der Begriff, vor allem für den Kontext dieser Arbeit erst geprüft und im Methodenkapitel das Vorgehen an sich definiert werden, um eine bewertbare Nachvollziehbarkeit zu ermöglichen.

Vibe Research bezeichnet in dieser Arbeit einen KI-gestützten Arbeitsmodus für die Aufbereitung, Strukturierung und Auswertung empirischer Daten im Rahmen menschzentrierter Entwicklung. Er beschreibt einen Arbeitsmodus, in dem Forschende Zielsetzung, Analyseplan und Bewertungsmassstäbe festlegen, während KI-Systeme unterstützend eingesetzt werden, um Datenmaterial aufzubereiten, Strukturierungs- und Kodiervorschläge zu erzeugen, Konsistenz über das Material zu prüfen, Zwischenergebnisse zu verdichten und Gegenperspektiven sichtbar zu machen. Begrifflicher Ausgangspunkt ist der Beitrag von Bertina (2025), der Vibe Research als Weiterentwicklung KI-gestützter Deep-Research-Ansätze beschreibt und definiert als:

„AI-augmented research workflow where autonomous agents (driven by LLMs, transformers, and ‚vibe-coding‘ techniques) identify knowledge gaps, formulate hypotheses, gather or generate data, test those hypotheses, and synthesize findings, with minimal human intervention beyond initial guidance and oversight“ (Bertina 2025)

Die Bezeichnung knüpft an das von Karpathy geprägte «Vibe Coding» an und überträgt die Logik einer auf hoher Abstraktionsebene gesteuerten, durch KI ausgeführten, Bearbeitung auf den Forschungsprozess. Bertina unterscheidet drei Modi, nämlich klassische, menschlich getriebene Forschung, KI-unterstützte Deep Research und KI-getriebene, agentenbasierte Vibe Research mit menschlicher Aufsicht. Sein Verständnis umfasst dabei den gesamten Forschungszyklus einschliesslich Datenerhebung, Datenanalyse und Synthese. Für die vorliegende Arbeit wird dieser Zyklus bewusst eingegrenzt. Massgeblich ist ein Teil der Datenerhebung und vor allem die Datenanalyse, da in der Arbeit empirische Daten aufbereitet, analysiert und ausgewertet werden.

3.4.1. Positionierung im Rahmen dieser Arbeit

Vibe Research wird in dieser Arbeit nicht im stark autonomen Sinn Bertinas verwendet, sondern als unterstützender Arbeitsansatz innerhalb eines reflektierten methodischen Vorgehens. Massgeblich ist somit nicht eine möglichst weitgehende Delegation des Forschungsprozesses, sondern der gezielte Einsatz generativer KI für kodifizierbare Teilaufgaben der Datenaufbereitung, Strukturierung und Verdichtung. Die KI dient nicht als eigenständige Evidenzquelle und liefert keine eigenständigen Befunde. Sie wird als Werkzeug eingesetzt, um umfangreiches Datenmaterial schneller zu erschliessen, Perspektiven zu erweitern, blinde Flecken sichtbar zu machen und die Entwicklung einer tragfähigen Argumentationsstruktur zu unterstützen. Zielsetzung, fachliche Beurteilung, Datenkritik, Interpretation und Verantwortung verbleiben durchgängig bei den Fachpersonen. Sämtliche aus KI-Ausgaben übernommenen Aussagen werden am Originalmaterial geprüft, fachlich plausibilisiert und eigenständig eingeordnet.

Der Bezug zur menschzentrierten Entwicklung ergibt sich aus der Funktion der analysierten Daten. Empirische Daten über Nutzende, Aufgaben und Nutzungskontext, etwa aus Interviews, Beobachtungen oder Befragungen, bilden die Grundlage menschzentrierter Gestaltungsentscheidungen. Fehler oder Verzerrungen in der Analyse wirken sich deshalb unmittelbar auf die Qualität dieser Entscheidungen aus. Daraus ergibt sich der hohe Stellenwert von Nachvollziehbarkeit und Ergebnisverlässlichkeit, der das hier definierte Verständnis von Vibe Research prägt.

3.4.2. Anschluss an den Forschungsstand

Der Kerngedanke von Vibe Research lässt sich entlang mehrerer Forschungsstränge stützen. Wang et al. (2023) zeigen, dass KI zunehmend entlang verschiedener Phasen des Forschungsprozesses eingesetzt wird, etwa bei der Hypothesenbildung, beim Experimentdesign, bei der Datenauswertung und bei der Interpretation grosser Informationsmengen. Ren et al. (2025) beschreiben in einem Überblick zu LLM-basierten Scientific Agents, dass grosse Sprachmodelle nicht nur Texte generieren, sondern in agentischen Systemen mit definierten Rollen, Werkzeugnutzung, Planung, Feedbackschleifen und Validierungsmechanismen verbunden werden können. Damit wird der Einsatz von KI als strukturierendes Element der

Datenanalyse methodisch anschlussfähig. Dieser Aspekt hat insofern eine Bedeutung für diese Arbeit, da hier ein Research-Tool für Nicht-Researcher erarbeitet wird, was, wenn auch nicht tatsächlich agentisch im technischen Sinne, zumindest konzeptionell im Sinne der agentischen Systeme eine Form von internem Validierungsmechanismus benötigt.

Konkrete Systeme illustrieren, was diese Anschlussfähigkeit in der Praxis bedeutet. Schmidgall et al. (2025) zeigen mit Agent Laboratory, dass LLM-basierte Agentensysteme für zusammenhängende Forschungsprozesse einsetzbar sind, die experimentelle und analytische Arbeitsschritte sowie Berichtserstellung umfassen, wobei menschliches Feedback die Ergebnisqualität nachweisbar verbessert. Lu et al. (2024b) berichten mit «The AI Scientist» über ein System, das den Anspruch erhebt, einen vollständigen Forschungszyklus weitgehend autonom auszuführen, und machen dabei zugleich die Grenzen eines vollautomatisierten Ansatzes sichtbar. Diese Arbeiten stützen den Einsatz von Vibe Research insbesondere dort, wo es um die Strukturierung umfangreichen Datenmaterials, die Erzeugung überprüfbarer Zwischenergebnisse und die iterative Verfeinerung mit menschlichem Feedback geht. Sie begründen zugleich, weshalb eine fachliche Letztprüfung nicht entfallen darf.

3.4.3. Risiken und methodische Grenzen

Der Einsatz von Vibe Research ist methodisch kritisch zu begrenzen. Generative KI kann plausible, aber falsche Aussagen erzeugen, Datenmaterial unvollständig erfassen, bestehende Verzerrungen reproduzieren und Scheinsicherheit vermitteln. Messeri und Crockett (2024) verweisen in diesem Zusammenhang auf das Risiko sogenannter «illusions of understanding». KI-Systeme können den Eindruck erzeugen, ein Datenbestand sei bereits verstanden, obwohl die zugrunde liegenden Zusammenhänge, Grenzen und Unsicherheiten nicht ausreichend geprüft wurden. Van Dis et al. (2023) betonen, dass der Einsatz von ChatGPT und verwandten Systemen in Forschungskontexten Transparenz, kritische Prüfung, Verantwortlichkeit und menschliche Kontrolle erfordert, und formulieren dafür fünf Prioritäten, die von der menschlichen Letztverantwortung bis zur transparenten Offenlegung des KI-Einsatzes reichen.

Auch Bertina (2025) selbst benennt in seinem Beitrag Risiken, welche für diese Arbeit relevant bleiben, nämlich Halluzinationen, Bias, Überabhängigkeit mit der Gefahr einer Erosion eigener Fachkompetenz, Reproduzierbarkeitsprobleme sowie ethische Implikationen. Diese Risiken bilden die Grundlage dafür, Vibe Research in dieser Arbeit eng zu führen, also begrenzt auf kodifizierbare Aufgaben, mit obligatorischer Verifikation jeder KI-Ausgabe und mit dokumentiertem Einsatz.

3.4.4. Konkretisierung im Kontext dieser Arbeit

Der Einsatz von Vibe Research in dieser Masterarbeit ist sachlich begründet, weil Gegenstand und Erkenntnisinteresse zur Methode passen. Die Arbeit untersucht einen KI-unterstützten Analyseprozess innerhalb des bestehenden M365-Evolution-Assessments und verortet sich damit in einem Themenfeld, das durch digitale, datenbasierte und KI-gestützte Arbeitsformen geprägt ist. Es ist folgerichtig, solche Entwicklungen nicht nur zum Untersuchungsgegenstand zu machen, sondern sie kontrolliert und reflektiert auch im eigenen Arbeitsprozess zu erproben. Vibe Research wird genutzt, um die im Verlauf der Arbeit erhobenen empirischen Daten, etwa aus Interviews und Evaluationsschritten, strukturiert aufzubereiten, zu verdichten und für die fachliche Interpretation vorzubereiten. Die methodische Verwendung erfolgt begrenzt und transparent. Der Einsatz von Vibe Research als Vorgehen wird strukturiert offengelegt (siehe [Kapitel 4](#)).

3.4.5. Ergebnisverlässlichkeit

Dies ist eine Ergänzung aus der KI-Forschung und umfasst zwei eng verbundene Dimensionen. Erstens die Vermeidung von Halluzinationen, theoretisch erklärt durch Kalai et al. (2025) und systematisch geordnet durch Huang et al. (2024). Zweitens die Reproduzierbarkeit über mehrfache Durchläufe, empirisch

belegt durch Chen et al. (2023) sowie Cui und Alexander (2026). Shneiderman (2020) bündelt diese Anforderungen unter dem Anspruch Reliable als zentralen HCAI-Grundwert.

Ergebnisverlässlichkeit ist in der vorliegenden Arbeit weder mit Validität noch mit Reliabilität deckungsgleich. Validität betrifft, ob das Richtige gemessen wird. Reliabilität betrifft die wiederholbare Messung unter denselben Bedingungen. Ergebnisverlässlichkeit verbindet beide Aspekte mit dem spezifischen Risiko probabilistischer Modellausgaben und macht sichtbar, dass eine plausible Ausgabe nicht zwingend eine korrekte Ausgabe ist.

3.4.6. Abgrenzung zu etablierten Auswertungsverfahren

Vibe Research ist kein eigenständiges Auswertungsverfahren und ersetzt keine etablierten Methoden der qualitativen oder quantitativen Datenanalyse. Es definiert vielmehr, wie generative KI innerhalb solcher Verfahren eingesetzt wird, also welche Teilaufgaben an die KI übertragen werden, wie deren Ausgaben am Originalmaterial geprüft werden und wie der Einsatz dokumentiert wird. Die Festlegung von Kategorien, die Bedeutungszuschreibung und die Interpretation der Ergebnisse bleiben eine fachliche Aufgabe der verantwortlichen Fachperson.

Mit dieser Abgrenzung wird auch das ausformulierte Vorgehen aufgenommen, wonach die Verifikation eines KI-Outputs annähernd dieselbe fachliche Kompetenz erfordert, wie seine Erstellung, weshalb der KI-Anteil bewusst begrenzt bleibt.

3.4.7. Abgrenzung zu User Research Democratization

Die Befähigung fachfremder Personen zur eigenständigen Auswertung empirischer Daten berührt das Feld der User Research Democratization, das User Research rollenunabhängig zugänglich und zugleich qualitätsgesichert machen will (Pernice 2022, Kaplan 2020). In dieser Arbeit wird das Konzept jedoch nicht als tragender Bezugsrahmen geführt. Zum einen zielt es in seiner ausgearbeiteten Form auf die organisationale Steuerung von Research-Bedarf, Kompetenzaufbau und Skalierung und damit auf eine Einführungsebene, die ausserhalb des Untersuchungsgegenstands liegt (siehe [Abschnitt 1.4](#)). Zum anderen sind die gegenstandsnahen Aspekte bereits durch präzisere und teils etabliertere Rahmen erfasst, namentlich Participatory Design (Spinuzzi 2005), die Mensch-KI-Aufgabenteilung mit Human-in-the-Loop sowie die Gütekriterien der qualitativen Inhaltsanalyse (Mayring 2022). Eine zusätzliche Rahmung über User Research Democratization würde diese Zusammenhänge daher lediglich neu benennen, ohne eigene Erklärungskraft. Das Konzept wird somit als verwandte, aber nicht leitende Perspektive eingeordnet.

3.4.8. Fazit

Vibe Research wird in dieser Arbeit als neuer, praxisnaher und reflektierter KI-gestützter Arbeitsansatz für die Analyse empirischer Daten verstanden. Seine methodische Tauglichkeit liegt nicht darin, fachliche Prüfung zu automatisieren oder zu ersetzen, sondern darin, umfangreiches Datenmaterial effizienter zu erschliessen, überprüfbare Zwischenergebnisse zu erzeugen und den Analyseprozess iterativ zu unterstützen. Die Qualität des Vorgehens hängt wesentlich davon ab, dass KI-generierte Ergebnisse nicht ungeprüft übernommen, sondern am Originalmaterial validiert, kritisch eingeordnet und transparent dokumentiert werden. Diese Grundhaltung ist konsistent mit der für die Inhaltsanalyse formulierten Anwendungslogik von Vibe Research im Methodenkapitel und bildet zugleich die Grundlage für die im weiteren Verlauf der Arbeit beschriebene Konzeption des KI-unterstützten Analyseprozesses.

3.5. Lernreflexion - Literaturvertiefung

Die Literaturvertiefung ist der Abschnitt der Arbeit, in dem sich das fachlich Neue am deutlichsten gezeigt hat. Die menschenzentrierte Gestaltung und die qualitativen Analyseverfahren lassen sich auf gefestigte Standardwerke stützen. Beim KI-Teil betrete ich dagegen ein Feld, das sich laufend verändert und für das es kaum kanonische Literatur gibt. Gerade dieser Teil hat mich am meisten gereizt, weil er unmittelbar an

mein Lernziel anschliesst, den menschenzentrierten Einsatz von KI in der Nutzerforschung mit seinen Potenzialen, Grenzen und Risiken besser zu verstehen. Rückblickend ist es zugleich der Teil, der mich am meisten verunsichert hat.

Zu Beginn habe ich Literaturrecherche vor allem als Suchen verstanden, also als das Finden möglichst vieler passender Beiträge. Diese Vorstellung hat sich im Verlauf verschoben. Der eigentliche Aufwand lag nicht im Finden, sondern im Prüfen, Ordnen und Einordnen. Ich musste Quellen beschaffen, nach einem einheitlichen Schema ablegen, gegen die zitierten Stellen abgleichen und vor allem einschätzen, wie belastbar sie sind. Diese Verschiebung vom Sammeln zum kritischen Verwalten war eine erste, eher unspektakuläre, aber prägende Lernerfahrung.

Unruhig wurde ich, als ich die KI-bezogenen Quellen genauer betrachtete. Für zentrale Begriffe wie Context Engineering oder Vibe Research gibt es noch keine gefestigte wissenschaftliche Grundlage. Verfügbar sind Preprints auf arXiv, technische Beiträge von Anbietern wie Anthropic und vereinzelt Blogbeiträge, etwa der namensgebende Text zu Vibe Research von Bertina (2025). Bei einer Belastbarkeitsprüfung musste ich feststellen, dass mehrere Quellen weder peer-reviewt waren, noch ein klares Publikationsdatum trugen, woraufhin ich einzelne wieder aussortierte. Verstärkt wurde das Unbehagen durch die Einsicht, dass selbst Aussagen über die Technik rasch veralten. Chen et al. (2023) zeigen, dass dasselbe Modell bei identischer Aufgabe innerhalb weniger Monate deutlich unterschiedliche Ergebnisse lieferte. Mir wurde unangenehm bewusst, dass ich auf einem Fundament argumentiere, das sich noch bewegt, und ich habe mich wiederholt gefragt, ob diese Quellenlage einer kritischen Prüfung standhält. Bis zu einem gewissen Grad sollte ich hier recht behalten, da während meiner Arbeit, nach Abschluss meiner offiziellen Recherche, im Frühling 2026 erste wissenschaftliche Dokumente zum Thema Vibe Research veröffentlicht wurden. Zwar habe ich mir diese angeschaut und fand es interessant, dass ich diese gerade für meine Definition und Vorgehen gebrauchen könnte, doch der Verhältnisaufwand, einer erneuten Quellenkritik, um diese auch wirklich zu nutzen, war schlichtweg zu gross, dafür, dass es sich um eine praktische Arbeit handelt.

Geholfen hat insgesamt, diese Unsicherheit nicht als Schwäche zu behandeln, sondern als etwas, das sich methodisch ordnen lässt. Ich habe die Quellen bewusst geschichtet, also etablierte Standardwerke für die zeitstabilen Grundlagen, aktuelle Beiträge und Preprints für das schnelllebige KI-Feld und Plattformquellen nur als Praxisbelege. Die Orientierung an Webster und Watson (2002) sowie an vom Brocke et al. (2009) hat mir geholfen, die Auswahl nicht als Lücke, sondern als begründete Fokussierung zu verstehen. Am meisten beruhigt hat mich der Entschluss, die Unsicherheit offen auszuweisen, statt sie zu verdecken. Ich halte den Preprint-Status transparent fest, ergänze Zugriffsdaten und einen eigenen Hinweis zur Quellenlage. Dass ich für einzelne Quellen noch mit Permalinks und Archivversionen ringe, führt mir vor Augen, dass diese Sorgfalt mühsam, aber notwendig ist.

Eine besondere Erfahrung war, dass die KI in diesem Abschnitt zugleich Werkzeug und Gegenstand war. Ich habe sie bewusst genutzt, um Zugänge zu Quellen zu finden, Open-Access-Alternativen zu suchen, Dateien zu ordnen und die Belastbarkeit einzelner Beiträge einzuschätzen. Gleichzeitig ist genau diese Art von System der Gegenstand, dessen Schwächen die Literatur beschreibt. Greifbar wurde das, als die Unterstützung Autorenangaben lieferte, die sie nur aus Dateinamen abgeleitet hatte, was mir erst bei der Prüfung auffiel. Die in der Arbeit zitierten Risiken, von den «stochastic parrots» bei Bender et al. (2021) bis zu den «illusions of understanding» bei Messeri und Crockett (2024), waren damit nicht mehr nur Theorie, sondern Erfahrung am eigenen Material. Ich habe somit bereits hier erneut erfahren dürfen, dass die Prüfung eines KI-Ergebnisses fast dieselbe fachliche Kompetenz verlangt wie seine Erstellung. Obwohl ich es aus meiner beruflichen Erfahrung heraus besser wissen sollte, habe ich anfangs plausibel klingenden Ausgaben zu schnell vertraut und mehr als einmal vermischte sich Material unbemerkt, bis ich eine striktere Materialdisziplin einführte. Die Skepsis, die ich in der Arbeit fordere, musste ich also zuerst an meiner eigenen Arbeitsweise einlösen. Dadurch entstand ein wohl etwas übertriebenes Literaturrecherche-Dokument, welches mir jedoch sehr half, die Kontrolle zu wahren.

Im Rückblick war die anfängliche Nervosität produktiv. Sie hat mich zu einem disziplinierteren und transparenteren Umgang mit Quellen gezwungen, als ich ihn in einem etablierten Themenfeld vermutlich gewählt hätte. Die wichtigste Erkenntnis liegt für mich darin, dass die Qualität der Quellenarbeit in einem jungen Feld nicht davon abhängt, ob kanonische Standardwerke vorliegen, sondern davon, wie nachvollziehbar, kritisch und offen ich mit dem umgehe, was verfügbar ist. Der menschenzentrierte Anspruch an Kontrolle und Verantwortung, den die Arbeit an den KI-unterstützten Analyseprozess stellt, gilt damit auch für meine eigene Auseinandersetzung mit der Literatur.

4. Vorgehen und Methoden

4.1. Vorgehensmodell

Die Gestaltung des KI-unterstützten Analyseprozesses stützt sich auf drei sich ergänzende Bezugsrahmen. Die DIN EN ISO 9241-210 liefert die normativen Grundsätze menschenzentrierter Gestaltung, Participatory Design den methodischen Zugang zum aktiven Einbezug der M365-Fachpersonen, und Garretts Five Elements of UX dienen als Strukturierungsrahmen für die prototypische Ausgestaltung. Die folgenden Abschnitte ordnen die drei Ansätze ein und begründen ihre Rolle in der vorliegenden Arbeit.

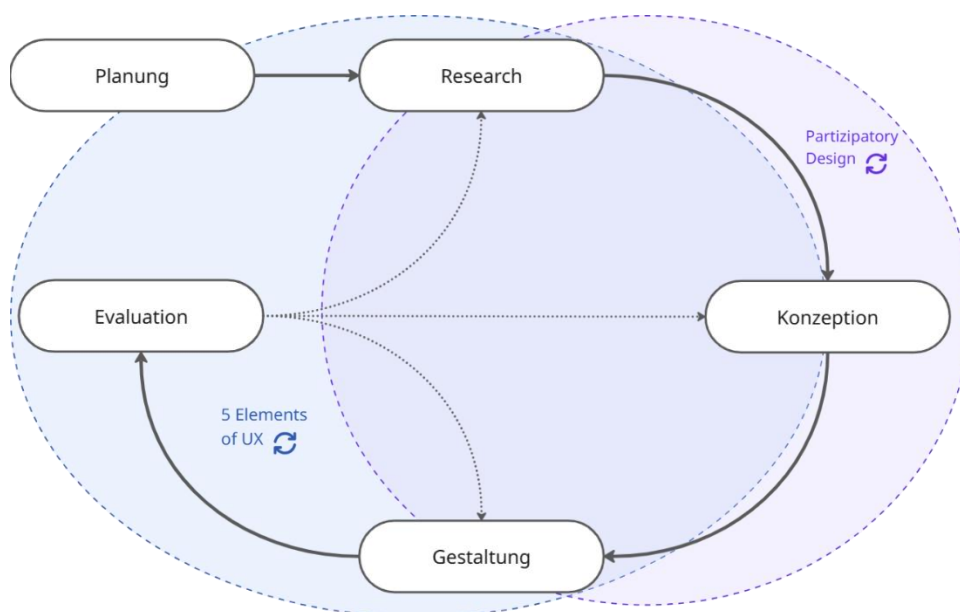


Abbildung 2: Kombiniertes Vorgehensmodell orientiert am HCD-Kreislauf (Eigene Darstellung)

4.2. Menschzentrierte Entwicklung als Vorgehensmodell

4.2.1. Der HCD-Kreis nach DIN EN ISO 9241-210

Begründung

Das grundlegende Vorgehensmodell dieser Arbeit orientiert sich an der DIN EN ISO 9241-210:2020 zur menschenzentrierten Gestaltung interaktiver Systeme. Diese Wahl ergibt sich aus dem Ziel der Arbeit. Der KI-unterstützte Analyseprozess soll nicht nur nach technischen Kriterien entwickelt, sondern mit Blick auf die Fachpersonen und im Hinblick auf Gebrauchstauglichkeit, Nachvollziehbarkeit, Verantwortungszuordnung und Passung zum konkreten Arbeitskontext gestaltet und bewertet werden. Die Norm eignet sich dafür besonders, weil sie einen klaren menschenzentrierten Bezugsrahmen vorgibt, ohne einen festen Methodenkanon vorzuschreiben. Sie lässt offen, was im Einzelnen erarbeitet wird und ist sowohl auf klassische

Software als auch auf Arbeitsprozesse anwendbar. An die Stelle fester Methodenvorgaben setzt sie eine methodische Ordnung, innerhalb derer konkrete Verfahren projektspezifisch gewählt und begründet werden. Diese Offenheit ist in der Norm ausdrücklich angelegt. Gemäss der ISO-Norm (DIN EN ISO 9241-210, Abschnitt 7.1) lassen sich die menschenzentrierten Gestaltungsaktivitäten in unterschiedliche Vorgehensansätze von der Wasserfall- bis zur agilen Entwicklung einbinden und die Gestaltung kann je nach Projekt an einem beliebigen Punkt ansetzen. Dieser Spielraum erlaubt es, mehrere methodische Zugänge flexibel zu kombinieren, damit der noch neuartige Untersuchungsgegenstand eines KI-unterstützten Analyseprozesses angemessen bearbeitet werden kann. Die Gestaltungsaktivitäten sind dabei nicht als starre Abfolge zu verstehen, sondern als wechselseitig abhängige Schritte, die iterativ durchlaufen werden (DIN EN ISO 9241-210, Abschnitt 5.5).

Adaption

In dieser Arbeit werden die Gestaltungsaktivitäten zu vier nummerierten Phasen adaptiert. Massgebend ist dabei nicht eine deckungsgleiche Zuordnung, sondern eine nachvollziehbare Dokumentationsstruktur. Jede Phase rückt eine gewählte leitende Methode in den Vordergrund und dokumentiert sie mit Begründung in einem geschlossenen Kapitel. Die Phasenstruktur dient damit als Ordnungsrahmen der Dokumentation, nicht als striktes Ablaufdiktat. Sie bleibt mit der iterativen Logik der Norm vereinbar, weil sich die Aktivitäten in der praktischen Durchführung überlagern und mehrfach aufeinander zurückwirken.

Tabelle 1: Aufteilung der Gestaltungsaktivitäten in angepasste Phasen

Gestaltungsaktivität (DIN EN ISO 9241-210)	Phase	Leitende Methode (Verortung)
Verstehen des Nutzungskontextes (7.2) und Festlegen der Nutzungsanforderungen (7.3)	Phase 1	Qualitative Interviews (Abschnitt 4.3), reguliert über Vibe Research (Abschnitt 4.4)
Erarbeiten von Gestaltungslösungen (7.4), Teilaktivität Gestalten (7.4.2)	Phase 2	Participatory Design (Abschnitt 4.2.2)
Erarbeiten von Gestaltungslösungen (7.4), Teilaktivität Konkretisieren (7.4.3)	Phase 3	Elements of UX nach Garrett (Abschnitt 4.2.3)
Evaluiieren der Gestaltung (7.5)	Phase 4	Usability Testing (Abschnitt 4.5)

Die Phasenstruktur folgt drei eigenen Gliederungsentscheidungen, welche die Nachvollziehbarkeit gewährleisten soll und welche jeweils mit der Norm vereinbar sind.

Erstens wird in Phase 1 alles zusammengeführt, was zum Vibe-Research-Prozess gehört, von den qualitativen Interviews bis zur Ableitung der Nutzungsanforderungen. Damit liegen die ISO-Aktivitäten "Verstehen des Nutzungskontextes" (DIN EN ISO 9241-210, Abschnitt 7.2) und "Festlegen der Nutzungsanforderungen" (DIN EN ISO 9241-210, Abschnitt 7.3) in derselben Phase. Diese Bündelung ist methodisch vertretbar, weil die Spezifikation der Nutzungsanforderungen den Nutzungskontext gemäss ISO-Norm (DIN EN ISO 9241-210, Abschnitt 7.3.3) als Bestandteil enthält. Sie erlaubt es, den Vibe-Research-Prozess (siehe [Abschnitt 4.4](#)) bis und mit dessen Ergebnissen und die zuspieldenden qualitativen Interviews (siehe [Abschnitt 4.3](#)) in einem gemeinsamen Kapitel zu dokumentieren.

Zweitens werden zwei verschiedene Sichten auf die Gestaltungslösung in eigenen Phasen verortet. Phase 2 dokumentiert den partizipativen Erarbeitungsprozess (siehe [Abschnitt 4.2.2](#)), Phase 3 die ausgestaltete Lösung entlang des angepassten Fünf-Ebenen-Modells nach Garrett (2011) (siehe [Abschnitt 4.2.3](#)). Damit wird die ISO-Aktivität "Erarbeiten von Gestaltungslösungen" (DIN EN ISO 9241-210, Abschnitt 7.4) auf zwei Phasen aufgeteilt. Diese Aufteilung ist mit der Norm vereinbar, weil sie zwischen dem Gestalten (DIN EN ISO 9241-210, Abschnitt 7.4.2) und dem Konkretisieren mit Modellen und Prototypen (DIN EN ISO 9241-210, Abschnitt 7.4.3) bereits selbst unterscheidet. Dass sich im Verlauf der partizipativen

Erarbeitung zusätzliche Nutzungsanforderungen ergeben können, ist von der Norm ausdrücklich vorgesehen (DIN EN ISO 9241-210, Abschnitt 7.4.1).

Drittens wird die Evaluation bewusst in einer eigenen Phase verortet und mit anderen Teilnehmenden durchgeführt als der partizipative Prozess. Damit entsteht eine Rückmeldung, die von der Erarbeitung unabhängig ist. Phase 4 entspricht damit inhaltlich der benutzerzentrierten Evaluierung nach der ISO-Norm (DIN EN ISO 9241-210, Abschnitt 7.5).

4.2.2. Participatory Design als ergänzender Methodenraum

Begründung

Als methodischer Zugang zur Beteiligung der Fachpersonen ergänzt Participatory Design die normativen Grundsätze der menschenzentrierten Gestaltung. Zum einen ist dieser Ansatz für die Arbeit relevant, weil der zu entwickelnde Analyseprozess stark von implizitem Praxiswissen, routinierten Deutungen und arbeitsbezogenen Erfahrungswerten geprägt ist. Solche Wissensbestände lassen sich durch distanzierte Beschreibung nur begrenzt erfassen. Sie werden besser sichtbar, wenn Nutzende nicht nur befragt, sondern wiederholt in die Ausarbeitung, Prüfung und Reflexion entstehender Artefakte einbezogen werden (Spinuzzi 2005). Zum anderen wurde das bestehende Evolution Assessment bereits partizipativ gestaltet. Das führte zu hoher Akzeptanz, schliesslich zum Bedürfnis der KI-unterstützten Weiterentwicklung und damit zu dieser Arbeit. Für die vorliegende Arbeit ist Participatory Design anschlussfähig, weil auch hier Erkenntnisgewinn und Gestaltung ineinandergreifen. Im konkreten Projektsetting wird diese Entscheidung zusätzlich durch die Risikoanalyse gestützt. Die Verankerung des Projekts in einem partizipativen Vorgehen dient somit als gezielte Risikoreduktion.

Adaption

Innerhalb des Vorgehensmodells wirkt Participatory Design als ergänzender Methodenraum in den Phasen 1 bis 3, von der Kontexterhebung über die Anforderungsdefinition bis zur Gestaltung. Der Co-Design-Landscape von Gaete Cruz et al. (2022) verbindet die Beteiligungsleiter mit dem Designzyklus und stellt damit ein Raster bereit, mit dem sich der Einbezug der Fachpersonen pro Phase differenziert beschreiben lässt. Im vorliegenden Projektsetting ist der Einbezug pragmatisch ausgelegt, weil die Projektgruppe begrenzt und der Zeitrahmen einer Masterarbeit klar umrissen ist. Die konkrete Anwendung wird exemplarisch in der Konzeptionsphase dokumentiert (siehe [Kapitel 7](#)).

Trotzdem wird die Arbeit nicht als voll ausgeprägte Action-Research-Studie ausgewiesen. Sie beansprucht keine umfassende organisationale Feldintervention mit fortlaufender Co-Steuerung des Gesamtprojekts. Participatory Design dient daher als ergänzender Zugang innerhalb des übergeordneten menschenzentrierten Rahmens, nicht als alleiniger Primärrahmen (Spinuzzi 2005).

4.2.3. Elements of UX als methodische Strukturierungshilfe

Begründung

Für die Konzeption und prototypische Ausgestaltung des Produkts dient Garretts Modell der Five Elements of User Experience als zusätzlicher heuristischer Strukturierungsrahmen (Garrett 2011). Es legt zwar seinen Fokus auf die Gestaltungsentscheidungen von abstrakten Zielsetzungen bis zur sichtbaren Ausgestaltung von visuellen Lösungen und Garrett weist selbst darauf hin, dass das Modell im Webkontext entstanden ist. In seiner Grundlogik lässt es sich aber auf andere Produkte und Services übertragen (Garrett 2011). Diese Übertragbarkeit wird in der vorliegenden Arbeit genutzt, jedoch in einer gezielt angepassten Form. Diese Anpassungen und die Reduktion sind methodisch begründbar. Garrett selbst räumt mittlerweile eine solche Adaption ein:

"If you work on the Web, everything in this book applies to you. If you work on other kinds of technology products, you'll see strong parallels to Familiar considerations. Even if you work on products or services that have nothing to do with technology, you can map these concepts to your own processes."
(Garrett 2011)

Adaption

Damit würde das Fünf-Ebenen-Modell für den Kontext meiner Arbeit so aussehen.



Abbildung 3: Adaptiertes Fünf-Ebenen-Modell in Anlehnung an Garrett (2011) (Eigene Darstellung)

Strategy

Die Strategy-Ebene bleibt im Kern unverändert und lässt sich ohne strukturelle Anpassung auf den prozessbezogenen und dialogbasierten Gegenstand übertragen. Sie umfasst die Produktziele aus Sicht der Organisation und die Bedürfnisse des Fachpersonals als künftige Nutzende des KI-unterstützten Analyseprozesses.

Scope

Im Kontext dieser Arbeit dient die Scope-Ebene der Festlegung der notwendigen Verarbeitungsschritte, die der Prozess leistet, sowie der inhaltlichen Elemente, die er erzeugt und verarbeitet. Dazu zählen insbesondere Codes, Evidenzen und Artefakte. Damit wird abgegrenzt, welche Aufgaben grundsätzlich Teil des Prozesses sind und welche bewusst ausserhalb bleiben.

Structure

Die Strukturebene überträgt sich auf den KI-unterstützten Analyseprozess als Gestaltung des generellen Prozessablaufs und der Aufgabenteilung zwischen Fachpersonal und KI. Ergänzend legt sie fest, wie Daten und Zwischenergebnisse entlang des Prozesses geordnet und nachvollziehbar geführt werden. Die konkrete Ausformulierung und Komposition werden erst auf der Skeleton-Ebene festgelegt.

Skeleton

Die Skeleton-Ebene ist aus ihrem klassischen Interface-Kontext heraus nur eingeschränkt anwendbar. Im Rahmen dieser Arbeit wird sie auf die konkrete Ausgestaltung der dialogbasierten Interaktion bezogen. Dazu zählen die vorgeschlagenen Prompts, die formulierten Texte des Agenten sowie die konkrete Komposition der Agenten und Flows, durch die der Light Assessment Agent als ausführbares Konstrukt realisiert wird.

Surface

Das finale Produkt ist ein Prozess, der über eine Chatinteraktion mit einem M365-Copilot-Agenten durchgeführt wird. Die Surface-Ebene umfasst bei Garrett die sinnliche Ausgestaltung und Präsentation der auf der Skeleton-Ebene festgelegten Anordnungen. Welche Sinneskanäle dabei gestaltet werden, ist nach Garrett vom Produkttyp abhängig. Im vorliegenden Fall ist die visuelle Ausprägung dieser Ebene durch die M365-Umgebung vorgegeben und liegt ausserhalb des Gestaltungsraums dieser Arbeit. Gestaltbar bleibt die Form, in der sich der Agent gegenüber den Nutzenden manifestiert: die sprachliche Ausgestaltung seiner Antworten, deren Struktur, Formatierung und Tonalität. Diese kommunikative Präsentation übernimmt funktional die Rolle, die Garrett der sinnlichen Gestaltung zuweist, bis hin zur Vermittlung von Persönlichkeit, die er selbst am Beispiel nicht visueller Kanäle beschreibt. Die oberste Ebene wird in dieser Arbeit deshalb im Sinne Garretts als Interface-Ebene verstanden, also als oberste Schnittstelle der Kommunikation zwischen Nutzenden und Agent. Der Begriff Surface bleibt dennoch erhalten, damit er nicht mit dem Element Interface Design der Skeleton-Ebene verwechselt wird, welches die Anordnung von Interaktionselementen betrifft. Somit bestimmt die Surface-Ebene, wie sich die Antworten des Agenten manifestiert und wie die Kommunikation mit den Nutzenden auszusehen hat.

4.3. Qualitative Interviews

4.3.1. Begründung

Die OST-Checkliste nennt Baxter et al. (2015) neben Garrett (2011), Steimle und Wallach sowie Cooper als zulässige HCID-Standardreferenz. Was Baxter im konkreten Projektkontext zur bevorzugten Wahl macht, ist die Kombination aus methodischer Breite, expliziter Stakeholder-Behandlung, Industry-Setting und konkretem Bias-Katalog. Diese vier Merkmale zusammen werden von keiner der anderen genannten HCID-Quellen in derselben Dichte angeboten.

Die Erhebung der Stakeholder-Sichten orientiert sich daher an Baxter et al. (2015), weil die methodische Logik des Buchs zum Kontext dieser Arbeit passt. Die Arbeit ist ein HCID-Projekt im privatwirtschaftlichen Beratungsumfeld, nicht eine sozialwissenschaftliche Studie. Baxter ist explizit für genau dieses Setting geschrieben. Die Befragten sind keine externen Studienteilnehmenden, sondern Arbeitskollegen und Stakeholder, die bereits in den Projektkontext eingebunden sind. Baxter behandelt diese Situation unter «Industry settings» und liefert für Stichprobengrösse, Frageformulierung und Interviewphasen praxisnahe Empfehlungen, die für eine kleine, projektnahe Gruppe von Fachpersonen tragfähig sind. Zudem ist der zentrale Untersuchungsgegenstand der Arbeit, der KI-gestützte Analyseprozess, von der Auftraggeberseite bereits als gewünschtes technologisches Mittel gesetzt. Eine offene Erhebungssituation im Sinne klassischer narrativer Interviewmethoden, in denen die Lösungsrichtung erst aus dem Material entstehen soll, würde diesen vorgegebenen Rahmen ausblenden. Baxter erlaubt es, einen gesetzten Lösungsraum offen, aber methodisch sauber zu adressieren.

4.3.2. Adaption

Im Generellen wurde sich an Baxter et al. (2015) orientiert. Durch die Eigenheiten des Projektkontexts, ergeben sich jedoch drei Abweichungen. Erstens ist der Untersuchungsrahmen vor der Erhebung gesetzt, weil die technologische Stossrichtung vorab definiert wurde und den Kern der Arbeit bildet. Zweitens sind die Befragten in das direkte, vorgängige Projekt, welches zu dieser Masterarbeit führte, eingebunden, also keine externen Studienteilnehmenden, sondern Arbeitskollegen mit bestehendem Vorwissen und eigener Fachsprache. Drittens haben die Interviews eine Doppelfunktion zwischen Datenerhebung und nachvollziehbarer Dokumentation des HCD-Prozesses. Daraus erklären sich die Setzung des KI-Themas, die Verwendung etablierter Geschäftsbegriffe, das gezielte Eröffnen von Spekulationen über KI-Einsatz, die mündliche statt schriftlicher Einleitung und der bewusste Verzicht auf künstliche Distanz für eine natürlichere Gesprächsführung.

4.4. Vibe Research als reguliertes Vorgehen

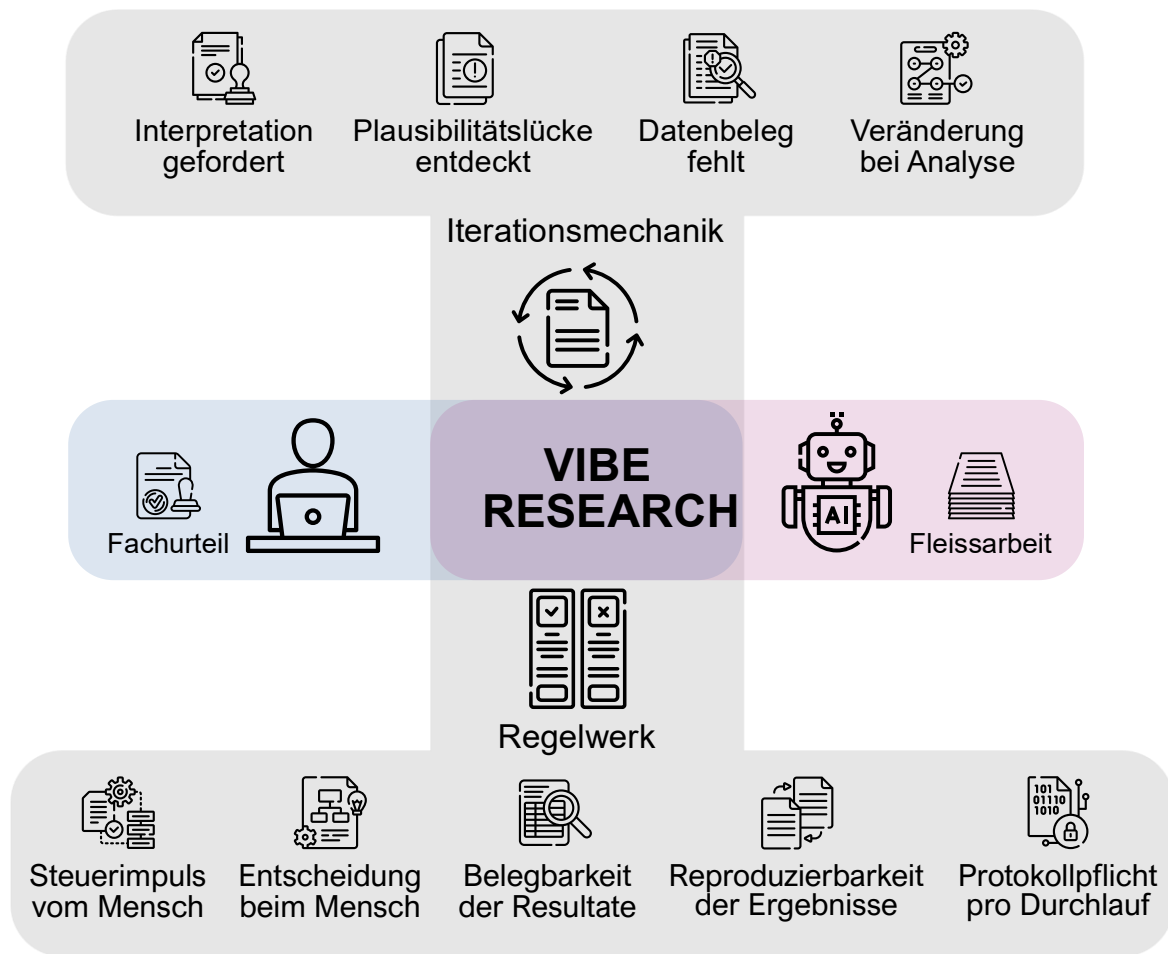


Abbildung 4: Übersichtsgrafik eigenes Vorgehen Vibe Research orientiert an Bertina (2025) (Eigene Darstellung)

4.4.1. Begründung

Der Einsatz von Vibe Research in dieser Arbeit ist gut begründbar. Zum einen besteht eine inhaltliche Passung zwischen Gegenstand und Arbeitsweise. Die Arbeit untersucht einen KI-unterstützten Analyseprozess innerhalb des M365-Evolution-Assessments, weshalb es folgerichtig ist, den KI-Einsatz bei der Analyse empirischer Daten nicht nur zu konzipieren, sondern ihn im eigenen Forschungsprozess kontrolliert und reflektiert zu erproben. Die dabei gewonnenen Erfahrungen fließen als zusätzliche Erkenntnisquelle in die Konzeption und in das Vorgehensmodell ein. Zusätzlich entlastet das Vorgehen bei kodifizierbarer Fleissarbeit. Die im Verlauf der Arbeit erhobenen empirischen Daten sind in Einzelarbeit aufzubereiten, zu strukturieren und auszuwerten, und die Übertragung überprüfbarer Teilaufgaben an die KI schafft Kapazität für das Fachurteil, also für Kategorienbildung, Bedeutungszuschreibung und Interpretation. Ausserdem bleibt der Einsatz methodisch abgesichert. Das Regelwerk mit Materialbindung, Verifikationspflicht und Protokoll stellt sicher, dass die Entlastung nicht zulasten der Nachvollziehbarkeit und der Ergebnisverlässlichkeit geht.

4.4.2. Operationalisierung im Analyseprozess

Aufgabenklassen als Steuerungsgrundlage

Das Vorgehen unterscheidet konsequent zwischen zwei Aufgabenklassen.

Fleissarbeit

Diese umfasst kodifizierbare Bearbeitungsschritte, deren Ergebnis an Originaldaten überprüfbar ist. Dazu zählen Datenreinigung, Formatangleichung, Transkriptaufbereitung, deskriptive Kennzahlen, Kodier- und Clustervorschläge, Konsistenzprüfung über das Material, Reformulierung sowie das Erzeugen von Gegenperspektiven.

Fachurteil

Dies umfasst nicht kodifizierbare Entscheidungen, die ein begründetes Verständnis von Gegenstand, Daten und Forschungsfrage voraussetzen. Dazu zählen Zielsetzung und Analyseplan, Festlegung von Variablen, Kategorien und Kodierregeln, Bedeutungszuschreibung, Quellen- und Datenkritik sowie die abschließende Interpretation.

Fleissarbeit wird der KI übertragen. Fachurteil verbleibt beim Forschenden. Diese Trennung folgt der von van Dis et al. (2023) formulierten Forderung nach menschlicher Letztverantwortung und greift Messeri und Crockett (2024) auf, die das Risiko einer scheinbaren Verstehensleistung durch KI als zentrale Begrenzung beschreiben.

Iterationsmechanik mit definierten Eingriffspunkten

Anstelle einer linearen Phasenfolge arbeitet Vibe Research in einer wiederkehrenden Schleife aus Steuerimpuls, KI-Verarbeitung, Prüfung und Übernahme oder Korrektur. Der Steuerimpuls definiert vor jedem KI-Einsatz Aufgabe, Datenausschnitt und erwartete Ausgabeform. Die KI bearbeitet die Aufgabe. Die Prüfung gleicht das Ergebnis mit den Originaldaten ab und entscheidet über Übernahme, Präzisierung oder Verwerfen. Die Schleife wird verlassen, sobald das Fachurteil gefällt ist. Sie wird zurückgesetzt, sobald einer von vier Eingriffspunkten erreicht ist.

Erstens

Wenn eine Plausibilitätslücke zwischen KI-Ausgabe und Datenmaterial sichtbar wird.

Zweitens

Wenn ein Datenbeleg fehlt oder eine Aussage nicht rückverfolgbar ist.

Drittens

Wenn eine fachliche Interpretation gefordert ist, die über reine Strukturierung hinausgeht.

Viertens

Wenn die Analyse einen methodischen Neueinsatz erfordert, etwa eine veränderte Forschungsfrage oder eine ergänzende Erhebung.

Diese Mechanik entspricht der von Schmidgall et al. (2025) beschriebenen Beobachtung, wonach iterative Verfeinerung mit menschlichem Feedback die Qualität KI-gestützter Analyseergebnisse nachweisbar verbessert. Lu et al. (2024b) machen umgekehrt sichtbar, warum eine Schleife ohne fachliche Letztprüfung nicht trägt.

Regelwerk

Fünf Regeln sichern das Vorgehen.

R1: Steuerimpuls vor jedem KI-Einsatz

Kontextverständnis, Aufgabe, Datenausschnitt, Auswertungslogik und erwartete Ausgabeform werden vor der KI-Bearbeitung schriftlich festgelegt.

R2: Bedeutungszuschreibung beim Forschenden

Die KI liefert keine eigenständigen Befunde, sondern Vorschläge. Befunde entstehen durch fachliche Entscheidung.

R3: Belegbarkeit und Materialbindung bei qualitativen Daten

Jede aus einer KI-Ausgabe übernommene Aussage kann auf eine identifizierbare Datenstelle rückverfolgt werden. Bei qualitativen Daten arbeitet die KI mit dem konkreten Material, nicht aus dem Modellwissen. Der Rückbezug auf Textstellen, Zeitstempel oder Kodiereinheiten bleibt erhalten.

R4: Reproduzierbarkeit quantitativer Bearbeitung

Bei quantitativen Daten werden Berechnung, Parameter, Filter und ausgeschlossene Fälle so dokumentiert, dass die Ergebnisse unabhängig reproduzierbar sind.

R5: Protokollpflicht

Datum, Modell und Version, Aufgabenbeschreibung, Prompt sowie Verifikationsschritt werden je Bearbeitungszyklus festgehalten. Die KI kann dazu angeleitet werden, entsprechende Informationen pro Zyklus mitzuliefern.

Nachweis und Offenlegung

Die Nachvollziehbarkeit wird durch das Zusammenspiel von Regelwerk und Protokoll gesichert. Das Protokoll macht entlang der Bearbeitungszyklen sichtbar, wo KI eingesetzt wurde, welche Steuerung gegeben war und wie die Ausgabe verifiziert wurde. Wang et al. (2023) und Ren et al. (2025) verorten KI als strukturierendes Element im Forschungsprozess. Die hier gewählte Nachweisform übersetzt diese Verortung in eine prüfbare Praxis und stützt damit die methodische Tragfähigkeit des Vorgehens.

4.5. Usability Testing

Begründung

Für das Prototyp-Testing wird Baxter et al. (2015) als methodische Referenz beibehalten. Diese Wahl ist nicht beliebig, sondern folgt zwei Überlegungen, die sich aus dem Projekt selbst ergeben. Erstens sichert die Beibehaltung derselben Methodenquelle für Erhebung und Evaluation die methodische Konsistenz innerhalb der Arbeit. Baxter behandelt Stakeholder-Interviews, Diary Studies, Surveys und Evaluation Methods in einem einheitlichen methodischen Bezugsrahmen. Damit lässt sich die vorliegende Hybridform aus aufgabenbasiertem Test, tagebuchartiger Selbstdokumentation und abschliessendem Fragebogen ohne Quellenbruch begründen. Zweitens passt Baxters Ausrichtung auf praxisnahe Industry-Settings zum Untersuchungsgegenstand. Der Prototyp wird mit einem Wizard-of-Oz-Verfahren unter möglichst realitätsnahen Arbeitsbedingungen erprobt, mit echten Projektdaten, einem vollständigen Arbeitsprozess und einer Zeitvorgabe, die der späteren produktiven Nutzung entspricht. Baxter erlaubt diesen praxisnahen Massstab und stellt zugleich die methodische Sprache bereit, in der sich die Beobachtungen entlang der Usability-Dimensionen Effektivität, Effizienz und Zufriedenheit nach DIN EN ISO 9241-11 strukturieren lassen.

Adaption

Das Testing folgt den zentralen Strukturelementen, die Baxter für Usability-Evaluationen beschreibt. Die Testpersonen bearbeiten repräsentative Aufgaben entlang des realen Arbeitsprozesses, eingeleitet durch eine schriftliche Kontextbeschreibung mit Erklärung des Wizard-of-Oz-Verfahrens und mit dem Hinweis, dass es keine richtigen oder falschen Antworten gibt. Pro Aufgabe werden formative Zwischenfragen gestellt, die Eindrücke unmittelbar einfangen. Am Ende folgt ein summativer Fragebogen, der die Usability-Dimensionen Effektivität, Effizienz und Zufriedenheit systematisch aufnimmt und damit Baxters Verständnis einer Endbewertung entspricht. Die Stichprobe von drei Personen bewegt sich im Bereich der Discount-Usability-Logik nach Nielsen, die Baxter explizit als legitimen Industrieansatz beschreibt. Methodisch weicht das Setup wegen des Projektkontextes an einigen Punkten von Baxters klassischer Vorlage,

begründet ab. So ist die Testdauer auf bis zu zwanzig Arbeitsstunden ausgelegt, weil ein vollständiger Arbeitsprozess simuliert wird und nicht eine einzelne Interaktion. Die Zeitvorgabe fungiert dabei selbst als Effizienzmassstab. Der Test wird asynchron und ohne durchgehende Moderation durchgeführt, um die Realbedingungen einer eigenständigen Bearbeitung möglichst wenig zu verzerren. Die bei Baxter zentrale Funktion der moderierenden Vertiefung wird strukturell durch die schriftlichen Zwischenfragen ersetzt, die als tagebuchartige Selbstdokumentation wirken. Damit werden im Setup drei Baxter-Methoden bewusst kombiniert, nämlich Usability Testing, Diary Study und qualitativer Survey. Diese Hybridisierung ist nicht eine Abweichung von Baxter, sondern eine bewusste Anwendung seiner Empfehlung, Methoden nicht isoliert einzusetzen, sondern gezielt zu verbinden, um unterschiedliche Aspekte der User Experience sichtbar zu machen.

4.6. Stakeholder Analyse

4.6.1. Begründung

Für die Stakeholder Analyse wird der Stakeholder Circle nach Bourne und Walker (2005) als methodische Referenz gewählt. Diese Wahl ergibt sich aus zwei Überlegungen, die zum konkreten Projektkontext passen. Erstens ist der Stakeholder Circle eine etablierte Methode, mit der projektbezogene Anspruchsgruppen identifiziert, priorisiert und hinsichtlich ihres Einflusses auf den Projektverlauf strukturiert sichtbar gemacht werden können. Diese Logik trägt eine zentrale Aufgabe der vorliegenden Arbeit, nämlich die nachvollziehbare Auswahl jener Rollen, deren Einbezug für die Konzeption und Evaluation des KI-unterstützten Analyseprozesses tatsächlich relevant ist. Zweitens passt der Charakter der Methode als visualisierendes Strukturierungsinstrument zum Untersuchungsgegenstand. Die Arbeit zielt nicht auf ein vollständiges Stakeholdermanagement im Sinne eines laufenden Projektportfolios, sondern auf eine fokussierte Standortbestimmung der relevanten Anspruchsgruppen für Konzeption, Evaluation und organisatorische Einbettung. Bourne und Walker bieten dafür ein Vokabular und eine Darstellungsform, die im HCID-Praxiskontext anschlussfähig sind und sich ohne methodische Reibung mit der menschenzentrierten Gestaltungslogik nach DIN EN ISO 9241-210 verbinden lassen.

4.6.2. Adaption

Die Stakeholder Analyse folgt den Grundprinzipien des Stakeholder Circle, nämlich der Identifikation relevanter Anspruchsgruppen, ihrer Priorisierung entlang konzentrischer Relevanzstufen und ihrer visualisierten Zusammenführung in einer Stakeholder-Map (siehe [Abschnitt 5.2](#), Abb. 4). Die daraus abgeleitete Map dient als Grundlage für die Auswahl der Interviewpartner und für die Verortung der späteren Nutzungsnähe einzelner Rollen zum Analyseprozess. Soweit erkennbar wird damit die methodische Grundstruktur eingehalten. Die Analyse wird dabei aus der Perspektive einer einzelnen verfassenden Person durchgeführt, die durch mehrjährige Praxis ein vertieftes Wissen zu Rollen, Interessen und Abhängigkeiten mitbringt. Bourne und Walker (2005) legen ein mehrperspektivisches Vorgehen nahe, das im Rahmen dieser Masterarbeit nicht realisierbar ist. Die Konsequenz ist eine bewusst akzeptierte pragmatische Reduktion der Perspektivenbreite, kombiniert mit erhöhter Sorgfalt bei der Begründung der getroffenen Einordnungen, was für den Rahmen einer solchen Arbeit methodisch vertretbar ist. Die konzentrischen Kreise sind nicht ausschliesslich als Distanzkategorien interpretiert. Sondern sie dienen als zusammengefasste Relevanzstufen, also Vital, Beteiligt und zu Berücksichtigen. Diese werden ergänzt, um die Unterscheidung zwischen firmeninternen und externen Personen und damit die Einteilung, noch mehr für den Projektkontext zu fokussieren. Damit werden Projektnähe, Betroffenheit und Bearbeitungspriorität gezielt zusammengeführt und fokussiert. Zudem wird innerhalb dieser Struktur eine nutzungsbezogene Einordnung in Primär-, Sekundär- und Tertiärnutzende ergänzt. Diese Erweiterung ist im Originalmodell nicht vorgesehen, knüpft aber direkt an die menschenzentrierte Gestaltungslogik der Arbeit an und bereitet die spätere Auswahl der Interviewpartner und die Gestaltung des partizipativen Vorgehens vor. Die methodische Grundstruktur des Stakeholder Circle bleibt von diesen Anpassungen unberührt.

4.7. Risikoanalyse

4.7.1. Begründung

Für die Risikoanalyse wird die Grundlogik eines qualitativen Risikoregisters in Anlehnung an ISO 31000 gewählt. Zum einen versteht sie die Risiken nicht als isolierte Einzelereignisse, sondern als Unsicherheiten, die die Erreichung von Zielen beeinflussen können. Diese zielbezogene Sichtweise trägt die Forschungslogik der Arbeit, weil sie die Verbindung zwischen identifizierten Unsicherheiten und konkreten Gestaltungsentscheidungen bereits methodisch anlegt. Zum anderen passt der Charakter eines qualitativen Registers zum Untersuchungsgegenstand. Die Arbeit zielt nicht auf ein operatives Risikomanagement im Sinne eines laufenden Geschäftsbetriebs, sondern auf eine fokussierte Standortbestimmung jener Unsicherheiten, die ein Projekt für eine Konzeption, prototypische Ausgestaltung und Evaluation eines KI-unterstützten Analyseprozesses betreffen. ISO 31000 bietet dafür den methodischen Rahmen mit den Phasen Identifikation, Analyse, Bewertung, Behandlung und Beobachtung, ohne ein vollständiges Risikomanagementsystem zu verlangen.

4.7.2. Adaption

Die Standardphasen der Identifikation, Analyse, Bewertung, Behandlung und Beobachtung werden eingehalten. Jedes Risiko wird mit einer einheitlichen Feldstruktur erfasst, die Ursache, Eintrittswahrscheinlichkeit, Auswirkung, Frühindikatoren, Gegenmassnahmen, Reaktion bei Eintritt, Status und Restrisiko abdeckt. Die Bewertung erfolgt qualitativ entlang der Stufen gering, mittel und hoch. Damit bleibt die methodische Grundstruktur der Methode unberührt. Dabei wird jedoch die Analyse in vereinfachter Form durchgeführt. Gegenstand sind nicht die Risiken eines vollständigen Assessments oder einer organisatorischen Einführung, sondern jene Unsicherheiten, die die Konzeption und Evaluation innerhalb des bestehenden Assessment-Kontexts betreffen. Diese Eingrenzung entspricht der Reichweite der Masterarbeit. Auch die Kategorisierung um die Unterscheidung zwischen Projektrisiken und Prozessrisiken wird ergänzt. Damit wird die Doppelnatur der Arbeit sichtbar gemacht, also einerseits der Projekterfolg im engeren Sinn und andererseits die methodische Qualität und der praktische Nutzen des entwickelten Vorgehens, der den wissenschaftlichen Kern der Arbeit bildet. Diese Differenzierung ist im Originalmodell nicht vorgesehen, knüpft aber direkt an die Forschungslogik einer HCID-Masterarbeit an. Das Register wird auch nur anlassbezogen aktualisiert, nicht fortlaufend bei jeder kleineren Änderung. Diese pragmatische Anpassung trägt dem Umstand Rechnung, dass die Arbeit von einer einzelnen Person verfasst wird und das Risikoregister methodisch wirksam bleiben muss, ohne in operative Routine zu kippen.

4.8. Qualitative Inhaltsanalyse nach Mayring als methodische Ergänzung zu Vibe Research

4.8.1. Begründung

Die Auswertung des qualitativen Interviewmaterials benötigt ein methodisch sauber verortetes, regelgeleitetes und gütegesichertes Verfahren. Die qualitative Inhaltsanalyse nach Mayring (2022) leistet genau das. Sie ist ein etabliertes, regelgeleitetes Verfahren mit definierten Analyseeinheiten, expliziter Belegbindung und einem ausgearbeiteten Satz inhaltsanalytischer Gütekriterien (Mayring 2022). Sie liefert damit die wissenschaftliche Verankerung, die Vibe Research als junger, nicht etablierter Ansatz für sich allein nicht bereitstellt. Mayring begründet, dass die Auswertung methodisch tragfähig ist und Vibe Research regelt, wie die KI darin kontrolliert eingesetzt wird. Die Wahl gerade dieses Verfahrens ergibt sich zusätzlich aus der Passung seiner Kodierlogik zur Funktionsweise generativer KI. Die deduktive Kategorienanwendung, bei Mayring die Strukturierung, trägt das Material über einen Kodierleitfaden mit Definition, Ankerbeispiel und Kodierregel an die Daten heran (Mayring 2022). Diese Struktur entspricht unmittelbar einer kontextbasierten Klassifikationsanweisung und damit einer operativen Kernstärke von LLMs. Die induktive

Kategorienbildung im Rahmen der zusammenfassenden Inhaltsanalyse verdichtet das Material auf einem festgelegten Abstraktionsniveau und entspricht der Verdichtungsleistung eines LLM, bleibt aber wegen der Risiken von Übergeneralisierung und Halluzination eng zu kontrollieren. Diese LLM-Affinität betrifft dabei nicht ausschliesslich, aber für den Kontext der Arbeit vorwiegend, die Kodierverfahren und nicht das vollständige Ablaufmodell, weshalb darauf der Fokus gelegt wird.

4.8.2. Adaption

Der Analyseprozess nutzt den analytischen Kern des allgemeinen Ablaufmodells ab Schritt 6, in einer Kombination aus deduktiver Strukturierung und zusammenfassender Inhaltsanalyse mit induktiver Kategorienbildung (Mayring 2022). Die Strukturierung steuert die Hauptkategorien, die zusammenfassende Inhaltsanalyse verdichtet das Material innerhalb der Hauptkategorien zu Subkategorien. Die Schritte 1 bis 5 sind im übergeordneten methodischen Rahmen der Arbeit abgedeckt und werden hier nicht eigens geführt, weil sie keinen kodifizierbaren KI-Schritt enthalten. Die Explikation und das vollständige formale Reduktionsverfahren der zusammenfassenden Inhaltsanalyse werden bewusst nicht dogmatisch mitgeführt, da sie unter dem Verifikationsregime von Vibe Research einen unverhältnismässigen Kontrollaufwand erzeugen würden, ohne den analytischen Ertrag für die Analyse im Kontext der Arbeit entsprechend zu erhöhen.

Konkret laufen die beiden Kodierverfahren in einer festen Logik. In der deduktiven Strukturierung bildet die forschende Person die Hauptkategorien. Die KI liefert auf Anfrage Vorschläge, woraus iterativ der Kodierleitfaden mit Definition, Ankerbeispiel und Kodierregel entsteht, gefolgt von Probedurchgang und endgültigem Materialdurchgang. In der anschliessenden induktiven Subkategorienbildung gilt pro Hauptkategorie das Selektionskriterium der Hauptkategorie selbst sowie ein mittleres Abstraktionsniveau. Die KI schlägt Subkategorien je Sinnabschnitt vor, die Entscheidung bleibt menschlich. Die Analyseeinheiten, also Kodier, Kontext- und Auswertungseinheit, werden vor dem ersten Materialdurchgang verbindlich festgelegt und sichern Präzision und Auditierbarkeit.

Die Aufgabenteilung folgt gemäss Vibe Research durchgängig dem Grundsatz, dass die KI kodifizierbare, am Material überprüfbare Fleissarbeit übernimmt und die forschende Person das Fachurteil, also Bedeutungszuschreibung, methodische Entscheidungen und Qualitätsprüfung, verantwortet. Mayrings Gütekriterien wirken dabei als Querschnitt und nicht erst am Schluss, insbesondere semantische Gültigkeit, Stabilität, Verfahrensdokumentation und Regelgeleitetheit, unter dem Leitsatz, dass Validität vor Reliabilität geht (Mayring 2022).

4.9. Lernreflexion – Vorgehen und Methoden

Die Frage der Methoden und des Vorgehens war für mich weniger eine Frage des Anwendens als des Auswählens, Begründens und Anpassens. Der Gegenstand der Arbeit, nämlich ein KI-unterstützter Analyseprozess, ist neuartig, und es gab kein fertiges Methodenset, das sich einfach zu meiner Zufriedenheit übernehmen liess. Rückblickend habe ich hier am meisten darüber gelernt und praktisch angewendet, wie man etablierte Methoden auf einen Gegenstand überträgt, für den sie nicht zwingend gedacht waren, und wo die Grenze zwischen begründeter Anpassung und methodischer Beliebigkeit verläuft. Dies war deswegen für mich so interessant, da ich doch bereits seit 2018 im HCD-Bereich unterwegs bin und mir dabei viel aus der Praxis von anderen abgeschaut oder im Rahmen von Projekten selbst erarbeitet habe. Nun damit konfrontiert zu sein, ein solches Projekt methodisch zu begründen und nicht einfach zu sagen, «dass ich das jetzt so mache», war besonders lehrreich für mich, da es mir sowohl meine persönlichen Stärken, als auch meine Schwächen aufzeigte.

In der Praxis in meiner Arbeitswelt scheiterte ich beim verständlich machen von UX Lösungen anhand des HCD Kreislaufes. Seitdem ich Garretts Fünf-Ebenen-Modell nutze, stosse ich auf mehr Verständnis. So habe ich in meiner Arbeitswelt eine Kombination geschaffen, in welcher die HCD-Schritte iterativ, je nach Bedürfnis entlang der fünf Ebenen im Hintergrund von mir geplant werden. Doch im Vordergrund

kommuniziere ich mit Kollegen und Kunden über das Garrett-Modell. Da mir dies so liegt und ich bereits reine Arbeitsprozesse damit analysiert habe, hätte ich beinahe vergessen, das Vorgehensmodell als Projektleitfaden zu bezeichnen und alles anhand des Garrett-Modells zu erklären, wie ich es in der Praxis gewohnt bin. Zum Glück wurde ich darauf hingewiesen. Ich hatte mich also an der ISO 9241-210 orientiert. Zum einen, da ich die entsprechende ISO-Zertifizierung habe und das Vorgehen schon seit Jahren einsetze, zum anderen aber, da mir die Methodenoffenheit gefällt. Es kommt nicht mit einem fertigen Methodenkatalog, sondern einem befüllbaren Gerüst.

Bei Garrett war meine grösste Sorge die verständliche Argumentation der Auswahl und Adaption. Denn sein Modell stammt aus dem Webdesign und beschreibt den Weg zu einer sichtbaren Oberfläche, während mein Produkt ein Prozess ohne klassisches Interface ist, der über einen Chat mit einem M365-Copilot-Agenten läuft, dessen Oberfläche ausserhalb meines Einflussbereiches liegt. Doch dass Garrett selbst die Übertragbarkeit seines Modells einräumt, gab mir den Anker, die Ebenen für einen prozessbezogenen Gegenstand schlichtweg neu zu besetzen. Eine eigene Unsicherheit zeigte sich, als ich die oberste Ebene von *Surface* zunächst provokativ weglassen wollte oder dann in *Interface* umbenennen wollte, weil mir *Surface* zu sehr nach sichtbarer Oberfläche klang. Im Verlauf der Arbeit wechselte sich das einige Male, doch nach kritischem Nachdenken, ergab sich für mich, dass diese Umbenennung mit Garretts eigener Begriffslogik kollidiert, in der *Interface* bereits zur darunterliegenden Ebene gehört. Ich habe die Umbenennung zurückgenommen und stattdessen den Inhalt der Ebene umgedeutet. Daraus habe ich einmal mehr an einem praktischen Beispiel gelernt und erfahren dürfen, dass eine Anpassung nur dann tragfähig ist, wenn sie die Bedeutung verändert und nicht nur das Etikett. In der Praxis meines Arbeitsalltages liegt beim Fünf-Ebenen-Modell jedoch der Fokus rein kommunikativ immer auf dem Produkt selbst und der Produktentwicklung. Die tatsächlichen Methoden rücken in den Hintergrund. Ein markanter Unterschied zu dieser Masterarbeit, bei welcher genau diese Methoden im Vordergrund stehen. Doch auch Garrett bietet nur ein Gerüst. Somit hatte ich zwei Gerüste und benötigte noch einen Methodenrahmen.

Im Gegensatz zur ISO und Garrett ergab sich der Ansatz des partizipativen Arbeitens komplett aus der Praxis, ohne bewussten methodischen Bezug. Hier möchte ich transparent erwähnen, dass vor dieser Arbeit Participatory Design oder Action Research, nicht bewusst oder präsent in meinem Methodenwissen vorhanden waren. Das ursprüngliche Evolution Assessment war bereits in firmenübergreifender Zusammenarbeit unter Anwendung partizipativer Grundsätze entstanden, schlichtweg da es mir als richtig erschien, so zu arbeiten. Genau diese, wie ich nun weiss, «partizipative» Praxis hatte meines Erachtens nach zu erhöhter Akzeptanz und schliesslich zum Wunsch nach der KI-gestützten Weiterentwicklung geführt. Participatory Design war damit weniger eine Methode, die ich von aussen wählte, als die nachträgliche Benennung einer Zusammenarbeit, die ohnehin stattfand und welche ich weiter pflegen wollte. Die methodische Begründung musste daher die gelebte Praxis verankern, da ich kein fremdes Verfahren darüberlegen wollte. Daher musste ich ehrlich bleiben und den Einbezug als pragmatisch ausweisen und somit nicht als vollständige Action Research, die er schlicht nicht war, sondern als selektive Auswahl eines Methodenrahmens.

Am meisten gefordert und zugleich am stärksten beschäftigt hat mich Vibe Research. Anders als bei den übrigen Methoden gab es hier keine etablierte, standardisierte Definition, an der ich mich hätte festhalten können. Eine kurze erneute Prüfung, während des Schreibens dieser Reflexion, bestätigt das. Der Begriff stammt aus einem Blogbeitrag (Bertina 2025), bezeichnet dort einen lockeren, weitgehend autonomen KI-Einsatz und wird in der Fachdiskussion bis heute eher kritisch und uneinheitlich behandelt. Ich stand damit vor der Aufgabe, ein Vorgehen, das ich in der Praxis längst für mich selbst entdeckt hatte (vor der Veröffentlichung von Bertina) und testweise schon angewandt habe, erst zu einer benennbaren und verantwortbaren Methodik zu formulieren und damit bis zu einem gewissen Grad über die Literatur hinaus zu gehen. Das war beängstigend, eben weil ich einen nicht gesicherten Begriff, aus einer nicht wissenschaftlichen Quelle, zu einer methodischen Grundlage der eigenen Arbeit erklärte und mich damit angreifbar machte. Es war zugleich faszinierend, weil hier zwei Stränge zusammenkamen: meine praktische Erfahrung im

Umgang mit KI und die Befunde aus der Literaturvertiefung zu Halluzinationen, Reproduzierbarkeit und Mensch-KI-Aufgabenteilung. Aus dieser Mischung ist schlussendlich das Modell entstanden. Sicherheit gewann ich erst, als ich Vibe Research bewusst eng führte, also auf kodifizierbare Fleissarbeit begrenzte, mit einem Regelwerk, einer Verifikationspflicht und einem Protokoll versah und an die etablierte Inhaltsanalyse nach Mayring (2022) ankoppelte. So wurde aus einem unscharfen Begriff ein Vorgehen, das regelt, wie die KI eingesetzt wird, ohne den Anspruch zu erheben, ein eigenständiges wissenschaftliches Verfahren zu sein. Mir ist bewusst, dass dieses Verfahren sicher Raum zur Diskussion und Verbesserung bietet und noch ausgereift werden kann, doch ich denke ich kann behaupten, dass es sich damit in bester Gesellschaft im aktuellen Diskurs dieser Thematik befindet. Jedenfalls werde ich dieses Vorgehen für mich selbst im Alltag hinterfragen und weiterentwickeln, da ich davon überzeugt bin.

Im Rückblick zieht sich durch das ganze Kapitel dieselbe Bewegung. Ich habe keine Methode unverändert übernommen, sondern jede an einen Gegenstand angepasst, für den sie nicht gemacht war, und musste lernen, diese Anpassungen offen zu begründen, statt sie einfach aus mangelndem Interesse der Adressaten unter den Teppich zu kehren. Das anfänglich Beängstigende, besonders bei Vibe Research, wurde tragbar, sobald ich es an Etabliertes ankoppelte und der Anpassung eine Begründungspflicht auferlegte. Und es fällt mir auf, dass ich die Aufgabenteilung zwischen Mensch und KI, die ich für das Produkt entworfen habe, beim Erarbeiten der Methoden selbst gelebt habe. Die KI hat mir zugearbeitet, geordnet und geprüft, die methodischen Entscheidungen aber sind meine geblieben.

5. Planung des Vorgehens

5.1. Projektplanung und Umgang mit Planabweichungen

5.1.1. Flexible Planung

Die Projektplanung dieser Arbeit ist bewusst flexibel angelegt. Das gewählte Vorgehensmodell verlangt, mehrere methodische Zugänge projektspezifisch zu kombinieren. Übergreifend kommt hinzu, dass es bei dieser koordinatorischen Ausgangslage im Verlauf der Arbeit zu Engpässen führen kann. Zudem ist die Arbeit organisatorisch an ein laufendes Pilot-Projekt gekoppelt. Im Rahmen der Gestaltung des Light Assessments wird dieses mit einer Gruppe aus fünf M365-Consultants aus drei verschiedenen Firmen und dem Autor für einen anonymen Kunden durchgeführt, mit dem Ziel der Neukonzeptionierung des Assessmentprozesses im laufenden Projekt und entsprechend flexiblem Zeitplan. Die Aktivitäten der Masterarbeit, also Erhebung, Konzeption, prototypische Umsetzung und Evaluation, sind in dieses Pilot-Projekt eingebettet und werden gemeinsam mit dessen Verlauf geplant. Die eigentliche operative Projektplanung des Pilot-Projekts, die Korrespondenzen und die vollständigen Kundendaten liegen ausserhalb des Untersuchungsrahmens dieser Arbeit.

Aus diesen Unsicherheiten folgt eine planerische Grundhaltung, die nicht nur Iteration und Versuchsordnung als festen Bestandteil des Vorgehens versteht, sondern im Zusammenhang damit Anpassungen am Plan nicht als Mangel, sondern als erwarteten Mechanismus behandelt. Tätigkeiten und Meilensteine werden so in möglichen Zeiträumen und nicht auf feste Daten geplant.

5.1.2. Phasenplanung

Planungsgrundlage ist das aus den methodischen Zugängen zusammengesetzte Vorgehen, das die vier Aktivitäten der DIN EN ISO 9241-210 mit den partizipativen Iterationen verbindet.

Es legt fest, wann welche Aktivität stattfindet. In der frühen Analysephase werden über leitfadengestützte Interviews und die Stakeholder-Analyse der Nutzungskontext erschlossen und erste Anforderungen abgeleitet. Daran schliesst der Übergang zur Gestaltung an, in welcher mit der Projektgruppe der neue Prozess erarbeitet und getestet wird. Darauf folgen drei Iterationen in der Projektgruppe, in denen die Ergebnisse der KI-gestützten Analyse besprochen, überprüft und angepasst werden. Diese wirken zugleich als formative Evaluation. In dieser Zeit finden entsprechend auch das Proto-Typing und technischer proof of concept statt. Den Abschluss der partizipativen Phase bildet die Präsentation der Ergebnisse und Feedback durch den Kunden, sowie der Reflexions-Workshop in der Gruppe. Die finale Phase überführt den Prozess in das Lieferobjekt und erprobt ihn abschliessend mit Nutzenden ausserhalb der Projektgruppe.

Meilensteine

Als inhaltliche Zwischenziele strukturieren mehrere Übergänge den Projektverlauf:

- Der Abschluss der Erhebung und Inhaltsanalyse mit den zentralen Befunden
- Die Ableitung des priorisierten Anforderungskatalogs
- Das frühe Testing neu gestalteter oder definierter Prozessschritte im Projektteam
- Der Abschluss der drei partizipativen Iterationen mit dem Reflexions-Workshop
- Die Finalisierung des Light Assessment Agent sowie die abschliessende Erprobung ausserhalb der Projektgruppe

5.2. Stakeholder Analyse

Im Vorfeld der weiteren Untersuchung wurden die für die Masterarbeit handlungsrelevanten Anspruchsgruppen über eine vereinfachte Stakeholder Analyse identifiziert und priorisiert. Die resultierende Map versteht sich als Arbeitsinstrument mit begrenzter Reichweite. Aufgrund der reduzierten Perspektivenbreite bildet sie nicht sämtliche möglichen Einflussnetzwerke und informellen Beziehungen vollständig ab. Ihr Wert liegt darin, die für die Masterarbeit handlungsrelevanten Anspruchsgruppen transparent zu ordnen und damit Entscheidungen über Rekrutierung, Einbezug und Fokus des Vorhabens zu unterstützen. Das bedeutet, dass die Map nicht nur dokumentiert, sondern auch zur Schärfung des Projektverständnisses beiträgt. Eine Version der Stakeholder Map inkl. Beschreibung pro Stakeholder befindet sich (siehe Anhang C, S. 118).

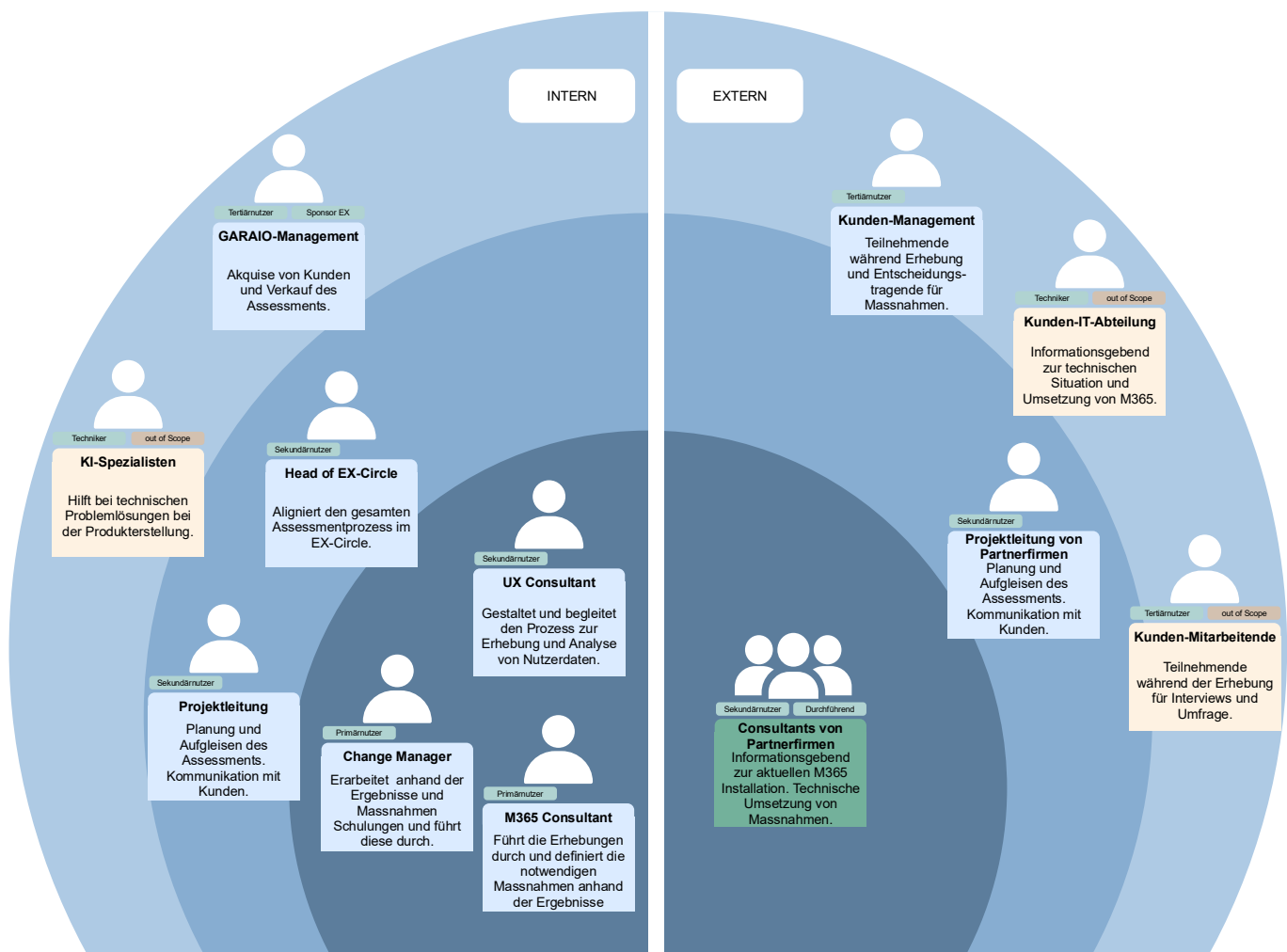


Abbildung 5: Stakeholder-Map der handlungsrelevanten Anspruchsgruppen (Eigene Darstellung)

5.2.1. Identifizierte Nutzende

Die Stakeholder-Map unterscheidet zwischen internen und externen Anspruchsgruppen und bündelt sie zugleich in drei nutzungsbezogene Kategorien.

Primärnutzende

Im Rahmen der Arbeit werden diese Nutzenden schlicht als Fachpersonen bezeichnet. Es sind jene Rollen, die den künftigen Analyseprozess voraussichtlich direkt anwenden sollen und generell als M365-Fachpersonen zusammengefasst werden können. Dazu zählen M365 Consultants und Change Manager des eigenen Unternehmens sowie in abstrahierter Form Consultants möglicher Partnerfirmen. Diese Bündelung externer Rollen zur Consultant-Sammelkategorie dient der pragmatischen Reduktion. Konkrete Funktionsbezeichnungen, organisatorische Einbettungen und Verantwortlichkeiten sind über Unternehmensgrenzen hinweg zu heterogen, um sie im Rahmen dieser Arbeit sinnvoll ausdifferenzieren zu können.

Sekundärnutzende

Das sind Rollen, welche den Prozess nicht primär bedienen, aber seine Ausgestaltung, Einführung oder Qualität wesentlich beeinflussen, insbesondere Projektleitungen sowie die Leitung des EX-Circle.

Tertiärnutzende

Dabei handelt es sich um Anspruchsgruppen, die vom Vorhaben mittelbar betroffen sind oder dessen Ergebnisse in weiterem Sinn bewerten. Hierzu zählen das Management im Hinblick auf strategische Passung und Vermarktbarkeit sowie kundenseitige Ansprechpersonen, sofern sie Ergebnisse oder daraus abgeleitete Entscheidungsgrundlagen erhalten.

Für einen späteren praktischen Einsatz wären weitere Rollen auf Kundenseite relevant. Da die Arbeit jedoch weder eine produktive Implementierung noch die vollständige Durchführung eines realen Assessments umfasst, bleiben diese Gruppen bewusst peripher.

5.2.2. Sonderregelungen

IT-Abteilung

Sie steht dem Vorhaben fachlich nahe, die technische Untersuchung der Systemlandschaft erfolgt jedoch über ein separates Verfahren und liegt ausserhalb des Fokus dieser Arbeit. Sie wird daher als relevante, aber periphere Anspruchsgruppe geführt.

UX Consultant

Hier ergibt sich eine besondere Konstellation, denn diese Rolle ist im vorliegenden Kontext doppelt eingebunden, da diese Rolle nur von einer Person abgedeckt wird. Zum einen tritt sie fachlich an der Weiterentwicklung des Assessments beteiligt auf, ist jedoch andererseits im Rahmen dieser Arbeit als methodenspezifische und konzeptionelle Schlüsselrolle tätig und zugleich Autor dieser Masterarbeit. Im Stakeholder Circle ist sie aufgrund ihrer fachlichen Nähe und ihres Einflusses zwischen vital und beteiligt verortet, in der nutzungsbezogenen Einordnung als Sekundärnutzer. Diese Doppelrolle hat unter anderem die Wahl des partizipativen Vorgehens mit beeinflusst.

5.3. Risikoanalyse

Ergänzend zur Stakeholder Analyse wurde eine projektbezogene Risikoanalyse durchgeführt. Sie unterscheidet zwischen Projektrisiken, die das im Projektkontext der Masterarbeit entstehende Ergebnis betreffen und Prozessrisiken, die die methodische Qualität und den praktischen Nutzen des entwickelten Vorgehens berühren. Die vollständige Risikoliste findet sich (siehe Anhang D, S. 119).

5.3.1. Einzelperson

Die Risikoanalyse erfolgt in Form eines Ein-Personen-Workshops. Dieses Vorgehen ermöglicht eine dichte Auseinandersetzung mit dem Projektkontext, bleibt jedoch anfällig für blinde Flecken. Trotz systematischer Verdichtung und wiederholter Überprüfung ist davon auszugehen, dass einzelne Risiken erst im Projektverlauf sichtbar werden. Als Vorsichtsmassnahme wurden projekterfahrene Stakeholder informell um Inputs gebeten, ohne dass dabei substantiell neue Risiken hinzukamen. Zur Erweiterung der Perspektiven kamen ergänzend KI-gestützte Recherche- und Strukturierungshilfen zum Einsatz, deren Mehrwert vor allem in der Ausdifferenzierung bekannter Risikofelder und in sprachlichen Präzisierungen lag. Die Begrenzung ist bei der gewählten Vorgehensweise eingeplant und wird in der weiteren Arbeit bewusst mitreflektiert.

Tabelle 2: Ausschnitt Risikoanalyse - Priorisierte Einzelrisiken der Risikoanalyse

Nr.	Typ	Kurzbezeichnung	Typ	Warsch.
R04	Projektrisiko	Unklare Abgrenzung zwischen Assessment und Analyseprozess	Fachlicher Bezugsrahmen ist das Assessment, Untersuchungsgegenstand aber nur der Analyseprozess	Mittel
R01	Projektrisiko	Unterschätzung des Gesamtaufwands	Breiter Projektumfang trotz 300h-Rahmen. Iteration, Prototyping, Testing und schriftliche Arbeit konkurrieren um Zeit	hoch
R16	Projektrisiko	Verzerrung bei Datenerhebung durch Doppelrolle im Unternehmenskontext.	Nähe zum Assessment, eigene Beteiligung am bestehenden Verfahren und zugleich Untersuchung desselben	mittel

5.3.2. Kernerkenntnis

Das wichtigste Ergebnis der Risikoanalyse ist, dass sich mehrere der priorisierten Risiken um ein gemeinsames Kernproblem gruppieren lassen: dem unzureichenden oder zu späten Einbezug relevanter Stakeholder, insbesondere der Primärnutzenden. Ein enger Austausch mit diesen Akteuren wirkt auf verschiedene Risiken gleichzeitig und ist deshalb nicht nur als einzelne Gegenmassnahme, sondern als übergreifendes Gestaltungsprinzip zu verstehen.

Tabelle 3: Ausschnitt Risikoanalyse - Priorisierte Risiken und Zuordnung zum Kernproblem

Nr.	Typ	Kurzbezeichnung	Typ	Warsch.
R02	Projektrisiko	Verzögerung durch Einzelarbeit	Keine Verteilung von Analyse-, Konzeptions- und Evaluationsaufgaben, reduzierte Perspektivenbreite.	hoch
R03	Projektrisiko	Eingeschränkter Zugang zu Testpersonen	Abhängigkeit von M365-Fachpersonen und weiteren Beteiligten für User-Testings und Feedback.	hoch
R06	Prozessrisiko	Ungeeignete Aufgabenverteilung zwischen Mensch und KI	Automatisierung wird zu weit oder an der falschen Stelle getrieben. Verantwortung bleibt unklar	hoch

R09	Usability-Risiko	Fehlende tatsächliche Effizienz- oder Zufriedenheitssteigerung	KI spart Einzelaufwand, erzeugt aber Zusatzaufwand für Vorbereitung, Kontrolle und Korrektur	hoch
R11	Usability-Risiko	Geringe Akzeptanz bei M365-Fachpersonen	Prozess passt nicht zu realen Arbeitslogiken, ist zu komplex oder wirkt nicht vertrauenswürdig	gering
R12	Usability-Risiko	Zu hoher Schulungs- oder Einführungsbedarf	Fachpersonen sind nicht primär auf Research spezialisiert. Zusätzlicher Prozess darf ihre Hauptarbeit nicht erschweren	mittel
R07	Prozessrisiko	Zu geringe Nachvollziehbarkeit der KI-Auswertung	LLM-basierte Verarbeitung verdichtet Daten und trifft Entscheidungen, ohne dass Ableitungsschritte ausreichend transparent sind.	Hoch
R14	Evaluations-risiko	Iterationen bleiben aus Zeitgründen unzureichend	HCD verlangt Rückmeldungen, Überarbeitung und erneute Prüfung. Zeitdruck verkürzt Schleifen	mittel

Dieser Befund hat die Methodenwahl unmittelbar geprägt. Die Verankerung des Projekts in einem partizipativen Vorgehen wurde als gezielte Risikoreduktion gewählt. Die methodische Ausarbeitung und Begründung dieser Wahl, einschliesslich des Bezugs zu Participatory Design nach Spinuzzi (2005) und der Co-Design-Landscape nach Gaete Cruz et al. (2022), erfolgt im Methodenkapitel (siehe [Kapitel 4](#)). Eine weitere methodische Konsequenz, die Behandlung der Quellenlage im KI-bezogenen Teil der Arbeit, wird im Abschnitt zu den methodischen Grenzen aufgegriffen.

5.4. Lernreflexion – Planungsphase

Rückblickend habe ich hier vor allem gelernt, wie unterschiedlich der Wert eines Planungsinstruments ausfällt, je nachdem ob man es als formale Pflicht oder als praktisches Steuerungsmittel begreift, und wie sehr eine Planung zur eigenen Arbeitsrealität passen muss.

Die Stakeholder-Analyse und die Risikoanalyse habe ich zu Beginn eher als geforderte Schritte gesehen, doch dann schnell verstanden, welche Relevanz sie für mich haben. Die Stakeholder-Analyse hat mir geholfen, den Fokus bewusst zu setzen, also zu entscheiden, wen ich einbeziehe und wen ich begründet peripher halte, und sie hat die Doppelrolle sichtbar gemacht, die ich als zugleich beteiligte Fachperson und verfassende Person einnehme. Es hat mir durch das Auslegen der Stakeholder eine Übersicht verschafft, welche mir ironischerweise genau deswegen fehlte, weil ich die Stakeholder eigentlich schon kannte. Doch ich habe sie mir nie klar vor Augen geführt. Das hat definitiv geholfen und am Ende auch zur weiteren Festigung der Methodenwahl geführt. Die Risikoanalyse war noch wertvoller. Sie hat mich nicht nur einzelne Gefahren auflisten lassen, sondern zu der Erkenntnis geführt, dass sich mehrere Risiken um ein gemeinsames Kernproblem gruppieren, dem zu späten oder zu geringen Einbezug der Fachpersonen. Diese Einsicht hat ebenfalls unmittelbar die Methodenwahl geprägt und das partizipative Vorgehen als gezielte Gegenmassnahme begründet. Beide Analysen waren damit keine Formalität, sondern haben das Projekt tatsächlich gesteuert.

An der Projektplanung lässt sich auf den ersten Blick Kritik üben. Sie war bewusst flexibel und in Teilen fragmentiert, und ich sehe, dass sie dadurch von aussen schwerer nachvollziehbar wirkt, als ein klassischer, durchgetakteter Plan. Ich halte das dennoch für eine vertretbare Entscheidung. Ich war die einzige Person, die mit dieser Planung tatsächlich arbeiten musste, und habe sie so angelegt, wie sie für mich brauchbar war. Hinzu kommt, dass die Arbeit an ein laufendes Pilot-Projekt gekoppelt war, dessen Verlauf ich nicht allein steuern konnte. Gerade die flexible Anlage hat es mir ermöglicht, auf Hindernisse zu

reagieren und Tätigkeiten zu verschieben, als sie auftraten. Was von aussen wie mangelnde Stringenz aussehen kann, war aus meiner Arbeitsperspektive die Voraussetzung dafür, handlungsfähig zu bleiben. Die Spannung zwischen nachvollziehbarer Dokumentation und praktischer Brauchbarkeit war mir von Anfang an bewusst und ich habe mich hier bewusst für die Brauchbarkeit entschieden.

Beim tatsächlichen Einsatz von Vibe Research hat sich bestätigt, was ich aus der eigenen Praxis und aus der Literaturvertiefung bereits erwartet hatte. Die KI-gestützte Auswertung hat insgesamt länger gedauert, als wenn ich sie von Hand gemacht hätte. Der Aufwand liegt nicht im Ausführen, sondern darin, die Anweisung stabil aufzubauen, sodass sie über das Material hinweg verlässlich funktioniert. An dieser Stelle gehört auch eine Korrektur hin, die ich zuerst unterschätzt habe. Ursprünglich wollte ich die qualitative Inhaltsanalyse nach Mayring (2022) in allen zehn Schritten als Grundlage nehmen und durchgehend mit KI unterstützen. In der Umsetzung wurde das so umständlich und unübersichtlich, dass ich das Vorgehen noch einmal überarbeitet habe. Ich habe nur die deduktive und die induktive Analyse übernommen und den Rest weitestgehend KI-gerecht an den Prozess angelehnt. Daraus habe ich gelernt, dass sich ein vollständiges, für die menschliche Bearbeitung gedachtes Verfahren nicht eins zu eins auf einen KI-gestützten Ablauf übertragen lässt und dass die ehrlichere Lösung darin bestand, gezielt zu reduzieren, statt die Methode vollständig mitzuschleppen.

Im Rückblick zeigt sich, dass sich Planung und Durchführung gegenseitig formen. Die wichtigste Einsicht ist für mich, dass methodische Vollständigkeit und praktische Brauchbarkeit nicht dasselbe sind und dass ein guter Teil der Arbeit darin bestand, beide in eine vertretbare Balance zu bringen. Fehler entstehen hier nur, wenn etwas nicht funktioniert und dann nicht im Rahmen des Projektes reagiert werden kann.

6. Phase 1: Nutzungskontext verstehen

6.1. Stakeholder Interviews

6.1.1. Konkrete Durchführung

Die Erhebung erfolgt über drei leitfadengestützte Einzelinterviews. Befragt wurden Personen aus dem EX Circle, die bereits in das vorgängige Evolution Assessment eingebunden waren. Die Auswahl folgt anhand der projektinternen Stakeholder-Map. Abgedeckt sind Primärnutzende durch operativ mit dem Assessment befassende Fachpersonen, sowie einer leitenden Person, in einer Doppelrolle als Sekundärnutzer (fachlicher Stakeholder) und Primärnutzer (potenziellem Anwender). Die Tertiärnutzenden auf Management- und Kundenseite wurden bewusst nicht in die Interviews einbezogen, da sie in der Stakeholder-Map als peripher ausgewiesen sind. Die geringe Stichprobengrösse ist im Sinne von Baxter et al. (2015) für eine kleine, projektnahe Gruppe von Fachpersonen angemessen. Die Befragten werden in der Auswertung pseudonymisiert als Person01, Person02 und Person03 geführt.

Die drei Interviews fanden an drei aufeinanderfolgenden Tagen Ende Januar 2026 statt und wurden vollständig als Online-Gespräche über Microsoft Teams geführt. Die Gespräche wurden mit der Aufzeichnungsfunktion von Teams aufgenommen und automatisch transkribiert. Die folgende Übersicht fasst die drei Interviews zusammen.

Tabelle 4: Übersicht der durchgeführten Stakeholder-Interviews

Befragte Rolle	Datum	Datum
Person 01 - Leitung (Sekundärnutzer, Primärnutzer)	26.01.2026	rund 37 Min.
Person 02 - Fachperson (Primärnutzer, Change Management)	27.01.2026	rund 37 Min.
Person 03 - Fachperson (Primärnutzer, User Adoption)	28.01.2026	rund 45 Min.

Die beiden Primärnutzenden umfassen zwei komplementäre Profile mit ähnlichen Aufgabenbereichen, eine Person auf User Adoption mit operativer Erfahrung und eine auf Change Management mit Schwerpunkt auf Schulung und Begleitung. Beide haben auch in den, wie in der Stakeholder Analyse (siehe [Abschnitt 5.2](#)) erklärten Rollen, Projektleitungserfahrung. Damit decken die Interviews sowohl die analytisch-durchführende, als auch die schulungs- und begleitungsnahe und Projektleitungssicht auf das Assessment ab. Die leitende Person hat vorwiegend den Fokus auf Projektleitung mit strategischer Übersicht. Die Gesprächsdauer lag zwischen rund 37 und 45 Minuten und bewegt sich damit in einem für projektnahe Fachgespräche angemessenen Rahmen.

Gesprächsleitfäden

Für die Gesprächsführung lagen zwei rollenspezifische Leitfäden vor, je einer für die beiden Consultants und einer für die Leitung. Diese Differenzierung trägt den unterschiedlichen Perspektiven Rechnung. Der Consultant-Leitfaden ist auf die operative Erfahrung mit dem Assessment ausgerichtet, der Leitfaden für die Leitung stärker auf die strategische und geschäftliche Einordnung. Beide Leitfäden folgen einer vergleichbaren Dramaturgie, die von der Schilderung der eigenen Rolle und des bestehenden Assessment-Prozesses über dessen Schmerzpunkte und Mehrwert zu den Erwartungen an das geplante Light Assessment, zur Haltung gegenüber dem KI-Einsatz und schliesslich zu den sogenannten Killerkriterien für die Durchführung führt. Die Fragen sind methodenüblich überwiegend offen formuliert und durch konkretisierende Unterfragen ergänzt, sodass der Leitfaden die thematische Abfolge steuert, zugleich aber Raum für narrative Schilderungen und situatives Nachfragen lässt. Beide Gesprächsleitfäden sind vollständig dokumentiert (siehe Anhang E, S. 120 und Anhang F, S. 121).

Gesprächsführung und Umsetzung der methodischen Adaptionen

Die Gespräche realisieren die im Methodenkapitel begründeten Adaptionen der Methode in konkreter Form. Der KI-Einsatz wurde nicht als offene Frage behandelt, sondern als gesetzter Untersuchungsrahmen eingeführt, da die technologische Stossrichtung von der Auftraggeberseite bereits definiert war und den Kern der Arbeit bildet. Innerhalb dieses Rahmens wurden die Befragten gezielt zu Spekulationen über einen möglichen KI-gestützten Ablauf eingeladen, etwa durch die Aufforderung, einen idealen Ablauf des vorgesehenen Light Assessments zu schildern. Etablierte Geschäfts- und Projektbegriffe wie Light Assessment, Massnahmenkatalog oder Killerkriterien wurden bewusst verwendet, da die Befragten als projektnahe Fachpersonen über eine gemeinsame Fachsprache verfügen. Auf eine künstliche Distanz wurde zugunsten einer natürlichen, kollegialen Gesprächsführung verzichtet, was sich im informellen Einstieg und in der durchgehenden Du-Form zeigt. Die halbstrukturierte Grundstruktur nach Baxter et al. (2015) bleibt davon unberührt.

Auf eine schriftliche Einleitung und eine formalisierte Einverständniserklärung wurde verzichtet. Stattdessen wurde das Einverständnis zur Aufzeichnung und zur Verwendung des Materials für die Masterarbeit zu Beginn jedes Gesprächs mündlich eingeholt und ist als Teil der Aufzeichnung dokumentiert. In einem Fall wurde dabei explizit festgehalten, dass keine gesonderte schriftliche Einverständniserklärung vorliegt. Dieses Vorgehen ist Ausdruck der bewusst an die kollegiale Nähe angepassten Gesprächsführung und wird in der Reflexion der Methode entsprechend transparent eingeordnet.

6.1.2. Ergebnisse

Als Rohmaterial liegt je Interview die Videoaufzeichnung sowie das automatisch durch Microsoft Teams erstellte Transkript vor. Die maschinellen Transkripte sind erwartungsgemäss fehlerbehaftet, etwa durch Hörfehler und verschobene Wortgrenzen, und bilden die Grundlage für die anschliessende qualitative Inhaltsanalyse. Die sprachliche Glättung der Transkripte erfolgt nicht im Rahmen der Erhebung, sondern als eigener, dokumentierter Arbeitsschritt der Auswertung, sodass die Trennung zwischen erhobenem Material und aufbereiteter Datengrundlage nachvollziehbar bleibt.

6.2. Inhaltsanalyse mit Vibe Research

Die Auswertung der drei Stakeholder-Interviews wird gemäss der Definition des Vorgehens Vibe Research aus dem Methodenkapitel durchgeführt. Operativ ist dieses Vorgehen in einer dreischichtigen Dokumentenstruktur umgesetzt. Ein Orchestrator legt Reihenfolge, Übergaben und Stoppbedingungen fest. Anweisungsdokumente steuern die einzelnen Arbeitsschritte und Kontextdokumente geben relevante Informationen und methodische Grundlagen vor. Pro Aufgabe wird daraus ein kuratierter Kontext-Pool zusammengestellt, der nur das für den jeweiligen Schritt nötige, am Material überprüfbare Bezugsmaterial enthält. Damit wird das Prinzip des Context Engineering unmittelbar wirksam, die KI-Bearbeitung bleibt an konkretes Material gebunden und im Verlauf konsistent steuerbar. Der Prozess verläuft in sechs nachvollziehbaren Schritten.

6.2.1. Durchgängige Steuerung

Jeder KI-Einsatz folgt demselben Bearbeitungszyklus. Dessen Phasen, die Verantwortung und die jeweils greifende Sicherung fasst die folgende Tabelle zusammen.

Tabelle 5: Phasen des KI-Bearbeitungszyklus mit Verantwortung und greifender Sicherung

Phase	Vorgehen	Verantwortung	Greifende Sicherung
Steuerimpuls	Aufgabe, Datenausschnitt und erwartete Ausgabeform werden vor der Bearbeitung festgelegt	Mensch (Fachurteil)	Steuerimpuls vor jedem KI-Einsatz
KI-Verarbeitung	Die KI erstellt Vorschläge oder regelbasierte Bearbeitungen am bereitgestellten Material	KI (Fleissarbeit)	Materialbindung, Arbeit am Material statt aus dem Modellwissen
Prüfung	Das Ergebnis wird mit dem Originalmaterial abgeglichen	Mensch (Fachurteil)	Belegbarkeit, Rückführbarkeit auf eine Datenstelle
Entscheidung	Übernehmen, Präzisieren oder Verwerfen	Mensch (Fachurteil)	Bedeutungszuordnung bleibt beim Menschen
Eingriffspunkt-Prüfung	Rückstellung bei Plausibilitätslücke, fehlendem Beleg, geforderter Interpretation oder methodischem Neueinsatz	Mensch (Fachurteil)	Die vier Eingriffspunkte
Protokoll	Aufgabe, Datenbasis, Steuerimpuls und Verifikationsschritt werden festgehalten, Datum und Modell durch die forschende Person	KI und Mensch	Protokollpflicht je Bearbeitungszyklus

Über alle Schritte hinweg ist jeder Bearbeitungszyklus lückenlos protokolliert. Massgeblich bleiben dabei zwei durchgängige Sicherungen, die Materialbindung, die jede übernommene Aussage auf eine identifizierbare Datenstelle rückführbar hält, und die Protokollpflicht, die je Zyklus festhält, wo die KI eingesetzt wurde und wie die Ausgabe verifiziert wurde.

6.2.2. Schritt 1: Materialaufbereitung

Im ersten Schritt werden die als Word-Dokumente vorliegenden Transkripte für die Analyse aufbereitet und in eine bereinigte, einheitlich strukturierte Markdown-Fassung überführt. Je Interview werden zunächst inhaltslose Zwischenbemerkungen entfernt und durch sie getrennte Aussagen desselben Sprechers zusammengeführt, danach wird der Text sprachlich geglättet, ohne inhaltlich relevante Informationen zu verändern.

KI-Aufgaben

Die KI führt die regelbasierte Bereinigung und sprachliche Glättung aus und überführt das Word-Transkript in die einheitliche Markdown-Struktur. Inhaltstragende Aussagen bleiben dabei unverändert, reine Hörsignale werden nur entfernt, wo dies die Aussage nicht verfälscht.

Mensch-Aufgaben

Die forschende Person trifft die Anonymisierung und prüft die bereinigten und geglätteten Fassungen, die sie am Entscheidungs-Gate freigibt.

6.2.3. Schritt 2: Deduktive Strukturierung der Hauptkategorien

Die Hauptkategorien werden deduktiv aus Forschungsfrage, Unterfragen, den beiden Gesprächsleitfäden und dem Projektkontext abgeleitet und danach in einen Kodierleitfaden mit Definition, Ankerbeispiel und Kodierregel je Kategorie überführt, der in einem Probedurchlauf am ersten Interview geschärft wird. Der vollständige Kodierleitfaden der Hauptkategorien findet sich (siehe Anhang G, S. 122). Es entsteht ein nominales System aus acht kodierten Hauptkategorien, während die reine Rollen- und Fallbeschreibung als Kontext ausserhalb des Kategoriensystems geführt wird und unkodiert bleibt. Die acht Hauptkategorien sind:

- HK2 Bedeutung und Funktion des Assessments
- HK3 Bestehender Analyseprozess, Ablauf und Werkzeuge
- HK4 Schwachstellen und Pain Points
- HK5 Positive Erfahrungen und Stärken
- HK6 Anforderungen, Killerkriterien und Soll-Vorstellungen
- HK7 KI-Einsatz, Haltung und Nutzung
- HK8 Aufgabenteilung Mensch und KI, Verantwortung und Ergebnisverlässlichkeit
- HK9 Geschäftliche und wirtschaftliche Rahmenbedingungen

Der vollständige Materialdurchgang erfolgt zweiphasig und pro Interview, sodass der Kodierakt je Interview abgeschlossen wird und sich nicht über die Transkripte fortträgt. Stellen, die inhaltlich auf mehrere Hauptkategorien abbilden, werden dabei mehrfach kodiert. Jede Zuordnung ist anhand der Definition im Kodierleitfaden begründet und auf ihre Belegstelle rückführbar.

KI-Aufgaben

Die KI leitet die Kandidaten für die Hauptkategorien mit je einer Arbeitsdefinition ab. Sie liefert im Materialdurchgang die Markierung der Fundstellen sowie die Zuordnungsvorschläge samt Begründung und Belegstelle. Sie prüft die Konsistenz über das Material und liefert damit Vorschläge, keine festgelegten Kategorien und keine Befunde.

Mensch-Aufgaben

Die forschende Person legt das Kategoriensystem fest, also die acht Hauptkategorien und die nominale Anlage, und entscheidet die Abgrenzungen, etwa die gebündelte Führung der Anforderungen, die eigenständige Führung der wirtschaftlichen Rahmenbedingungen und die Schärfung der Grenze zwischen HK2 und HK5. Sie entscheidet über jede Zuordnung und gibt den Materialdurchgang je Interview frei.

6.2.4. Schritt 3: Induktive Subkategorienbildung

Innerhalb jeder Hauptkategorie werden die Subkategorien induktiv am Material gebildet. Als Selektionskriterium dient die Definition der jeweiligen Hauptkategorie. Das Abstraktionsniveau wird durchgängig auf ein mittleres Niveau festgelegt. Nach jeder Hauptkategorie wird ein Revisionspunkt gesetzt und das Zwischenergebnis freigegeben. Wie eng das Fachurteil steuert, zeigt sich an der umfangreichsten Hauptkategorie, den Anforderungen und Soll-Vorstellungen. Das Material bündelt sich inhaltlich entlang von Effizienz, Entlastung, Schlankheit, Integration, Qualität und weiteren Aspekten. Daher werden deren formbasierte Aufteilung, zunächst am Gate auf eine inhaltsbasierte Aufteilung korrigiert wird.

KI-Aufgaben

Die KI identifiziert die einschlägigen Sinnabschnitte innerhalb der jeweils aktiven Hauptkategorie, bildet Subkategorienvorschläge nahe am Material auf dem festgelegten Abstraktionsniveau und liefert die Belegstellen. Bei jeder weiteren Stelle schlägt sie vor, ob sie unter eine bestehende Subkategorie fällt oder eine neue zu bilden ist.

Mensch-Aufgaben

Die forschende Person legt Selektionskriterium und Abstraktionsniveau fest und entscheidet je Stelle über Subsumtion oder Neubildung. Sie nimmt die inhaltliche Korrektur der Aufteilung bei den Anforderungen vor und gibt das Subkategoriensystem je Hauptkategorie frei.

6.2.5. Schritt 4: Konsistenzprüfung und Aufbereitung

Nach Abschluss der Subkategorienbildung wird das gesamte Set an Subkategorien über alle Hauptkategorien hinweg auf Konsistenz geprüft. Frühere Kodierungen werden nachgeführt. Dabei werden inhaltliche Überschneidungen bereinigt und einzelne Subkategorien zusammengeführt oder einer treffenderen Hauptkategorie zugeordnet. So entsteht aus dem ursprünglichen Entwurf ein konsolidiertes, in sich stimmiges System, das die Zahl der Subkategorien leicht reduziert (von 45 auf 41). Subkategorien, die nur einmal belegt sind, werden bewusst beibehalten, weil bei einer kleinen, gezielt gewählten Teilnehmerzahl auch einzelne Aussagen relevant sind. Anschliessend werden die Sinneinheiten je Subkategorie zusammengetragen und auf einem einheitlichen Abstraktionsniveau paraphrasiert. Eine abschliessende Repräsentationsprüfung je Hauptkategorie bestätigt, dass die verdichtete Fassung das ursprüngliche Material weiterhin abbildet.

KI-Aufgaben

Die KI prüft die Konsistenz des Subkategoriensets, schlägt Nachführungen und Zusammenführungen vor und meldet Überschneidungen. Für die Aufbereitung trägt sie die Sinneinheiten je Subkategorie zusammen und erzeugt die Paraphrasen auf dem einheitlichen Abstraktionsniveau, unter Beibehaltung des Bezugs zur Belegstelle.

Mensch-Aufgaben

Die forschende Person entscheidet über die Zusammenführungen und Streichungen, beschliesst die Beibehaltung der einmal belegten Subkategorien und prüft die inhaltliche Korrektheit der Verdichtung. Sie gibt das konsistente Kategoriensystem und die Aufbereitung frei.

6.2.6. Schritt 5: Interpretation und Befundverdichtung

Die strukturierten und verdichteten Kategorien bilden die Grundlage für die Interpretation. Die Befunde werden je Unterfrage geordnet und mit den einschlägigen Hauptkategorien verknüpft. Auch hier greift das Fachurteil korrigierend ein, indem eine erste Fassung, die einzelne Unterfragen vorschnell einschränkt, verworfen und die Strukturierung über das volle Spektrum der Unterfragen wiederholt wird. Über die Hauptkategorien hinweg werden Bezugspaare und Spannungsfelder herausgearbeitet. So etwa die Spannung zwischen Effizienz und Qualität oder zwischen Automatisierung und menschlicher Unersetzlichkeit,

ergänzt um Gegenperspektiven. Jede Kernaussage wird durch Belege gestützt, und es wird aktiv nach Gegenbelegen gesucht. Eine Aussage bleibt dabei ohne direkten Gegenbeleg und wird als offener Punkt markiert.

KI-Aufgaben

Die KI trägt die Befunde je Unterfrage strukturierend zusammen, verdichtet sie und erzeugt die Vorschläge für Bezugspaare, Spannungen und Gegenperspektiven. Zu jeder Kernaussage stellt sie die stützenden Belege bereit und sucht aktiv nach Gegenbelegen. Sie bleibt damit auf strukturierende Vorschläge begrenzt.

Mensch-Aufgaben

Die forschende Person deutet die nun strukturierten Daten, entscheidet, welche Befunde welche Unterfrage beantworten, und verantwortet die Synthese. Sie verwirft bei Bedarf die Einschränkung der Unterfragen und veranlasst die vollständige Wiederholung, und sie wägt Belege und Gegenbelege ab.

6.2.7. Schritt 6: Prüfung der Ergebnisverlässlichkeit

Abschliessend wird die Ergebnisverlässlichkeit in ihren zwei Dimensionen geprüft, der Vermeidung von Halluzinationen und der Reproduzierbarkeit, wobei alle drei Interviews vollständig einbezogen werden. Die Belegbarkeitsprüfung bestätigt, dass alle Fundstellen, Paraphrasen und Befunde rückführbar sind, es werden keine nicht belegten Stellen und keine Halluzinationsfunde festgestellt. Zur Reproduzierbarkeit wird eine Wiederholungskodierung mit der Erstkodierung verglichen. Die Übereinstimmung ist hoch und liegt auf der Ebene der Hauptkategorien bei geschätzt rund 95 Prozent, unter Einbezug der vollständigen Mehrfachkodierungen bei geschätzt rund 85 bis 90 Prozent. Diese Werte sind als Schätzungen ausgewiesen, und es wird offen vermerkt, dass die Wiederholung von derselben Modellinstanz im selben Arbeitskontext erfolgt. Daher messen sie primär die Robustheit des Kodierleitfadens an den Grenzfällen. Die wenigen festgestellten Abweichungen konzentrieren sich auf Mehrfachkodierungen und bekannte Grenzpaare benachbarter Kategorien.

KI-Aufgaben

Die KI prüft die Belegbarkeit aller übernommenen Aussagen, führt die Wiederholungskodierung ausschliesslich anhand von Kodierleitfaden und Textsegmenten aus und vergleicht das Ergebnis mit der Erstkodierung. Sie stellt Übereinstimmungen und Abweichungen mit Belegstelle dar.

Mensch-Aufgaben

Die forschende Person bewertet die Übereinstimmung und die Belegbarkeit. Sie beurteilt die Abweichungen als nachvollziehbar und entscheidet gegen eine Anpassung der Erstkodierung. Sie benennt den Nachschärfungsbedarf, also eine präzisere Mehrfachkodierungsregel, geschärfte Kodierregeln an den Grenzpaaren und eine unabhängige Zweitkodierung. Zusätzlich trägt sie die methodische Letztverantwortung.

6.2.8. Einordnung der Durchführung

Der Durchlauf zeigt das Vorgehen in seiner praktischen Anwendung. Die Trennung von Fleissarbeit und Fachurteil bleibt über alle sechs Schritte gewahrt. Sie kommt zu Tragen an den Stellen, an denen die forschende Person die KI-Vorschläge korrigiert. Etwa bei der inhaltsbasierten Aufteilung der Anforderungen oder beim Verwerfen der vorzeitigen Einschränkung in der Interpretation. Die Materialbindung ist durchgängig eingelöst. Alle Fundstellen und die daraus verdichteten Befunde sind rückführbar. Die Ergebnisverlässlichkeit wird mit ihren zwei Dimensionen geprüft und in ihren Grenzen offen ausgewiesen. Das Protokoll macht über alle Bearbeitungszyklen hinweg nachvollziehbar, wo die KI eingesetzt wurde, wie sie gesteuert war und wie die Ausgabe verifiziert wurde.

6.3. Herleitung der Kernbedürfnisse

Die Kernbedürfnisse werden in drei Stufen aus dem Interviewmaterial verdichtet, von den wörtlichen Sinnabschnitten über Paraphrasen und Nutzerbedürfnisse bis zu den Kernbedürfnissen (Mayring 2022).

Tabelle 6: Verdichtungsstufen von Sinnabschnitten zu Kernbedürfnissen (Anzahl je Ebene)

Ebene	Anzahl
Sinnabschnitte	102
Paraphrasen	44
davon ohne Nutzerbedürfnis-Bezug	7
Nutzerbedürfnisse	12
Kernbedürfnisse	5

Zuerst werden die Sinnabschnitte zu Paraphrasen zusammengefasst, indem bedeutungsgleiche Aussagen über Sprecher und Interviewstellen hinweg gebündelt werden. Die meisten Paraphrasen fassen mehrere Sinnabschnitte zusammen, einige beruhen auf einem einzelnen.

Anschliessend werden die Paraphrasen gruppiert und in Nutzerbedürfnisse übertragen. Jedes Nutzerbedürfnis ist also durch mehrere Paraphrasen belegt. Ein Teil der Paraphrasen bleibt ohne direkte Übertragung, da er geschäftliche Rahmenbedingungen, das Setup des Ist-Prozesses oder den allgemeinen Umgang mit KI betrifft.

Schliesslich werden die Nutzerbedürfnisse zu den Kernbedürfnissen zusammengefasst. Die folgende Tabelle zeigt die Gruppierung und die Zahl der stützenden Paraphrasen je Nutzerbedürfnis. Damit lässt sich jedes Kernbedürfnis über die zugeordneten Nutzerbedürfnisse und Paraphrasen bis zu den wörtlichen Sinnabschnitten und ihrem Beleg zurückverfolgen. Die vollständige Liste der Nutzerbedürfnisse mit Schilderungen befindet sich (siehe Anhang H, S. 126).

Tabelle 7: Gruppierung der Nutzerbedürfnisse zu Kernbedürfnissen mit Zahl stützender Paraphrasen

Kernbedürfnis	Nutzerbedürfnis	Paraphrasen
KB-01 Tragfähige Analyse mit vertretbarem Aufwand	NB-01 Assessmentdaten in eine tragfähige Einschätzung überführen	15
	NB-02 Die eigene Aufmerksamkeit auf fachlich wertvolle Arbeit richten	12
	NB-12 Das Assessment auch in begrenzten Projektrahmen einsetzen können	11
KB-02 Orientierung und fachliche Verdichtung	NB-03 In heterogenen Daten rasch Orientierung gewinnen	8
	NB-04 Den Ist-Zustand mit Zielbild und Framework verbinden	5
KB-03 Fachliche Verantwortung und Beurteilbarkeit	NB-06 Ergebnisse fachlich vertreten können	5
	NB-07 Die eigene Fachrolle im Analyseprozess erhalten	6
	NB-08 Unsicherheit vor der Weitergabe klären	3
KB-04 Verlässlichkeit und ausreichende Perspektivenbreite	NB-09 Zu vergleichbaren Einschätzungen gelangen	3
	NB-10 Relevante Perspektiven ausreichend berücksichtigen	6
KB-05 Handlungsorientierung und Anschlussfähigkeit der Ergebnisse	NB-05 Aus den Daten begründete Entwicklungsschritte ableiten	9
	NB-11 Ergebnisse in Kundensituationen anschlussfähig machen	10

6.4. Kernbedürfnisse

6.4.1. Geschäftsziele in den Bedürfnissen

Aus den Befunden der Analyse werden manuell zwölf mögliche Nutzerbedürfnisse identifiziert. Aufgrund der Gesprächsleitfäden sind auch Geschäftsziele erhoben worden. Dementsprechend lassen sich Geschäftsziele als Bedürfnisse in den Befunden identifizieren. Auch dies ist eine bewusste Entscheidung des Autors in seiner Doppelrolle. Wie im Projektkontext erklärt, decken sich die Geschäftsbedürfnisse mit den Nutzerbedürfnissen. Explizit sind dies beispielsweise die Wirtschaftlichkeit und der flexiblere Einsatz in kleineren Projekten. Wird dies ermöglicht, dann können sowohl Nutzerbedürfnisse als auch Geschäftsbedürfnisse erfüllt werden. Dabei bleibt jedoch die Gewichtung bei der Umsetzung auf den Nutzenden. Konkret bedeutet dies, dass ein Geschäftsbedürfnis zur Effizienzsteigerung nicht berücksichtigt wird, wenn es zur wahrgenommenen Unzufriedenheit oder Benachteiligung der Fachpersonen führt.

6.4.2. Die Kernbedürfnisse

Die zwölf Nutzerbedürfnisse lassen sich zu fünf Kernbedürfnissen verdichten. Diese Verdichtung reduziert die Detailtiefe der Nutzerbedürfnisse nicht, sondern bündelt sie entlang wiederkehrender Nutzungsmuster und Spannungsfelder. Im Zentrum stehen effiziente Analysearbeit, fachliche Verantwortung, Qualitätserhalt, planbare Prozessnutzung und die Verwertbarkeit der Ergebnisse. Damit bilden die Kernbedürfnisse eine Brücke zwischen der empirischen Bedarfsanalyse und den später abzuleitenden Anforderungen an den KI-unterstützten Soll-Prozess.

Tabelle 8: Liste der verdichteten Kernbedürfnisse

Nr.	Kernbedürfnis	Schilderung
KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, Assessmentdaten so auszuwerten, dass eine fachlich belastbare Einschätzung entsteht, ohne dass der Analyseaufwand den Projektrahmen unverhältnismässig belastet.
KB-02	Orientierung und fachliche Verdichtung	M365-Fachpersonen brauchen eine Auswertung, die heterogene Daten in eine verständliche fachliche Ordnung bringt, zentrale Muster und Spannungen sichtbar macht und den Ist-Zustand mit Zielbild und Framework in Beziehung setzt.
KB-03	Fachliche Verantwortung und Beurteilbarkeit	M365-Fachpersonen brauchen Analyseergebnisse, die sie prüfen, einordnen, klären und gegenüber der Kundschaft fachlich vertreten können.
KB-04	Verlässlichkeit und ausreichende Perspektivenbreite	M365-Fachpersonen brauchen eine Analyse, die über unterschiedliche Datengrundlagen hinweg zu konsistenten Einschätzungen führt und dabei genügend relevante Perspektiven berücksichtigt.
KB-05	Handlungsorientierung und Anschlussfähigkeit der Ergebnisse	M365-Fachpersonen brauchen Analyseergebnisse, aus denen sich begründete Reifegradeinschätzungen, Handlungsfelder und nächste Entwicklungsschritte ableiten lassen und die sich in Kundengesprächen, Entscheidungsprozessen und Massnahmenplanung sinnvoll weiterverwenden lassen.

6.4.3. Bündelung der Nutzerbedürfnisse zu Kernbedürfnissen

Die Zuordnung der Nutzerbedürfnisse zu den Kernbedürfnissen folgt dem jeweils verbindenden Nutzungsmuster. Massgeblich ist nicht die Automatisierung als Selbstzweck, sondern der gemeinsame fachliche Anspruch hinter den gebündelten Bedürfnissen.

KB-01 bündelt NB-01, NB-02 und NB-12

Verbindendes Muster ist die Spannung zwischen fachlich belastbarer Auswertung und vertretbarem Aufwand. Alle drei betreffen die Tragfähigkeit der Analyse unter Aufwands- und Projektrestriktionen, nicht den Wunsch nach Automatisierung.

KB-02 bündelt NB-03 und NB-04

Beide zielen auf Orientierung und fachliche Einordnung, indem sie das rasche Erkennen wesentlicher Muster mit der Einordnung des Ist-Zustands in Zielbild und Framework verbinden.

KB-03 bündelt NB-06, NB-07 und NB-08

Gemeinsamer Nenner ist die fachliche Verantwortung. Prüfbarkeit der Ergebnisse, erhaltenes Fachurteil und die Klärung von Unsicherheit sorgen dafür, dass die Verantwortung bei der Fachperson bleibt.

KB-04 bündelt NB-09 und NB-10

Beide betreffen die Belastbarkeit der Auswertung trotz Verschlinkung. Verlässlichkeit über unterschiedliche Datengrundlagen und ausreichende Perspektivenbreite bilden gemeinsam diese Grundlage.

KB-05 bündelt NB-05 und NB-11

Verbindendes Muster ist die Verwertbarkeit der Ergebnisse, die aus der Analyse Handlungsschritte entstehen lässt und diese in Kundensituationen anschlussfähig macht.

6.5. Ergebnisse

6.5.1. Anforderungen

Die Anforderungen sind je Kernbedürfnis nummeriert (NA1.x für KB-01, NA2.x für KB-02 und so weiter), womit die Rückführbarkeit auf die Kernbedürfnisse KB-01 bis KB-05 und damit auf die empirischen Befunde der Bedarfsanalyse erhalten bleibt.

Tabelle 9: Liste der Nutzeranforderungen

Nr.	Anforderungsname	Beschreibung	Prüfbarkeit
NA 1.1	Durchgängige Datenverarbeitung	M365-Fachpersonen müssen die Daten aus Interviews, Umfrage und technischen Abklärungen in einem durchgängigen Arbeitsablauf weiterverarbeiten können, ohne Inhalte manuell zwischen nicht verbundenen Einzelwerkzeugen zu übertragen.	Vergleich der Anzahl manueller Übertragungsschritte zwischen Ist-Prozess und Soll-Prozess.
NA 1.2	Delegation der Fleissarbeit	M365-Fachpersonen müssen wiederkehrende Verarbeitungsschritte, namentlich Aufbereitung, Strukturierung und Kernaussagen-Extraktion, an die KI-Unterstützung übergeben können.	Nachweis im Prozessprototyp, welche Verarbeitungsschritte ohne manuelle Ausführung durch die Fachperson erfolgen.
NA 1.3	Einhaltung der Aufwandsgrenze	M365-Fachpersonen müssen den gesamten Analyseprozess eines Assessments innerhalb der unternehmensseitigen Zielgrösse von maximal rund zwanzig Arbeitsstunden abschliessen können.	Effizienz-Evaluation gemäss Teilzielen, also Vergleich der dokumentierten Arbeitsstunden vergangener Assessments mit dem im Testing erhobenen Aufwand.
NA 2.1	Prüfung und Freigabe an Eingriffspunkten	M365-Fachpersonen müssen jedes KI-generierte Zwischenergebnis an definierten Eingriffspunkten prüfen, korrigieren und freigeben können, bevor es weiterverwendet wird.	Nachweis der Eingriffspunkte im dokumentierten Prozessablauf sowie Beobachtung im Usability-Testing, ob Korrekturen tatsächlich vorgenommen werden können.
NA 2.2	Rückführbarkeit auf das Ausgangsmaterial	M365-Fachpersonen müssen jede KI-generierte Aussage auf die zugrunde liegende Stelle im Ausgangsmaterial zurückführen können.	Belegbarkeitsprüfung im Rahmen der Evaluation der Ergebnisverlässlichkeit, Anteil der rückführbaren Aussagen.
NA 2.3	Nachvollziehbare Aufgabenteilung	M365-Fachpersonen müssen für jedes Ergebnis nachvollziehen können, welche Arbeitsschritte die KI ausgeführt hat und welche Entscheide bei der Fachperson lagen.	Vorliegen eines vollständigen Protokolls je Analysedurchlauf.

NA 3.1	Zielbild- und Reifegradbezug	M365-Fachpersonen müssen Analyseergebnisse erzeugen können, die den Bezug zum Zielbild der Kundenorganisation und zum Reifegradmodell des Assessments explizit herstellen.	Fachliche Beurteilung der Ergebnisartefakte im Testing daraufhin, ob Zielbild- und Reifegradbezug vorhanden sind.
NA 3.2	Kontrolle der Verdichtung	M365-Fachpersonen müssen überprüfen können, ob die verdichteten Ergebnisse das zugrunde liegende Material hinreichend abbilden und keine relevanten Perspektiven verloren gegangen sind.	Vorhandensein einer Repräsentationsprüfung im Soll-Prozess und deren Anwendung im Testing.
NA 3.3	Steuerbare Analysetiefe	M365-Fachpersonen müssen die fachliche Tiefe und Breite der Auswertung fallbezogen anpassen können, sodass die Verschlinkung gesteuert erfolgt und nicht pauschal zulasten der Aussagekraft geht.	Nachweis im Prozessprototyp, dass der Detaillierungsgrad je Assessment angepasst werden kann.
NA 4.1	Definierter Prozessablauf	M365-Fachpersonen müssen den Analyseprozess entlang eines definierten, schrittweisen Ablaufs mit klaren Verantwortlichkeiten und Übergaben zwischen Fachperson und KI durchführen können.	Vorliegen des dokumentierten Prozessablaufs und Beobachtung der Schrittfolge im Usability-Testing.
NA 4.2	Delegierbarkeit	M365-Fachpersonen müssen den Analyseprozess nach einer kurzen Einführung selbstständig anwenden können, ohne auf die Unterstützung der HCD-Fachperson angewiesen zu sein.	Usability-Testing, in dem Teilnehmende den Prozess ohne Eingreifen der HCD-Fachperson durchlaufen.
NA 4.3	Kalkulierbarer Zeitbedarf	M365-Fachpersonen müssen den Zeitbedarf eines Analysedurchlaufs vor Beginn auf Basis des definierten Ablaufs abschätzen können.	Vergleich von geschätztem und tatsächlich erhobenem Aufwand im Testing.
NA 5.1	Anschlussfähiges Summary	M365-Fachpersonen müssen am Ende des Analyseprozesses ein finales Summary erzeugen können, das ohne umfangreiche manuelle Nachbearbeitung als Lieferobjekt oder als Grundlage für die Abschlusspräsentation dient.	Fachliche Beurteilung des erzeugten Summaries im Testing, insbesondere des Umfangs der erforderlichen Nachbearbeitung.
NA 5.2	Ableitbare Handlungsfelder	M365-Fachpersonen müssen aus den Analyseergebnissen Handlungsfelder und mögliche Massnahmen für die Anschlussarbeit ableiten können.	Beurteilung im Testing, ob die Ergebnisartefakte Handlungsfelder und Massnahmen erkennbar machen.

6.5.2. Bezug zur Forschungsfrage (Teilfrage 1) und Zielsetzung (Teilziel 1)

Mit dieser Bedarfsanalyse sind die Schwachstellen, Anforderungen und Automatisierungspotenziale des bestehenden Analyseprozesses systematisch verortet und das Teilziel der IST- und Bedarfsanalyse erreicht. Der bestehende, manuell durchgeführte Analyseprozess überzeugt fachlich, bindet aber Ressourcen ausserhalb des fachlichen Kerns der M365-Fachpersonen. Als Schwachstellen treten ein hoher Auswertungsaufwand, eine manuelle Werkzeugkette mit Doppelarbeit über nicht verbundene Werkzeuge und teilweise nicht verrechenbare Aufwände hervor. Dieser wird verschärft, durch den nötigen Einbezug zusätzlicher UX-Fachpersonen. Als Resultat liegen die relevanten Prozessschritte, die fünf verdichteten Nutzerbedürfnisse mit den zugehörigen Anforderungen sowie die verorteten Einsatzpotenziale für KI-Unterstützung vor.

6.6. Lernreflexion – Phase 1

Das war der Teil der Arbeit, in dem das geplante Vorgehen zum ersten Mal vollständig auf die Praxis trifft. Hier habe ich die Interviews geführt, das KI-gestützte Auswertungsverfahren tatsächlich angewendet und aus dem Material Bedürfnisse und Anforderungen abgeleitet. Rückblickend ist es das Kapitel, in dem ich am meisten über die Differenz zwischen einem sauber geplanten Verfahren und seiner tatsächlichen Durchführung gelernt habe. Besonders, da das Verfahren so neu und von mir selbst erarbeitet wurde. Gleichzeitig habe ich dadurch die KI-Unterstützung zum ersten Mal nicht nur konzipiert, sondern am eigenen Material erprobt.

Die Interviews waren für mich ebenfalls ein neuartiger Fall. Ich führe seit Langem Gespräche zur Erhebung empirischer Daten und bin darin recht routiniert. Gerade das hat es hier schwierig gemacht. Meine Interviewpartner waren zwangsläufig Arbeitskollegen, die schon mehrere Evolution Assessments mit mir durchgeführt haben und meine Arbeitsweise kennen. Ein künstlich distanziertes, methodisch stures Interview zu führen, hätte sich in dieser Konstellation falsch angefühlt. Ich halte es sogar für den grösseren Methodenfehler, die Distanz zu forcieren, als bewusst ein angepasstes Gespräch zu führen, das im Kern seinen Zweck erfüllt. Denn am Ende geht es darum, Meinungen und Erfahrungen einzuholen. Methodisch blieb das Vorgehen halbstrukturiert mit einem Gesprächsleitfaden. Doch da ich ohnehin nur drei Personen befragte, habe ich die Gespräche bewusst individuell gehalten. Eine Quantifizierung, wie oft ein Thema genannt wird, war nie mein Ziel. Bei einer so kleinen und gezielt ausgewählten Gruppe ist jede einzelne Aussage potenziell wichtig und genau darauf habe ich die Auswertung ausgerichtet.

Den grössten Erkenntnisgewinn brachte die tatsächliche Durchführung von Vibe Research. Was im Methodenkapitel sauber als Vorgehen beschrieben ist, war in der Anwendung deutlich mühsamer, und der Aufwand lag weniger im Auswerten, als im stabilen Aufbau der Anweisungen, der Übergabepunkte und der Kontrollschritte. Wichtig ist mir die Einordnung, dass ich diese Durchführung weniger als reinen Analyseschritt verstanden habe, sondern als eine Art Übungsfläche. Ich habe die Interviews genutzt, um das Verfahren am Interviewmaterial zu erproben. So konnte ich verstehen lernen, wo die KI-Unterstützung wirklich trägt und wo sie eng kontrolliert werden muss, um daraus konkrete Erkenntnisse für die Gestaltung des eigentlichen Light Assessment Prozesses abzuleiten. In diesem Sinn war die Inhaltsanalyse für mich doppelt nützlich, einmal als Auswertung der Interviews und einmal als Testlauf für das Produkt, das ich anschliessend gestalten wollte. Am Ende meiner Probierphasen war es mir dann möglich, die gesamte Inhaltsanalyse in einem Tag gemäss Vorgehen durchzuführen und diese Ergebnisse habe ich dann verwendet.

Ähnlich verlief die Herleitung der Bedürfnisse und Anforderungen. Auch hier habe ich die KI-Unterstützung ausprobiert und bin am Ende fast vollständig zu manueller Überarbeitung zurückgekehrt. Meine gewohnte Herangehensweise, in einer Excel-Tabelle Spalte für Spalte zu verdichten, kommt den Fähigkeiten eines Sprachmodells zwar grundsätzlich entgegen, doch nur, solange konsequent auf Halluzinationen geachtet wird, und genau dieser Kontrollbedarf hat mich am Ende ohnehin wieder selbst an die Verdichtung gebracht. Inspiriert von der Gegenprüfung aus der Inhaltsanalyse habe ich dabei einen eigenen kleinen Versuch eingebaut. Nach dem Verdichten und Überarbeiten habe ich jeweils einen separaten Auftrag zur rückwärtigen Zuteilung erstellt, bei dem die KI den verdichteten Inhalt aus einer eigenen Tabelle erneut seinem Ursprung zuordnen musste. Diese Zuordnung habe ich gegen die ursprüngliche Verdichtung gehalten und die Unterschiede verglichen. Ehrlicherweise war das am Ende eher eine Spielerei zum besseren Verständnis, wo die Stärken und Schwächen tatsächlich liegen, als ein echter Prozessschritt, und ich erwähne es nur der Transparenz halber. Die Ergebnisse dieses spontanen Testversuchs habe ich verworfen. Das eigentliche Ergebnis dieses Tests war ein anderes, nämlich der für das Light Assessment angepasste Kodierleitfaden und wie sich die KI unter Umständen selbst kontrollieren kann in einem Analyseprozess, der geschlossen funktionieren und am Ende dennoch nachvollziehbar sein muss.

Mein persönliches Fazit aus dieser Phase ist, dass der Wert der KI-Erprobung weniger in den unmittelbaren Resultaten lag, als in dem, was ich dabei über das Verfahren gelernt habe. Vieles, was ich mit KI probiert habe, habe ich am Ende manuell überholt oder verworfen, und trotzdem war keiner dieser Versuche umsonst. Sie haben mir gezeigt, wo die Aufgabenteilung zwischen mir und der KI tatsächlich sinnvoll verläuft, wie viel Kontrolle nötig bleibt und wie eine stabile Anweisung aussehen muss. Genau diese Erfahrung, und nicht die einzelnen Zwischenergebnisse, ist in die Gestaltung des Light Assessment Prozesses eingeflossen. Rückblickend war Phase 1 für mich weniger eine reine Erhebung, als ein praktisches Lernfeld, auf dem ich mein eigenes Vorgehen für den späteren Prozess geschärft habe.

7. Phase 2: Konzeption

7.1. Participatory Design

Phase 2 überführt die in Phase 1 erhobenen Bedürfnisse (siehe [Kapitel 6](#)) und die daraus abgeleiteten Anforderungen in die Konzeption des KI-unterstützten Soll-Prozesses. Tragender methodischer Zugang dieser Phase ist Participatory Design. Das vorliegende Kapitel dokumentiert, wie die Beteiligung der Fachpersonen konkret umgesetzt wurde und verortet dazu die partizipativen Aktivitäten des gesamten Projektverlaufs im Einordnungsraster nach Gaete Cruz et al. (2022).

7.1.1. Begründung der Wahl

Die Verankerung des Participatory Design als ergänzender Methodenraum der Phasen 1 bis 3 ist im Methodenkapitel begründet (siehe [Kapitel 4](#)), ebenso die Festlegung, die konkrete Anwendung exemplarisch in der Konzeptionsphase zu dokumentieren. Diese Verortung hat zwei Gründe. In dieser Phase fallen die gestaltungsprägenden Entscheide der Arbeit, also der Zuschnitt des Soll-Prozesses, die Aufgabenteilung zwischen Fachperson und KI sowie die fachliche Kodierung der Analyse. Genau diese Entscheide hängen von implizitem Praxiswissen ab, das sich weniger über distanzierte Befragung als im gemeinsamen Arbeiten an konkreten Artefakten zeigt (Spinuzzi 2005). Die Beteiligung der Projektgruppe ist in der Konzeption deshalb am dichtesten und erreicht hier mit der gemeinsamen Erarbeitung der Kodierung die höchste Kollaborationsstufe des Projekts.

Hinzu kommt, dass sich die Beteiligung nur als zusammenhängender Verlauf sinnvoll beurteilen lässt. Die partizipativen Aktivitäten setzen bereits während der Kontextphase ein und reichen bis in die Gestaltung und die formative Evaluation hinein. Damit der Projektverlauf dennoch als ein gemeinsam getragener Co-Design-Prozess nachvollziehbar bleibt, führt das vorliegende Kapitel alle Beteiligungsaktivitäten in einem Raster zusammen und beschreibt ihren Ablauf im Zusammenhang.

7.2. Konkrete Durchführung

7.2.1. Setting und Beteiligte

Den partizipativen Rahmen bildet das im Projektkontext abgegrenzte Pilot-Projekt für eine pseudonymisierte Kundenorganisation, das zwischen Januar und April 2026 durchgeführt wurde. Die Projektgruppe umfasst fünf M365-Consultants aus drei Firmen. Dadurch decken die Teilnehmenden im Rahmen der Möglichkeiten die in der Stakeholder Analyse (siehe [Abschnitt 5.2](#)) erwähnte Problematik des unter Fachpersonen zusammengefassten breiten Primärnutzerfeldes ab. Ebenfalls Teil des Teams ist der Autor. Im partizipativen Kontext hat dieser die Rolle der HCD-Fachperson und wird im diesem Kapitel auch so benannt. Die Beteiligung wurde nicht als eigenständige Workshopreihe neben dem Projekt organisiert, sondern in die reale Projektarbeit eingebettet. Die Projektgruppe erhob echte Kundendaten, verantwortete ein verbindliches Lieferobjekt und prüfte die entstehenden Teilprozesse unmittelbar an diesem Material. Dieses Setting entspricht dem Grundverständnis des Participatory Design, wonach Nutzende als Fachpersonen

ihrer eigenen Arbeitspraxis an der Gestaltung mitwirken, während die gestaltende Person die Auseinandersetzung strukturiert und moderiert (Sanders und Stappers 2008).

Als Beteiligungsinfrastruktur dienten die ohnehin genutzten M365-Arbeitsmittel, namentlich ein gemeinsamer Teams-Arbeitsbereich mit geteilten Ablagen, kommentierbare Arbeitsdokumente sowie regelmässige Arbeits- und Reviewtermine. Die Zusammenarbeit verband damit synchrone Besprechungen mit asynchronen Kommentar- und Reviewrunden und blieb für alle Beteiligten im gewohnten Arbeitsumfeld niederschwellig zugänglich.

7.2.2. Einordnung der Beteiligung

Zur differenzierten Beschreibung des Einbezugs dient das im Methodenkapitel eingeführte angepasste Raster von Gaete Cruz et al. (2022), welches die Beteiligungsleiter mit dem Designzyklus verbindet. Konsultation bezeichnet darin das beratende Einbringen von Wissen und Rückmeldungen ohne geteilte Entscheidungsbefugnis, Partizipation die Mitentscheidung in einzelnen Teilfragen und Kollaboration die wiederkehrende gemeinsame Entscheidungsfindung (Gaete Cruz et al. 2022). In der untersten Stufe Information werden keine Aktivitäten ausgewiesen, weil die vorwiegend informativen Tätigkeiten sich auf Themen beziehen, welche sich nicht im Scope dieser Arbeit befinden. In der vorliegenden Adaption ist die Achse des Designzyklus an den Aktivitäten des Vorgehensmodells ausgerichtet, also Research, Konzeption, Gestaltung und Evaluation. Dadurch kann der angewandte iterative Ansatz des HCD-Kreislaufes innerhalb des Methodenraums übersichtlicher verstanden werden.

7.3. Der partizipative Prozess

7.3.1. Die Kerntätigkeiten während dem Participatory Design

Insgesamt lassen sich neun Beteiligungsaktivitäten unterscheiden, die im Raster als T-1 bis T-6 verortet sind und im Folgenden in ihrem zeitlichen Verlauf beschrieben werden.

Tabelle 10: Kerntätigkeiten im Participatory Design nach Design Cycle und Collaboration Level

Nr.	Design Cycle	Collaboration level	Tätigkeit
T-1	research	Consultive	Grundlegenden Assessment-Prozess erarbeiten gemäss Bedürfnissen und Erwartungen
T-2	research	Participative	Identifikation und Massnahmen für Hürden KI (Human in the Loop und Datenqualität)
T-3	konzeption	Collaborative	Kodierung für Inhaltsanalyse
T-4.a	evaluation	Participative	Teilprozess «Transkript» mit Projektgruppe Testen
T-4.b	evaluation	Consultive	Anpassungen Teilprozess «Transkript»
T-5.a	konzeption	Participative	Teilprozess «Analyse» mit Projektgruppe Testen
T-5.b	evaluation	Participative	Besprechung und Reflexion von Analyseergebnissen
T-5.c	gestaltung	Consultive	Anpassungen Teilprozess «Analyse»
T-6	gestaltung	Participative	Teilprozess «Präsentation» mit Projektgruppe Testen

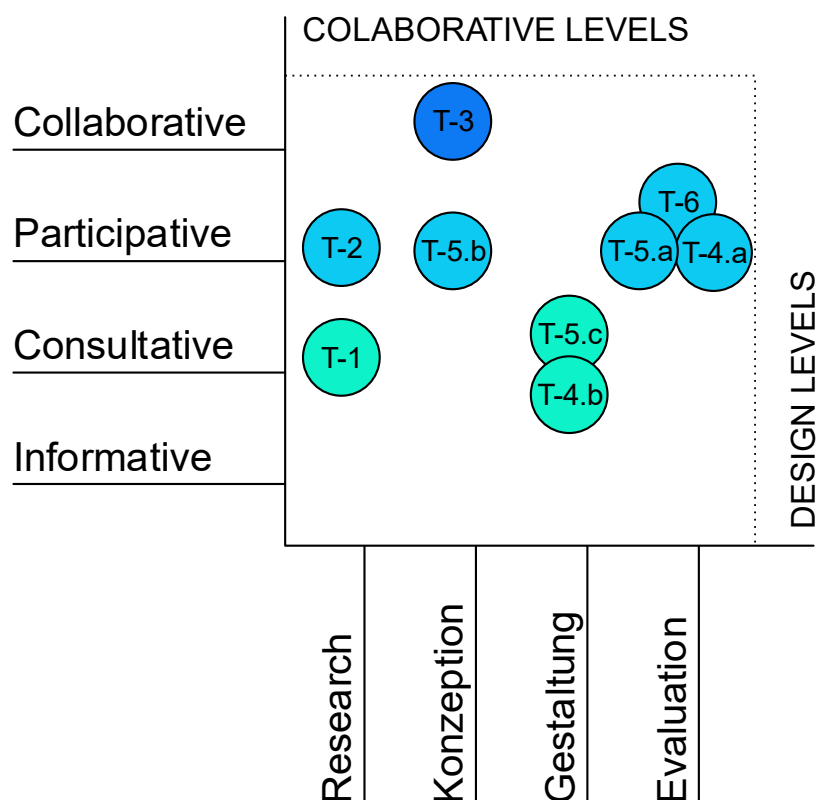


Abbildung 6: Adaptierte Co-Design Landscape Matrix zur Verortung kollaborativer Arbeit nach Collaborative Levels und Design Levels, in Anlehnung an Gaete Cruz et al. (2022) (Eigene Darstellung)

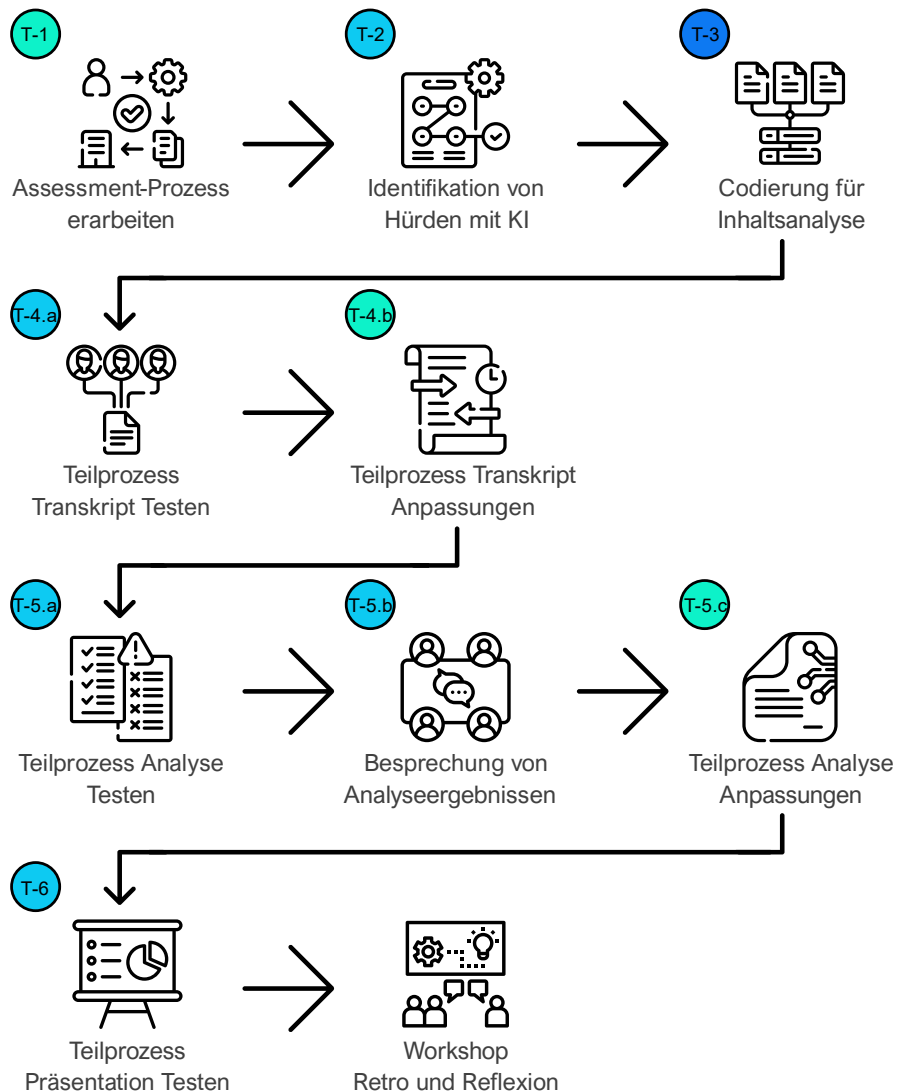


Abbildung 7: Übersicht des partizipativ durchgeführten Prozesses (Eigene Darstellung)

7.3.2. Von der Prozessgrundlage zur gemeinsamen Kodierung (T-1 bis T-3)

T-1: Den Ausgangspunkt bildete die konsultative Erarbeitung des grundlegenden Assessment-Prozesses gemäss den Bedürfnissen und Erwartungen der Fachpersonen. Auf Basis der Befunde aus Phase 1 (siehe [Kapitel 6](#)) und der Erwartungen aus der Projektgruppe entwarf die HCD-Fachperson die Grundstruktur des Analyseprozesses mit den drei Teilprozessen Transkript, Analyse und Präsentation, dem in der Abgrenzung definierten Scope folgend, also Aufbereitung der erhobenen Daten, Analyse und Erstellung des finalen Summaries. Die Consultants brachten ihr Praxiswissen beratend ein, die konzeptionellen Entscheide lagen in dieser frühen Aktivität bei der HCD-Fachperson. Im Raster ist die Aktivität dem Research zugeordnet, weil sie primär Wissen und Erwartungen zusammenführte und den Lösungsraum noch nicht festlegte.

T-2: Darauf aufbauend identifizierte die Projektgruppe partizipativ die Hürden eines KI-Einsatzes im Analyseprozess und legte entsprechende Massnahmen fest. Im Zentrum standen die Verlässlichkeit der KI-Ausgaben und die Qualität der Eingangsdaten, namentlich die Fehleranfälligkeit automatischer Transkripte und die eingeschränkte Aussagekraft automatisiert erzeugter Umfrageauswertungen, deren Standardvisualisierungen ohne fachliche Einordnung Fehlinterpretationen begünstigen. Gemeinsam beschlossen wurde

daraus, dass die Daten für die Kontrolle vollumfänglich und nachvollziehbar zur Verfügung gestellt werden sollen. Ausserdem müssen die Transkripte einmal von einer Fachperson inhaltlich korrigiert werden, insbesondere die Prüfung auf Worttreue vor der sprachlichen Glättung.

Diese Punkte wurden in einer offenen Diskussion besprochen, in welcher die Fähigkeiten und Limitationen von LLM-basierten Systemen sowie persönliche Erfahrungen der Fachpersonen im Umgang mit diesen besprochen wurden.

T-3: Das technische Kernstück der Konzeption bildet die Kodierung für die Inhaltsanalyse. Sie ist die einzige Aktivität auf der Stufe Kollaboration. Das liegt daran, dass sie einen konstanten Gegenstand der Diskussionen darstellt und inhaltlich in der Verantwortung der Fachpersonen liegt. Die HCD-Fachperson dient rein mit methodischer Beratung. Grundlage ist die Building-Block-Struktur des M365-Evolution-Modells, die in eine Kodierungslogik überführt wurde. Je Baustein wurden Zielzustände, mögliche Lücken und zugehörige Empfehlungen ausformuliert, sodass die spätere Auswertung empirischer Daten den Bezug zum Zielbild und zum Reifegradmodell des Assessments explizit herstellen kann, wie es die Anforderung NA3.1 verlangt. Die Erarbeitung verlief als wiederkehrende gemeinsame Entscheidungsfindung. Über Arbeitstreffen und Kommentarrunden hinweg entschied die Projektgruppe wiederholt über den Zuschnitt einzelner Bausteine, die Formulierung der Zielzustände und die Abgrenzung benachbarter Themenfelder. Die Kodierung wurde in mehreren Sitzungen vor dem Projekt definiert. Diese genaue Kodierung wurde dann durch die forschende Person für das LLM-basierte System aufbereitet. Die hohe Kollaborationsstufe ist inhaltlich begründet. Die Kodierung legt fest, entlang welcher fachlicher Begriffe die KI das Material strukturiert und die Fachpersonen die Ergebnisse deuten. Sie bündelt damit genau jenes implizite Praxiswissen, dessen Erfassung den Einbezug der Fachpersonen begründet (Spinuzzi 2005) und eine ohne geteilte Entscheidungsfindung festgelegte Kodierung würde die fachliche Anschlussfähigkeit und die Akzeptanz des gesamten Analyseprozesses gefährden. Der gefüllte Kodierleitfaden kann aus, wie zu Vertraulichkeitsgründen (siehe [Abschnitt 1.6](#)), nicht im Detail gezeigt werden. Daher reduziert sich die Information hierzu rein auf das verwendete Schema (siehe [Abschnitt 10.3](#)).

7.3.3. Frühes Testen des Teilprozesses Transkript (T-4)

T-4.a: Der erste neu definierte Teilprozess, die Bereinigung des Transkriptes, kam früh in die Erprobung. Drei Teilnehmende der Projektgruppe testeten den, als «Teilprozess Transkript» bezeichneten, Schritt an realen Interviewaufzeichnungen aus dem Pilot-Projekt. Sie prüften, ob die KI-gestützte Bereinigung und Konsolidierung der automatisch erzeugten Transkripte inhaltstreue und weiter verarbeitbare Arbeitsfassungen lieferten. Die Mitglieder der Projektgruppe kontrollierten die Worttreue gegenüber den Aufzeichnungen und entschieden mit, nach welchen Regeln bereinigt wird, etwa welche Bestandteile entfernt werden dürfen und welche Aussagen unverändert erhalten bleiben müssen. Diese Aktivität ist der Evaluation zugeordnet, weil sie ein bereits definiertes Prozessstück aus Nutzungsperspektive überprüfte, und sie entspricht dem in der Phasenplanung vorgesehenen frühen Testing neu definierter Prozessschritte.

T-4.b: Die anschliessenden Anpassungen der Bereinigungs- und Konsolidierungsregeln nahm die HCD-Fachperson konsultativ vor, die Entscheidung über die konkrete Umsetzung lag bei ihr. Exemplarische Beispiele zur Veranschaulichung werden gezeigt (siehe [Abschnitt 7.4](#)).

7.3.4. Iterationen am Teilprozess Analyse (T-5)

T-5.a: Den grössten Anteil der gemeinsamen Arbeit beanspruchte der Teilprozess Analyse. Er wurde gemäss Phasenplanung in drei Iterationen mit der Projektgruppe erprobt, besprochen und angepasst, die zugleich als formative Evaluation wirkten. Im Testen des Teilprozesses wertete die HCD-Fachperson die konsolidierten qualitativen und quantitativen Daten des Pilot-Projekts KI-gestützt aus und stellte die Ergebnisse der Gruppe als kommentierbare Arbeitsdokumente zur Verfügung. Zunächst als vollständige Auswertung, später ergänzt um eine priorisierte Fassung mit Summary. Die Bereitstellung war ausdrücklich mit der Aufforderung verbunden, Fehler zu markieren, einzelne Evidenzen stichprobenartig am Material zu

prüfen und die vorgeschlagenen Priorisierungen kritisch zu hinterfragen. Die Analyseergebnisse wurden damit nicht als finaler Wurf, sondern als überprüfbare Arbeitsartefakte behandelt. Die Gruppe wirkte unmittelbar am Testen mit und entschied in Teilfragen mit. So etwa, welche Befunde als tragfähig gelten und an welchen Stellen der Teilprozess nachzubessern ist. Auch wie die Artefakte besser aufbereitet werden können, wurde von ihnen mitgestaltet und entschieden. Die Kodierung kam an dieser Stelle ebenfalls wieder auf und Vorschläge zur Verbesserung wurden aufgenommen. Deshalb ist diese Aktivität insgesamt als partizipativ ausgewiesen.

T-5.b: In der Besprechung und Reflexion der Analyseergebnisse prüfte die Projektgruppe die Befunde in einem dedizierten Reviewtermin und in asynchronen Kommentarrunden. Die Rückmeldungen betrafen nicht nur einzelne Formulierungen, sondern die Begründungstiefe der qualitativen Befunde, die Verknüpfung der quantitativen Evidenzen, die Strukturierung entlang der Bausteine sowie die Übersetzung von Befunden in priorisierte Handlungsfelder und Massnahmen. Über Priorisierung und Massnahmenrichtung entschied die Gruppe gemeinsam.

T-5.c: Zwischen den Iterationen passte die HCD-Fachperson den Teilprozess an. Konsultativ gestützt auf die Rückmeldungen und stichprobenartigen Prüfungen aus der Gruppe. Die Überarbeitungen betrafen unter anderem die belegbare Begründung qualitativer Befunde, die systematischere Einbindung quantitativer Evidenzen in die Argumentation und die Stabilität der Verarbeitung über den gesamten Durchlauf. Im Raster ist diese Umsetzungsarbeit der Gestaltung zugeordnet, weil sie den Teilprozess in seiner konkreten Ausgestaltung veränderte.

7.3.5. Teilprozess Präsentation und Abschluss der partizipativen Phase (T-6)

Die letzte Aktivität galt dem Teilprozess Präsentation, also der Überführung der priorisierten Befunde in das kundengerechte Ergebnisartefakt. Die Projektgruppe entschied dabei mit, nach welcher Logik die Ergebnisse dargestellt werden, namentlich die Verdichtung auf eine reduzierte Zahl von Empfehlungen, denen jeweils konkrete Massnahmen zugeordnet sind, sowie die Konsistenz zwischen dem analytischen Ergebnisdokument und der Präsentation. Die Entwürfe wurden iterativ in der Gruppe abgestimmt und überarbeitet. Dabei entfiel die Befüllung und Gestaltung der Präsentationsunterlagen aus dem Scope der Arbeit. Die Projektgruppe entschied, die Automatisierungen zur Befüllung wären nicht erwünscht. Die Kontrolle über das Erstellen der Präsentationsunterlagen solle gänzlich manuell bleiben, wenn auch mit Hilfe einer Vorlage. Mit dieser Entscheidung entfiel die Verantwortung der HCD-Fachperson für die Präsentationsunterlagen und diese half nur noch unterstützend bei der Gestaltung. Den Abschluss der partizipativen Phase bildeten die Ergebnispräsentation bei der Kundenorganisation, mit anschliessendem strukturiertem Feedback sowie ein interner Reflexions-Workshop im Anschluss des Projektes, unter Leitung der HCD-Fachperson.

Die anschliessende Überführung in den Prototypen und die Erprobung mit Nutzenden ausserhalb der Projektgruppe sind Gegenstand der Phase Evaluation. Da planmässig an dieser Stelle keine der teilnehmenden Personen des Pilot-Projektes mehr am finalen Nutzertest teilgenommen hat, wird diese Phase strikt nicht als Teil des Partizipativen Designs betrachtet.

7.4. Exemplarische Vertiefung: Transkript Bereinigung

Gerade da es bei Fedosov et al. (2026) ein Thema ist, werden an dieser Stelle praktische Beispiele zur Transkriptbereinigung gezeigt. Sie veranschaulichen wie wichtig der verantwortungsvolle Umgang, bezüglich Kontrolle mit den Rohdaten ist. Eine Haltung, welche sich natürlich durch den Prozess zieht.

Inkoherente Sprache

Menschen reden, besonders wenn sie nachdenken, nicht so kohärent, wie wir es glauben. Der Kopf filtert das Gesprochene und setzt es für uns verständlich zusammen. Bei einer Transkription wird genau das aufgeschrieben, was gesagt wird, aber auch nur so weit, wie es verstanden wird.

Inhaltliche und kontextuelle Korrekturen

Eine der essenziellsten Punkte ist der Umgang mit der Art von Fehlern, welche als Fehler identifiziert werden müssen, doch nicht offensichtlich falsch sind.

Die Reihenfolge der Bereinigung

Man kann argumentieren, dass es doch sinnvoll ist, die KI erst das Skript bereinigen zu lassen und anschliessend von einem Menschen kontrollieren zu lassen. So müsste es nur einmal korrigiert werden. Doch es ist nun mal so, dass manche Fehler dann bereits verarbeitet oder weggelassen werden, sodass es schwieriger ist, sie zu korrigieren. Ausserdem verändern sich durch die notwendige sprachliche Glättung teilweise ganze Satzstrukturen und Aussagen. Es ist für die Fachpersonen leichter, dem nicht bereinigten, aber exakt dem Audio entsprechenden, Skript entlang einzelne Begriffe zu korrigieren und den Satzbau dabei ignorieren zu können, als die umgestellten Sätze, welche vom Gehörten abweichen, zu bearbeiten und umzuschreiben.

7.4.1. Praktische Beispiele zur Notwendigkeit der Inhaltskontrolle.

Diese folgenden Beispiele stammen aus den echten Projektdaten. Sie sind zwecks Datenschutzes teilweise angepasst oder zensiert und nur auf das Nötigste beschränkt. Doch die Kernprobleme sind unverändert. Das erste Beispiel enthält einen Ausschnitt aus einem Transkript. Zunächst ist der Text ohne menschliche Korrektur, einmal roh und einmal bereinigt. Dann der selbe Text mit menschlicher Korrektur und bereinigt.

Tabelle 11: Veranschaulichung von Transkriptionsfehlern und inhaltlicher Korrektur am Beispiel eines ganzen Abschnitts

Rohtext unkorrigiert	Durch KI bereinigt
«[...]und ich habe einfach die Kabel mein mein Computer kabisiert und kein Video, kein Ton und das war neu hab ich nicht verstanden warum [...] Mutter hat mir gesagt ja vielleicht poleme ist einfach der Kabel und ich ein Computer [...] Die US Bekabel und wir haben die Kirsche ich glaube, die Informatik. Es gibt wie eine Informatik unternehmen, die für die ganze Zahlen verantwortlich ist, oder und ich habe ihn gerne gerufen wie gesagt ja, ich muss schauen, dass damals noch 34 Tage und am Schluss ja, sie haben 100 Meter neue Kabel machen müssen.»	«[...]und ich habe einfach die Kabel an meinen Computer angeschlossen, aber es gab kein Video und keinen Ton. Das war neu für mich, ich habe nicht verstanden, warum. [...] Meine Mutter hat mir gesagt, ja, vielleicht liegt es einfach am Kabel oder am Computer [...] Die USB-Kabel – und wir haben dann, ich glaube, die Informatik kontaktiert. Es gibt ja so ein Informatikunternehmen, das für das Ganze zuständig ist, oder? Und ich habe dort angerufen und gesagt, ja, ich muss schauen [...] damals waren es noch etwa vierunddreissig Tage und am Schluss haben, sie dann hundert Meter neue Kabel machen müssen.»
Inhaltlich korrigiert	Durch KI bereinigt
«[...] Und ich habe einfach die Kabel mein mein Computer kalibriert und kein Video, kein Ton und das war neu hab ich nicht verstanden warum[...] Und er hat mir gesagt ja vielleicht Problem ist einfach der Kabel und nicht der Computer[...] Die USB kabel und wir haben die [Bestimmte Firma] ich glaube, die Informatik ist das. Es gibt wie eine Informatik unternehmen, die für die ganze Sache verantwortlich ist, oder und ich habe ihn gerne gerufen wie gesagt ja, ich muss schauen,[...] damals waren es noch 3-4 Tage und am Schluss, haben sie 100 Meter neues Kabel legen müssen.»	«[...] Ich habe die Kabel an meinem Computer angeschlossen und kalibriert, aber es gab weder Bild noch Ton. Das war neu für mich und ich habe nicht verstanden, warum[...] Er meinte dann, dass das Problem vielleicht bei den Kabeln liegt und nicht beim Computer.[...] Es ging um die USB-Kabel, und wir haben dafür [Bestimmte Firma] glaube ich. Das ist die Informatik. Es gibt ein Informatikunternehmen, das für die ganze Sache verantwortlich ist. Ich habe ihn dann gerufen und gesagt, dass ich das anschauen muss, [...] es blieben damals noch drei bis vier Tage. Am Ende mussten sie dann hundert Meter neue Kabel verlegen.»

Diese Unterschiede wirkten sich auf die Auswertung aus. So hat die KI in diesem Falle erklärt: «[,...] hat ein Mitarbeiter seine Mutter um Hilfe gefragt.» Oder «In einem Fall dauerte es vierunddreissig Tage[...]

» Zwei grundlegend falsche Aussagen. Hier kann ebenfalls auch im Kleinen schon gesehen werden, wie sehr sich die sprachlich geglättete Variante vom Originaltranskript unterscheiden kann. Wenn diese bereinigten Texte inhaltliche korrigiert werden müssen, ist dies bei fünf Transkripten, je einstündiger Interviews

mühsamer, die Sätze noch umzubauen, als im Rohtranskript nur die richtigen Worte zu tauschen. Was dieses Beispiel nämlich auch zeigt, ist die relativ gute sprachliche Glättung von der KI. Kleinere Grammatik- oder Satzfehler, wirken sich nicht problematisch auf die Analyse aus, so lange sie nicht sinnesverfälschend sind. Somit kann sich die Fachperson nur auf den Inhalt fokussieren und nicht auf den Satzbau.

Als nächstes folgen Beispiel, Transkriptausschnitte mit dem Fokus auf einzelne Worte, welche von der KI nicht als falsch erkannt werden. Obiges Beispiel lässt erahnen, welche falschen Aussagen durch diese falschen Worte provoziert werden könnten.

Tabelle 12: Veranschaulichung von Transkriptionsfehlern und inhaltlicher Korrektur am Beispiel einzelner Worte

Rohtext	Inhaltlich korrigiert
„Sehr gerne also ich bin fast spezialisiert Kommunika- tion bei der [Firma]...“	„Sehr gerne also ich bin Fachspezialisiert Kommunika- tion bei der [Firma]...“
Finde ich bald alles, was intern ist darf man glaub ich relativ grosszügig auch in Intimes freischalten...	...darf man glaub ich relativ grosszügig auch in Teams freischalten...
...haben wir da diese Twitter Lösung quasi gehabt mit Teams [und anderem], dass wir trotzdem mit externen [arbeiten konnten]...	...diese Zwitter Lösung quasi gehabt mit Teams [und anderem], dass wir trotzdem mit externen [arbeiten konnten]...
...also mit Bildschirmschälern geht es natürlich sehr gut... Komplizieren gemein. Sam Dokumente bearbei- ten.	...mit Bildschirm teilen geht es natürlich sehr gut... kommunizieren gemeinsam Dokumente bearbeiten.
Es verfügt, ab wann man auf Tipps kompliziert und ab wann Skype nicht mehr funktioniert.	...ab wann man auf Teams kommuniziert und ab wann Skype nicht mehr funktioniert.
Also, wir haben eine Internetseite .	Also, wir haben eine Intranetseite .
...das PDF mit dieser. Das hab ich mir angehört und das denke, das machen auch andere...	...Das hab ich mir angewöhnt und das denke, das ma- chen auch andere...

7.5. Reflexionsworkshop



Abbildung 8: Reflexionsworkshop des partizipativen Projekts (Foto: Eigene Quelle)

Der Reflexionsworkshop bildet den Abschluss des partizipativen Vorgehens und wird als gemeinsame Auswertung des im Pilotprojekt erprobten Analyseprozesses durchgeführt. Er ist nicht als Test einzelner Artefakte angelegt, sondern als zusammenhängende Reflexion der erlebten Prozess Erfahrung. Als strukturierende Heuristik dient die 4L-Methode mit den Feldern Liked, Learned, Lacked und Longed for.

Zu Beginn wird geklärt, dass der Gegenstand auf Wunsch des Projektteams ebenfalls das Projekt an sich war und der Workshop somit auch als eine Art Retro dient, doch der Fokus des Workshops auf dem gemeinsam erarbeiteten Prozess liegt. Die Ergebnisse sollen im weiteren Verlauf helfen, die nächste Iteration zu ermöglichen. Die Beteiligten werden dabei nicht als Befragte adressiert, sondern als Mitinterpretierende ihrer eigenen Erfahrung. Anschliessend werden die vier Felder nacheinander bearbeitet. Die Beiträge werden je Feld gesammelt, gemeinsam besprochen und verdichtet:

- Unter **Liked** werden positive Erfahrungen oder Meinungen gesammelt.
- Unter **Learned** werden persönliche, fachliche oder organisatorische Erkenntnisse festgehalten.
- Unter **Lacked** wird benannt, was gefehlt hat, sowohl um den Prozess tragfähiger oder wirksamer zu machen, als auch im Verlauf des Projektes selbst.
- Unter **Longed for** werden Punkte gelistet, welche nachträglich betrachtet noch wünschenswert gewesen wären.

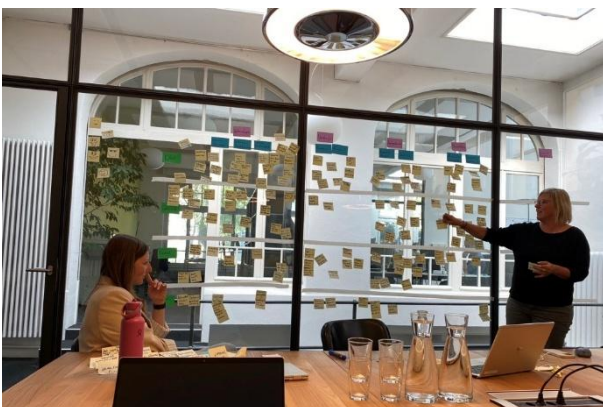


Abbildung 9: Affinity-Wand mit den Themen des Reflexionsworkshops (Foto: Eigene Quelle)

Themen, welche der Projektgruppe wichtig sind:

Zum Abschluss werden die zentralen Befunde zu Themen konsolidiert und für die Weiterentwicklung des Analyseprozesses übernommen. Damit schliesst der Workshop die partizipative Reflexionsschleife innerhalb des iterativen, menschenzentrierten Gestaltungsprozesses. Ein Foto der Workshopwand dokumentiert die Durchführung (siehe Anhang I, S. 138).

Workshop Ergebnisse

Die Ergebnisse des Reflexionsworkshops konnten auf sechs Themen konsolidiert werden. Die vollständigen, konsolidierten Workshop-Ergebnisse finden sich (siehe Anhang J, S. 139). Grob gefasst sind dies folgende

Thema 1: Die Datenerhebung muss verlässlich organisiert und fachlich fokussiert sein, mit verbindlich abgesicherten Interviewterminen, dem bewährten niederschweligen Umfragezugang, einem verständlichen und auf das Wesentliche begrenzten Leitfaden und Fragenkatalog sowie einer bewusst heterogenen Stichprobe, deren Umfang sich an der inhaltlichen Sättigung orientiert.

Thema 2: Das Building-Block- und Kodierungsgerüst muss fachlich trennscharf und kohärent sein, indem überladene Bausteine feiner gegliedert werden, etwa Digital Workplace in Communicate, Collaborate und Automate, und Interview- wie Umfragefragen konsistent auf Building Blocks, Massnahmen und Endergebnisse abgebildet sind, damit die Kodierung tragfähig bleibt.

Thema 3: Die Transkripte müssen inhaltstreu aufbereitet werden, mit einem verbindlichen fachlichen Kontrolldurchgang, der die Worttreue vor der sprachlichen Glättung prüft und möglichst von der interviewenden Person durchgeführt wird.

Thema 4: Die Analyse muss belegbare und verständliche Ergebnisse erzeugen, indem Evidenzen redundanzarm den Building Blocks zugeordnet und quantitative Kennzahlen fachlich eingeordnet werden, wobei die Priorisierung als manueller Fachschritt menschlich verantwortet bleibt und nicht an die KI delegiert wird.

Thema 5: Die Präsentation muss die Ergebnisse fokussiert und empfehlungsorientiert im vorgegebenen Zeitrahmen vermitteln, wertneutral als Momentaufnahme gerahmt, in Ziel und Rollen vorab abgestimmt, durch ein anschlussfähiges Summary begleitet und um einen moderierten Austausch mit dem Kunden ergänzt.

Thema 6: Der Gesamtprozess muss steuerbar bleiben durch klar definierte Rollen und einem geklärten internen Informationsfluss, bewährte Arbeitsformen wie Co-Moderation, Feedbackschleifen und digitale Werkzeuge, wiederverwendbare Standardbausteine sowie eine fest nach der Präsentation verankerte Reflexion.

7.6. Ergebnisse

Zentrales Ergebnis der Konzeptionsphase ist der gemeinsam mit der Projektgruppe erarbeitete KI-unterstützte Light-Assessment-Prozess. Er überführt die in Phase 1 (siehe [Kapitel 6](#)) erhobenen Bedürfnisse und Anforderungen in einen durchgängigen Analyseablauf und setzt in der Auswertungsphase an, nachdem die Datenerhebung über Interviews und Umfrage abgeschlossen ist.

Innerhalb des vorgesehenen Zeitrahmens werden die erhobenen qualitativen und quantitativen Daten kontrolliert, aufbereitet, analysiert und zu einer fachlich verantworteten Einschätzung verdichtet. Der Ablauf gliedert sich in vier aufeinander aufbauende Schritte: die Aufbereitung der Interviewtranskripte, die Aufbereitung der Umfragedaten, die Inhaltsanalyse mit anschließender fachlicher Priorisierung sowie die Erstellung von Summary und Präsentation. Diese Schritte konkretisieren die drei Teilprozesse Transkript, Analyse und Präsentation, wobei die Aufbereitung der Umfragedaten die quantitative Datenbasis für die Analyse bereitstellt.

Durchgängiges Gestaltungsprinzip ist der Human-in-the-Loop-Ansatz. Jeder KI-gestützte Schritt erzeugt ein abgegrenztes, prüfbares Zwischenergebnis, das die Fachperson an einem definierten Eingriffspunkt kontrolliert, bei Bedarf korrigiert und freigibt, bevor es weiterverwendet wird (NA2.1). Die KI übernimmt die strukturierende, vergleichende und verdichtende Fleissarbeit (NA1.2), während Bedeutungszuschreibung, fachliche Bewertung und Priorisierung bewusst beim Fachpersonal verbleiben. Dieses Prinzip geht aus der gemeinsamen Hürdenanalyse hervor (T-2) und wird durch die Befunde des Reflexionsworkshops gestützt, in denen die fachliche Kontrolle der KI-Ausgaben durchgängig als notwendig beschrieben wird. Als äußerer Rahmen ist zudem die unternehmensseitige Aufwandsgrenze von rund zwanzig Arbeitsstunden je Assessment eingeschrieben (NA1.3).

7.6.1. Bezug zur Forschungsfrage (Teilfrage 3) und Zielsetzung (Teilziel 2)

Damit ist die Teilfrage beantwortet, wie der KI-unterstützte Soll-Prozess zu gestalten ist, und das Teilziel der iterativen Konzeption und prototypischen Ausgestaltung erreicht. Die Konzeption erfolgt iterativ und partizipativ mit einer Projektgruppe aus fünf M365-Consultants, mit bewusst abgestufter Beteiligung und dem kollaborativ erarbeiteten Kodierungsgerüst als fachlichem Kern. Der Soll-Prozess ist als Arbeitssystem zu verstehen, das die Auswertung in vier aufeinander aufbauende Schritte gliedert und dadurch planbar und delegierbar wird, ohne an einzelne erfahrene Personen gebunden zu sein. Entscheidend ist, dass er mit den Fachpersonen entsteht und nicht für sie. Als Resultat liegt ein durchgängiger Light-Assessment-

Prozess vor, dessen Teilprozesse vor ihrer Verfestigung an realem Material erprobt wurden, ergänzt um einen konzeptionellen Prototyp, der Abfolge, Verantwortlichkeiten und Übergaben sichtbar macht.

7.7. Lernreflexion – Phase 2

Anders als in der Erhebung oder der späteren Evaluation habe ich hier nicht über die Fachpersonen, sondern mit ihnen gearbeitet. Dies geschah am realen Material eines laufenden Pilot-Projekts und an einem verbindlichen Lieferobjekt. Rückblickend habe ich in dieser Phase weniger über Methoden als über Zusammenarbeit mit Nutzenden gelernt. Einige der wichtigsten Erkenntnisse haben sich erst aus konkreten Reibungspunkten mit der Gruppe ergeben.

Die partizipative Arbeit hatte für mich eine besondere Qualität, weil die Beteiligten nicht in eine künstliche Studiensituation kamen, sondern Team-Kollegen waren, die mit mir an einem echten Kundenprojekt arbeiteten. Das machte die Zusammenarbeit unmittelbar und ehrlich, brachte aber auch eine Spannung mit sich. Als HCD-Fachperson und mit meiner Erfahrung und Erarbeitung bezüglich der Nutzung von KI bei Analyseprozessen, hatte ich eine Vorstellung davon, wie der Prozess fachlich sauber aussehen sollte und zugleich war es nicht meine Aufgabe, diese Vorstellung durchzusetzen, sondern den Prozess mit der Gruppe gemeinsam zu entwickeln. Diese Doppelrolle bewusst auszuhalten, war eine der eigentlichen Lernerfahrungen dieser Phase.

Am deutlichsten wurde das bei der Kodierung. Sie war das fachliche Kernstück der Konzeption und der Punkt, an dem die Gruppe am stärksten mitentschied. Hier habe ich mich bewusst getraut, mehr Codes zu akzeptieren, als ich fachlich für sinnvoll hielt. Die Gruppe wollte einzelne Bausteine und Unterscheidungen behalten, obwohl ich davor gewarnt hatte, dass die Kodierung dadurch zu umfangreich und zu wenig trennscharf wird. Statt das im Gespräch zu erzwingen, habe ich ihren Wunsch übernommen und die Analyse genauso erstellt, wie sie es wollten, mit dem Ergebnis einer viel zu umfangreichen Auswertung. Diese Entscheidung war keine Nachlässigkeit, sondern Absicht. Ich wollte, dass die Gruppe an der konkreten Erfahrung selbst erkennt, wie wichtig eine dedizierte und reduzierte Kodierung ist, statt es ihr nur zu sagen. Genau das ist eingetreten und der Befund, dass das Kodierungsgerüst trennscharfer werden muss, kam am Ende aus der Gruppe selbst und nicht von mir. Rückblickend war dieser Umweg über die überladene Variante einer der wirksamsten Momente der ganzen Zusammenarbeit.

Eine ähnliche Erfahrung machte ich beim Teilprozess «Präsentation». Ich hatte vorgesehen, auch die Erstellung der Präsentationsunterlagen mit KI zu unterstützen, doch die Gruppe entschied, dass dieser Schritt bewusst manuell und vollständig in ihrer Kontrolle bleiben soll. Damit fiel die Präsentation aus dem Automatisierungsfokus der Arbeit heraus und ich habe diese Entscheidung mitgetragen, statt sie zu hinterfragen. Für mich war das eine wichtige Lektion darüber, was Partizipation tatsächlich bedeutet. Wenn ich die Beteiligten ernst nehme, muss ich auch akzeptieren, dass sie den Automatisierungsgrad mitbestimmen und an einer Stelle, die mir konzeptionell gepasst hätte, die menschliche Kontrolle bewusst vorziehen.

Rückblickend fällt mir auch auf, dass sich nicht jede Aktivität sauber einer Phase zuordnen lässt. Besonders die Besprechung und die Anpassung der Analyseergebnisse, im Raster als T-5.b und T-5.c geführt, liessen sich ebenso gut der Konzeption wie der Gestaltung zurechnen. Mir war bei der Durchführung ein guter Arbeitsfluss wichtiger als die spätere methodische Trennschärfe, und so lassen sich beide Schritte an beiden Orten verargumentieren. Im Rahmen eines solchen Projektes halte ich diese pragmatische Unschärfe für vertretbar und ich weise sie hier lieber offen aus, als sie nachträglich zu glätten.

Über den gesamten Verlauf betrachtet folgte die Beteiligung keinem linearen Durchlauf des Designzyklus. Die Aktivitäten pendelten zwischen Konzeption, Evaluation und Gestaltung, und genau darin hat sich die im Methodenkapitel angelegte Verankerung der Iteration im Participatory Design konkret eingelöst. Eine gute Entscheidung war für mich auch, dass ich die Beteiligungsintensität bewusst variiert habe. Von den neun Aktivitäten verliefen fünf partizipativ, drei konsultativ und eine kollaborativ. Konsultativ blieben jene, in denen Rückmeldungen effizient in Umsetzungsarbeit zu überführen waren, partizipativ die Tests und

Reflexionen an realem Material, kollaborativ der fachliche Kern der Kodierung. Damit folgt der Verlauf der Grundfigur des Participatory Design nach Spinuzzi (2005), von der Exploration der bestehenden Praxis über die gemeinsame Aushandlung von Prozessgrundlage, KI-Hürden und Kodierung bis zum iterativen Erproben an realem Material. Alle drei Teilprozesse waren vor ihrer prototypischen Verfestigung durch die Projektgruppe getestet, obwohl wir bei Prototypen von bestimmten Handlungsschritten und Dokumentenstrukturen reden. Im Rückblick ist das für mich der Kern dieser Phase, der Soll-Prozess wurde nicht für die Consultants, sondern mit ihnen definiert. Dass der abschliessende halbtägige Workshop zugleich als Retrospektive angelegt war, hat ihn zu einem wertvollen eigenen Beitrag gemacht, dessen Inhalte ich für die Weiterentwicklung des Prototypen und die Relativierung von Bedürfnissen und Nutzungsanforderungen nutzte. An dieser Stelle hätte auch ein finales Testing mit den Nutzenden stattfinden können und ich hatte es auch einmal in Erwägung gezogen, doch abgesehen von dem zeitlichen Aufwand, schien es mir schlichtweg falsch, das Team, welches den Prozess miterarbeitet hat, um ihre Meinung zum Prozess zu fragen, statt um ihre gewonnenen Erfahrungen aus dem Projekt über den Workshop.

Aus diesen Beobachtungen ziehe ich für mich eine zentrale Erkenntnis. In der Zusammenarbeit mit den Nutzenden lag der Wert nicht darin, sie von meiner Lösung zu überzeugen, oder zu einer Entscheidung von mir zu befragen, sondern darin, ihnen Entscheidungen zu überlassen und die Konsequenzen gemeinsam zu erleben. Mehrere der tragfähigsten Ergebnisse, von der Einsicht in die Kodierung bis zur klaren Rollenverteilung bei der Präsentation, sind nicht durch meine Steuerung entstanden, sondern dadurch, dass ich Steuerung an den richtigen Stellen abgegeben habe. Zugleich habe ich gelernt, dass diese Offenheit ihre Grenze dort hat, wo die methodische Verlässlichkeit auf dem Spiel steht. Die Gruppe durfte über den Zuschnitt der Kodierung entscheiden, aber die Belegbarkeit, die fachliche Kontrolle und die Verantwortung für die Ergebnisse blieben gesetzt. Diese Balance zwischen echter Mitbestimmung und gewahrter fachlicher Substanz ist für mich das eigentliche Learning aus der Konzeptionsphase.

8. Phase 3: Gestaltung

8.1. Five Elements of UX

Die Gestaltungsphase überführt die in den vorangehenden Phasen (siehe [Kapitel 6](#) und [7](#)) erarbeiteten Bedürfnisse, Anforderungen und konzeptionellen Entscheide in das eigentliche Produkt, den KI-unterstützten Light-Assessment-Prozess und seinen Prototypen. Als Strukturierungsrahmen dieser Überführung dient Garretts Modell der Five Elements of User Experience (Garrett 2011), dessen fünf Ebenen die Gestaltung von den strategischen Zielen bis zur konkreten Ausgestaltung der Agenteninteraktion ordnen.

8.1.1. Begründung der Wahl

Die Wahl des Modells und seine Anpassung an den prozessbezogenen, dialogbasierten Gegenstand sind im Methodenkapitel begründet (siehe [Abschnitt 4.2.2](#)). Über den Projektverlauf wirken die fünf Ebenen als Strukturierungs- und Übersetzungshilfe, die Aussagen, Anforderungen und Entscheide jeweils jener Abstraktionsebene zuordnet, auf der sie am wirksamsten bearbeitet werden. Exemplarisch verankert ist die Anwendung in der Gestaltungsphase, weil hier das Produkt selbst entsteht. Während die unteren Ebenen vor allem in Planung, Research und Konzeption tragen, werden die oberen Ebenen Skeleton und Surface erst mit der prototypischen Realisierung des Light Assessment Agenten wirksam. Die Gestaltungsphase ist damit die einzige Phase, in der sich das Modell über alle fünf Ebenen hinweg einlösen lässt.

8.2. Konkrete Durchführung

Die konkrete Durchführung folgt den fünf Ebenen von Strategy bis Surface. Während die unteren Ebenen die in den vorangehenden Phasen erarbeiteten Grundlagen aufnehmen und für die Gestaltung verfügbar

machen, rücken die oberen Ebenen die eigentliche Lösungsumsetzung in den Vordergrund, also den Prozess selbst und seine prototypische Realisierung als Light Assessment Agent.

8.2.1. Strategy

Die Strategy-Ebene wird in der Gestaltungsphase nicht separat erarbeitet, sondern fügt sich in Planungs- und Researchphase (siehe [Kapitel 5](#) und [6](#)) ein und steht der Gestaltung als gesetzte Grundlage zur Verfügung. Die organisationale Perspektive ist über die Stakeholder-Analyse erschlossen, die nutzungsbezogene über die Stakeholder-Interviews (siehe [Abschnitt 5.2](#) und [Abschnitt 6.1](#)). Beide Perspektiven verdichten sich in den fünf Kernbedürfnissen KB-01 bis KB-05, in denen die Geschäftsziele bewusst als Bedürfnisse mitgeführt werden, ohne die Gewichtung von der Nutzungsperspektive zu lösen (siehe [Abschnitt 6.4](#)). Für die Gestaltung bedeutet dies, dass Produktziele und Bedürfnisse des Fachpersonals nicht mehr gegeneinander abgewogen werden müssen, sondern als konsistente Zielvorgabe in die folgenden Ebenen eingehen.

8.2.2. Scope

Auf der Scope-Ebene werden die strategischen Vorgaben in greifbare Anforderungen an den Prozess überführt. Den funktionalen Rahmen bilden die in Phase 1 (siehe [Abschnitt 6.5.1](#)) abgeleiteten Anforderungen NA1.1 bis NA5.2, die festlegen, welche Verarbeitungsschritte, Kontrollpunkte und Übergaben der Prozess leisten muss und welche Aufgaben bewusst ausserhalb bleiben, etwa die Datenerhebung und die Erstellung der Präsentationsunterlagen, die bereits in der Abgrenzung nicht zum Analyseprozess gezählt werden (siehe [Abschnitt 1.4](#)). Inhaltlich umfasst die Scope-Ebene jene Elemente, die der Prozess erzeugt und weiterverarbeitet, also die Codes des Kodierungsrahmens, die qualitativen und quantitativen Evidenzen sowie die Artefakte, die von den Personen-Fragmenten über den konsolidierten Bericht bis zum finalen Summary entstehen. Damit ist abgegrenzt, woraus der Prozess besteht, bevor seine Abfolge auf der Structure-Ebene festgelegt wird.

8.2.3. Structure

Auf der Structure-Ebene erhält der Prozess seine Ablauflogik. Hier wird festgelegt, in welcher Reihenfolge die Verarbeitungsschritte ineinandergreifen und wie die Aufgaben zwischen Fachperson und KI verteilt sind. Leitend ist der Human-in-the-Loop-Ansatz. Jeder KI-gestützte Schritt erzeugt ein abgegrenztes, prüfbares Zwischenergebnis, das die Fachperson an einem definierten Eingriffspunkt kontrolliert, bei Bedarf korrigiert und freigibt, bevor es weiterverwendet wird (NA2.1). Initial hatte die KI einen experimentellen, eigenständigen Bewertungsschritt für die Einschätzungsfelder. Diese sollten als Vorschlag dienen. Doch nach der Überarbeitung in der partizipativen Gruppe wurde beschlossen, dass diese Vorschläge irritierend sind und die Priorisierung gänzlich den Fachpersonen vorbehalten sein sollen. In den reflektierten Begründungen wurden Sorgen geschildert, die Vorschläge schlicht zu akzeptieren. Insgesamt schilderte die Gruppe durch ihre selbstreflektierte Haltung Neigungen, welche in Richtung des «Fallens in eine Subordinationsrolle» gehen, wie es auch bei Matter et al. (2024) beobachtet wurde.

Die KI übernimmt die strukturierende und verdichtende Fleissarbeit, während Bedeutungszuschreibung, fachliche Bewertung und Priorisierung beim Fachpersonal verbleiben.

Der inhaltliche Ablauf dieser Schrittfolge, von der Aufbereitung der Transkripte und Umfragedaten über die Inhaltsanalyse und fachliche Priorisierung bis zu Summary und Präsentation, ist dargestellt (siehe [Abschnitt 8.3](#)). Ergänzend ordnet die Structure-Ebene die Daten und Zwischenergebnisse entlang fester Übergaben, sodass jede spätere Aussage auf das geprüfte Ausgangsmaterial zurückführbar bleibt (NA2.2). Diese Übergaben verlaufen von den einzelnen Personen-Fragmenten über den konsolidierten Bericht und seine quantitative Anreicherung bis zu den synthetisierten Befunden und dem finalen Lieferobjekt.

8.2.4. Skeleton

Die Skeleton-Ebene überführt den auf der Structure-Ebene festgelegten Ablauf in ein ausführbares Konstrukt. Aus Sicht des Fachpersonals tritt der Prozess als einheitlicher M365-Copilot-Agent in Erscheinung, der die internen Übergaben steuert, den jeweils nächsten Schritt vorschlägt und nach jedem Schritt ein abgegrenztes Zwischenergebnis zur Sichtung zurückgibt. Hinter dieser geschlossenen Interaktion liegt eine bewusst getrennte Komposition zweier Verarbeitungsarten. Sprachmodellläufe übernehmen jene Aufgaben, in denen Bedeutung erkannt, verdichtet oder zwischen Aussagen und Daten vermittelt werden muss. Deterministische Skripte übernehmen jene Aufgaben, in denen Reihenfolge, Format und Vollständigkeit mechanisch sicherzustellen sind, etwa den Zusammenbau des konsolidierten Berichts. Diese Aufteilung trennt die Verantwortlichkeiten innerhalb der KI, bleibt im Anwendungserlebnis aber unsichtbar.

Die Texte, die das Verhalten des Agenten festlegen, sind in einem verbindlichen Steuerelement zusammengeführt. Es hält die methodischen Vorgaben der deduktiven Inhaltsanalyse fest und schreibt den Ablauf der einzelnen Läufe vor. Besonders ausgeprägt ist diese Steuerung in der Synthese, in der pro Code drei aufeinander folgende Läufe in fester Reihenfolge ablaufen, ein Analyse-, ein Validierungs- und ein Kurzform-Lauf. Diese strikte Lauf-Disziplin verhindert, dass Schreiben, Prüfen und Verdichten vermischt werden, und macht die Entstehung jeder Aussage nachvollziehbar. Die fachliche Grundlage dieser Steuerung bildet der kollaborativ erarbeitete Kodierungsrahmen, der dem Agenten als fixe Vorgabe übergeben wird (siehe [Kapitel 7](#)). Eine ausführliche Darstellung der einzelnen Verarbeitungsschritte und der zugehörigen Agententexte findet sich im technischen Teil (siehe [Kapitel 10](#)).

8.2.5. Surface

Diese Ebene bestimmt, wie sich der auf der Skeleton-Ebene festgelegte Ablauf in der konkreten Kommunikation mit den Nutzenden manifestiert. Da die visuelle Oberfläche durch die M365-Umgebung vorgegeben ist und ausserhalb des Gestaltungsraums liegt, betrifft die Gestaltung hier die sprachliche Form der Agentenantworten, ihre Struktur, Formatierung und Tonalität. Im Erleben des Fachpersonals stellt sich der Prozess als geführte, schrittweise Konversation dar, in der der Agent die nötigen Eingaben anfordert, Zwischenergebnisse bestätigt und notwendige Wiederholungen, etwa nach einem Validierungsfehler, offenlegt. Diese Transparenz ist nicht nur kommunikativ, sondern macht nachvollziehbar, weshalb eine Aussage in ihrer finalen Form so ausfällt, wie sie ausfällt.

Die Tonalität ist durchgehend sachlich und wertneutral gehalten. Der Agent beschreibt empirisch und nimmt keine Erfolgsbewertung vor, was der fachlichen Verantwortung des Menschen entspricht und einer vorzeitigen Deutung durch die KI vorbeugt. Auch die Formatierung folgt dieser Logik. Quantitative Evidenzen werden tabellarisch dargestellt, qualitative als nummerierte Liste mit hervorgehobener Personen-Kennzeichnung, sodass die Herkunft jeder Aussage sichtbar bleibt. Den Abschluss bildet die Übergabe in ein Word-Dokument, dessen leere Einschätzungstabellen mit den fünf Stufen Hoch, Mittel, Niedrig, Zu Erwähnen und Ignorieren die zentrale Schnittstelle für die abschliessende fachliche Bewertung bilden. In dieser Form manifestiert sich das Gestaltungsprinzip des gesamten Prozesses noch einmal an der Oberfläche, die KI bereitet strukturiert vor, die Entscheidung trifft das Fachpersonal.

8.3. Ergebnis: Der Light Assessmentprozess

Das Ergebnis der Gestaltungsphase ist der konkret ausgestaltete Light-Assessment-Prozess. Er setzt nach der Datenerhebung an und führt die qualitativen und quantitativen Daten in sieben aufeinander folgenden Schritten zu einer fachlich verantworteten Einschätzung zusammen. Kennzeichnend ist der durchgängige Wechsel zwischen Fachperson und KI. Jeder Schritt schliesst mit einer definierten Übergabe ab, an der ein abgegrenztes Zwischenergebnis weitergereicht wird, sodass Verantwortung und Kontrolle stets eindeutig zugeordnet bleiben. Die nachfolgende Abbildung stellt diesen Ablauf mit seinen sieben Schritten und den zugehörigen Übergaben dar.

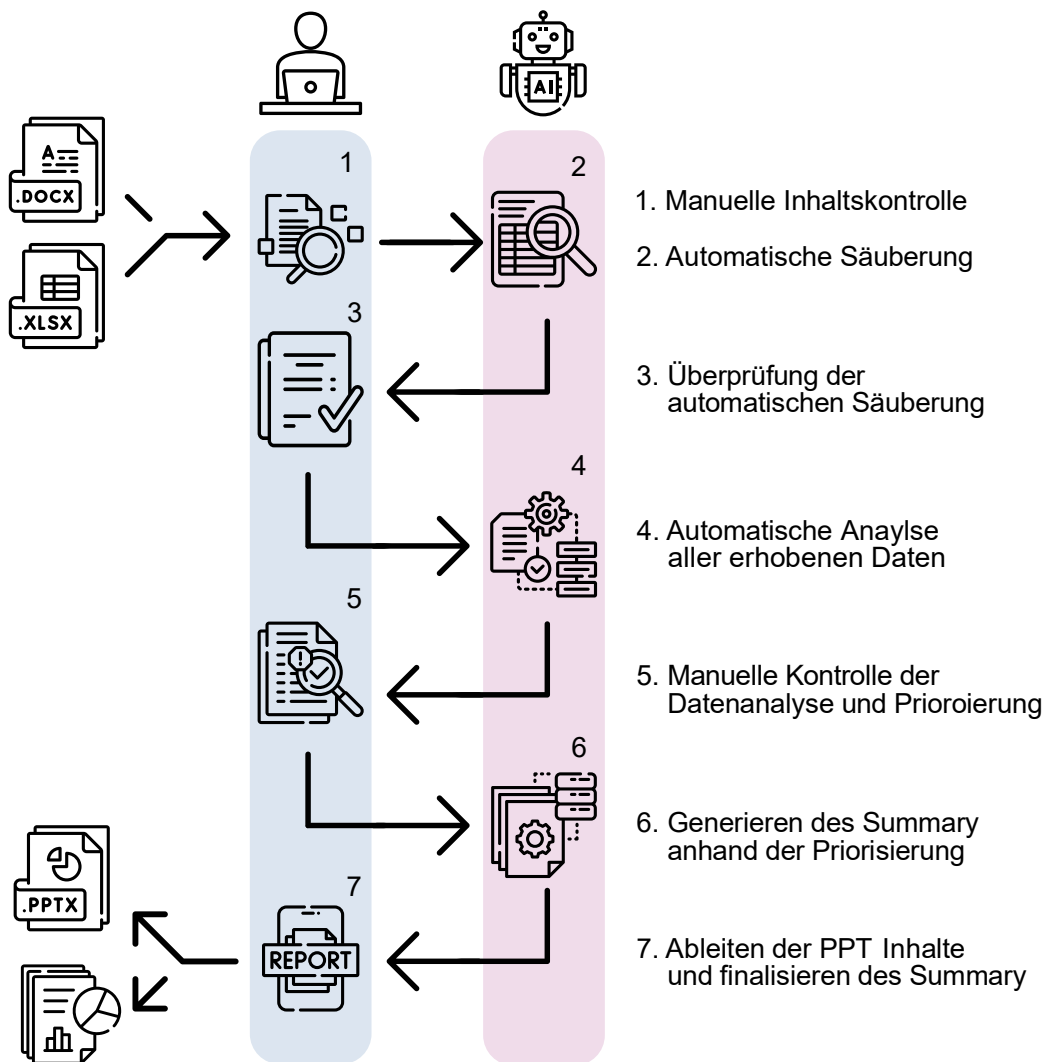


Abbildung 10: Übersicht des Light Assessment ohne out of Scope Bereiche (Eigene Darstellung)

Schritt 1 (Mensch): Manuelle Inhaltskontrolle der Daten

Der Prozess beginnt mit einem fachlich verantworteten Kontrolldurchgang. Die Fachperson prüft die automatisch erzeugten Interviewtranskripte gegen die Aufzeichnungen. Im Zentrum steht die inhaltliche Worttreue, also die Korrektheit von Begriffen, Fachbegriffen und Akronymen sowie das Erkennen von Aussagen, die das automatische Transkribieren verfälscht haben. Formfehler wie Grammatik, Satzzeichen oder Sprechunflüssigkeiten bleiben bewusst unberücksichtigt, da sie die inhaltliche Aussage nicht verändern. Diese Kontrolle ist der KI bewusst entzogen, weil ein Sprachmodell kontextabhängige oder akronymbezogene Fehler nicht zuverlässig erkennt, solche Fehler aber zu falschen analytischen Schlüssen führen können. Der Schritt greift ein im Reflexionsworkshop bestätigtes Bedürfnis auf, wonach die Prüfung der

Worttreue vor der sprachlichen Glättung erfolgen und möglichst eng an die interviewende Person gekoppelt bleiben soll, die den Gesprächsverlauf kennt (Thema 3). Die kontrollierten Daten werden anschließend an die KI übergeben.

Schritt 2 (KI): Automatische Bereinigung

Auf Grundlage der kontrollierten Daten übernimmt die KI die sprachliche Glättung der Transkripte, also die Bereinigung von Füllwörtern, Doppelungen, Satzbau und Sprechabbrüchen, sodass klar lesbare Evidenzen entstehen. Parallel bereitet die KI die quantitativen Umfragedaten auf. Sie säubert die Rohdaten und bereitet sie lesbar und grafisch auf, wobei nicht verwertbare Datensätze ausgeschlossen werden, etwa abgebrochene oder unvollständige Teilnahmen, unrealistisch kurze Bearbeitungszeiten oder auffällige Antwortmuster wie durchgängig gleiche Angaben. Die bereinigten Daten gehen zurück an die Fachperson.

Schritt 3 (Mensch): Überprüfung der bereinigten Daten

Die Fachperson kontrolliert das Ergebnis der Bereinigung. Bei den geglätteten Transkripten erfolgt eine Stichprobenkontrolle, bei den Umfragedaten die Prüfung der erzeugten Dokumente und Grafiken. Ziel ist eine verständliche und belastbare Grundlage, die in der Analyse fachlich eingeordnet werden kann. Der Schritt reagiert auf einen wiederkehrenden Befund, wonach automatisiert erzeugte Umfrageauswertungen ohne fachliche Einordnung Fehlinterpretationen begünstigen und einzelne Kennzahlen wie der Net Promoter Score als schwer verständlich beschrieben wurden (T-2, Thema 4). Damit bleiben spätere Befunde auf das geprüfte Ausgangsmaterial rückführbar (NA2.2). Die vollständig kontrollierten Daten werden an die KI übergeben.

Schritt 4 (KI): Automatische Analyse aller erhobenen Daten

Die KI führt die deduktive Inhaltsanalyse über die qualitativen und quantitativen Daten durch (siehe [Kapitel 4](#)). Grundlage ist das kollaborativ erarbeitete Kodierungsgerüst, das die Building-Block-Struktur des M365-Evolution-Modells in Ziele, mögliche Lücken und zugehörige Empfehlungen überführt und so den Zielbild- und Reifegradbezug der Ergebnisse sichert (NA3.1). Die KI ordnet die Interviewaussagen den passenden Ziel- und Lücken-Codes zu, gruppiert die Evidenzen je Code nach Personen und führt sie über die Interviews hinweg zusammen. Passende Umfrageitems ergänzen die quantitativen Evidenzen, jeweils mit einer Einordnung, wie eng das Item zur Code-Definition passt und wie vorsichtig der Befund zu lesen ist. Aus beiden Evidenzarten verdichtet die KI je Code einen Befund und eine Kurzform, abgesichert durch den auf der Skeleton-Ebene beschriebenen Validierungslauf. Es entsteht ein vollständiger Analysebericht, der zu jedem Ziel und jeder Lücke die qualitativen und quantitativen Evidenzen, einen Befundtext und eine Kurzform enthält, ergänzt um offene Felder für die Priorisierung. Dieser Bericht geht an die Fachperson.

Schritt 5 (Mensch): Kontrolle und Priorisierung der Befunde

Die Fachperson kontrolliert die Befunde im Bericht und nimmt die Priorisierung vor. Sie stuft je Ziel und Lücke ein, welche Relevanz der Punkt für die konkrete Kundenorganisation hat, von hoher Handlungsrelevanz bis zu bewusst ignorierten Punkten, und wählt darüber hinaus mögliche Empfehlungen aus und priorisiert sie. Die Bewertung bleibt der KI entzogen, da der vielschichtige und teils politische Kontext einer Kundenorganisation ein fachliches Urteil erfordert (NA2.1). Fällt dabei ein Widerspruch auf, prüft die Fachperson die betreffende Stelle in den geglätteten Interviews oder im Umfragebericht und korrigiert sie. Der Schritt verdichtet mehrere Workshop-Themen, wonach die Priorisierung der Handlungsfelder ein menschliches Fachurteil bleiben und die Evidenzführung verständlich und klar den Building Blocks zugeordnet sein muss (Thema 4). Die fachliche Trennschärfe des Kodierungsgerüsts war bereits der kollaborative Kern der Konzeption und wurde in den Analyse-Iterationen wiederholt nachgeschärft (T-3, T-5). Der vollständige Analysebericht mit den kontrollierten und priorisierten Befunden geht zurück an die KI.

Schritt 6 (KI): Erstellung des priorisierten Berichts mit Summary

Aus dem priorisierten Bericht erstellt die KI ein verdichtetes Summary als schriftlichen Kundenbericht mit den relevanten Erkenntnissen, den fachlich eingeordneten Daten und konkreten Handlungsempfehlungen.

Die vollständigen Rohdaten, also unbearbeitete Transkripte, Umfrageexporte oder interne Analysedokumente, werden dem Kunden nicht übergeben, sodass der Bericht fokussiert bleibt und sensible Detaildaten geschützt sind. Das Summary erfüllt damit die Anforderung an ein anschlussfähiges Lieferobjekt und macht Handlungsfelder ableitbar (NA5.1, NA5.2). Es wird an die Fachperson übergeben.

Schritt 7 (Mensch): Finalisierung des Summaries und Ableitung der Präsentation

Die Fachperson finalisiert das Summary und leitet daraus die Abschlusspräsentation ab. Die Präsentation verdichtet die zentralen Erkenntnisse weiter und fokussiert auf die Punkte mit dem grössten fachlichen und organisatorischen Gewicht. Die Präsentationsunterlagen werden bewusst manuell aus einer Vorlage erstellt, wobei die Texte des Summaries als Ausgangspunkt dienen. Diese Festlegung geht auf eine Entscheidung der Projektgruppe zurück, die eine automatisierte Befüllung als nicht erwünscht beurteilte und die Kontrolle über das Kundenartefakt vollständig beim Fachpersonal behalten wollte (T-6). Wo möglich werden auch positive Aspekte der Organisation sichtbar gemacht. Die Ausgestaltung nimmt damit Workshop-Themen auf, namentlich die verdichtete, wertneutrale Darstellung mit dem Summary als Begleitdokument sowie den Wert eines moderierten Austauschs in der Ergebnisbesprechung (Thema 5).

8.3.1. Einordnung

Mit dem Light-Assessment-Prozess liegt ein durchgängiger, schrittweise definierter Analyseablauf vor (NA4.1). Er ist nicht für die Fachpersonen, sondern mit ihnen entstanden, seine Teilprozesse wurden entlang der realen Projektarbeit erprobt (T-3 bis T-6) und die im Reflexionsworkshop verdichteten Bedürfnisse sind in die Ausgestaltung der einzelnen Schritte eingeflossen. Als Ergebnis der Gestaltungsphase bildet der Prozess die Grundlage für die anschliessende Evaluation.

8.3.2. Bezug zur Forschungsfrage (Teilfrage 2)

Damit ist die Teilfrage beantwortet, welche Aufgaben sich für eine KI-Unterstützung eignen und welche aus Gründen der Qualität, Nachvollziehbarkeit und Verantwortung beim Fachpersonal verbleiben. An die KI delegierbar sind die wiederkehrenden Verarbeitungsschritte, also die sprachliche Glättung der zuvor inhaltlich geprüften Transkripte, die Bereinigung und grafische Aufbereitung der Umfragedaten, die Zuordnung der Aussagen zu den vorgegebenen Ziel- und Lücken-Codes, die Zusammenführung der Evidenzen sowie die Verdichtung je Code zu Befund, Kurzform und Summary. Beim Fachpersonal verbleiben die urteilenden Schritte, insbesondere die der KI bewusst entzogene inhaltliche Worttreuekontrolle der Transkripte und die fachliche Priorisierung der Handlungsfelder, ebenso die bewusst manuell gehaltene Kundenpräsentation. Getragen wird die Aufgabenteilung durch den Human-in-the-Loop-Ansatz mit Prüfung und Freigabe an jedem Eingriffspunkt und durchgängiger Materialbindung.

8.4. Lernreflexion – Phase 3

Das eigentliche Bauen des Prototyps hat mir einige wichtige Erkenntnisse gebracht, weil sich hier zeigte, was von meinen Vorstellungen trägt und was nicht. Vieles, was ich konzeptionell für selbstverständlich gehalten hatte, musste ich in der Umsetzung erst neu lernen, und die Gestaltung des Light-Assessment-Prozesses war über weite Strecken ein Vorgang des Ausprobierens, Scheiterns und Nachschärfens. Dabei fing der Bau oder zumindest die konzeptionellen Ansätze schon früh in der Planung und im Research an. Auch wenn viele dieser Tätigkeiten eher technisch waren. Daher sind sie streng genommen nicht Teil des Scopes aber die Erfahrungen und Learnings dieser Versuche, schreibe ich gerne hier noch auf.

Die ersten Versuche und zentrale Erkenntnis betrafen die Frage, woher die Verlässlichkeit des Prozesses eigentlich kommt. Da dieser Prozess von Nicht-Researchern durchgeführt wird, müssen diese Ergebnisse funktionieren und das machte mich doch wieder sehr nervös. Auch wenn hier ein rein theoretisches Konzept, welches von KI-Fachpersonen abgesegnet wurde, reichen würde, wollte ich den Prozess auf einem Prinzip aufbauen, welches nachgewiesen verlässlich ist und welches ich verstehe, um die methodische Hoheit darüber zu wahren und nicht einfach mal KI machen lassen oder KI-Fachpersonen irgendwie bauen

lassen. Zu Beginn habe ich, zusammen mit der Definition des Vibe Researches und dem Ausarbeiten des Vibe

Research Vorgehensmodells, testweise versucht, möglichst viel der KI zu überlassen, um zu sehen, was passiert. So auch beim Zusammenführen, Formatieren und Vervollständigen der Zwischenergebnisse. Dabei wurde mir schnell wieder klar, warum genau diese Aufgaben das grösste Risiko bergen, weil das Modell beim Schreiben unbemerkt etwas zusammenfasst, vergisst oder glättet. Daher habe ich schnell wieder entschieden, die Verarbeitung bewusst zu trennen, doch dieser Exkurs half mir genau dabei, diese Trennung besser zu identifizieren. Das Sprachmodell übernimmt also nur dort, wo Bedeutung zu verdichten ist, während deterministische Skripte für Reihenfolge, Format und Vollständigkeit sorgen, wie geplant. Rückblickend war das der wichtigste Lernschritt der Gestaltung, da er eine konzeptionelle Grundlage geschaffen hat. Ich habe die Vertrauensfrage vom nicht überprüfbareren Modell auf ein einfaches, nachprüfbares Skript verschoben und erst damit wurde der Prozess für mich belastbar.

Eine ähnliche Lernerfahrung machte ich bei der Kodierung selbst. Mein erster Ansatz, das Modell alle Codes in einem einzigen Durchgang zuordnen zu lassen, lieferte unbefriedigende Ergebnisse. An einzelnen Stellen behauptete die KI schlicht, es gebe keine Evidenz, obwohl ich wusste, dass es welche geben musste. Zuerst habe ich das für ein inhaltliches Problem gehalten, bis ich verstand, dass es ein Problem der Überforderung war. Bei zu vielen Codes auf einmal vermischt das Modell die breite Suche mit der strengen Prüfung und gibt zu früh auf. Die Lösung lag nicht in einem besseren Werkzeug, sondern in der Aufteilung. Ich habe die Aufgabe in kleinere Einheiten zerlegt und die Suche bewusst von der Endprüfung getrennt. Diese Einsicht, dass die Qualität weniger vom Modell als von der Grösse und dem Zuschnitt der Aufgabe abhängt, hat meine Sicht auf den ganzen Prozess geprägt und half mir später sogar dabei, mit älteren Modellen gute Ergebnisse zu erzielen, als es um Datensicherheit ging, denn wenn man im Moment bspw. mit einem GPT Modell Daten auf Schweizer Servern verarbeiten will (was relevant für den Auftraggebenden ist), kann man nur maximal GPT 4.1. haben und zum Zeitpunkt dieses Schreibens ist die aktuellste Version GPT 5.5 vor kurzem rausgekommen.

Während der Gestaltung in der partizipativen Gruppe wurde mir auch deutlich, wo die Grenze der KI bewusst liegen muss. Die Gruppe hatte schlussendlich die ursprüngliche Funktion einer für vorgeschlagene Priorisierung und Empfehlungsabgabe, nach dem Testen klar zurückgewiesen. Die Priorisierung der Befunde ist ein kundenspezifisches Fachurteil, welches die KI nicht treffen kann und nicht treffen sollte. Im Gespräch äusserte die Gruppe ihre Bedenken, weil sie das Gefühl hatten, im Gegensatz zu den Analyse-daten, die Priorisierung einfach hinnehmen zu wollen, statt sie zu prüfen. Dies war äusserst faszinierend, besonders da ich auf Grund meiner Forschungen erkannte was gerade passiert, denn ich beobachtete eine Diskussion des «Double Chinese Room» wie von Matter et al. (2024) beschrieben (Siehe [Abschnitt 3.2.4](#)). Der Partizipative Prozess, legte die Gefahr der Assistenzrolle jedoch in der Diskussion offen. Diese Episode hat mir schlichtweg nochmals bestätigt, dass die menschliche Kontrolle kein nachträglicher Zusatz ist, sondern als fester Eingriffspunkt in den Prozess eingebaut sein muss, wie auch Matter erwähnt.

Man kann nicht einfach die KI irgendetwas machen lassen und dann sagen «Der Mensch kontrolliert jetzt.» Der Punkt wann und die Entscheidung warum müssen bewusst und im Sinne der betroffenen Nutzenden gemacht werden. Das ist für mich das wichtigste und seltsam überraschendste Learning dieser Arbeit. In Garretts Modell liess sich das gut verorten, denn gerade weil mein Produkt kein klassisches Interface hat, wurde die Stelle, an der das Fachpersonal entscheidet, zur eigentlichen Oberfläche des Prozesses. Nämlich die Interaktion: Was bekommen die Nutzenden, an welchem Punkt und für welchen Zweck?

Eine weitere Erkenntnis schälte sich erst im Zuge mehrerer Überarbeitungen heraus und betrifft den Rhythmus des Prozesses. Zu Beginn hatte ich den Ablauf vor allem als Abfolge inhaltlicher Schritte gedacht. Erst als ich ihn konsequent als durchgängigen Wechsel zwischen Mensch und KI anlegte, mit sieben klar getrennten Schritten und definierten Übergaben, wurde an jeder Stelle eindeutig, wer gerade verantwortlich ist. Jeder KI-Schritt endet an einem Punkt, an dem die Fachperson kontrolliert und freigibt, bevor der nächste beginnt. Diese abwechselnde Abfolge war für mich weniger eine technische als eine gestalterische

Entscheidung. Sie hat mir erlaubt die Erkenntnis einzusetzen, dass die Klarheit darüber, wer wann verantwortlich ist, selbst ein Gestaltungsmerkmal ist und nicht nur eine organisatorische Begleiterscheinung.

Schliesslich musste ich lernen, ehrlich zu bleiben, was der Prototyp bzw. der Prozess zu dem Zeitpunkt tatsächlich ist. Es ist eine Ansammlung nummerierter Markdown-Dokumente, welche der Reihe nach mit entsprechenden Daten und Dokumenten einem Copilot-Chat eingefügt werden. Das funktioniert, war aber noch nicht stabil und verlangte viel Kontrolle. Im Verlauf war ich versucht, ihn als fertigen, einheitlichen Agenten zu beschreiben, doch in dieser Form war er nicht durchgängig realisiert oder eine Umsetzung zu erzwingen. Ich habe mich aber entschieden, ihn klar als konzeptionellen Proof-of-Concept einzuordnen und auch die schwächeren Stellen offen zu benennen, etwa die sprachliche Ausgestaltung der Agentenantworten, für die mir noch klare Vorgaben fehlten. Diese Offenheit fiel mir zuerst schwer, weil sie nach einer Lücke aussieht. Rückblickend halte ich sie für die methodisch sauberere Lösung, weil ein ehrlich ausgewiesener Stand mehr wert ist als ein geglättetes Bild, das einer Prüfung nicht standhält. Hier hatte mich beinahe meine «Dem Kunden recht gemacht und im Hintergrund fertiggemacht» Mentalität überholt. Zu dieser Denkweise wurde ich leider in der Arbeitspraxis seit Jahren mehr oder weniger gezwungen. Solange es nicht mehr kostet und funktioniert, kümmert es niemanden, wenn hinter der Fassade noch gehämmert wird und die Farbe noch trocknet. Aber eben: Hier zählt es und es war eine Erfahrung, die gut getan hat.

Mein Fazit aus der Gestaltungsphase ist, dass die Qualität eines KI-unterstützten Prozesses nicht im Modell selbst liegt, sondern in der Art, wie der Prozess gebaut ist. Verlässlichkeit entsteht durch prüfbare Skripte, durch den richtigen Zuschnitt der Aufgaben, durch einen klaren Wechsel zwischen Mensch und KI mit eindeutigen Eingriffspunkten und durch eine ehrliche Einordnung dessen, was der Prozess leisten kann. Garretts Modell hat mir geholfen, diese Entscheidungen den richtigen Ebenen zuzuordnen. Am meisten gelernt habe ich aber nicht über das Modell, sondern über mich als gestaltende Person, nämlich dass gute KI-Unterstützung vor allem aus Zurückhaltung, Kontrolle und einer sauberen Architektur besteht.

9. Phase 4: Evaluation

9.1. Konkrete Durchführung

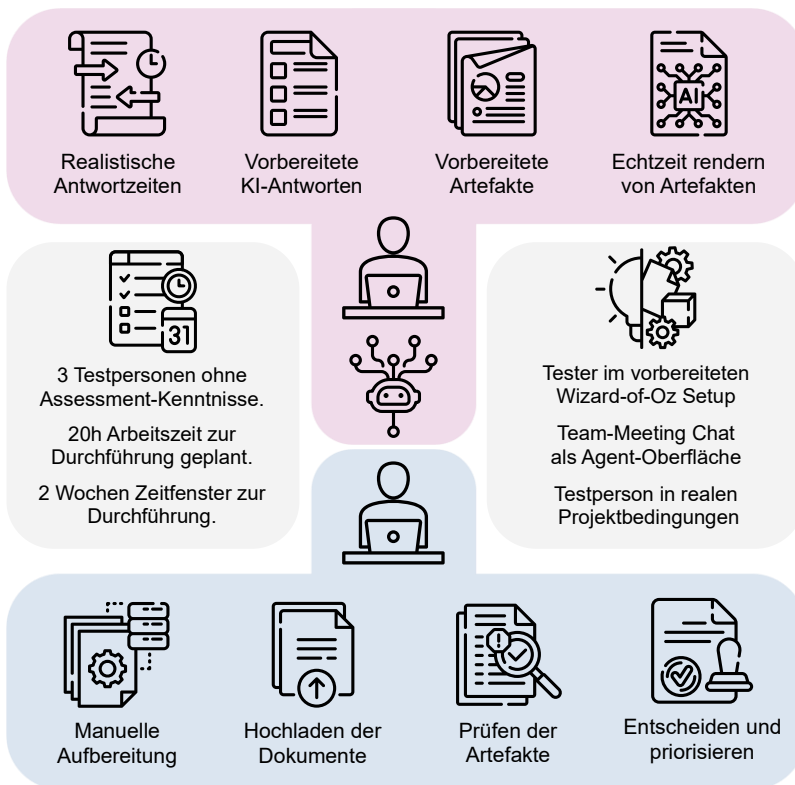


Abbildung 11: Grafische Veranschaulichung des Testing Setups
(Eigene Darstellung)

Gegenstand des Tests ist der Analyseprozess des Light Assessments in der Auswertungsphase, also die Strecke vom kontrollierten Rohmaterial bis zum bewertungsbereiten Summary. Genau hier fällt im bestehenden Assessment die manuelle Strukturierungs- und Verdichtungsarbeit an, die der konzipierte Prozess entlasten soll. Die drei Testpersonen durchlaufen diesen Ausschnitt eigenständig anhand von realen Kundendaten mit anonymisierten, zentral bereitgestellten Projektdaten. Als Ausgangslage liegen fünf Interviewtranskripte und ein Umfragedatensatz vor, die Datenerhebung gilt als abgeschlossen. Die Bearbeitung beginnt damit an dem Punkt, an dem im realen Assessment die aufwendige Analysearbeit einsetzt.

Die Durchführung erfolgte asynchron über ein knapp zweiwöchiges Fenster. Pro Testperson war ein Zeitbudget mit einem Richtwert von zwanzig Arbeitsstunden vorgesehen. Diese Vorgabe ist nicht allein organisatorisch begründet, sondern dient als Effizienzmassstab, denn eine zu starke Überschreitung wird ausdrücklich als Mangel des Prozesses und nicht als Leistung der Testperson gewertet. Da ein vollständiger Arbeitsprozess über mehrere Tage simuliert wird, ist eine durchgehende Beobachtung nicht möglich. An ihre Stelle treten eine eigenständige Bearbeitung über einen Aufgabenkatalog, mit tagebuchähnlichen Selbstdokumentationen, ergänzt durch einen begleitenden Rückfragekanal.

9.1.1. Aufgabenblatt und Selbstdokumentation

Den Rahmen bildet ein Aufgabenblatt (siehe Anhang K, S. 149), welches den Test- und den Aufgabenkontext einführt. Zu Beginn gibt es ein Kick-off, in welchem das Konzept des Light Assessments, das grundlegende Vorgehen und das damit verbundene Aufgabenblatt erklärt wird. Es legt eingangs offen, dass im Test keine Live-KI eingesetzt wird. Anschliessend gliedert es die Bearbeitung in vier aufeinander aufbauende Schritte: nämlich die Kontrolle und sprachliche Glättung der Transkripte, die Aufbereitung der Umfragedaten, die Inhaltsanalyse mit anschliessender Priorisierung sowie die Erstellung des Summaries. Jede Teilaufgabe benennt die konkrete Handlung, bietet Kontext- und Hintergrundinformationen und macht die vorgesehene Aufgabenteilung sichtbar.

Nach jeder Teilaufgabe beantwortet die Testperson dieselben drei offenen Fragen, welche tagebuchähnlich ausgefüllt werden können. Diese fragen nach der Nachvollziehbarkeit des Schritts, auf positive und negative Eindrücke sowie auf weitere Anmerkungen. Diese wiederkehrende Struktur hält die Eindrücke unmittelbar fest, solange der jeweilige Schritt präsent ist und übernimmt im asynchronen Setting die vertiefende Funktion, die sonst der moderierten Nachfrage zukommt.

9.1.2. Simulation der KI-Unterstützung

Die KI-Unterstützung wurde über einen Chat eines Meeting-Termins in Microsoft Teams bereitgestellt, in welchem mit dem simulierten Light Assessment Agent geschrieben wurde. Als Rückfragekanal dienten persönliche Nachrichten. Der Light Assessment Agent war während Geschäftszeiten und bestimmten Sonderzeiten zugänglich, um die spontane Nutzung bestmöglich zu simulieren. Die Testpersonen luden ihre Dokumente in den Chat und richteten ihre Arbeitsanweisungen an den Light Assessment Agent, dessen Antworten und Lieferdokumente eine Person im Hintergrund ausgab. Um den vorgesehenen Ablauf realitätsnah abzubilden, wurden die Antworten mit vorbereiteten realistischen Textbausteinen und mit Angaben zur Bearbeitungsdauer versehen, etwa einer simulierten Rechenzeit von rund zwanzig Minuten für die sprachliche Glättung. Die vollständigen Chatverläufe der drei Testpersonen sind dokumentiert (siehe Anhang L bis N, S. 163–169).

Die ausgegebenen Artefakte stammen durchgängig aus echten KI-Läufen und nicht aus manueller Nachbildung. Es wurden jedoch beim Layout des Analyseberichts manuelle Anpassungen vorgenommen, um bereits bekannte Probleme anzugehen und neue Ideen zu testen. Die Inhalte waren davon jedoch nicht betroffen. Schritte mit hoher Bearbeitungszeit, namentlich die Glättung der Transkripte und die Kodierung, wurden vorab erzeugt und zum passenden Zeitpunkt ausgegeben, um die asynchrone Durchführung praktikabel zu halten. Das Summary hingegen wurde während der Interaktion live mit der KI erzeugt und zurückgesandt. Auf diese Weise erlebten die Testpersonen Inhalte und Interaktionsform des geplanten Prozesses, ohne dass dieser vollständig entwickelt sein musste.

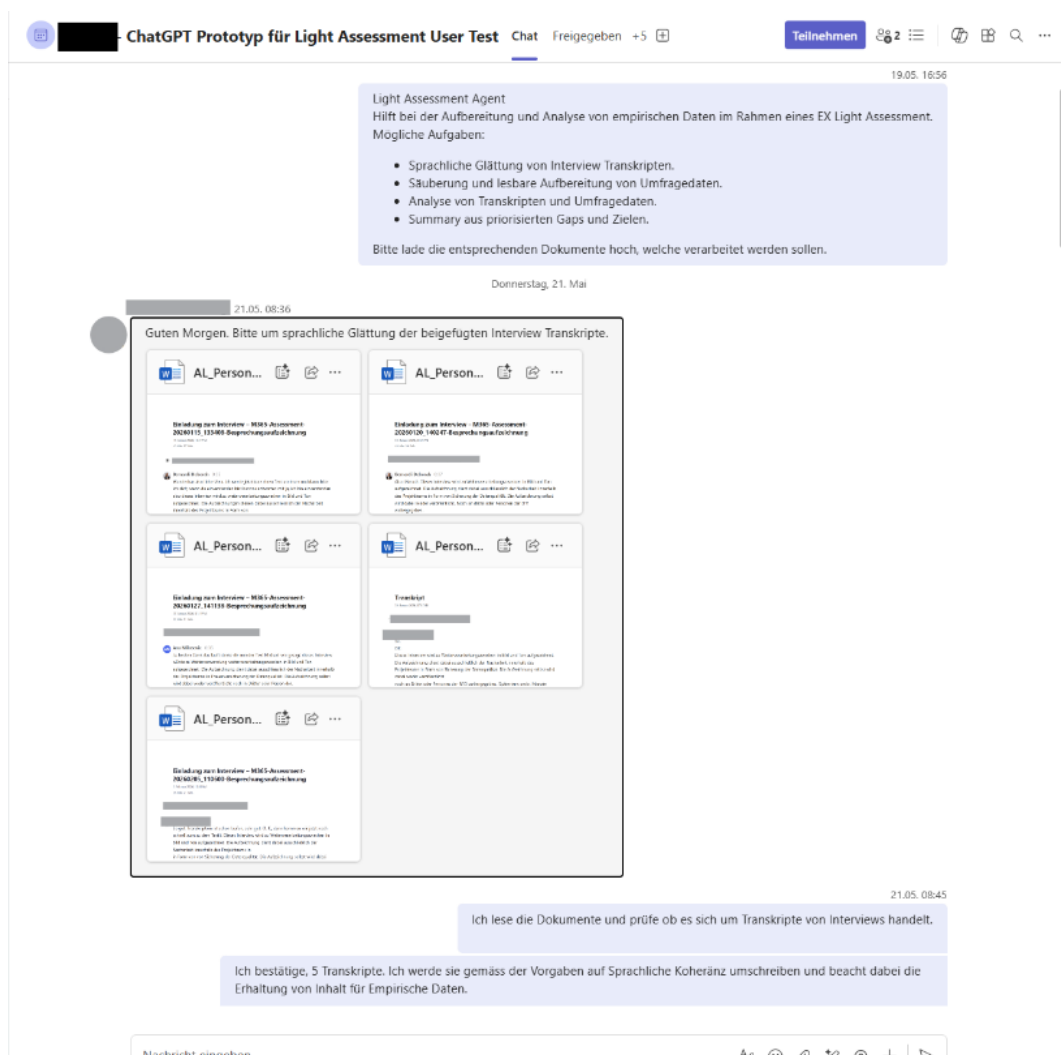


Abbildung 12: Screenshot Beispiel 1 des prototypisierten Light Assessment Agents (Eigene Quelle)

9.1.3. Abschliessender Fragebogen

Nach Abschluss der Bearbeitung folgte ein Fragebogen (siehe Anhang O, S. 172), der den gesamten Prozess rückblickend erfasst. Er ist entlang der Usability-Dimensionen aufgebaut und nimmt in offenen Fragen die Effektivität der Ergebnisse, die erlebte Effizienz sowie die Zufriedenheit während und nach der Durchführung auf, ergänzt um den tatsächlich benötigten Zeitaufwand. Während die Zwischenfragen die einzelnen Schritte begleiten, richtet sich der Fragebogen auf das Gesamtbild und bildet die summative Endbewertung des Tests.

9.2. Ergebnisse

Die Rückmeldungen der drei Testpersonen werden im Folgenden entlang der Usability-Dimensionen Effektivität, Effizienz und Zufriedenheit verdichtet dargestellt. Den Abschluss bilden exemplarische Verbesserungsvorschläge, die sich aus den wiederkehrenden Schwachstellen ergeben. Die verdichteten Rückmeldungen zu den einzelnen Aufgaben und zum abschliessenden Fragebogen finden sich (siehe Anhang P, S. 174 und Anhang Q, S. 176).

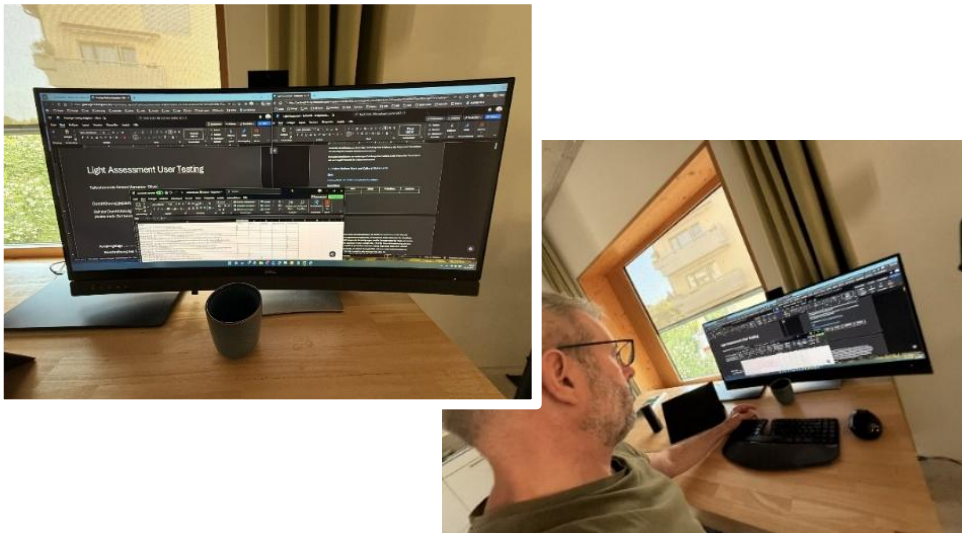


Abbildung 13: Versuchsaufbau des asynchronen Usability-Tests am Light Assessment (Foto: Teilnehmende Person)

9.2.1. Effektivität

Hinsichtlich der Effektivität, verstanden als das vollständige Erreichen des angestrebten Arbeitsergebnisses, fällt das Bild überwiegend hoch, aber wie zu erwarten, nicht vollständig ohne Mängel aus. Die Zwischenergebnisse und das Summary werden als weitgehend korrekt eingestuft, «auf einem hohen Niveau, aber nicht perfekt» (Person 02), und gelten als «ohne Zweifel weiterverwendbar» (Person 02). Sie erscheinen weitgehend vollständig und verlässlich, wenn auch mit Ausnahmen und manuellen Anpassungen, welche jedoch nur die Notwendigkeit einer fachlichen Prüfung demonstrieren, inhaltlich fehlt «nichts» und es sind «nur kleinere Anpassungen» nötig (Person 03), während der Mensch «weiterhin die finalen Entscheidungen und Einschätzungen» treffen soll (Person 02). Die Quellenangaben im Bericht stützen die Nachvollziehbarkeit, durch sie erhält das Endergebnis «eine maximale Nachvollziehbarkeit» (Person 02). Abstriche zeigen sich vor allem beim Lieferobjekt für den Kunden. Der Bericht wird in der vorliegenden Form als zu lang und zu wenig aufbereitet beschrieben, als «zu umfangreich und zu wenig selbsterklärend» (Person 02) und «eindeutig zu lange» (Person 03), und ist ohne ergänzende Präsentation nicht direkt weitergebbar, eine Weitergabe müsste «zwingend in Form einer Präsentation» erfolgen (Person 03). Das Vertrauen in die Ergebnisse wird als gestützt geschildert, «ganz klar gestützt, aber ebenso klar die Grenzen aufgezeigt» (Person 01), bleibt aber an die menschliche Kontrolle und an die Qualität der Datengrundlage gebunden, da die KI «nur so gut» ist, «wie wir sie auch beliefern» (Person 03).

ChatGPT Prototyp für Light Assessment User Test Chat Freigegeben +5 **Teilnehmen**

Ich habe die Transkripte gemäss den Vorgaben geglättet und dabei darauf geachtet den Inhalt nicht zu verändern. Bitte überprüfe die Dokumente auf ihre Richtigkeit. Bei Fehlern, bitte Korrekturen vornehmen und bei gravierenden Fehlern, kontaktiere die Verantwortliche Person für den Agent.

Guten Tag. Säubere die beigefügten Umfragedaten und bereite sie lesbar und grafisch auf.

	A	B	C	D
1	400401.000000	400401.000000	anonymisiert	
2	400401.000000	400401.000000	anonymisiert	
3	400401.000000	400401.000000	anonymisiert	
4	400401.000000	400401.000000	anonymisiert	
5	400401.000000	400401.000000	anonymisiert	
6	400401.000000	400401.000000	anonymisiert	
7	400401.000000	400401.000000	anonymisiert	
8	400401.000000	400401.000000	anonymisiert	
9	400401.000000	400401.000000	anonymisiert	
10	400401.000000	400401.000000	anonymisiert	
11	400401.000000	400401.000000	anonymisiert	
12	400401.000000	400401.000000	anonymisiert	
13	400401.000000	400401.000000	anonymisiert	
14	400401.000000	400401.000000	anonymisiert	

21.05. 10:31

Ich bestätige es handelt sich um Umfragedaten für das Light Assessment.
Ich prüfe die Excel-Datei auf typische Datenqualitätsprobleme: sehr kurze Bearbeitungsdauer, auffällige Antwortmuster wie Straightlining, viele fehlende Werte und gegebenenfalls Duplikate oder technisch leere Fälle. Danach liefere ich eine Gesäuberte Exceldatei und eine lesbare Version als Worddokument.
Einen Moment.

(Vorgang Läuft)

Im Datensatz wurden mehrere qualitativ auffällige Fälle identifiziert, die durch extrem kurze Bearbeitungszeiten, monotone Antwortmuster, fehlende Kernangaben oder widersprüchliche Antwortlogiken auffallen und daher aus der Auswertung ausgeschlossen werden sollten.
Ich bereite nun die Daten gemäss Anweisungen lesbar auf.

(Vorgang Läuft)

21.05. 10:39

Hier sind die aufbereiteten Umfragedaten, sowie eine gesäuberte Version des Datensatzes als Excel.

Guten Tag. Führe eine Inhaltsanalyse an allen beigefügten Transkripten durch.

Abbildung 14: Screenshot Beispiel 2 des prototypisierten Light Assessment Agents (Eigene Quelle)

9.2.2. Effizienz

Der Ablauf wird überwiegend flüssig und hindernisarm erlebt. Die Prozessschritte werden als logisch, konsistent und klar empfunden und die Rückmeldungen des Agenten als verständlich, sodass der jeweils nächste Schritt erkennbar bleibt. Eine Testperson beschreibt das Zusammenspiel der Werkzeuge als «sehr logisch und konsistent», die Rückmeldungen als «sehr verständlich» und hält fest: «Ich wusste zu jedem Zeitpunkt, was der genaue Schritt ist» (Person 03). Eine weitere nennt den Ablauf «einfach und klar» (Person 01). Das Aufwand-Ergebnis-Verhältnis wird überwiegend als sehr gut bezeichnet, einmal schlicht als «sehr gut» (Person 02), allerdings mit dem Hinweis, dies gelte vor allem für «Personen die einen Projekt- resp. Kundenbezug haben» (Person 01). Eine empfundene Zeitersparnis, eine «klare Zeitersparnis» (Person 02), wird auch darauf zurückgeführt, dass sich während der Rechenzeit der KI andere Aufgaben erledigen lassen, «dass man in der Zeit, wo die KI arbeitet, andere Tasks erledigen kann» (Person 03). Belastender und langsamer fällt vor allem der Einstieg beim Überprüfen der Interviews aus, solange Themen- und Kontextwissen zum Kunden fehlen, die Belastung wird «vor allem zu Beginn doch als sehr hoch» empfunden (Person 03). Dies ist dem Testkontext geschuldet, da bereits geführte Interviews eines anonymisierten Kunden zur Verfügung gestellt wurden und daher eigentlich nicht gegen den Prozess zu werten. Es findet dennoch Erwähnung, da dieses Empfinden nochmals die Notwendigkeit demonstriert, die Erhebungen zum positiven Nutzungsempfinden der Fachpersonen nicht zu vollautomatisieren. Dabei steigt die empfundene Effizienz mit zunehmender Einarbeitung deutlich, was wiederum der tatsächlichen Zeitersparnis durch die KI geschuldet ist, da die Durchführungszeit im Vergleich zur manuellen Analyse drastisch reduziert werden konnte. Eine Testperson hätte «das selber nie so schnell geschafft» (Person 03) und mit Blick auf den umfangreichen Analysebericht «bräuchte man wahrscheinlich Wochen» (Person 02).

9.2.3. Zufriedenheit

Während der Durchführung ergibt sich ein insgesamt zufriedenes Bild mit einzelnen Reibungspunkten. Orientierung und Handlungssicherheit werden als sehr zufriedenstellend und besser als erwartet beschrieben, einmal als «sehr zufriedenstellend» (Person 02), einmal als «besser als erwartet, bestätigt durch den jeweiligen Output» (Person 01), verbunden mit dem Gefühl, «sehr gut geführt und sicher» gewesen zu sein (Person 03). Die Beanspruchung gilt als gut handhabbar, auch dank einer Arbeitsweise im Stop-and-go-Verfahren, das «in meinem Arbeitsumfeld sehr relevant ist» (Person 01). Die weiteren Prozessschritte werden als «nicht sehr anstrengend» und «in einem mehr als vertretbaren Rahmen» erlebt (Person 03). Positiv hervorgehoben werden die sprachliche Glättung der Transkripte. «Ich bin sehr beeindruckt, wie gut die KI die Glättung vorgenommen hat» (Person 03), sowie die klar und angenehm erlebte Interaktion mit dem Agenten samt seinen Zwischenzusammenfassungen. Sie wirkt «logisch und klar» (Person 01) und «angenehm» (Person 02), wobei «jeweils eine kurze Zusammenfassung kam, was hochgeladen wurde und was nun der nächste Schritt ist» (Person 03). Auch hier kann wieder der themenferne Einstieg als belastender gelten, mit teils schlechter Aufnahmequalität. Der Einstieg ist «zu Beginn sehr anstrengend, da ich nicht im Thema drin war» und «die Aufnahmequalität nicht sehr gut» (Person 03), sowie wiederkehrend der Priorisierungsschritt (Aufgabe 3.b), bei dem es «einige Unsicherheiten» gab (Person 01). Diese Mängel sind auch wieder dem Testkontext geschuldet, doch ebenfalls wichtige Beispiele zur Priorisierung von Bedürfnissen für spätere Iterationen.

Das rückblickende Gesamtbild fällt durchweg positiv aus, auch dort, wo einzelne Schritte während der Nutzung als anstrengend erlebt wurden. Es ist ein «positives Gesamtbild, durch nachvollziehbare Schritte und einen logischen Ablauf» (Person 01) und «sehr positiv», da der Prozess «bei korrektem Einsatz wirklich einen Mehrwert für uns und die Kunden bieten» kann (Person 02). Die Aufgabenverteilung zwischen Mensch und KI wird einhellig als richtig beurteilt, mit der Fleissarbeit bei der KI und den Entscheidungen beim Menschen. «Die Fleissarbeit macht die KI und die Leitplanken sowie Entscheidungen trifft der Mensch» (Person 02), denn «es gibt einfach Dinge, die kann die KI nicht übernehmen» (Person 03). Ein Einsatz in realen Kundenprojekten wird grundsätzlich befürwortet und für jedes Assessment-Projekt für

vorstellbar gehalten. «Grundsätzlich kann ich mir den Einsatz bei jedem Projekt vorstellen, bei dem ein Assessment durchgeführt wird» (Person 03), einmal explizit mit der Anregung, ihn als Standardvorgehen zu verankern, «gemessen am Resultat verstehe ich das als neuen Standard bei Projekten» (Person 01). Als Bedingung gilt das nötige Verständnis der Methoden samt interner Schulung. Eine Testperson hält «intern Schulungen für den richtigen Weg» und betont, «dass die Methoden verstanden werden» (Person 02), sowie die datenschutzrechtliche Klärung mit dem Kunden, «ob die Informationen durch KI analysiert werden dürfen» (Person 03).

9.2.4. Bezug zur Forschungsfrage (Teilfrage 4) und Zielsetzung (Teilziele 3 und 4)

Mit den oben dargestellten Rückmeldungen ist die Teilfrage beantwortet, inwieweit der Prototyp Nutzbarkeit und Effizienzgrad verbessert, und sind die Teilziele der Nutzbarkeits- und Effizienzevaluation erreicht. Insgesamt zeigt sich eine hohe Nutzbarkeit bei klar benannten Schwachstellen. Den Effizienzgewinn stützt der Vergleich mit der bisherigen Praxis. Die manuelle Vollanalyse beansprucht näherungsweise fünfzig bis sechzig Arbeitsstunden, während der im Testing erhobene Aufwand bei dreizehn, zwanzig und einundzwanzig Stunden liegt, wobei die höheren Zeiten zunächst auch nicht zu wertende Arbeiten wie das Ausfüllen des Fragebogens und den Kick-off enthielten. Damit gelten zwanzig Stunden zur Durchführung als plausibel und im nächsten Iterationsschritt erreichbar und bleiben im Rahmen der unternehmensseitigen Zielgrösse.

9.3. Exemplarische Verbesserungsvorschläge

Aus den wiederkehrenden Schwachstellen lassen sich mehrere konkrete Ansatzpunkte ableiten, von denen vier der wichtigsten hier exemplarisch erörtert werden.

9.3.1. Der Priorisierungsschritt 3.b

Schwachstelle

Dies ist der einzige Schritt, den die Testpersonen übereinstimmend überarbeiten würden. Er gilt als zu komplex und zu wenig unterteilt, die Anleitung lässt offen, welche Felder auszufüllen sind und Gap-Titel sind ohne zugehörige Evidenz teils nicht verständlich. Angeregt werden eine stärkere Unterteilung des Schritts, eine klarere Ausfüllanleitung sowie eine zusätzliche, abgespeckte Berichtsvariante, deren Aufwand sich nach der Kundengrösse richtet und in der nur noch bestehende Empfehlungen und Einschätzungen geprüft werden.

Verbesserung

Das war anhand der Reflexion am Ende des Participatory Designs bereits zu erwarten. Der Umfang des Analyseberichts ist auf Grund der vielen vorgegebenen Codes zu gross und die Ziele oder Gaps sind nicht so selbsterklärend, wie angenommen.

Die Kodierung (und damit die Gaps und Ziele) wird durch Teammitglieder des Participatory Designs und zusätzlichen neuen Fachpersonen nochmals überarbeitet und reduziert. Diese Verbesserung ist bereits eingeleitet, noch während dem Schreiben dieser Arbeit. Eine Verbesserung der Selbsterklärung, durch weitere und bessere beschreibende Text, ist eine weitere Anpassung zu diesem Schritt.

Diese Reduktion und Verbesserung der Selbsterklärung zählt auch auf eine weitere Reduktion der Stunden ein und in den Umfang des Lieferobjekts.

9.3.2. Das Summary 4.a

Schwachstelle

Das Lieferobjekt erreicht in der vorliegenden Form noch nicht die Kundenreife und ist immer noch zu umfangreich. Zudem wird erwähnt, es sei zu wenig visuell aufbereitet beschrieben. Vorgeschlagen werden

ein kundengerechtes Layout, gezielte Grafiken und Highlights, etwa ein Kreisdiagramm zum Anteil festgestellter Handlungsbedarfe, sowie die durchgängige Weitergabe über eine begleitende Präsentation.

Verbesserung

Auch hier ist der schiere Umfang ein Priorisierungsproblem. Es wurden sehr viele Codes hoch priorisiert. Zum einen, da das genaue Ausmass der Bedeutung, was es heisst zu priorisieren, nicht ausreichend erklärt war. Zum anderen gab es Unsicherheiten, da der Kunde nicht bekannt war, wodurch vorsichtshalber mehr hoch priorisiert wurde.

Während das Problem mit dem unbekannten Kunden, dem Prototypkontext geschuldet, hier ausser Acht gelassen werden kann, werden durch die Anpassungen am Analysebericht wichtige Schwachstellen im Kern auch hier gelöst.

Eine weitere wichtige Anpassung ist die visuelle Aufbereitung, welche einen bewussten nächsten Schritt in der Weiterentwicklung darstellt. In diesem Kontext wird auch der Umgang mit den Summary Daten und Übergang zur Präsentation angegangen. Dieser Schritt war bis jetzt nur am Rande des Projektes beachtet worden und gilt ausserhalb des Scopes dieser Arbeit als offener Punkt, der noch mit Fachpersonen nach dem Erstellen des Analyseprozesses erarbeitet werden muss. Dass diese beiden Punkte als Mangel wahrgenommen wurden, bestätigt die Notwendigkeit dieses Schrittes.

9.3.3. Sensible Daten

Schwachstelle

Beim Datenschutz zeigt sich, dass der Bericht zwar anonymisierte Personenkürzel verwendet, die Ausführungen jedoch Hinweise auf Arbeitskollegen oder Abteilungen enthalten, die Rückschlüsse erlauben. Als Lösungsansatz wird erwogen, die Entfernung solcher Bezüge bereits als Teilschritt der Transkriptkontrolle zu verankern und die datenschutzrechtliche Zulässigkeit vorab mit dem Kunden zu klären. Da dies prozessgeschuldet ist und im Verständnis der Unterteilung von KI-Mensch-Aufgaben eine wichtige Rolle spielt, stellt dies im Kontext des Prozesses im weitestgehenden Sinne ein Usability Problem dar.

Verbesserung

Die KI kann und soll nicht eigenständig und unkontrolliert Zensuren von Personen vornehmen, da bestimmte Personen wichtige zu identifizierende Rollen besitzen. Diese Zensur muss unter rein manueller Kontrolle bleiben. Jedoch wurde dies nicht im Prozessschritt der manuellen Bereinigung erörtert. Hier muss eine Sensibilisierung und Aktualisierung der Prüfpunkte vorgenommen werden.

9.3.4. KI-Mensch Interaktion

Schwachstelle

Bei den länger laufenden KI-Schritten fehlt eine Rückmeldung über den Bearbeitungsstand. Bei der Inhaltsanalyse mit einer Laufzeit von rund zwei Stunden entsteht ohne Fortschrittsanzeige Unsicherheit, ob der Agent noch arbeitet. Angeregt werden eine sichtbare Verarbeitungsanzeige sowie die Angabe einer durchschnittlichen Bearbeitungsdauer je Schritt.

Dieses Problem kann als reines «Prototyp-Kontext-Problem» betrachtet werden, da wegen der Chat-Umgebung keine aktive Arbeitsanzeige stattfand. Es zeigt jedoch einen wichtigen Hinweis für das Empfinden von Vertrauen und Sicherheit gegenüber der KI-Lösung während des eigentlichen Prozesses.

Verbesserung

Es ist ein Standard bei LLM-basierten Systemen, stetige Hinweise über den aktuellen Stand der Tätigkeit und Denkvorgänge zu geben. Durch die schiere technische Umsetzung ist das Problem theoretisch gelöst. Erwähnt wird diese Schwachstelle allerdings, da je nach tatsächlicher Umsetzung und Oberfläche, diese Statusanzeigen entweder dennoch entfallen oder nicht hilfreich sein können. Ein bewusster Umgang mit

klaren Statusrückmeldungen, welche vorgegeben und nachvollziehbar sind, kann hier eine höhere Zufriedenheit erreichen, ähnlich wie ein optimiertes Design einer Oberfläche.

9.3.5. Weitere Erkenntnisse und Verbesserungen

Während des Testings versuchten die Testpersonen, Fragen an den Agenten zu stellen, an welche bis dahin nicht gedacht wurde. Aus diesen Interaktionen ergaben sich Erkenntnisse hinsichtlich der Nutzung des Agenten, welche zu konkreten Anpassungsvorschlägen für weitere Iterationen führen.

Human-Fallback Meldung

Aufgrund dessen, wie die KI zur Sicherung der Verlässlichkeit eingeschränkt ist, gibt es schlichtweg Dinge, zu welchen sie keine Aussagen machen darf. Gewisse Themen, besonders hinsichtlich Interpretationsfragen oder Datenverarbeitung ausserhalb des Harness Engineering, dürfen nicht durchgeführt werden. Solche Dinge müssen mit dem Owner des Assessment und der damit fachlich verantwortlichen Person geklärt werden.

Im Testing wurde im ersten dieser Momente spontan eine Meldung mit den Kontaktdaten geschaffen, welche darauf hinweist, sich bei grösseren Problemen oder Fehlern an die zuständige Fachperson zu wenden. Dann wurde in einem persönlichen Chat das Problem behoben. Eine solche Fallbackmeldung muss in späteren Iterationen noch eingebaut werden.

Bedeutungsabfrage

Konkret wurden Definitionsfragen zu den Codes (Ziele und Lücken) gestellt. Dabei sollte beantwortet werden, wie genau ein bestimmter Code zu verstehen ist.

Daraus entsteht die notwendige Anpassung einer Bedeutungsabfrage. Die KI darf nicht von selbst die Bedeutung interpretieren. Es müssen klare Definitionen geschaffen werden, auf welche die KI zurückgreifen kann.

Isolierte Rückprüfung

In einem anderen Fall wurde eine Nachprüfung gefordert, als in einem Punkt die Validität der Daten angezweifelt wurde. Da es gedacht und kommuniziert war, diese an der Stelle komplett manuell zu prüfen oder sich mit dem Owner des Assessments in Verbindung zu setzen, der fachlich unterstützen kann, war das Testsetting nicht darauf ausgelegt. Daraus hat sich jedoch die Idee ergeben, dass eine abrufbare isolierte Rückprüfung noch in nächsten Iterationen eingebaut werden sollte.

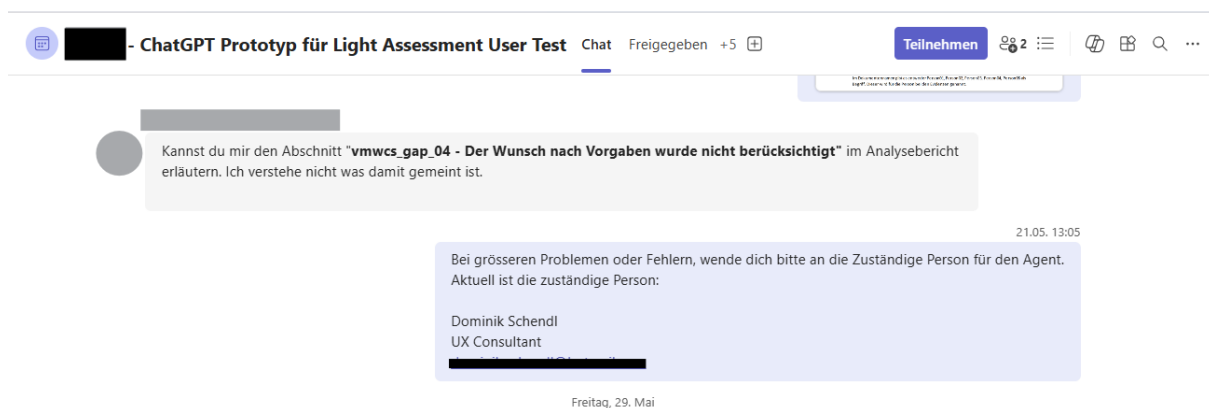


Abbildung 15: Screenshot Beispiel 3 mit spontanem Admin-Text im prototypisierten Light Assessment Agent (Eigene Quelle)

9.4. Lernreflexion – Phase 4

Das finale User-Testing ist für mich eine der aussergewöhnlicheren Stücke der ganzen Arbeit gewesen. Bis dahin hatte ich vor allem konzipiert, gebaut und im engeren Kreis erprobt. Zum einen, wie ich es vom User-Testing kenne, wurde in der finalen Evaluation der Prozess zum ersten Mal Personen überlassen, die nicht Teil der Projektgruppe waren. Das Besondere war aber, dass schon der Aufbau dieses Tests Entscheidungen verlangte, die ich so bei den mir vertrauten regulären Testings noch nie getroffen hatte.

Am ungewohntesten war das Setup selbst. Ein klassischer Usability-Test ist eine kurze, moderierte Sitzung, bei der ich der Person über die Schulter schaue. Das war hier nicht möglich, weil der zu testende Prozess ein vollständiger Arbeitsablauf über mehrere Tage ist und nicht eine einzelne Interaktion. Ich musste den Test deshalb unmoderiert und asynchron über zwei Wochen als Projekt anlegen, mit einem Zeitbudget von zwanzig Stunden pro Testperson und obwohl ich natürlich während dieser Zeit jederzeit erreichbar war, musste ich die fehlende Beobachtung durch ein Aufgabenblatt mit tagebuchartigen Zwischenfragen ersetzen. Wobei diese Zwischenfragen zunächst spezifisch pro Aufgabe und eher ein elaborierter Fragebogen waren. Dann habe ich die Tagebuch-Idee endlich aufgegriffen und die Fragen so weit reduziert, dass ich nur dieselben drei wichtigsten offenen Grundfragen für jede Aufgabe stellen konnte. Dafür habe ich dann am Schluss noch einen Fragebogen für den gesamten Prozess ausfüllen lassen. Ich habe mich hierbei gegen ein nachträgliches offizielles Interview entschieden. Ich hielt es schlichtweg für richtig, da die Testpersonen auf diese Weise nach der Durchführung und in eigener Geschwindigkeit den Fragebogen ausfüllen konnten. Dabei konnten sie auch Tagebucheinträge durchgehen und anhand dieser nachgedacht antworten, anstatt erst vielleicht Tage später, bis ein Termin zustande kommt, in ungefährender Erinnerung spontan zu antworten. Diese Kombination aus Usability-Test, Tagebuchstudie und Fragebogen habe ich bewusst gewählt, weil keine einzelne Methode, welche ich dafür in Erwägung ziehen konnte, den Prozess in meinen Augen hätte abbilden können. Hinzu kam, dass der Prototyp noch kein fertiger oder zumindest irgendwie einsatzfähiger Agent war, wie ich es eigentlich gehofft hatte, weshalb ich die KI im Wizard-of-Oz-Verfahren simulieren musste. Im Nachhinein bin ich sehr froh, dass ich das getan habe und ich glaube, ich würde es wieder so tun, einfach etwas vorbereiteter, denn ich habe den Stress deutlich unterschätzt. Im Hintergrund, im richtigen Moment die vorbereiteten und teils live erzeugten Ergebnisse auszugeben kann etwas hektisch werden und über zwei Wochen auch etwas auf die Psyche drücken. Das erst recht, wenn man den Anspruch hat, während der gesamten Geschäftszeiten immer abrufbar zu sein und KI-gerecht schnell zu reagieren, um die Flexibilität der Nutzung realitätsnah zu simulieren. Das hatte ich so nicht gedacht und schlichtweg unterschätzt und müsste ich für ein nächstes Mal nochmals besser durchdenken.

Die Durchführung selbst hat dennoch besser funktioniert, als ich erwartet hatte. Das Wizard-of-Oz-Setup, das mir etwas Sorge bereitet hatte, ging auf. Weil die ausgegebenen Artefakte und Antworten aus echten KI-Läufen stammten, erlebten die Testpersonen den Prozess in Inhalt und Interaktionsform real. Mich hat überrascht, wie selbstverständlich die Beteiligten den Ablauf annahmen und wie klar sie sich in den Dokumenten orientieren konnten. Besonders gefreut hat mich, dass gerade die Schritte, bei denen ich am unsichersten war, etwa die sprachliche Glättung der Transkripte, ausdrücklich gelobt wurden.

Nicht alles lief reibungslos und einige Schwächen traten erst in der eigenständigen Nutzung zutage. Der Priorisierungsschritt erwies sich übereinstimmend als zu komplex und zu wenig angeleitet und der erzeugte Bericht war den Testpersonen schlicht zu lang. Beides hatte ich nach dem Reflexionsworkshop zwar schon geahnt, doch im finalen Test wurde unübersehbar, wie deutlich diese Punkte stören. Wichtig war für mich, sauber zu trennen, was ein echter Prozessmangel ist und was nur dem Testaufbau geschuldet war. Etwa, dass die Testpersonen mit den Daten eines ihnen unbekannten, anonymisierten Kunden arbeiten mussten. Der Punkt hat mich dennoch sehr beschäftigt und ich nutzte ihn für die firmeninterne Kommunikation um ihn bei den Diskussionen um die Automatisierung der Erhebung einzusetzen.

Besonders lehrreich war, dass die Testpersonen den Prozess eigenständig nutzten und ihn dabei an Stellen forderten, an die ich nicht gedacht hatte. Es ist immer toll, wenn man bei einem Testing Momente hat, in

welchen sich Unerwartetes ergibt. Irgendwie hatte ich mir aber in den Kopf gesetzt, dass es solch einen Moment hier wahrscheinlich nicht geben würde. Aber solche Momente zu bieten, da ist einfach ein Verlass auf User-Testings. Die Testpersonen stellten dem Agenten Fragen, die im Konzept gar nicht vorgesehen waren und ich musste spontan eine Rückfallmeldung erfinden, die auf die zuständige Fachperson verweist. Meiner Meinung nach sind, gerade weil der Test nicht moderiert war, diese Situationen überhaupt erst aufgetreten. Denn wäre ich daneben gesessen hätte ich hier sowas gesagt wie: «Der Agent würde jetzt XY schreiben», und hätte das nicht so sehr gespürt. Damit haben mir diese Momente konkrete Verbesserungen geliefert, die ich am Schreibtisch kaum gefunden hätte. Damit gehören, wie auch in regulären Usability-Testings, einmal mehr diese ungeplanten Momente für mich im Nachhinein zu den wertvollsten Ergebnissen der ganzen Evaluation.

Am meisten überrascht hat mich die Nachwirkung des Tests. Ich hatte mit nüchternem Feedback gerechnet, doch die Begeisterung der Testpersonen hielt an und alle äusserten sich im Rahmen normaler kollegialer Gespräche etwa zwei Wochen später noch als Fürsprecher des Prozesses. Eine Person bezeichnete ihn sogar als möglichen neuen Standard für Assessmentprojekte. Dass eine eigenständige, über zwei Wochen verteilte und durchaus anstrengende Auseinandersetzung mit empirischen Daten am Ende nicht in Ermüdung, sondern in echter Überzeugung mündete, hatte ich so nicht erwartet und es hat mir mehr über den Wert des Prozesses gesagt als jede einzelne Kennzahl.

Rückblickend war die Evaluation für mich weniger eine Prüfung des Prozesses, als eine Prüfung meiner eigenen methodischen Beweglichkeit. Ich musste akzeptieren, dass mein Gegenstand kein Standardverfahren erlaubt und mir aus mehreren Methoden ein tragfähiges und ehrliches Setup selbst zusammenstellen. Dass dieses ungewöhnliche und aufwändige Vorgehen am Ende belastbare und zugleich begeisternde Ergebnisse lieferte, ist die Erkenntnis, die ich aus dieser Phase mitnehme. Eine Evaluation muss zum Gegenstand passen, ihre Grenzen offen ausweisen und Raum für das Unerwartete lassen, wenn nötig auch in der freien, eigenständigen Nutzung.

10. Technischer Aspekt des KI-Einsatzes

10.1. Abgrenzung

Vollständigkeitshalber wird das Proof of Concept des Light Assessments im Folgenden auch technisch geschildert. Der Schwerpunkt der Arbeit liegt auf der menschenzentrierten Gestaltung, doch erscheint ein Einblick in die KI-gesteuerten Prozessanteile lohnend, da dieses Zusammenspiel bislang wenig erschlossen ist und einzelne Grundlagen später in das Vorgehensmodell einfließen. Zur Veranschaulichung dienen konkrete Beispiele aus dem Light Assessment, dessen technische Teile mit KI-Fachpersonen aus drei Unternehmen, darunter Microsoft, fachlich abgestützt wurden.

10.2. Awareness für Harness Engineering

Während Context Engineering auf die inhaltliche und strukturelle Bereitstellung des relevanten Bezugsmaterials zielt, bezeichnet Harness Engineering die Gestaltung der technischen Ausführungs-, Steuerungs- und Kontrollumgebung, in der ein LLM-basiertes oder agentisches System arbeitet. Dazu gehören insbesondere die Einbindung von Werkzeugen, Zustands- und Speicherstrukturen, Ablaufsteuerung, Berechtigungen, Prüfmechanismen, Protokollierung, Feedbackschleifen und menschliche Eingriffsmöglichkeiten (Zhong und Zhu 2026). Für die technische Umsetzung und die Funktionalitäten des Prototyps ist diese Ebene relevant, da sie bestimmt, wie das System Kontext verarbeitet, Werkzeuge nutzt, Zwischenergebnisse absichert und Bearbeitungsschritte kontrollierbar macht.

Der analytische Fokus dieser Arbeit liegt jedoch bewusst auf Context Engineering. Diese Eingrenzung ist methodisch begründet: Für die Konzeption von KI-Prozessen, auf dem Level, wie es HCD-Fachpersonen potenziell machen, in Analogie zu Prototypen, steht nicht die technische Architektur im Zentrum. Es ist die Frage, wie relevantes Bezugsmaterial so ausgewählt, strukturiert und situativ bereitgestellt werden kann, dass KI-gestützte Bearbeitungen an konkretes Material gebunden, am Original überprüfbar und im Zeitverlauf konsistent steuerbar bleiben, welche im Zentrum stehen muss. Dabei sollte jedoch Harness Engineering berücksichtigt werden, ähnlich wie beim Prototypen die Größe des Screens, Positionierungen, native Interaktionselemente, CI-Vorgaben usw.

Im Zusammenspiel mit Vibe Research bildet Context Engineering die Grundlage dafür, dass Steuerimpulse, Belegbarkeit und iterative Weiterentwicklung verlässlich ineinandergreifen. Wenn diese als saubere Prozessreihe in funktionierenden Dokumenten in klarer Auszeichnungssprache abliegen, können diese mit Harness Engineering sauber übernommen werden. Wie ein guter Prototyp von Programmierern.

10.3. Kodierleitfaden für LLM-basierte Systeme

Im Kontext der vorliegenden Arbeit ist die strikte deduktive und deterministische Anwendung des Kodierleitfadens eine zentrale Voraussetzung für das Funktionieren des Light Assessment Prozesses. Dieser zielt nicht darauf ab, qualitative Daten explorativ auf neue oder induktiv entstehende Themen hin zu untersuchen, sondern dient der systematischen Identifizierung vorab definierter Themen im Interviewmaterial. Das LLM wird daher nicht als exploratives Analyseinstrument eingesetzt, sondern als regelgebundene Kodierinstanz, die prüft, ob, und in welcher Form, die festgelegten Haupt- und Subkategorien im Material nachweisbar sind.

Aus dieser Anforderung leitet sich die doppelte Funktion des erweiterten Kodierleitfadens ab: Er operationalisiert die deduktiv festgelegten Kategorien im Sinne der qualitativen Inhaltsanalyse und hält zugleich die Kodierung des LLM innerhalb der Logik des Light Assessment Prozesses. Das Modell darf ausschliesslich vorgegebene Code-IDs verwenden, keine Kategorien umbenennen und keine neuen oder induktiven

Kategorien bilden. Lässt sich ein Segment keiner bestehenden Subkategorie eindeutig zuordnen, wird es nicht neu kategorisiert, sondern als nicht kodierbarer oder prüfbedürftiger Fall markiert.

Damit wird der Kodierleitfaden zum zentralen Kontrollinstrument des Light Assessment Prozesses. Die Qualität der LLM-gestützten Kodierung bemisst sich folglich nicht daran, ob das Modell möglichst viele neue Bedeutungen aus dem Material ableitet, sondern daran, ob es die vorgegebenen Kategorien konsistent, regelkonform und am Text belegbar anwendet.

10.3.1. Herantasten an den genauen Kodierleitfaden

Um einen verlässlichen Kodierleitfaden für die KI zu erstellen, wird dieser anfangs bewusst sehr knapp gehalten und dann durch Testen erweitert. In einer frühen Phase werden dann bereits die Empfehlungen im Verlauf der partizipativen Gestaltung aus den Subkategorien entfernt und nur noch Zielzustände und Mangelzustände erhoben. Wobei hier eher von der Strukturierung des Codebuches gesprochen werden muss, denn gemäss Mayring (2022) gehören beispielsweise Forschungsfrage, Materialbasis, Analyseform, Auswertungseinheit, Kontexteinheit oder Kodiereinheit dazu. Solche Felder befinden sich jedoch noch in den Anweisungen für die KI, nicht im Kodierleitfaden. Denn bei der Art und Weise, wie die initialen Content Engineering Ansätze durchgeführt werden, sorgt dieser Ansatz für mehr Stabilität und bessere Ergebnisse.

Tabelle 13: Rudimentärer Kodierleitfaden für die rein deduktive Inhaltsanalyse im Light Assessment

Feld	Zweck
Hauptkategorie	Beschreibt das eigentliche Hauptthema
Subkategorien	Beschreibt ein konkretes Ziel- oder Mangelzustand
Id	Eine eindeutige Zuordnung für die Mensch und KI
Bezeichnung	Eine genaue Formulierung für Mensch und KI
Definition	Eine erweiterte Erklärung explizit angepasst für die KI
Kontext	Ergänzende kontextuelle Informationen zum besseren Verständnis für die KI.

Mit genau diesem vereinfachten Leitfaden werden im durchgeführten Reliabilitätstest sehr gute und in der Validität gute Werte erreicht. Wenn man nun weiter in Richtung Harness Engineering geht und zuverlässigere deterministische Denkprozesse erzeugen kann, ist es wieder möglich und auch sinnvoll, den Leitfaden auszuweiten, da die einzelnen Felder zu bestimmten Zeitpunkten unter bestimmten Logiken zum Einsatz kommen können.

Durch Trial-and-Error, wird dabei folgender Kodierleitfaden, als zuverlässig und mit hoher Validität auch bei niedrigeren LLM-Modellen eingeschätzt, was im nächsten Abschnitt thematisiert wird. Zum Zeitpunkt dieser Arbeit steht jedoch die vollständige technische und Validitätprüfung noch aus.

Tabelle 14: Erweiterter Kodierleitfaden für die rein deduktive Inhaltsanalyse im Light Assessment

Feld	Zweck
Forschungsfrage	Legt fest, welche inhaltliche Fragestellung mit der Kodierung beantwortet wird.
Materialbasis	Beschreibt, welches Material kodiert wird.
Analyseform	Legt das methodische Vorgehen fest.
Auswertungseinheit	Definiert, welche Materialteile in welcher Ordnung ausgewertet werden.
Kontexteinheit	Definiert den maximalen Kontext zur Interpretation.
Kodiereinheit	Definiert die kleinste Einheit, die kodiert werden darf.
Hauptkategorie-ID	Eindeutige Kennung der Hauptkategorie.
Hauptkategorie-Name	Benennung des übergeordneten analytischen Bereichs.
Subkategorie-ID	Eindeutige Kennung der kodierbaren Unterkategorie.
Subkategorie-Name	Kodierbare Kategorie auf der operativen Ebene.
Definition	Legt den Bedeutungsumfang der Kategorie fest.
Ankerbeispiel	Zeigt eine typische Textstelle für die Kategorie.
Kodierregel	Regelt die Anwendung der Kategorie in Grenzfällen.
Geschlossenheitsregel	Verhindert induktive Kategorienbildung.
Änderungsverbot	Schützt Codebook und Kategorienstruktur.
Erlaubte Code-IDs	Beschränkt die Ausgabe technisch auf gültige Subkategorien.
Einschlusskriterien	Operationalisiert, wann ein Code vergeben wird.
Ausschlusskriterien	Verhindert Überkodierung und Fehlzuordnung.
Gegenbeispiel	Schärft die Abgrenzung zu ähnlichen, aber nicht passenden Textstellen.
Abgrenzung zu Nachbarkategorien	Unterscheidet nahe verwandte Kategorien.
Mehrfachkodierungsregel	Regelt, ob und wann mehrere Subkategorien vergeben werden dürfen.
Prioritätsregel	Legt Vorrang bei konkurrierenden Codes fest.
Konfliktregel	Regelt, was bei widersprüchlichen Segmenten geschieht.
Evidenzpflicht	Macht Kodierung auditierbar.
Begründungspflicht	Verknüpft Kodierentscheidung mit Regelwerk.
Unsicherheitsregel	Markiert Grenzfälle für menschliche Prüfung.
NO_CODE-Regel	Definiert, was bei nicht passendem Material geschieht.
Outputschema	Standardisiert die technische Ausgabe.
Versionierung	Sichert Nachvollziehbarkeit des verwendeten Leitfadens.

10.4. Modellwahl nach Wirtschaftlichkeit und Datensicherheit

Die ursprünglichen Testläufe für das Proof of Concept wurden in direkter Copilot-Umgebung mit GPT Modellen 5.2 und später auf 5.4 durchgeführt. Der Gedanke war, dass ein besseres Modell bessere Logik für die komplexen Aufgaben besitzt. In der Realität müssen aber verschiedene Aspekte berücksichtigt werden. Auf oberflächlicher Ebene sind dies exemplarisch:

- Welche Datensicherheit soll gewährleistet werden?
- Wie viel Tokens verwendet werden können?
- Wie häufig oder gleichzeitig soll eine Aufgabe geschehen?

So muss es im Falle des Light Assessments möglich sein, die Daten auf Schweizer Servern zu halten. Damit ist im M365 Kontext nur noch GPT 4.1 zur Verfügung. Dies wirkt sich komplex der Anweisungen auf das Context Engineering und auf die Menge und Art aus, wie Hintergrundinformationen gegeben werden können. Beim Erstellen eines solchen Tools, ist dies eine Form von Prototyping, da man die Analyse-Logik und den Prozess entsprechend anpassen muss. Zudem müssen beim Harness Engineering die Auswahl und das Zusammenspiel von Tools und die Grösse einzelner Denkprozesse berücksichtigt werden. Für HCD-Fachpersonen, welche in diesem Feld Berührung haben, gilt die Auswahl des Modells nicht als eine rein technische Entscheidung, sondern eine konzeptionelle anhand des Nutzungskontextes. Ein vergleichbarer Gedankengang, der im HCD-Bereich geläufiger ist, wäre die Wahl, ob eine Lösung vorwiegend via Smartphone oder PC genutzt wird und somit auf Mobile-Screen konzipiert wird oder auf Widescreen. An dieser Stelle spielt ein gut angepasster Kodierleitfaden eine wichtige Rolle, um auch mit älteren Modellen gute Ergebnisse zu erzielen.

10.5. Ergebnisverlässlichkeit: Lauf-zu-Lauf-Stabilität

10.5.1. Verlässlichkeitsprüfung der KI-Analyse

Der Prozess wird fünfmal auf demselben Fall ausgeführt, bei identischem Codebuch, identischen Transkripten und gleichen Einstellungen, sodass zwischen den Läufen allein der Nichtdeterminismus von GPT-4.1 variiert. Verglichen werden nur die fünf Durchläufe. Die 111 Codes und fünf befragten Personen spannen die Analyseseinheit auf.

Die Auswertung erfolgt auf zwei Ebenen. Auf der Code-Ebene (Ebene 1) wird je Kombination aus Code und Person binär erfasst, ob mindestens eine Evidenz zugeordnet wurde, woraus ein über alle Läufe identisches Raster aus 555 Einheiten mit je fünf Werten entsteht. Auf der feineren Passagen-Ebene (Ebene 2) ist die einzelne Sprecherpassage die Einheit, identifiziert über Person und Zeitstempel. Sie unterscheidet, ob eine Passage konstant als Evidenz gewählt wird (2a Extraktionsstabilität) und ob sie beim selben Code landet (2b Zuordnungsstabilität).

Als Masse dienen prozentuale Übereinstimmung, Einstimmigkeit (alle fünf Läufe gleich) sowie die zufalls-korrigierten Koeffizienten Fleiss Kappa (Fleiss 1971), Krippendorffs Alpha (Krippendorff 2019) und Gwets AC1 (Gwet 2008). Da die meisten Felder leer bleiben, ergänzt das Prävalenzrobuste AC1 das Kappa und federt dessen Paradoxon ab. Die Einordnung folgt Landis und Koch (1977): ab 0.61 substantiell, ab 0.81 nahezu perfekt.

Reliabilität

Der KI-unterstützte Analyseprozess ordnet Sinnabschnitte deduktiv den vordefinierten Codes zu. Da grosse Sprachmodelle nicht deterministisch arbeiten, prüft die Untersuchung die Reliabilitätskomponente der Ergebnisverlässlichkeit: ob wiederholte Läufe auf identischer Datengrundlage hinreichend ähnliche Resultate liefern. Geprüft wird die Stabilität der Zuordnungen über Wiederholungsläufe, nicht ihre fachliche Richtigkeit.

Validität

Die Validität wurde rein durch menschliche Prüfung sichergestellt. Dabei gilt es anzumerken, dass unter dem Modell GPT4.1, welches getestet wurde, wenn auch eine akzeptable, doch eine verminderte Validität, im Vergleich zu neueren Modellen festgestellt wurde. Dabei kann dieser Mangel auf Probleme zurückgeführt werden, welche sowieso im Zuge der Weiterentwicklung offenstehen oder automatisch mit der verbesserten Reliabilität von gewissen Optimierungen profitieren kann. Ein wichtiger Schritt hierbei ist schon die Anpassung des Kodierleitfadens, wie in Tabelle 14 gezeigt (siehe [Abschnitt 10.3.1](#)).

10.6. Ergebnisse

Auf der Code-Ebene ist die Stabilität hoch. In rund 84 Prozent aller 555 Felder entscheiden alle fünf Läufe identisch, die paarweise Übereinstimmung liegt bei 92 Prozent. Die zufallskorrigierten Masse bestätigen dies mit Werten im Bereich substanziell bis nahezu perfekt. Auf der Passagen-Ebene nimmt die Stabilität ab, und zwar umso stärker, je feiner die betrachtete Entscheidung ist.

Tabelle 15: Lauf-zu-Lauf-Stabilität je Ebene mit zufallskorrigierten Massen (Eigene Darstellung)

Ebene	Gemessene Entscheidung	Leitmasse	Werte	Einordnung
Ebene 1 (Code x Person)	Wird ein Code bei einer Person belegt	Fleiss / AC1	0.84 / 0.86	substanziell bis nahezu perfekt
Ebene 2a (Passage)	Wird eine Passage als Evidenz gewählt	AC1 / Alpha	0.74 / 0.46	substanziell bzw. moderat
Ebene 2b (Passage zu Code)	Landet eine Passage beim selben Code	Jaccard / Alpha	0.75 / 0.34	moderat bzw. ausreichend

Ergänzend zeigt die Aufschlüsselung, wo die Streuung konzentriert ist. Von 186 über alle Läufe zitierten Passagen erscheinen 115 in allen fünf Läufen, 29 hingegen nur in einem einzigen. Inhaltlich bündelt sich die Instabilität im Block Digital Workplace Building Blocks (Fleiss 0.77 gegenüber 0.84 bis 0.88 in den übrigen Blöcken) und dort in Codes mit unscharfer oder überlappender Abgrenzung, etwa zu Kommunikationskanälen, Tool-Landschaft und Use Cases. Ein Teil der Variation ist daher nicht reines Modellrauschen, sondern Ausdruck von Mehrdeutigkeit im Codebuch. Schränkt man die Auswertung auf die Evidenzen einzelner Personen ein, bewegt sich die Lauf-zu-Lauf-Stabilität zwischen Fleiss 0.75 und 0.91. Auch diese Aufschlüsselung vergleicht nicht Personen miteinander, sondern zeigt, für welche Person die fünf Durchläufe stabiler übereinstimmen.

10.6.1. Beurteilung reine Reliabilität

Die Resultate weichen von einer Durchführung zur nächsten nicht grundlegend ab. Das Szenario komplett unterschiedlicher Ergebnisse wird durch die Befunde klar widerlegt. Auf der Ebene, auf der die Assessment-Resultate gelesen und interpretiert werden, nämlich welche Ziele und Gaps pro Person mit Evidenz hinterlegt sind, ist die Übereinstimmung substanziell bis nahezu perfekt. In dieser Hinsicht ist die Ergebnisverlässlichkeit in ihrer Reliabilitätskomponente gegeben und Fachpersonen ohne Methodenhintergrund können darauf vertrauen, dass das inhaltliche Gesamtbild reproduzierbar ist.

Einzuschränken ist die Aussage auf der Detailebene. Welche konkrete Textstelle als Beleg herangezogen wird und welchem einzelnen Code sie zugeordnet wird, ist nur moderat stabil. Für die Nachvollziehbarkeit

der einzelnen Belegstelle bedeutet dies, dass sich die Evidenzbasis im Detail zwischen den Läufen verschieben kann, ohne dass sich das übergeordnete Ergebnis wesentlich ändert.

Insgesamt trägt der Prozess damit ein verlässliches Gesamtergebnis, benötigt auf der Belegebene jedoch eine absichernde Kontrolle. Die Vertrauenswürdigkeit für Anwendende ohne Methodenhintergrund ist auf der Ergebnisebene gegeben, auf der Einzelbeleg-Ebene bedingt.

10.6.2. Schlussfolgerung und Empfehlungen

Die Lauf-zu-Lauf-Stabilität des KI-unterstützten Analyseprozesses ist auf der ergebnisrelevanten Ebene gewährleistet. Die Daten weichen zwischen den Durchführungen nicht wesentlich ab und das analytische Gesamtbild ist reproduzierbar. Die verbleibende Unschärfe betrifft die feinkörnige Zuordnung einzelner Belegstellen. Zur weiteren Absicherung der Ergebnisverlässlichkeit bieten sich drei Massnahmen an:

- Die finale Kodierung kann aus einem Konsens mehrerer Läufe gebildet werden, etwa über einen Mehrheitsentscheid, was die Detailstabilität erhöht.
- Es sollten die unscharf abgegrenzten Codes des Digital Workplace Building Blocks präzisiert oder bewusst als Mehrfachkodierung geführt werden.
- Es sollte die ohnehin vorgesehene menschliche Prüfung gezielt auf die als instabil ausgewiesenen Codes und Passagen fokussiert werden.

10.6.3. Bezug zur Forschungsfrage (Teilfrage 5) und Zielsetzung (Teilziel 5)

Mit der vorstehenden Lauf-zu-Lauf-Stabilitätsprüfung ist die Teilfrage beantwortet, inwieweit der Prozess über unterschiedliche Datensätze hinweg hinreichend verlässliche Ergebnisse liefert, und das Teilziel die Prüfung der Ergebnisverlässlichkeit erreicht. Auf der entscheidungsrelevanten Ebene sind die Resultate reproduzierbar und auch ohne methodische Vorkenntnis belastbar, während die feinere Zuordnung einzelner Belegstellen stärker schwankt und durch die fachliche Kontrolle abgesichert wird. Da die Beurteilung auf einem Assessmentfall beruht, ist die Absicherung über weitere Datensätze als nächster Schritt angelegt.

11. Ergebnisse

11.1. Beantwortung der Forschungsfragen

11.1.1. Die übergreifende Leitfrage

Wie kann ein KI-unterstützter Analyseprozess zur systematischen Auswertung qualitativer und quantitativer empirischer Daten im Rahmen des M365-Evolution-Assessments im Microsoft-365-Ökosystem konzipiert und evaluiert werden, so dass der Analyseaufwand reduziert, die Ergebnisverlässlichkeit gewahrt und die fachliche Verantwortung beim Fachpersonal belassen wird?

Ein solcher Prozess lässt sich realisieren, indem die Auswertung nicht als technische Automatisierung, sondern als menschenzentriertes Arbeitssystem konzipiert wird, das partizipativ mit den M365-Fachpersonen entsteht und die Auswertung in einen schrittweisen Human-in-the-Loop-Ablauf gliedert. Den Analyseaufwand senkt dieser Ablauf, indem er die wiederkehrende Fleissarbeit, also Aufbereitung, Strukturierung und Verdichtung, an die KI delegiert und die Daten durchgängig ohne manuelle Übertragung zwischen Einzelwerkzeugen verarbeitet. Die Ergebnisverlässlichkeit wahrt er, indem jede KI-Aussage an ihre Belegstelle im Ausgangsmaterial gebunden bleibt und eine Prüfung auf Belegbarkeit und Reproduzierbarkeit fest im Ablauf verankert ist. Die fachliche Verantwortung bleibt beim Fachpersonal, weil vor jedem KI-Einsatz ein steuernder Impuls steht, jedes Zwischenergebnis an einem Eingriffspunkt geprüft und freigegeben wird und die deutenden Schritte, namentlich Worttreuekontrolle, Interpretation und Priorisierung, bewusst beim Menschen bleiben.

Evaluiert wird der Prozess, indem drei M365-Fachpersonen ihn eigenständig an echten Kundendaten durchlaufen, während die KI in einem Wizard-of-Oz-Aufbau durch eine Person im Hintergrund simuliert wird. Dieser Aufbau ist bewusst gewählt, weil das Ergebnis ein konzeptioneller Prozessprototyp und keine fertige Systemoberfläche ist. Er bleibt aussagekräftig, weil die ausgegebenen Artefakte aus echten KI-Läufen stammen, sodass die Testpersonen Inhalt und Interaktionsform des geplanten Prozesses real erleben.

11.1.2. Die Teilfragen

Die einzelnen Teilfragen werden in den Ergebnisabschnitten der jeweiligen Kapitel beantwortet. Die Schwachstellen und Anforderungen des bestehenden Prozesses ergeben sich aus der Bedarfsanalyse in Phase 1 (siehe [Abschnitt 6.5](#)), die geeignete Mensch-KI-Aufgabenteilung und die Gestaltung des Soll-Prozesses aus den Phasen 2 und 3 (siehe [Abschnitte 7.6](#) und [8.3](#)), die Verbesserung von Nutzbarkeit und Effizienz aus der Evaluation in Phase 4 (siehe [Abschnitt 9.2](#)) und die Reproduzierbarkeit der Ergebnisse aus der Verlässlichkeitsprüfung im technischen Teil (siehe [Abschnitt 10.6](#)). Die Ableitung von Gestaltungsprinzipien und Vorgehensmodell erfolgt im anschliessenden Transferkapitel (siehe [Kapitel 12](#)).

11.2. Erreichen der Zielsetzung

11.2.1. Das übergeordnete Projektziel

Entwicklung und empirische Evaluation eines KI-unterstützten Analyseprozesses für die Auswertung von M365-Evolution-Assessments im Microsoft-365-Ökosystem, der den Analyseaufwand reduziert und zugleich eine hohe Ergebnisverlässlichkeit sichert, sowie die Ableitung eines Vorgehensmodells zur KI-unterstützten empirischen Datenanalyse im M365-Ökosystem.

Das Ziel ist weitgehend erreicht. Der Analyseprozess liegt konzipiert, prototypisch ausgestaltet und evaluiert vor, die angestrebte Aufwandsreduktion ist im Testing belegt und die fachliche Verantwortung verbleibt strukturell und empfunden beim Fachpersonal. Die Ergebnisverlässlichkeit konnte im Rahmen von technischen Tests ebenfalls bestätigt werden und ein Vorgehensmodell wurde definiert.

Bezüglich der Ableitung eines Vorgehensmodells, Wird jedoch einen Schritt weitergegangen. Auf Grund des Anspruches zum flexiblen Einsatz durch den Auftraggeber entstanden Vorgehen, welche sich nicht auf das M365-Ökosystem beschränken.

Die vorliegende Arbeit geht dabei über eine reine Darstellung der bestehenden Literatur hinaus, indem sie die dort beschriebenen Konzepte auf einen konkreten Anwendungskontext überträgt, ihre Passung kritisch prüft und daraus ein begründetes Vorgehen für die Konzeption und Evaluation des KI-unterstützten Analyseprozesses ableitet.

So entstanden im Zuge der Arbeit unweigerlich zwei Vorgehensmodelle. Zum einen ist in dieser Arbeit Vibe Research so definiert worden, dass es an sich als «Vorgehensmodell zur KI-unterstützten empirischen Datenanalyse» zumindest für HCD-Fachpersonen fungieren kann (Siehe [Abschnitt 4.4](#)). Zum anderen ist das «Vorgehensmodell zur Erstellung eines Assessments für soziotechnische Systeme» entstanden (Siehe [Kapitel 12](#)). Dieses erlaubt es anderen HCD-Fachpersonen, in ähnlichen Situationen ein solches KI-unterstütztes Analyse-Tool gemeinsam mit System-Fachpersonen zu bauen.

11.2.2. Die Teilziele

Die einzelnen Teilziele sind jeweils in den zugehörigen Kapiteln nachgewiesen, die IST- und Bedarfsanalyse in Phase 1 (siehe [Abschnitt 6.5](#)), die iterative Konzeption und prototypische Ausgestaltung in Phase 2 und Phase 3 (siehe [Abschnitte 7.6](#) und [8.3](#)), die Evaluation von Nutzbarkeit und Effizienz in Phase 4 (siehe [Abschnitt 9.2](#)) und die Prüfung der Ergebnisverlässlichkeit im technischen Teil (siehe [Abschnitt 10.6](#)).

11.3. Lernziele

11.3.1. Vertiefung des Verständnisses für die Analyse empirischer Daten unter Einbezug KI-gestützter Auswertungsschritte und deren systematischer Überprüfung.

Dieses Lernziel ist erreicht. Die praktische Durchführung der KI-gestützten Inhaltsanalyse, von der deduktiven und induktiven Kodierung über die Verdichtung bis zur abschliessenden Verlässlichkeitsprüfung, hat das Verständnis für die Möglichkeiten und Grenzen einer solchen Auswertung deutlich vertieft. Besonders die systematische Überprüfung der Ergebnisse hat sich als zentral erwiesen. Die Auseinandersetzung mit Belegbarkeit, Materialbindung und Reproduzierbarkeit sowie die Prüfung der Lauf-zu-Lauf-Stabilität mit Konsistenzmassen wie Fleiss Kappa und Gwets AC1 haben gezeigt, dass eine einzelne KI-gestützte Analyse nur begrenzt belastbar ist und erst durch Wiederholung, Vergleich und fachliche Kontrolle tragfähig wird. Als wesentliche Erkenntnis lässt sich festhalten, dass die Verlässlichkeit einer KI-gestützten

Auswertung weniger vom eingesetzten Modell, als von der Art ihrer methodischen Einbettung abhängt, etwa von der Aufgabengranularität, der Trennung deterministischer und sprachmodellbasierter Verarbeitungsschritte und der durchgängigen Rückführbarkeit jeder Aussage auf das Ausgangsmaterial.

11.3.2. Vertiefung des Verständnisses für den menschenzentrierten Einsatz von KI-Systemen in der Nutzerforschung, insbesondere im Hinblick auf Potenziale, Grenzen und Risiken.

Dieses Lernziel ist erreicht. Die Verbindung von Literaturvertiefung und praktischer Anwendung hat ein differenziertes Verständnis für den menschenzentrierten Einsatz von KI in der Nutzerforschung aufgebaut. Die Potenziale zeigen sich vor allem in der Entlastung bei kodifizierbarer Fleissarbeit, die Kapazität für das Fachurteil schafft. Die Grenzen und Risiken wurden ebenso konkret erfahrbar, namentlich die Anfälligkeit für Halluzinationen, die eingeschränkte Reproduzierbarkeit, die fehlende eigentliche Verstehensleistung der Modelle sowie die Gefahr einer Scheinsicherheit gegenüber plausibel wirkenden Ausgaben. Daraus ist die Einsicht erwachsen, dass KI Fachpersonen unterstützen, aber weder Entscheidungshoheit noch fachliche Verantwortung übernehmen darf. Die menschliche Kontrolle ist deshalb nicht als nachträgliche Prüfung, sondern als fest eingebauter Bestandteil des Prozesses zu verstehen, mit klar definierten Eingriffspunkten und einer bewussten Aufgabenteilung zwischen Fachperson und KI. Zudem hat sich deutlich gezeigt, und damit auch der Vertiefung beigetragen, dass KI genauso in menschenzentrierte Bahnen gelenkt werden muss, wie jede andere Form von interaktivem Produkt oder Prozess auch.

11.3.3. Aufbau praktischer Erfahrung in der Konzeption und prototypischen Ausgestaltung eines Analyseprozesses auf Basis von LLM-Schnittstellen sowie in der Planung, Durchführung und Auswertung entsprechender User-Testings.

Dieses Lernziel ist erreicht. Über den gesamten Projektverlauf ist umfangreiche praktische Erfahrung im Aufbau eines KI-gestützten Analyseprozesses entstanden. In der Konzeption wurde der Soll-Prozess partizipativ mit der Projektgruppe erarbeitet, in der prototypischen Ausgestaltung folgte die Umsetzung auf Basis von LLM-Schnittstellen, einem verbindlichen Steuerelement, deterministischen Skripten und einem M365-Copilot-Agenten. Dabei hat sich gezeigt, dass die Tragfähigkeit eines solchen Prozesses weniger von der reinen Modellnutzung, als von einer sauberen Architektur und Steuerung abhängt. Auch in der Planung, Durchführung und Auswertung von User-Testings ist praktische Erfahrung aufgebaut worden. Weil der zu testende Prozess ein vollständiger, mehrtägiger Arbeitsablauf ist, was bei vergleichbaren Aufgaben und Produkten durchaus nochmals vorkommen kann. Dieser Umstand verlangte von der Evaluation eine bewusste Kombination mehrerer Methoden zu einem asynchronen Wizard-of-Oz-Aufbau, welcher so noch nicht vom Autor durchgeführt wurde. Gerade diese methodische Anpassung an einen ungewöhnlichen Gegenstand zählt zu den wertvollsten praktischen Lernerfahrungen der Arbeit hinsichtlich der gewünschten Vertiefung.

12. Vorgehensmodell zur Erstellung eines Assessments für soziotechnische Systeme

Aus dem durchgeführten Projekt lässt sich ein übertragbares Vorgehensmodell ableiten. Es überführt den einmaligen Projektverlauf, mit dem das Light Assessment entstanden ist, in eine geordnete Abfolge von Schritten, die sich auf vergleichbare Vorhaben anwenden lässt. Das Modell beschreibt damit nicht, was in diesem einen Fall getan wurde, sondern wie eine Klasse ähnlicher Vorhaben systematisch bearbeitet werden kann.

12.1. Geltungsbereich und Voraussetzungen

Das Modell ist für umfangreiche soziotechnische Systeme gedacht, die einen starken Einfluss auf den Arbeitsalltag von Mitarbeitenden haben. Es ermöglicht die menschenzentrierte Gestaltung eines Werkzeugs, mit dem sich ein solches System nutzerorientiert durch System-Fachpersonen einführen oder optimieren lässt. Die sinnvolle Anwendung setzt vier Bedingungen voraus:

- Das System verfügt über einen hohen optionalen Funktionsumfang, eine ausgeprägte Modularität oder ein breites Einsatzgebiet in einer Arbeitsumgebung.
- Es benötigt für seine Einführung oder Optimierung die fachliche Begleitung durch System-Fachpersonen.
- Eine Fachperson der menschenzentrierten Gestaltung (HCD-Fachperson) begleitet diese Fachpersonen durch den Prozess der Werkzeuggestaltung.
- Es steht eine KI-Fachperson als situative Unterstützung bereit, oder das Team (System-Fachpersonen und HCD-Fachperson) verfügt selbst über die nötigen KI-Kenntnisse.

12.2. Aufbau des Modells

Das Vorgehen gliedert sich in sieben aufeinander aufbauende Schritte, von der Analyse des Systems bis zur Definition der auszuliefernden Artefakte. Jeder Schritt verfolgt einen klar umrissenen Zweck und liefert ein definiertes Zwischenergebnis, das die Grundlage des nächsten Schritts bildet.

Tabelle 16: Übersicht der aufbauenden Schritte des Vorgehensmodells zur Gestaltung eines Assessments

Schritt	Zweck
1 Das System analysieren	Den Gegenstand als soziotechnisches System verstehen und für die KI aufbereiten
2 Codebook und Empfehlungskatalog erstellen	Den fachlichen Referenzrahmen für die deduktive Analyse schaffen
3 Standard-Erhebung planen	Eine vergleichbare, auswertbare Datenerhebung vorbereiten
4 Aufbereitung der Daten und Testen	Das Rohmaterial inhaltstreu in eine verlässliche Grundlage überführen
5 Aufbereitung der Daten definieren	Festlegen, wie die Daten analysiert und im Bericht aufbereitet werden
6 Technischer Test der Inhaltsanalyse	Die Reproduzierbarkeit der KI-Analyse prüfen
7 Zielartefakte definieren	Aus dem Bericht das kundengerechte Lieferobjekt ableiten

Durchgehendes Prinzip ist die Aufgabenteilung zwischen Mensch und KI. Gemäss Vibe Research (Siehe [Abschnitt 4.4](#)) Die KI übernimmt kodifizierbare, am Material überprüfbare Fleissarbeit, während Bedeutungszuschreibung, fachliche Bewertung und Verantwortung beim Fachpersonal verbleiben. Die einzelnen Schritte konkretisieren diese Teilung.

12.3. Schritt 1: Das System analysieren

Der erste Schritt erschliesst den Untersuchungsgegenstand als soziotechnisches System, bestimmt seine Soll-Zustände und bereitet das Systemverständnis für die KI auf.

Systemumfang und Soll-Zustände

Den Ausgangspunkt bildet eine Auslegeordnung. Das System wird in seine Funktionsbereiche zerlegt und je Bereich werden die angestrebten Soll-Zustände bestimmt. Der Fokus liegt bewusst auf der soziotechnischen Perspektive, also auf der tatsächlichen Nutzung, den Aufgaben und den organisationalen Bedingungen, und nicht allein auf dem technischen Funktionsumfang. Im zugrunde liegenden Fall bildet das dreistufige M365-Evolution-Modell diesen Bezugsrahmen, das danach fragt, wie Mitarbeitende die digitale Umgebung tatsächlich nutzen. Entscheidend ist, bereits hier in nutzungsnahen Bausteinen zu denken und von Beginn an trennscharf zu schneiden, da sich überladene Bausteine später als problematisch erweisen und die Stabilität der KI-gestützten Kodierung schwächen.

Kontext-Dokument

Aus diesem Systemverständnis entsteht ein strukturiertes Kontextdokument, das System, Fachbegriffe, Zielbild und Rahmenbedingungen für die KI beschreibt. Es bildet den kuratierten Bezugsraum, auf den die KI in den späteren Analyseschritten zurückgreift, und setzt damit das Prinzip des Context Engineering um, also die bewusste und aufgabenbezogene Bereitstellung des relevanten Materials (Mei et al. 2025). In das Dokument gehören nur relevante, fachlich tragfähige und am Material überprüfbare Inhalte. Die Qualität dieses Dokuments bestimmt die Stabilität der KI-Ergebnisse massgeblich und ermöglicht brauchbare Resultate auch mit älteren oder besonders datensicheren Modellen.

12.4. Schritt 2: Codebook und Empfehlungskatalog erstellen

Der zweite Schritt schafft den fachlichen Referenzrahmen, an dem die KI deduktiv kodiert und aus dem sich Handlungsempfehlungen ableiten lassen.

Target-Gap-Liste

Je Baustein werden die Zielzustände und die möglichen Mangelzustände einander gegenübergestellt. Diese Ziel-Lücken-Paare bilden die Grundlage der späteren Codes. Im Fall wurde die Bausteinstruktur des Systems in eine solche Kodierungslogik überführt, sodass die Auswertung den Bezug zum Zielbild und zum Reifegrad explizit herstellt. Über die Trennschärfe dieser Paare entscheidet sich die spätere Verlässlichkeit. Wenige, klar abgegrenzte Paare sind wertvoller als viele überlappende, und die Erarbeitung erfolgt am besten gemeinsam mit den Fachpersonen.

Codebook

Zu jedem Ziel-Lücken-Paar entsteht ein Codebook-Eintrag mit eindeutiger Kennung, Bezeichnung, Definition und kontextueller Erklärung für die KI. Das Codebook wird strikt deduktiv angewendet. Das Sprachmodell wirkt dabei nicht als exploratives Analyseinstrument, sondern als regelgebundene Kodierinstanz, die nur vorgegebene Kennungen verwendet, keine neuen Kategorien bildet und nicht eindeutig zuordenbare Stellen als prüfbedürftig markiert. In der Praxis empfiehlt es sich, mit einem knappen Leitfaden zu beginnen und ihn iterativ zu schärfen. Zu viele Codes auf einmal überfordern das Modell, das dann vorzeitig keine Evidenz mehr findet. Mit zunehmender technischer Stabilität lässt sich der Leitfaden um Felder wie Ankerbeispiel, Kodierregel und Ausschlusskriterien erweitern.

Empfehlungskatalog

Zu den Lücken werden konkrete Massnahmen formuliert und auf die Codes abgebildet, sodass aus einem Befund unmittelbar Handlungsoptionen ableitbar sind. Die Auswahl und Priorisierung dieser Empfehlungen bleibt jedoch ein kundenspezifisches Fachurteil und wird nicht an die KI delegiert.

12.5. Schritt 3: Standard-Erhebung planen

Der dritte Schritt standardisiert die Datenerhebung so, dass vergleichbares und auswertbares Material entsteht und die Erhebung auf die Codes ausgerichtet ist.

Casting-Liste

Vorab wird bestimmt, welche Mitarbeitentypen einbezogen werden, und zwar bewusst heterogen nach Expertise, Erfahrung und Einstellung, nicht nur Enthusiasten und Leistungsträger. Der Umfang orientiert sich an der inhaltlichen Sättigung statt an einer fixen Teilnehmerzahl. Bei einer kleinen, gezielt gewählten Gruppe ist jede einzelne Aussage relevant. Eine heterogene Auswahl verhindert ein verzerrtes Bild, und die Termine sind verbindlich abzusichern.

Interview-Leitfaden

Für etwa einstündige Gespräche entsteht ein halbstrukturierter Leitfaden mit offenen Fragen und konkretisierenden Unterfragen, dramaturgisch geführt von der eigenen Rolle und dem Ist-Prozess über die Schmerzpunkte zu den Soll-Vorstellungen. Die Orientierung an einem etablierten Vorgehen für praxisnahe Industry-Settings stützt diese Anlage (Baxter et al. 2015). Der Leitfaden bleibt verständlich und auf das Wesentliche begrenzt, und Interview- wie Survey fragen werden konsistent auf die Bausteine und Codes ausgerichtet, damit die qualitative und die quantitative Sicht zusammenpassen.

Survey-Fragen

Aus dem Leitfaden werden quantifizierbare Fragen abgeleitet, die Tendenzen messbar machen und die qualitative Sicht ergänzen. Die Umfrageitems ergänzen später die Codes, jeweils mit einer Einordnung, wie eng ein Item zur Code-Definition passt. Zu beachten ist, dass Standardvisualisierungen ohne fachliche Einordnung Fehlinterpretationen begünstigen. Einzelne Kennzahlen werden ohne Erläuterung leicht als schwer verständlich erlebt, weshalb die Items eng an die Codes zu koppeln sind.

Testdaten

Für die Kalibrierung der späteren Schritte werden mindestens ein bis zwei echte Interviews und eine kleine echte Umfrage durchgeführt, statt Daten künstlich generieren zu lassen. Nur reale Daten enthalten die typischen Transkriptionsfehler und kontextuellen Eigenheiten, an denen sich Bereinigung und Analyse überhaupt kalibrieren lassen. Künstliche Daten bilden diese Stolpersteine nicht ab.

12.6. Schritt 4: Aufbereitung der Daten und Testen

Der vierte Schritt überführt das Rohmaterial inhaltstreu in eine verlässliche, analysierbare Grundlage und kalibriert die Bereinigung am realen Material.

Transkriptions-Tool und Bereinigung

Das gewählte Transkriptionswerkzeug wird getestet, seine typischen Fehler und Eigenheiten werden identifiziert, und die Bereinigungsanweisungen werden darauf ausgerichtet. Maschinelle Transkripte sind erwartungsgemäss fehlerbehaftet, weshalb die sprachliche Glättung als eigener, dokumentierter Schritt geführt wird. Wichtig ist die Reihenfolge. Zuerst erfolgt die inhaltliche Kontrolle, danach die Glättung, denn glättet die KI zuerst, sind einzelne Fehler bereits verarbeitet und schwerer zu korrigieren.

Kontext-Bereinigung

Die System-Fachpersonen erstellen eine Liste schwer erkennbarer oder spezifischer Begriffe, Akronyme und kontextueller Eigenheiten, die als Erkennungshilfe für die Bereinigung dient. Ein Sprachmodell erkennt kontextabhängige oder akronymbezogene Fehler nicht zuverlässig, solche Fehler führen aber zu falschen analytischen Schlüssen. Die inhaltliche Worttreuekontrolle bleibt deshalb bewusst beim Menschen, möglichst bei der interviewenden Person, die den Gesprächsverlauf kennt. An derselben Stelle werden auch datenschutzrelevante Hinweise auf Personen oder Abteilungen manuell entfernt.

Testen und Anpassen

Die Bereinigung wird mehrfach durchgeführt und die Ergebnisse werden verglichen bis die Anweisungen über das Material hinweg stabil inhaltstreue Resultate liefern. Der eigentliche Aufwand liegt nicht im Ausführen, sondern im stabilen Aufbau der Anweisung. Der Teilprozess wird früh mit der Projektgruppe an realen Aufzeichnungen getestet, die Worttreue gegen die Aufnahme kontrolliert und die Bereinigungsregeln werden gemeinsam festgelegt.

12.7. Schritt 5: Aufbereitung der Daten definieren

Der fünfte Schritt legt fest, wie die qualitativen und quantitativen Daten analysiert und in einen prüfbaren Analysebericht überführt werden.

Analysebericht-Struktur

Die Zielstruktur des Berichts wird grob festgelegt, mit den qualitativen und quantitativen Evidenzen je Code, einem Befundtext, einer Kurzform und offenen Feldern für die Priorisierung. Quantitatives wird tabellarisch, Qualitatives als Liste mit Personenkennzeichnung dargestellt. Die Struktur ist so anzulegen, dass die Fachperson kontrollieren und priorisieren kann und jede Aussage auf ihre Belegstelle rückführbar bleibt. Da der Bericht leicht zu umfangreich wird, ist früh auf Fokussierung zu achten.

Analyseprozess definieren

Die KI-Anweisungen für die deduktive Inhaltsanalyse werden festgelegt, ausgerichtet auf das Codebook und die Berichtsstruktur. Methodisch lehnt sich das Vorgehen an die qualitative Inhaltsanalyse an (Mayring 2022), bleibt aber bewusst auf den analytischen Kern reduziert. Generative und deterministische Schritte werden getrennt, und eine feste Lauf-Disziplin wird definiert, etwa getrennte Durchläufe für Analyse, Validierung und Kurzform. Das Sprachmodell verdichtet dabei Bedeutung, während deterministische Skripte Reihenfolge, Format und Vollständigkeit sichern. Die Verlässlichkeit wird so nicht vom Modell erwartet, sondern von einer prüfbaren Architektur. Aufgaben werden klein geschnitten und das Suchen wird vom Prüfen getrennt.

Testen und Kalibrieren

Die System-Fachpersonen prüfen die Datenqualität. Die Inhaltsanalyse wird am Testmaterial wiederholt und die KI-Anweisungen werden angepasst, bis die gewünschte Validität erreicht ist. Die Validität wird dabei durch menschliche Prüfung sichergestellt, die Kalibrierung verläuft iterativ und gemeinsam mit den Fachpersonen. Eine bewusst zu umfangreich gehaltene Kodierung macht Probleme früh sichtbar und lässt die Gruppe die Notwendigkeit der Reduktion selbst erkennen. Es gilt der Grundsatz, dass Validität vor Reliabilität geht.

12.8. Schritt 6: Technischer Test der Inhaltsanalyse

Der sechste Schritt prüft die Reproduzierbarkeit der KI-Analyse, bevor der Prozess als verlässlich gilt.

Reliabilitätsprüfung

Der Prozess wird mehrfach, etwa fünfmal, auf demselben Fall mit identischem Codebook und Material ausgeführt, und die Stabilität der Zuordnungen wird mit Konsistenzmassen geprüft. Da Sprachmodelle

nicht deterministisch arbeiten, ist diese Prüfung notwendig. Im Fall wurden Fleiss Kappa, Krippendorffs Alpha und Gwets AC1 herangezogen und nach einer etablierten Skala eingeordnet (Landis und Koch 1977). Auf der entscheidungsrelevanten Ebene, also welche Ziele und Lücken pro Person belegt sind, ist die Übereinstimmung substanziell bis nahezu perfekt, während die Zuordnung einzelner Belegstellen stärker schwankt. Die Instabilität konzentriert sich auf unscharf abgegrenzte Codes. Dort ist nachzuschärfen, oder die finale Kodierung wird aus dem Konsens mehrerer Durchläufe gebildet.

12.9. Schritt 7: Zielartefakte definieren

Der siebte Schritt leitet aus dem Analysebericht das kundengerechte Lieferobjekt ab und gestaltet den Übergang von der Analyse zur Auslieferung.

Aufbereitung des Analyseberichts

Es wird festgelegt, wie die Berichtsergebnisse priorisiert, eingeschränkt und auf den konkreten Fall zugeschnitten werden, einschliesslich einer schlankeren Variante je nach Kundengrösse. Die Priorisierung je Ziel und Lücke bleibt ein menschliches Fachurteil und ist der KI bewusst entzogen. Sie nicht zu delegieren ist auch deshalb wichtig, weil Nutzende sonst leicht in eine Subordinationsrolle gegenüber dem System geraten und Vorschläge unkritisch übernehmen (Matter et al. 2024). Zugleich ist klar zu erklären, was priorisieren bedeutet, da sonst vorsichtshalber zu viel als hoch relevant eingestuft wird.

Lieferobjekt

Aus dem priorisierten Bericht wird das Lieferobjekt abgeleitet, etwa ein verdichtetes Summary als Kundenbericht und eine darauf aufbauende Präsentation, ohne dass Rohdaten weitergegeben werden. Die KI erstellt das Summary, während Finalisierung und Präsentation bewusst manuell beim Fachpersonal bleiben. Das Lieferobjekt bleibt fokussiert und wertneutral und die Rohdaten werden geschützt. Die visuelle Aufbereitung und der Übergang zur Präsentation sind ein bewusst offen gehaltener nächster Schritt.

12.10. Einordnung

Die sieben Schritte führen von der Analyse des Systems bis zum fertigen Lieferobjekt und machen den Weg nachvollziehbar, auf dem aus einem soziotechnischen System ein KI-unterstütztes Analysewerkzeug entsteht. Ihre Reihenfolge ist nicht streng linear zu verstehen, da insbesondere die Schritte vier bis sechs iterativ und am realen Material kalibriert werden. Tragend bleibt über alle Schritte hinweg die bewusste Aufgabenteilung, bei der die KI die überprüfbare Fleissarbeit übernimmt und das Fachpersonal die Deutung, die Bewertung und die Verantwortung behält.

13. Leistungsreflexion

13.1. Eigene Reflexion

13.1.1. Persönliche Meinung

Insgesamt habe ich erreicht, was ich mir vorgenommen hatte. Die übergreifende Forschungsfrage ist beantwortet und das Forschungsziel mit seinen Teilzielen eingelöst. Dass dies trotz der besonderen Rahmenbedingungen gelungen ist, werte ich als Hinweis darauf, dass die methodischen Grundentscheidungen tragfähig waren. Zugleich bin ich mir bewusst, dass die Flexibilität, welche ich mir zunutze gemacht habe, an einzelnen Stellen zu Umwegen geführt hat, etwa bei den frühen und später verworfenen KI-Experimenten, deren eigentlicher Ertrag weniger im Resultat als im geschärften Verständnis des Verfahrens lag. Doch am Ende zählt für mich das Ergebnis und damit bin ich sehr zufrieden.

13.1.2. Einschätzung der eigenen Leistung

Im Verhältnis zur ursprünglichen Zielsetzung bewerte ich die Arbeit als weitgehend erfolgreich. Aus der Analyse des bestehenden, aufwendigen Assessmentprozesses ist ein KI-unterstützter Soll-Prozess entstanden, der konzipiert, prototypisch ausgestaltet und entlang von Nutzbarkeit, Effizienz und Ergebnisverlässlichkeit evaluiert wurde. Aus den Befunden lässt sich zudem ein übertragbares Vorgehensmodell ableiten. Damit liegen alle Bausteine vor, die das Forschungsziel vorgesehen hatte. Als wesentliche Erfolgsfaktoren sehe ich den durchgängigen Praxisbezug, die partizipative Konzeption mit den M365-Consultants und die konsequente Aufgabenteilung zwischen Mensch und KI. Besonders trägt rückblickend die Entscheidung, Verlässlichkeit nicht vom Sprachmodell, sondern von einer prüfbaren Prozessarchitektur mit deterministischen Skripten und klar gesetzten Eingriffspunkten zu erwarten.

Nicht alle Teile sind zu meiner vollen Zufriedenheit ausgefallen. Auch wenn die Inhaltsanalyse des Light Assessments an Daten von vergangenen Evolution Assessments getestet wurde, stützt sich die tatsächliche Prüfung der Ergebnisverlässlichkeit bislang auf einen einzelnen Assessmentfall. In Verbindung mit den zu erwartenden Anpassungen zur Erhöhung der Validität, so wie die Zusammenstellung unterschiedlicher KI-Modelle zur optimierten Nutzung, muss noch mehrmals ein Reliabilitätstest durchgeführt werden. Denn in Bezug auf die Validität, auch wenn die Ergebnisse nicht grundsätzlich falsch sind, schwankt auf der Ebene der einzelnen Belegstellen die KI-gestützte Zuordnung im Rahmen der engeren technischen Vorgaben stärker als von mir erwünscht. Doch das, so wurde mir von KI-Fachpersonen versichert, gehört zum Prozess und muss schlichtweg kalibriert werden. Abseits dieser technischen Aspekte bleibt der Prototyp zum Ende dieser Arbeit, bewusst ein Proof of Concept ohne produktive Implementierung oder vollständige technische Umsetzung.

Die erhobenen Effizienzwerte beruhen auf wenigen Testdurchläufen und haben eine grössere Spannweite als erwünscht, doch befinden sich im plausiblen Rahmen. Einige der Grenzen sind teils dem Umfang einer Masterarbeit, teils dem frühen Entwicklungsstand des Gegenstands und teils schlichtweg der Konfrontation mit dem neuen Themenfeld geschuldet. Sie schmälern die Aussagekraft der Ergebnisse vor allem dort, wo es um Verallgemeinerung und damit um die allgemeine Anwendbarkeit geht, nicht aber im untersuchten Anwendungsfall.

Besonders hervorzuheben ist für mich in dieser Arbeit der pragmatisch geprägte Arbeitsmodus, in welchem die Methoden als das verstanden wurden, was sie im Kern sind, nämlich Mittel zum Zweck. Hier empfinde ich eine persönliche, aber auch fachliche Zufriedenheit, besonders in Bezug auf die Definition und das Methodisieren des Vibe Research. Eine eigene systematische methodische Fundierung nachzuforschen, zu formulieren, zu begründen und erfolgreich einzusetzen, ist im Rahmen einer solchen Arbeit mutig, doch hat sich rückblickend gelohnt und wird auch über diese Arbeit hinaus weiter von mir angepasst und bearbeitet.

13.2. Beiträge und Mehrwert

Die Arbeit leistet einen Beitrag auf drei Ebenen, einen wissenschaftlichen, einen fachtechnischen und einen mehrwertgenerierenden.

Wissenschaftlich

Die zum Zeitpunkt der Arbeit noch nicht wissenschaftlich etablierte Idee des Vibe Research nach Bertina (2025) wird als Begriff aufgenommen und definiert. In Verbindung mit ausgewählten Elementen der qualitativen Inhaltsanalyse nach Mayring (2022) wird daraus ein reguliertes Vorgehensmodell für eine KI-unterstützte Analyse von empirischen Daten.

Mehrwertgenerierend

Das Vibe Research Vorgehensmodell wird angewandt und überprüft, dann für den konkreten Arbeits- und Analysekontext iterativ adaptiert und bildet eine Grundlage für die Entwicklung eines KI-

unterstützten Analyseprozesses für die Auswertung von M365-Evolution-Assessments im Microsoft-365-Ökosystem. Im konkreten Projektkontext bestätigt der Auftraggeber den praktischen Nutzen des entwickelten Light Assessments für die eigene Beratungspraxis.

Fachtechnisch

Der entwickelte Prozess wird zusätzlich in ein thematisch neutrales Vorgehensmodell überführt, das HCD-Fachpersonen ermöglichen soll, vergleichbare Assessments für andere soziotechnische Systeme zu gestalten.

Der Beitrag der Arbeit liegt damit nicht nur in der begrifflichen Einordnung eines neuen Ansatzes, sondern auch in dessen praxisbezogener Überführung in ein konkret eingesetztes Werkzeug und in ein für vergleichbare Fragestellungen anschlussfähiges Vorgehensmodell.

13.3. Weitere Einschätzungen

Folgende Statements sind von Seitens GARAIO zum Light Assessment abgegeben worden.

Auftraggeber

Aus Sicht des Auftraggebers fällt die Beurteilung der Arbeit durchwegs positiv aus. Das entwickelte Light Assessment entspricht den Erwartungen, die GARAIO an eine schlanke und wirtschaftlich tragfähige Weiterentwicklung des bestehenden Assessments gestellt hat. Besonders geschätzt wird die Flexibilität des Prozesses, der sich an unterschiedliche Projektgrößen und Kundensituationen anpassen lässt, ohne die fachliche Qualität der Beurteilung aufzugeben. Vorgesehen ist, das Light Assessment künftig regelmässig in Kundenprojekten einzusetzen und auf dieser Grundlage iterativ weiterzuentwickeln.

Darüber hinaus wird geprüft, den Ansatz so weit zu einem Produkt auszugestalten, dass er in Zusammenarbeit mit Microsoft zu einem firmenintern einsetzbaren Werkzeug für die fortlaufende Standortbestimmung bei grösseren Unternehmen und Grossunternehmen ausgebaut werden kann. Insgesamt sieht der Auftraggeber in der Arbeit eine unmittelbar verwertbare Grundlage für die weitere Entwicklung seines Assessmentangebots.

(Von Adrian Rieder, EX-Circle Lead)

Interner KI-Spezialist

Ich habe das Konzept und den Prototyp aus fachlicher Sicht durchgesehen und halte den Ansatz für schlüssig und tragfähig. Die Übertragung von Mayrings deduktiver qualitativer Inhaltsanalyse in eine mehrstufige Agenten-Pipeline ist methodisch sauber gedacht: Jeder Schritt hat eine klar umrissene Aufgabe, die Aufteilung folgt einer nachvollziehbaren analytischen Logik, und Designentscheidungen wie der deterministische Merge und die Validierungsschleife zeugen von echtem technischem Verständnis. Auch die Ergebnisse überzeugen: Gegen den manuell erstellten Gold Standard erreicht die Pipeline durchweg belastbare Werte. Das Konzept ist plausibel, der Prototyp belegt überzeugend seine Umsetzbarkeit, und die Arbeit zeigt, dass hier ein durchdachtes, praxisnahes System entstanden ist. Eine rundum gelungene und fachlich fundierte Umsetzung.

(Von Alex Chilton, SW Fullstack Engineer.)

14. Persönliche Reflexion

Diese Arbeit ist in einer ungewöhnlichen Nähe zum eigenen Gegenstand entstanden. Formal stand mir mit Adrian Rieder ein Praxispartner zur Seite, faktisch war ich für das Assessment, das hier weiterentwickelt wird, selbst mitverantwortlich und bis zu einem gewissen Grad auch mein eigener Auftraggeber. Dies verschaffte mir Zugang, Daten- und Kontextverständnis und Entscheidungsfreiheit, verlangte aber, dass ich eine Distanz zu meinem eigenen Werk herstellte, die in einer fremden Umgebung von selbst entstanden wäre und eigentlich für viele Methoden von Vorteil ist. Manches, was im Ergebnis nach methodischer Disziplin aussieht, war zuerst der schlichte Versuch, mir selbst nicht zu sehr zu vertrauen und mit gewissen Schritten eine Art künstliches Kontrollorgan zu schaffen.

Hierbei spielte nämlich der Punkt mit, dass ich das Projekt in einem interdisziplinären Design-Kontext als Einzelperson bearbeitet habe, in welchem vorwiegende Gruppenarbeiten vorgesehen sind. Dabei stellte das Projekt selbst für mich kein wirkliches Problem dar, denn im Grunde hatte ich eine Gruppe und interdisziplinären Austausch, sowohl im Rahmen des partizipativen Arbeitens als auch im informellen Austausch mit Arbeitskollegen und Fachpersonen. Das grösste Problem und jenes, welches ich uneingeschränkt als Nachteil sehe, wenn man eine solche Arbeit allein schreibt, ist der schriftliche Teil. Mehr Augen sehen mehr und oftmals fehlte mir schlichtweg eine andere qualifizierte Gegenmeinung oder Austausch, wenn es darum ging, wie viel etwas noch an Information braucht und wann es reicht oder wie nun etwas formuliert, gegliedert oder beschrieben gehört. Hier habe ich mir fast schon verzweifelt versucht aus allen Ecken Meinungen zusammenzusuchen und es hat mich definitiv am meisten verunsichert.

Rückblickend war für mich, wie schon in den beiden CAS, die Arbeit weniger das Erlernen völlig neuer Dinge, als ein Überarbeiten und vor allem Hinterfragen dessen, was ich seit Jahren aus der Praxis kenne. Insofern habe ich hier durch das allein Arbeiten am meisten persönlich profitieren können, da ich damit auch mehr gezwungen war, mich selbst mehr beim Arbeiten reflektieren zu müssen. Ich bewege mich seit 2018 nach einem CAS an der FHNW explizit im menschenzentrierten Umfeld und hatte mir vieles in der Praxis unter erfahrenen UX-Fachpersonen mit eher pragmatischen Ansätzen angeeignet. Erst im MAS und besonders nun in dieser Arbeit war ich gezwungen, dieser natürlich gewachsenen Arbeitsweise mehr Begriffe und Begründungen zu geben. Denn diesmal ist es nicht «egal», wie gearbeitet wird, solange die Daten und Ergebnisse stimmen. Diesmal ist «genau das» was zählt. Dass die partizipative Zusammenarbeit, der das ursprüngliche Assessment seinen Erfolg verdankt, einen eigenen Methodennamen trägt, ist mir erst im Verlauf dieser Arbeit bewusst geworden, und ähnlich erging es mir mit dem KI-gestützten Arbeiten, das ich längst aus pragmatischem Ausprobieren heraus praktizierte, bevor ich es als ein aktuelles, benanntes Themenfeld fassen konnte. Diese Verschiebung vom Können zum «begründen müssen» war ungewohnt und zeitweise unbequem, weil sie auch die Lücken in meinen eigenen Selbstverständlichkeiten offenlegte. Genau sie ist aber das, was ich persönlich am stärksten aus der Arbeit mitnehme. Auffällig ist für mich, wie oft ich mich in diesem Projekt mit Absicht in Situationen der Unsicherheit begeben habe. So habe ich mich auch an mehreren Stellen angreifbar gemacht.

Beispielsweise in dem ich einen unscharfen Begriff zur methodischen Grundlage ausgearbeitet habe. Oder in einem Gestaltungsprozess Ideen umgesetzt habe, von denen ich wusste, dass sie später noch Probleme verursachen konnten. Das, weil ich die Integrität der Mitbestimmung von Nutzenden wahren wollte. Diese wiederkehrende Unsicherheit habe ich gewählt, weil sie mir die meiste Entwicklung geboten hat. Heute halte ich sie für den eigentlichen Arbeitsmodus dieser Arbeit, denn sie hat mich jedes Mal zu mehr Sorgfalt und zu einer ehrlicheren Einordnung gezwungen, als wenn ich sie in vertrautem Gelände gewählt hätte. Diese Entscheidung hat mir gutgetan.

Verändert hat sich durch die Arbeit auch mein Verhältnis zur künstlichen Intelligenz selbst. Ursprünglich sehr skeptisch, bin ich der Technik gegenüber grundsätzlich aufgeschlossener und nutze sie im Alltag selbstverständlicher und gezielter für Fleissarbeiten. Aus einem Werkzeug, das ich gern und schnell

einsetze, ist im Lauf der Arbeit etwas geworden, dem ich bewusster und kontrollierter begegne und in dem ich Chancen für die Weiterentwicklung meines Berufsfeldes sehe, sofern sie, wie jede andere neue Technologie jemals, für die Effizienz, Effektivität und Zufriedenstellung des Nutzenden eingesetzt wird. Das ist vielleicht die nüchternste, aber haltbarste persönliche Veränderung, die ich diesem Projekt verdanke.

Fazit

Betrachte ich die Arbeit als Ganzes, bleibt für mich mehr als ein abgeschlossenes Studienprojekt. Der Prozess, den ich entworfen habe, wird in der Praxis weitergeführt und auch das Vorgehen, das ich für mich benannt habe, werde ich über diese Arbeit hinaus weiter hinterfragen und anpassen. Dass eine Auseinandersetzung, die mich phasenweise verunsichert und zeitweise belastet hat, am Ende nicht in Erschöpfung, sondern in Überzeugung mündete, hätte ich zu Beginn nicht erwartet. Was ich persönlich mitnehme, ist vor allem die Erfahrung, dass es sich lohnt, das eigene Tun auch dann zu hinterfragen, wenn es sich längst bewährt hat, und dass Kontrolle, Zurückhaltung und Offenheit beim Einsatz neuer Technik kein Gegensatz zum Fortschritt sind, sondern seine Bedingung.

15. Literatur- und Quellenverzeichnis

Die Referenzierung folgt dem ACM-Author-Date-Stil gemäss den OST-Vorgaben.

AMERSHI, S., WELD, D., VORVOREANU, M., FOURNEY, A., NUSHI, B., COLLISSON, P., SUH, J., IQBAL, S., BENNETT, P. N., INKPEN, K., TEEVAN, J., KIKIN-GIL, R., UND HORVITZ, E. 2019. Guidelines for Human-AI Interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI 2019), Glasgow, Scotland UK. ACM, New York, NY, USA.

ANTHROPIC. 2025. Effective Context Engineering for AI Agents. Anthropic, Engineering Blog. Abgerufen am Dezember 2025 von <https://www.anthropic.com/engineering/effective-context-engineering-for-ai-agents>. Geprüft und Archiviert am 29.06.26: [<https://web.archive.org/web/20260629194734/https://www.anthropic.com/engineering/effective-context-engineering-for-ai-agents>].

BAEK, J., JAUHAR, S. K., CUCERZAN, S., UND HWANG, S. J. 2024. ResearchAgent: Iterative Research Idea Generation over Scientific Literature with Large Language Models. arXiv Preprint arXiv:2404.07738. KAIST und Microsoft Research. Hinweis: arXiv-Preprint, nicht peer-reviewt.

BAXTER, K., COURAGE, C., UND CAINE, K. 2015. Understanding Your Users: A Practical Guide to User Research Methods. 2. Auflage, Morgan Kaufmann (Elsevier), Waltham, MA, USA.

BENDER, E. M., GEBRU, T., MCMILLAN-MAJOR, A., UND SHMITCHELL, S. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT 2021), 610-623. ACM, New York, NY, USA.

BERTINA, A. 2025. Vibe-Research instead of Deep-Research: AI Agents and the Future of Scientific Discovery. Medium, 25. Mai 2025. Zuletzt Abgerufen am Dezember 2025 von <https://medium.com/@abertina/vibe-research-instead-of-deep-research-ai-agents-and-the-future-of-scientific-discovery-4d561248f3e2> Geprüft und Archiviert am 29.06.26: [<https://web.archive.org/web/20260629200257/https://medium.com/@abertina/vibe-research-instead-of-deep-research-ai-agents-and-the-future-of-scientific-discovery-4d561248f3e2>] Hinweis: populärwissenschaftlicher Online-Beitrag, nicht peer-reviewt.

BOMMASANI, R. ET AL. 2021. On the Opportunities and Risks of Foundation Models. Center for Research on Foundation Models (CRFM), Stanford University. arXiv:2108.07258.

BOURNE, L. UND WALKER, D. H. T. 2005. Visualising and mapping stakeholder influence. Management Decision 43, 5, 649-660.

CHEN, L., ZAHARIA, M., UND ZOU, J. 2023. How Is ChatGPT's Behavior Changing over Time? Stanford University und UC Berkeley. arXiv:2307.09009.

COHEN, J. 1960. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 20, 1, 37-46.

CUI, J. UND ALEXANDER, R. 2026. Same Prompt, Different Outcomes: Evaluating the Reproducibility of Data Analysis by LLMs. Working Paper, University of Toronto.

DEUTSCHES INSTITUT FÜR NORMUNG (DIN). 2020. DIN EN ISO 9241-210:2020-03 Ergonomie der Mensch-System-Interaktion. Teil 210: Menschzentrierte Gestaltung interaktiver Systeme (ISO 9241-210:2019). Deutsche Fassung EN ISO 9241-210:2019. Beuth Verlag, Berlin.

- FEDOSOV, A., KERN, D., UND BARTH, C.-M. 2026. LLM-powered Workflow to Support Qualitative Data Analysis: Lessons Learned from an Applied Research Project. Workshop-Beitrag, AlpCHI 2026. FHNW University of Applied Sciences and Arts Northwestern Switzerland und Universität Zürich.
- FLEISS, J. L. 1971. Measuring Nominal Scale Agreement Among Many Raters. *Psychological Bulletin* 76, 5, 378-382.
- GAETE CRUZ, M., ERSOY, A., CZISCHKE, D., UND VAN BUEREN, E. 2022. Towards a framework for urban landscape co-design: Linking the participation ladder and the design cycle. *CoDesign* 19, 3, 233-252.
- GARRETT, J. J. 2011. *Die Elemente der User Experience: Anwender-zentriertes (Web)Design*. Addison Wesley, München. Deutsche Ausgabe, Pearson Education Deutschland.
- GWET, K. L. 2008. Computing Inter-Rater Reliability and Its Variance in the Presence of High Agreement. *British Journal of Mathematical and Statistical Psychology* 61, 1, 29-48.
- HONG, S. UND PARK, W. 2025. Developing user-centered system design guidelines for explainable AI: a systematic literature review. *Artificial Intelligence Review* 58, 386.
- HUANG, L., YU, W., MA, W., ZHONG, W., FENG, Z., WANG, H., CHEN, Q., PENG, W., FENG, X., QIN, B., UND LIU, T. 2024. A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions. *ACM Transactions on Information Systems*.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). 2018. *ISO 31000:2018 Risk management – Guidelines*. 2. Auflage, ISO, Genf.
- KALAI, A. T., NACHUM, O., VEMPALA, S. S., UND ZHANG, E. 2025. Why Language Models Hallucinate. *arXiv Preprint arXiv:2509.04664*. OpenAI und Georgia Tech. Hinweis: arXiv-Preprint, nicht peer-reviewt.
- KAPLAN, K. 2020. *ResearchOps 101*. Nielsen Norman Group, 16. August 2020. Abgerufen am Juli 2026 von <https://www.nngroup.com/articles/research-ops-101/>.
- KRIPPENDORFF, K. 2019. *Content Analysis: An Introduction to Its Methodology*. 4. Auflage, SAGE Publications, Los Angeles und London.
- LANDIS, J. R. UND KOCH, G. G. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1, 159-174.
- LANGCHAIN. o. J. *Context Engineering in Agents*. LangChain Documentation. Abgerufen am Dezember 2025 von <https://docs.langchain.com/oss/python/concepts/context>. Geprüft und Archiviert am 29.06.26: [<https://web.archive.org/web/20260629194324/https://docs.langchain.com/oss/python/concepts/context>].
- LEW, G. UND SCHUMACHER, R. M. JR. 2020. *AI and UX: Why Artificial Intelligence Needs User Experience*. Apress, Berkeley, CA, USA.
- LU, C., LU, C., LANGE, R. T., FOERSTER, J., CLUNE, J., UND HA, D. 2024b. The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery. *arXiv Preprint arXiv:2408.06292*. Sakana AI, FLAIR/Oxford, University of British Columbia und Vector Institute. Hinweis: arXiv-Preprint, nicht peer-reviewt.

- LU, Y., YANG, Y., ZHAO, Q., ZHANG, C., UND LI, T. J.-J. 2024a. AI Assistance for UX: A Literature Review Through Human-Centered AI. arXiv Preprint arXiv:2402.06089. University of Notre Dame, Cornell University, University of Washington und Georgia Institute of Technology.
- MATTER, P., WIEDEMANN, A., LEIKERT-BÖHM, N., UND PLEGER, L. E. 2024. The Reverse Turing Test in the Double Chinese Room: Observing GenAI Use in Practical Settings. In *Proceedings of DIGIT 2024*, AIS Electronic Library.
- MAYRING, P. 2022. *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. 13., überarbeitete Auflage, Beltz Verlag, Weinheim und Basel.
- MEI, L., YAO, J., GE, Y., WANG, Y., BI, B., CAI, Y., LIU, J., LI, M., LI, Z.-Z., ZHANG, D., ZHOU, C., MAO, J., XIA, T., GUO, J., UND LIU, S. 2025. A Survey of Context Engineering for Large Language Models. arXiv Preprint arXiv:2507.13334. Abgerufen am Dezember 2025 von <https://arxiv.org/abs/2507.13334> Geprüft und Archiviert am 29.06.26: [<https://web.archive.org/web/20260608073756/https://arxiv.org/abs/2507.13334>]. Hinweis: arXiv-Preprint, nicht peer-reviewt.
- MESSERI, L. UND CROCKETT, M. J. 2024. Artificial intelligence and illusions of understanding in scientific research. *Nature* 627, 49-58.
- MOSQUEIRA-REY, E., HERNANDEZ-PEREIRA, E., ALONSO-RIOS, D., BOBES-BASCARAN, J., UND FERNANDEZ-LEAL, A. 2023. Human-in-the-loop machine learning: a state of the art. *Artificial Intelligence Review* 56, 3005-3054.
- PAGE, M. J., MCKENZIE, J. E., BOSSUYT, P. M., BOUTRON, I., HOFFMANN, T. C., MULROW, C. D., ET AL. 2021. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 372, n71.
- PERNICE, K. 2022. Democratize User Research in 5 Steps. Nielsen Norman Group, 12. Juni 2022. Abgerufen am Dezember 2025 von <https://www.nngroup.com/articles/democratize-user-research/>. Geprüft und Archiviert am 29.06.26: [<https://web.archive.org/web/20260629195840/https://www.nngroup.com/articles/democratize-user-research/>].
- REN, S., XIE, C., JIAN, P., REN, Z., LENG, C., UND ZHANG, J. 2025. Towards Scientific Intelligence: A Survey of LLM-based Scientific Agents. arXiv Preprint arXiv:2503.24047. Chinese Academy of Sciences. Hinweis: arXiv-Preprint, nicht peer-reviewt.
- SANDERS, E. B.-N. UND STAPPERS, P. J. 2008. Co-creation and the new landscapes of design. *CoDesign* 4, 1, 5-18.
- SCHMIDGALL, S., SU, Y., WANG, Z., SUN, X., WU, J., YU, X., LIU, J., LIU, Z., UND BARSOUM, E. 2025. Agent Laboratory: Using LLM Agents as Research Assistants. arXiv Preprint arXiv:2501.04227. AMD, Johns Hopkins University und ETH Zurich. Hinweis: arXiv-Preprint, nicht peer-reviewt.
- SHNEIDERMAN, B. 2020. Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-Centered AI Systems. *ACM Transactions on Interactive Intelligent Systems* 10, 4, Article 26.
- SPINUZZI, C. 2005. The Methodology of Participatory Design. *Technical Communication* 52, 2, 163-174.

- STIGE, A., ZAMANI, E. D., MIKALEF, P., UND ZHU, Y. 2024. Artificial Intelligence (AI) and User Experience (UX) Design: A Systematic Literature Review and Future Research Agenda. *Information Technology & People*.
- VAN DIS, E. A. M., BOLLEN, J., ZUIDEMA, W., VAN ROOIJ, R., UND BOCKTING, C. L. 2023. ChatGPT: Five priorities for research. *Nature* 614, 224-226.
- VOM BROCKE, J., SIMONS, A., NIEHAVES, B., RIEMER, K., PLATTFAUT, R., UND CLEVEN, A. 2009. Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In *Proceedings of the 17th European Conference on Information Systems (ECIS 2009)*, Verona, Italien.
- WANG, H., FU, T., DU, Y., ET AL. 2023. Scientific discovery in the age of artificial intelligence. *Nature* 620, 47-60.
- WEBSTER, J. UND WATSON, R. T. 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly* 26, 2, xiii-xxiii.
- ZHONG, H. UND ZHU, S. 2026. AI Harness Engineering: A Runtime Substrate for Foundation-Model Software Agents. *arXiv Preprint arXiv:2605.13357*. Abgerufen am Dezember 2025 von <https://arxiv.org/abs/2605.13357>. Geprüft und Archiviert am 29.06.26: [<https://web.archive.org/web/20260629195055/https://arxiv.org/abs/2605.13357>]. Hinweis: arXiv-Preprint, nicht peer-reviewt.

16. Nutzung von KI-Tools in der Arbeit

16.1. Deklaration der Nutzung von KI-Tools

Die folgende Übersicht deklariert die KI-Werkzeuge, die bei der Erstellung der schriftlichen Arbeit eingesetzt wurden, mit einer groben Angabe ihrer Verwendung. Aufgrund der Thematik dieser Arbeit wird der Vollständigkeit halber angemerkt, dass die fachliche, in dieser Arbeit untersuchte KI-Nutzung innerhalb des Analyseprozesses nicht Gegenstand dieser Deklaration ist, sondern in den jeweiligen Kapiteln beschrieben wird.

Werkzeug	Art der Verwendung beim Verfassen der Arbeit
Claude (Anthropic), eingesetzt in einer Desktop-Arbeitsumgebung mit direktem Zugriff auf die Projektdaten Modelle: - Claude Opus 4.7 - Claude Opus 4.8	Hauptwerkzeug der Textarbeit. Unterstützung durch Empfehlungen mit Fokus auf Redigieren und Lektorieren von Kapitel- und Abschnittsentwürfen und zugewiesenem Material. Die Empfehlungen wurden anschliessend geprüft und bei Bedarf übernommen. Unterstützung bei Quellenrecherche und Quellenkritik, Formatierung und Konvertierung von Tabellendaten.
ChatGPT (OpenAI), eingesetzt in der Webumgebung. Modelle: - GPT-5.5 Thinking - GPT-5.4 Thinking	Generell als intelligente Suchmaschine. Vorgelagerte methodische Recherche und zum Spiegeln von Ideen. Erstellung des Titelbilds.

16.2. Reflexion zur Nutzung von KI-Tools

Die Nutzung von KI beim Verfassen dieser Arbeit war durchgehend und intensiv, zugleich aber bewusst eng geführt. Auffällig ist für mich rückblickend, wie stark die Haltung, die ich im Untersuchungsgegenstand selbst entwickelt habe, auf meine eigene Schreibpraxis abgefärbt hat. Das in dieser Arbeit als Vibe Research benannte Prinzip, kodifizierbare Fleissarbeit an die KI zu delegieren und Fachurteil, Kontrolle und Verantwortung beim Menschen zu belassen, habe ich beim Schreiben unwillkürlich auf den Text selbst angewendet. Ich schreibe relativ schnell und so, wie ich rede, doch dabei kann es zu Redundanzen und Inkonsistenzen kommen. Der Text ist dann inhaltlich korrekt, aber nicht angenehm zu lesen. Die KI prüfte auf Konsistenz, machte Vorschläge zu Kürzungen und Umformulierungen aber die inhaltlichen und sprachlichen Entscheidungen sind durchgehend meine geblieben.

Damit hat sich meine Arbeitsweise im Lauf der Arbeit spürbar verändert. Wo ich KI früher eher frei und unmittelbar eingesetzt habe, bin ich dazu übergegangen, sie auch beim Schreiben regelgebunden zu führen. Konkret bedeutet das, dass ich basierend auf Context Engineering auf eine feste, für KI nachvollziehbar begründete Regelung bestanden habe, welche auf strikter Inhaltstreue statt freier Umformulierung festlegte und auf einer Materialdisziplin aufbaute, bei welcher explizite Kontextinformationen genutzt werden und nicht mehr. Dadurch konnten Passagen mit zugewiesenen Quellen abgeglichen werden, um auf falsche Zitate hinzuweisen und Stellen im Text zur Gegenprüfung mit dem Quellmaterial zu markieren. Solche Vorschläge der KI habe ich mir erst zur Durchsicht vorlegen lassen, teilweise farblich markiert, bevor sie in die Arbeit eingingen. Da eine Farbkodierung je nach «Status» des Textes (Rot: Grobe Notizen, Blau:

Textentwürfe, Grün: Inhaltlich ok aber ggf. falsch verortet) meiner gewohnten Arbeitsweise entspricht. So hatte ich neue Farben für «KI-Vorschläge». Die Letztentscheidung lag bei jedem Absatz bei mir.

Positiv war vor allem die Entlastung bei der eigentlichen Fleissarbeit des Schreibens. Wie erwähnt, fällt es mir leichter, längere Texte frei herauszuschreiben. Beim Kürzen und Zusammenfassen dieser Texte benötige ich dann die meiste Zeit. Ich habe bspw. eine A4-Seite in Kürze vollgeschrieben und brauche dann Stunden, daraus einen sauberen Absatz zu formulieren. Denn meine Texte sind manchmal durch dieses schnelle Schreiben, wie die KI gerne meinte «Inhaltlich korrekt aber akademisch angreifbar formuliert.» Auch über ein umfangreiches Dokument hinweg liessen sich die Fleissarbeiten rund um die Quellen Formalia deutlich beschleunigen. Gerade das iterative Arbeiten über viele Fassungen hinweg, mit immer neuen Überarbeitungsrounds, wäre von Hand in diesem Tempo kaum möglich gewesen.

Sprachlich formuliert die KI sicher, neigt aber dazu, Aussagen zu glätten, Inhalte unbemerkt zu verschieben oder Quellen- und Begriffsbezüge zu überdehnen. Aus dieser Beobachtung habe ich mich bewusst dafür entschieden, die KI beim Schreiben eher wie ein Lektorat einzusetzen, dessen Vorschläge ich prüfe und von Hand nachtrage, statt Textteile unmittelbar zu übernehmen. Ausschlaggebend war dabei weniger der Kontrollaufwand als die Sorge um die eigene Stimme. Das grösste Kompliment der Testlesenden lautete «Man hört dich reden, wenn man es liest» und «Es macht Spass zu lesen». Und gerade beim zweiten Satz, greift für mich meine Überzeugung, im Sinne der menschenzentrierten Entwicklung einen gewissen Lesespass mit dem eher technischen Anspruch, den Mindeststandard, die wissenschaftliche Schreibweise einzuhalten. Um diese Balance herzustellen, war die KI erneut lektorierend im Einsatz.

Hinzu kamen technische Reibungen, etwa wiederkehrende Probleme mit der korrekten Darstellung von Umlauten oder mit der Synchronisation der Dateien. Die Erfahrung, dass die Prüfung eines KI-Ergebnisses fast dieselbe Sorgfalt verlangt wie seine Erstellung, hat sich beim Schreiben ebenso bestätigt wie in der fachlichen Analyse.

Überrascht hat mich im Nachhinein, wie konsequent ich dieselbe kritische Distanz, die ich von einem KI-unterstützten Analyseprozess verlange, am Ende auch mir selbst und meinem Schreibwerkzeug abverlangen musste. Die KI hat mir die Arbeit erleichtert, sie aber an keiner Stelle abgenommen. Inhalt, Argumentation, Auswahl und Verantwortung für die Aussagen dieser Arbeit liegen vollständig bei mir. Die KI war Werkzeug und nicht Autor.

17. Anhang

Anhangsverzeichnis

Anhang A: Fachglossar	103
Anhang B: Literaturrecherche und Quellenreview	108
Anhang C: Stakeholder-Analyse-Map.....	118
Anhang D: Risikoanalyse.....	119
Anhang E: Gesprächsleitfaden Stakeholder-Interview Consultant.....	120
Anhang F: Gesprächsleitfaden Stakeholder-Interview Leitung EX Circle	121
Anhang G: Kodierleitfaden Hauptkategorien	122
Anhang H: Herleitung Bedürfnisse und Anforderungen	126
Anhang I: Participatory-Design-Workshop: Foto der Workshopwand	138
Anhang J: Workshop-Ergebnisse Reflexionsworkshop	139
Anhang K: Prototyp-Testing: Aufgabenblatt.....	149
Anhang L: Prototyp-Chatverlauf Person01	163
Anhang M: Prototyp-Chatverlauf Person02.....	166
Anhang N: Prototyp-Chatverlauf Person03	169
Anhang O: Prototyp-Fragebogen.....	172
Anhang P: Aufgaben und Feedback verdichtet	174
Anhang Q: Fragebogen Abschluss verdichtet	176

Anhang A: Fachglossar

Fachglossar M365 und KI

Das Glossar erklärt zentrale M365-, KI- und LLM-Begriffe, die in dieser Arbeit wiederholt verwendet werden und für das Verständnis des Analyseprozesses relevant sind.

M365-spezifische Begriffe

Begriff	Definition
Azure OpenAI	Cloud-Dienst von Microsoft, der grosse Sprachmodelle bereitstellt und in Unternehmensanwendungen wie Microsoft 365 Copilot eingebunden werden kann. In dieser Arbeit ist Azure OpenAI vor allem relevant, weil Copilot auf GPT-basierten Foundation Models aus dem Microsoft-Cloud-Kontext aufbauen kann
Building Blocks	Strukturelemente des M365-Evolution-Modells, die thematische Bereiche des digitalen Arbeitsplatzes in der Analyse systematisch abbilden. Sie dienen im Light Assessment als fachliche Grundlage für Kodierung, Befundstrukturierung und spätere Empfehlungen
Conditional Access	Sicherheitsmechanismus in Microsoft 365, mit dem Zugriffe auf Systeme und Daten anhand definierter Bedingungen gesteuert werden können. Im Kontext dieser Arbeit ist Conditional Access Teil des Governance- und Sicherheitsrahmens von Microsoft 365.
Copilot Agent / Microsoft-365-Copilot-Agent	Konfigurierbare, dialogbasierte Schnittstelle innerhalb von Microsoft 365 Copilot, die einem definierten Ablauf folgt und Nutzende schrittweise durch Aufgaben oder Prozesse führt. In dieser Arbeit bildet ein solcher Agent die konzeptionelle Interaktionsform des Light Assessment Prozesses
Data Loss Prevention (DLP)	Richtlinienbasierter Schutzmechanismus, der verhindern soll, dass vertrauliche Daten unkontrolliert weitergegeben oder ausserhalb definierter Grenzen genutzt werden. DLP ist Teil des im Dokument beschriebenen Governance-Rahmens von Microsoft 365.
EU Data Boundary	Richtlinie innerhalb der Microsoft-Cloud, die sicherstellt, dass bestimmte Daten innerhalb geografischer Grenzen der EU gespeichert und verarbeitet werden. In der Arbeit wird sie als Teil der Compliance-Zusagen im Microsoft-365-Kontext genannt.
Light Assessment	Fokussierte Weiterentwicklung des M365-Evolution-Assessments mit reduziertem Erhebungs- und Auswertungsaufwand. Ziel ist es, die Analyse auch in kleineren Kundenprojekten wirtschaftlich tragfähig zu machen, ohne die fachliche Qualität der Beurteilung aufzugeben.
Light Assessment Agent	Konzeptionell definierter Copilot-Agent, der den Analyseprozess des Light Assessments unterstützt und einzelne Arbeitsschritte KI-gestützt vorbereitet. Er dient im Prototyp als dialogbasierte Schnittstelle zwischen Fachperson und KI-gestütztem Prozess.
Kollaboration (im M365-Kontext)	Kollaboration bezeichnet die koordinierte Zusammenarbeit von Personen in Microsoft 365, bei welcher Inhalte, Kommunikation und Aufgaben in integrierten Anwendungen (z. B. Teams, SharePoint, OneDrive) gemeinsam erstellt, geteilt und bearbeitet werden.

Begriff	Definition
M365-Evolution-Assessment / Evolution Assessment	Strukturiertes Analyseverfahren zur Bestimmung des digitalen Reifegrads einer Organisation im Umgang mit Microsoft 365. Es macht Handlungsfelder sichtbar und dient als Grundlage für die Ableitung konkreter Entwicklungsschritte.
M365-Evolution-Modell	Projektspezifisches Reifegradmodell zur Bewertung der Nutzung und Verankerung von Microsoft 365 in Organisationen. Es unterscheidet die Entwicklungsstufen Work, Boost und Perform.
M365-Fachpersonen / M365-Consultants	Fachpersonen mit Expertise in Microsoft 365, die Organisationen bei Einführung, Nutzung und Optimierung der Plattform beraten und begleiten. In dieser Arbeit bilden sie die zentrale Nutzergruppe des konzipierten Analyseprozesses.
Microsoft 365 (M365)	Cloudbasierte Plattform von Microsoft, die Produktivitäts-, Kollaborations- und Sicherheitsdienste in einer integrierten Systemumgebung bündelt. In der Arbeit bildet Microsoft 365 die vorgegebene Umsetzungsumgebung des Analyseprozesses.
Microsoft 365 Copilot	KI-basierte Orchestrierungsplattform innerhalb von Microsoft 365, die Sprachmodelle mit organisationalen Daten und Anwendungen verbindet. In der Arbeit ist Copilot relevant, weil er den Rahmen für die KI-gestützte Bearbeitung innerhalb der Microsoft-365-Umgebung bildet.
Microsoft Graph	Zentrale Schnittstelle innerhalb von Microsoft 365, über die organisationsbezogene Inhalte und Berechtigungen zugänglich gemacht werden können. Im Copilot-Kontext dient der Graph als Vermittlungsschicht zwischen Modell, Nutzenden und unternehmensinternen Daten.
Microsoft Purview	Microsoft-Plattform für Datenverwaltung, Compliance und Monitoring. In der Arbeit wird Purview im Zusammenhang mit Audit-Logs und Nachvollziehbarkeit im Governance-Rahmen von Microsoft 365 genannt.
Microsoft Solutions Partner	Zertifizierungsstatus für Unternehmen, die nachweislich Expertise in bestimmten Microsoft-Technologiebereichen besitzen. Im Projektkontext beschreibt der Begriff die Microsoft-nahe Positionierung des Praxispartners.
Microsoft Teams	Kollaborationsplattform innerhalb von Microsoft 365, die Chat, Besprechungen, Aufzeichnung, Transkription und Zusammenarbeit in Teams unterstützt. In der Arbeit wird Teams unter anderem für Interviews, Transkripte und das Test-Setup verwendet.
Microsoft-365-Ökosystem	Gesamtheit der Dienste, Anwendungen, Datenflüsse und Integrationen innerhalb von Microsoft 365. Der Begriff beschreibt den technischen und organisatorischen Rahmen, in dem der Analyseprozess konzipiert wird.
Modern Workplace / Modern Work	Konzept für digitale Arbeitsumgebungen, das flexible und cloudbasierte Zusammenarbeit ermöglicht. In der Arbeit bildet Modern Work den fachlichen Kontext, in dem das Evolution Assessment den digitalen Reifegrad untersucht.
Tenant	Mandant innerhalb von Microsoft 365, der eine organisatorische Einheit mit eigener Datenhaltung, Benutzerverwaltung und Zugriffslogik repräsentiert. Der Begriff ist wichtig für das Verständnis von Datenbindung, Berechtigungen und Governance im Copilot-Kontext.

Begriff	Definition
Viva	Microsoft-Produktfamilie für Employee Experience, Kommunikation, Lernen und Wissensmanagement. Im Dokument wird Viva als Beispiel für digitale Werkzeuge im Microsoft-365-Kontext genannt.
Word / Excel	Office-Anwendungen innerhalb von Microsoft 365 zur Erstellung und Bearbeitung von Textdokumenten beziehungsweise Tabellen und Datenanalysen. In der Arbeit erscheinen sie als Teil der bestehenden und künftigen Werkzeugumgebung des Analyseprozesses.
Work, Boost, Perform	Drei Reifegradstufen des M365-Evolution-Modells. Work beschreibt grundlegende Arbeitsfähigkeit, Boost Effizienzsteigerung und Perform eine produktivitätsorientierte Nutzung von Microsoft 365.

KI- und LLM-spezifische Begriffe

Begriff	Definition
KI / künstliche Intelligenz	Sammelbegriff für Systeme, die Aufgaben unterstützen oder ausführen, die üblicherweise menschliche kognitive Fähigkeiten wie Mustererkennung, Sprachverarbeitung, Klassifikation oder Entscheidungsunterstützung erfordern. In dieser Arbeit bezeichnet KI vor allem unterstützende Verfahren zur Aufbereitung, Strukturierung und Verdichtung empirischer Daten.
Generative KI / GenAI	Form künstlicher Intelligenz, die neue Inhalte wie Texte, Zusammenfassungen, Klassifikationen oder Vorschläge erzeugen kann. In dieser Arbeit wird generative KI als unterstützendes Werkzeug verstanden, nicht als eigenständige Evidenzquelle oder als Ersatz fachlicher Interpretation.
LLM / Large Language Model	Grosses Sprachmodell, das auf umfangreichen Textdaten trainiert wurde und Sprache verarbeiten, erzeugen, strukturieren und klassifizieren kann. In der Arbeit werden LLMs für semantische Aufgaben wie Kodierung, Verdichtung und Zusammenfassung eingesetzt.
LLM-basierte Systeme	Systeme, deren zentrale Funktion auf einem grossen Sprachmodell beruht. Sie können mit Prompts, Kontextmaterial, Werkzeugen, Abläufen und Kontrollmechanismen kombiniert werden und bilden damit mehr als nur ein einzelnes Modell.
Foundation Model	Breit vortrainiertes Modell, das mit vergleichsweise wenig Anpassung für unterschiedliche Aufgaben eingesetzt werden kann. In der Arbeit wird der Begriff zur Einordnung moderner LLM-basierter Systeme verwendet.
Prompt	Eingabe oder Anweisung an ein KI-System, mit der Aufgabe, Kontext, Regeln und erwartete Ausgabeform beschrieben werden. Prompts sind in dieser Arbeit ein zentrales Mittel zur Steuerung KI-gestützter Arbeitsschritte.
Prompt Engineering	Gestaltung wirksamer Anweisungen für ein KI-System. Der Fokus liegt auf der Formulierung der Aufgabe, der Regeln, des Kontexts und der erwarteten Antwortstruktur.
Context Engineering	Bewusste Auswahl, Strukturierung und situative Bereitstellung des Bezugsmaterials, das ein LLM für eine konkrete Aufgabe benötigt. Der Begriff

Begriff	Definition
	geht über Prompt Engineering hinaus, weil nicht nur die Anweisung, sondern auch das verwendete Material methodisch gestaltet wird.
Harness Engineering	Gestaltung der technischen Ausführungs-, Steuerungs- und Kontrollumgebung eines LLM-basierten oder agentischen Systems. Dazu gehören unter anderem Werkzeuganbindung, Speicher, Ablaufsteuerung, Berechtigungen, Prüfmechanismen, Protokollierung und menschliche Eingriffsmöglichkeiten.
Human-in-the-Loop	Gestaltungsprinzip, bei dem Menschen an definierten Punkten in einen KI-gestützten Prozess eingebunden bleiben, um Zwischenergebnisse zu prüfen, zu korrigieren, freizugeben oder fachlich zu beurteilen. In dieser Arbeit ist Human-in-the-Loop zentral für die Sicherung fachlicher Verantwortung.
Mensch-KI-Aufgabenteilung	Bewusste Verteilung von Aufgaben zwischen Mensch und KI. In dieser Arbeit übernimmt die KI vor allem strukturierende und verdichtende Fleissarbeit, während Fachurteil, Verantwortung und Priorisierung beim Menschen verbleiben.
Fleissarbeit	Kodifizierbare, wiederkehrende und am Material überprüfbare Arbeitsschritte, die an KI delegiert werden können. Dazu gehören etwa Datenbereinigung, Strukturierung, Kodier- oder Clustervorschläge, Konsistenzprüfung und Reformulierung.
Fachurteil	Nicht delegierbare fachliche Entscheidung, die Kontextverständnis, methodische Einordnung und Verantwortung voraussetzt. Dazu gehören Interpretation, Bedeutungszuschreibung, Kategorienentscheidungen, Quellenkritik und abschliessende Bewertung.
Materialbindung	Prinzip, wonach KI-generierte Aussagen auf konkrete Stellen im Ausgangsmaterial zurückführbar sein müssen. Dadurch wird verhindert, dass Ergebnisse nur aus Modellwissen oder unprüfbaren Annahmen entstehen.
Belegbarkeit	Anforderung, dass jede übernommene Aussage durch eine identifizierbare Datenstelle, einen Textausschnitt, einen Zeitstempel oder eine andere nachvollziehbare Evidenz gestützt wird. Belegbarkeit ist eine zentrale Sicherung gegen unkontrollierte oder halluzinierte KI-Ausgaben.
Vibe Research	In dieser Arbeit definierter KI-gestützter Arbeitsmodus zur Aufbereitung, Strukturierung und Auswertung empirischer Daten. Die KI unterstützt kodifizierbare Teilaufgaben, während Zielsetzung, Interpretation, Datenkritik und Verantwortung beim Menschen verbleiben.
Halluzination	Fehlerart generativer KI, bei der plausible, aber falsche, unbelegte oder nicht durch das Ausgangsmaterial gedeckte Aussagen erzeugt werden. In der Arbeit wird Halluzination als zentrales Risiko für Ergebnisverlässlichkeit behandelt.
Bias	Verzerrung in Daten, Modellen oder Ausgaben, die zu einseitigen oder systematisch verzerrten Ergebnissen führen kann. Bias wird in der Arbeit als methodisches Risiko beim Einsatz generativer KI berücksichtigt.
Nichtdeterminismus	Eigenschaft eines KI-Modells, bei gleichen Eingaben und Einstellungen unterschiedliche Ergebnisse erzeugen zu können. Diese Eigenschaft begründet in der Arbeit die Notwendigkeit von Wiederholungsläufen und Stabilitätsprüfungen.

Begriff	Definition
Reproduzierbarkeit	Fähigkeit, bei wiederholter Durchführung unter gleichen Bedingungen vergleichbare Ergebnisse zu erhalten. In der Arbeit ist Reproduzierbarkeit eine zentrale Dimension der Ergebnisverlässlichkeit KI-gestützter Analysen.
Ergebnisverlässlichkeit	In dieser Arbeit verwendeter Begriff für die Kombination aus inhaltlicher Korrektheit, Belegbarkeit und Stabilität KI-gestützter Ergebnisse. Sie umfasst insbesondere die Vermeidung von Halluzinationen und die Prüfung der Reproduzierbarkeit.
Reliabilität	Wiederholbarkeit beziehungsweise Stabilität eines Ergebnisses unter gleichen Bedingungen. In der Arbeit wird die Reliabilitätskomponente der Ergebnisverlässlichkeit über wiederholte KI-Läufe geprüft.
Validität	Frage, ob ein Ergebnis fachlich das Richtige erfasst oder misst. In der Arbeit wird Validität von Reliabilität unterschieden und durch menschliche Prüfung sowie fachliche Einordnung abgesichert.
Lauf-zu-Lauf-Stabilität	Mass dafür, wie stark Ergebnisse mehrerer KI-Durchläufe bei gleichem Datensatz, gleichem Codebuch und gleichen Einstellungen übereinstimmen. Sie wird in der Arbeit zur Beurteilung der Reliabilitätskomponente verwendet.
Fleiss Kappa	Zufallskorrigiertes Mass zur Bewertung der Übereinstimmung mehrerer Rater oder Durchläufe bei kategorialen Entscheidungen. In der Arbeit wird Fleiss Kappa zur Prüfung der Lauf-zu-Lauf-Stabilität eingesetzt.
Krippendorffs Alpha	Reliabilitätsmass zur Bewertung der Übereinstimmung von Kodierungen. In der Arbeit wird es ergänzend zur Beurteilung der Stabilität auf verschiedenen Analyseebenen verwendet.
Gwets AC1	Prävalenzrobustes Übereinstimmungsmass, das als Ergänzung zu Kappa verwendet werden kann. In der Arbeit wird es eingesetzt, weil viele Felder leer bleiben und dadurch klassische Kappa-Werte verzerrt wirken können.
Kodierleitfaden	Regelwerk, das Kategorien, Definitionen, Ankerbeispiele und Kodierregeln festlegt. Für LLM-basierte Systeme dient der Kodierleitfaden zugleich als methodische Grundlage und als Begrenzung des Handlungsspielraums der KI.
Deduktive Kodierung	Kodierung anhand vorab definierter Kategorien. In der Arbeit wird sie genutzt, um Interviewmaterial systematisch auf festgelegte Ziele, Lücken und Themen zu prüfen.
Evidenz	Konkrete Datenstelle, Aussage, Passage oder quantitative Kennzahl, die eine analytische Aussage stützt. Im Prozess müssen Evidenzen nachvollziehbar den Codes und Befunden zugeordnet sein.
Wizard-of-Oz-Verfahren	Evaluationsverfahren, bei dem eine scheinbar automatisierte Systemfunktion durch eine Person im Hintergrund simuliert wird. In dieser Arbeit wird dieses Verfahren genutzt, um den geplanten KI-gestützten Prozess realitätsnah zu testen, obwohl noch kein vollständig implementiertes System vorliegt.
Proof of Concept	Technischer Nachweis, dass ein kritischer Teil einer Lösung grundsätzlich funktionsfähig ist. In der Arbeit dient der Proof of Concept der Veranschaulichung und Erprobung einzelner KI-gestützter Prozessschritte, ohne eine produktionsreife Umsetzung zu beanspruchen.

Anhang B: Literaturrecherche und Quellenreview

MAS HCID Masterarbeit | Quellenreview (kompakt)

Literaturrecherche und Quellenreview

Kompakte Übersicht der verwendeten Literatur

MAS HCID Masterarbeit

Konzeption und Evaluation eines KI-unterstützten Analyseprozesses im Microsoft-365-Ökosystem

Dominik James, OST Ostschweizer Fachhochschule

Stand: Mai 2026

KI HINWEIS: Dieses Dokument wurde mit Opus 4.7 nach Anweisungen des Studierenden über Monate erstellt, um die Übersicht in der Literaturvertiefung zu vereinfachen. Statt Citativ oder ähnlicher Programme.

1 Ziel der Quellenreview

Diese Übersicht zeigt, welche Quellen die Masterarbeit stützen und wie sie inhaltlich zusammenhängen. Sie dokumentiert das Vorgehen der Literatursuche und ordnet jede Quelle dem Konzept der Arbeit zu, das sie trägt.

Die Forschungslogik der Arbeit folgt drei Linien. Die menschenzentrierte Gestaltung liefert den Rahmen für die Konzeption. Die Forschung zu grossen Sprachmodellen klärt Chancen und Risiken. Die anwendungsnahe Literatur zeigt, wie KI in vergleichbaren Arbeitssituationen genutzt wird. Webster und Watson (2002) sowie vom Brocke et al. (2009) bilden die methodische Grundlage für den Aufbau dieser Review.

2 Vorgehen der Literatursuche

Die Suche fand zwischen Anfang August 2025 und Ende Januar 2026 statt. Genutzt wurden Google Scholar und Academia.edu. ChatGPT diente als Filterhilfe für Suchbegriffe und Vorauswahl, nicht als Quelle. Alle inhaltlichen Aussagen sind am Originaltext belegt.

Die Suche folgte dem dreistufigen Vorgehen von Webster und Watson (2002), also Schlagwortsuche, Rückwärts- und Vorwärtssuche. Aufgenommen wurden Quellen, die thematisch passen, wissenschaftlich belastbar sind und Praxisbezug haben. Schwerpunkt der KI-bezogenen Quellen liegt auf Beiträgen ab 2019. Klassische methodische Grundlagen wie Mayring oder Sanders und Stappers sind bewusst in den Originalfassungen einbezogen.

Die Suche ist pragmatisch und thematisch eingegrenzt. Sie ist kein vollständig systematischer Review nach PRISMA. Für eine Masterarbeit ist das angemessen und folgt der Empfehlung von vom Brocke et al. (2009), wonach ein fokussierter Review aussagekräftiger sein kann als eine breite, aber unscharfe Übersicht.

3 Konzeptmatrix

Die Matrix zeigt, welche Quelle welches Konzept der Arbeit trägt. Eine Markierung bedeutet, dass die Quelle einen substanziellen Beitrag zu diesem Konzept leistet.

Quelle	Menschz. Gestaltung	Quali. Analyse	LLM-Grundl. und Risiken	Mensch-KI-Interakt.	KI in UX und Analyse	Methodik Review	Vibe Research
Webster und Watson 2002						X	
vom Brocke et al. 2009						X	
DIN EN ISO 9241-210	X						
Garrett 2011	X						

MAS HCID Masterarbeit | Quellenreview (kompakt)

Sanders und Stappers 2008	X						
Spinuzzi 2005	X						
Gaete Cruz et al. 2022	X						
Bourne und Walker 2005	X						
Mayring 2022		X					
Baxter et al. 2015	X	X					
Bommasani et al. 2021			X				
Bender et al. 2021			X	X			
Huang et al. 2024			X				
Kalai et al. 2025			X				
Chen et al. 2023			X				
Cui und Alexander 2026			X		X		
Amershi et al. 2019	X			X			
Shneiderman 2020	X			X			
Mosqueira-Rey et al. 2023			X	X	X		
Hong und Park 2025	X			X			
Matter et al. 2024				X	X		
Lew und Schumacher 2020	X			X	X		
Lu et al. 2024a (UX)				X	X		
Stige et al. 2024					X		
Fedosov et al. 2026		X			X		
Bertina 2025							X

Wang et al. 2023			X				X
van Dis et al. 2023			X	X			X
Messeri und Crockett 2024			X	X			X
Baek et al. 2024			X				X
Lu et al. 2024b (AI Scientist)			X	X			X
Schmidgall et al. 2025			X	X			X
Ren et al. 2025			X				X
Page et al. 2021						X	

4 Quellen im Überblick

4.1 Methodik der Literaturarbeit

Webster und Watson (2002) sind der Klassiker für strukturierte Literaturreviews. Sie schlagen eine konzeptzentrierte Darstellung mit Konzeptmatrix sowie eine dreistufige Suche aus Datenbankrecherche, Rückwärts- und Vorwärtssuche vor. Dieses Vorgehen leitet die vorliegende Review.

vom Brocke et al. (2009) ergänzen Webster und Watson um die Forderung nach einer transparent dokumentierten Suche. Sie fordern klare Angaben zu Datenbanken, Suchbegriffen, Suchzeitraum und Filterkriterien. Diese Angaben sind in Kapitel 2 dieser Review enthalten.

4.2 Menschzentrierte Gestaltung

DIN EN ISO 9241-210 definiert die Grundsätze der menschenzentrierten Gestaltung, also Nutzungskontext verstehen, Stakeholder einbeziehen, iterativ entwickeln und aus Nutzungsperspektive bewerten. Diese Grundsätze machen den Analyseprozess zu einem Arbeitssystem, nicht zu reiner Automatisierung.

Garrett (2011) liefert mit den fünf Ebenen der User Experience das Modell, mit dem sich strategische Ziele, fachliche Anforderungen, Struktur, Konzeption und Oberfläche sauber unterscheiden lassen.

Sanders und Stappers (2008) prägen den Begriff des Co-Designs und beschreiben den Übergang vom user-centred design zum gemeinsamen Gestalten mit Nutzenden. Spinuzzi

(2005) ordnet Participatory Design wissenschaftlich ein und unterscheidet so methodisch tragenden Einbezug von rein beratenden Formen.

Gaete Cruz et al. (2022) verbinden die Beteiligungsleiter mit dem Designzyklus. Damit lässt sich der Einbezug der M365-Fachpersonen pro Phase differenziert beschreiben.

Bourne und Walker (2005) stellen den Stakeholder Circle vor. Er unterscheidet Stakeholder nach Macht, Nähe und Dringlichkeit. Das hilft, wenn eine Rolle mehrere Funktionen gleichzeitig hat, wie es bei den M365-Fachpersonen der Fall ist.

4.3 Qualitative Forschungsmethoden

Mayring (2022) ist der zentrale Bezugspunkt für die qualitative Inhaltsanalyse. Das Verfahren strukturiert Texte mit induktiver Kategorienbildung, deduktiven Codieranleitungen und strukturierten Zusammenfassungen. Mayring liefert den Massstab, an dem sich die KI-gestützten Analyseschritte messen lassen.

Baxter, Courage und Caine (2015) ergänzen Mayring auf der Erhebungsseite. Das Buch ist eine etablierte Methodensammlung des HCID-Felds und behandelt unter anderem Interviews, Tagebuchstudien, Surveys, Card Sorts, Fokusgruppen und Feldstudien jeweils mit Vorbereitung, Durchführung, Auswertung und Berichterstattung. Für die Stakeholder-Interviews dieser Arbeit liefert Kapitel 9 zu Interviews den direkten methodischen Bezugspunkt.

4.4 Grosse Sprachmodelle: Grundlagen, Chancen, Risiken

Bommasani et al. (2021) prägen den Begriff der Foundation Models. Diese Quelle liefert die technologische Einordnung der Modelle hinter Microsoft 365 Copilot.

Bender et al. (2021) kritisieren grosse Sprachmodelle als stochastic parrots. Sie warnen vor Verzerrungen, trügerischer Plausibilität und unklarer Verantwortlichkeit.

Huang et al. (2024) liefern eine Übersicht zu Halluzinationen, also Aussagen ohne Faktenbasis. Kalai et al. (2025) erklären formal, warum die heutige Bewertungspraxis das Raten gegenüber Unsicherheit belohnt. Folge für die Arbeit: Mit Falschaussagen ist zu rechnen, die Architektur muss Sicherungen vorsehen.

Chen et al. (2023) und Cui und Alexander (2026) zeigen empirisch, dass dieselben Modelle bei identischen Aufgaben unterschiedliche Ergebnisse liefern. Daraus folgt: Eine einmalige LLM-gestützte Analyse reicht nicht. Wiederholte Läufe und Vergleichsanalysen sind Standard.

4.5 Mensch-KI-Interaktion

Amershi et al. (2019) formulieren auf der CHI 2019 achtzehn Leitlinien für Mensch-KI-Interaktion. Sie fordern Transparenz über Fähigkeiten und Fehler, einfache Korrektur und vorsichtige Kommunikation von Updates. Sie sprechen gegen volle Automatisierung sensibler Aufgaben.

Shneiderman (2020) übersetzt Ethik-Leitlinien in konkrete Praxisempfehlungen auf Team-, Organisations- und Industriebene. Dazu gehören Audit-Trails, Bias-Tests, erklärbare Schnittstellen und externe Kontrolle.

Mosqueira-Rey et al. (2023) ordnen die Forschung zu Human-in-the-Loop. Sie unterscheiden Active Learning, Interactive Machine Learning und Machine Teaching nach der Frage, wer die Kontrolle über den Lernprozess hat. Dieses Vokabular hilft bei der Gestaltung der Aufgabenteilung im Analyseprozess.

Hong und Park (2025) zeigen, wie sich erklärbare KI nutzerzentriert gestalten lässt, mit fünf Phasen von der Kontextanalyse bis zur Evaluation.

Matter et al. (2024) beobachten in fünf Schweizer Projekten, dass Nutzende oft in eine Assistenzrolle gedrängt werden. Die wahrgenommene Verantwortung schwankt zwischen Frustration und Resignation. Das Bild des Double Chinese Room bringt das auf den Punkt.

Lew und Schumacher (2020) argumentieren in einer praxisnahen Monografie, dass KI-Systeme eine gute Nutzererfahrung brauchen, um angenommen zu werden.

4.6 KI im UX-Design und in der qualitativen Analyse

Lu et al. (2024a) analysieren 359 Beiträge zur KI in der UX-Praxis aus Human-Centered-AI-Perspektive. Sie zeigen: KI übernimmt heute eng umrissene Aufgaben, die analytische und empathische Tiefe bleibt beim Fachpersonal.

Stige et al. (2024) liefern eine Literaturübersicht zu KI und UX-Design mit Effizienz- und Innovationspotenzialen sowie einer Forschungsagenda.

Fedosov et al. (2026) dokumentieren einen dreistufigen LLM-Workflow für die qualitative Analyse von Interviewdaten. Sie berichten von Zeitgewinnen und hohem Aufwand für Prompts und Validierung. Diese Quelle ist Referenzpunkt für die Evaluation des Prototyps.

4.7 KI im Forschungsprozess und Vibe Research

Vibe Research bezeichnet eine Arbeitsweise, in der Forschende Frage, Fachwissen und Urteil einbringen, während LLM-Werkzeuge Recherche, Strukturierung und Verdichtung übernehmen. Der Begriff ist sehr jung, das Feld der KI-gestützten Forschung breiter. Die folgenden Quellen ordnen diese Praxis ein, von kritischen Frühpositionen bis zu konkreten Agentensystemen.

Bertina (2025) liefert einen frühen Medium-Beitrag und markiert damit den Beginn des öffentlichen Diskurses zum Begriff. Methodisch ist der Text nicht belastbar, aber als zeitlicher Marker brauchbar. Hinweis: populärwissenschaftlicher Beitrag, nicht peer-reviewt.

Wang et al. (2023) liefern in einer Nature-Übersicht eine breite Bestandsaufnahme zu KI in der wissenschaftlichen Entdeckung. Sie zeigen, wo KI Daten verarbeitet, Hypothesen unterstützt und Experimente plant. Der Beitrag ist peer-reviewt und liefert die Verortung der Diskussion im Feld AI for Science.

van Dis et al. (2023) formulieren in einem Nature-Comment fünf Prioritäten für den Forschungseinsatz von LLMs: menschliche Verifikation halten, klare Regeln für Autorenschaft, Investitionen in eigene Modelle, kritische Auseinandersetzung und eine breite Debatte. Diese Quelle stützt das Argument, dass KI-Output durch Fachpersonen geprüft werden muss.

Messeri und Crockett (2024) warnen in einer Nature-Perspective vor Illusionen des Verstehens beim KI-Einsatz in der Forschung. KI verspricht Objektivität und Produktivität, kann aber Monokulturen erzeugen, in denen weniger verstanden und mehr produziert wird. Diese Quelle ist ein zentraler Bezugspunkt für die Risikoreflexion.

Baek et al. (2024) stellen mit ResearchAgent ein LLM-System vor, das Forschungsideen aus wissenschaftlicher Literatur iterativ generiert. Es kombiniert Wissensgraphen mit mehreren Review-Agenten. Damit zeigt diese Quelle ein konkretes Beispiel für KI-gestützte Ideenfindung.

Lu et al. (2024b) gehen mit The AI Scientist einen Schritt weiter und versuchen einen vollautomatisierten Forschungsprozess von Idee bis Paper. Diese Quelle markiert das vollautonome Ende des Spektrums und dient als Kontrast zur Vibe-Research-Position, die den Menschen bewusst eingebunden lässt.

Schmidgall et al. (2025) stellen mit Agent Laboratory ein LLM-Framework vor, das Literatursuche, Experimente und Berichtschreiben durchläuft und an jeder Stufe menschliches Feedback einholt. Diese Quelle ist relevant, weil sie die Mensch-KI-Aufgabenteilung konkret operationalisiert.

Ren et al. (2025) liefern eine Übersicht zu LLM-basierten wissenschaftlichen Agenten, mit Differenzierung nach Aufgaben und Validierungsmechanismen. Die Quelle ordnet das schnell wachsende Feld systematisch und ist Hinweisgeber für weitere Anschlussliteratur.

Page et al. (2021) sind nicht direkt eine Vibe-Research-Quelle, aber methodisch der Goldstandard für systematische Reviews. Die PRISMA-2020-Anforderungen liefern den Massstab, gegen den die vorliegende, bewusst pragmatische Review sich abgrenzt. Diese Abgrenzung ist transparent zu dokumentieren.

5 Fazit

Die Literatur stützt die Forschungslogik der Arbeit in drei Punkten. Erstens liefert sie den Rahmen menschenzentrierter Gestaltung. Zweitens klärt sie die Grenzen und Risiken grosser Sprachmodelle so, dass sich daraus Designprinzipien ableiten lassen. Drittens zeigt sie an anwendungsnahen Beispielen, wie KI heute in vergleichbaren Kontexten genutzt wird.

Daraus ergeben sich vier Konsequenzen für die Arbeit. Generative und deterministische Komponenten klar trennen. Halluzinations- und Reproduzierbarkeitsrisiken durch wiederholte Läufe und Kontrollpunkte abfangen. Die Mensch-KI-Aufgabenteilung auch in Bezug auf

Verantwortung gestalten. Den Prototyp an Nutzbarkeit, Effizienz und Ergebnisverlässlichkeit messen.

Die Diskussion um KI im Forschungsprozess und um vibe research liefert ergänzend einen Rahmen, um das eigene Vorgehen in der IST-Analyse präziser zu fassen. Sie spannt einen Bogen von kritischen Frühpositionen (van Dis et al. 2023, Messeri und Crockett 2024) über konkrete Agentensysteme (Baek et al. 2024, Schmidgall et al. 2025) bis hin zur vollautonomen Variante (Lu et al. 2024b). Die vorliegende Arbeit positioniert sich bewusst im mittleren Feld, mit menschlicher Kontrolle an definierten Stellen.

6 Literaturverzeichnis

Referenzierung im ACM Author-Date-Stil gemäss OST-Vorgaben.

AMERSHI, S., WELD, D., VORVOREANU, M., FOURNEY, A., NUSHI, B., COLLISSON, P., SUH, J., IQBAL, S., BENNETT, P. N., INKPEN, K., TEEVAN, J., KIKIN-GIL, R., UND HORVITZ, E. 2019. Guidelines for Human-AI Interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI 2019), Glasgow, Scotland UK. ACM, New York, NY, USA.

BAEK, J., JAUHAR, S. K., CUCERZAN, S., UND HWANG, S. J. 2024. ResearchAgent: Iterative Research Idea Generation over Scientific Literature with Large Language Models. arXiv Preprint arXiv:2404.07738. KAIST und Microsoft Research. Hinweis: arXiv-Preprint, nicht peer-reviewt.

BAXTER, K., COURAGE, C., UND CAINE, K. 2015. Understanding Your Users: A Practical Guide to User Research Methods. 2. Auflage, Morgan Kaufmann (Elsevier), Waltham, MA, USA.

BENDER, E. M., GEBRU, T., MCMILLAN-MAJOR, A., UND SHMITCHELL, S. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT 2021), 610-623. ACM, New York, NY, USA.

BERTINA, A. 2025. Vibe-Research instead of Deep-Research: AI Agents and the Future of Scientific Discovery. Medium, 25. Mai 2025. <https://medium.com/@abertina/vibe-research-instead-of-deep-research-ai-agents-and-the-future-of-scientific-discovery-4d561248f3e2> (geprüft 29.05.2026). Hinweis: populärwissenschaftlicher Online-Beitrag, nicht peer-reviewt.

BOMMASANI, R. ET AL. 2021. On the Opportunities and Risks of Foundation Models. Center for Research on Foundation Models (CRFM), Stanford University. arXiv:2108.07258.

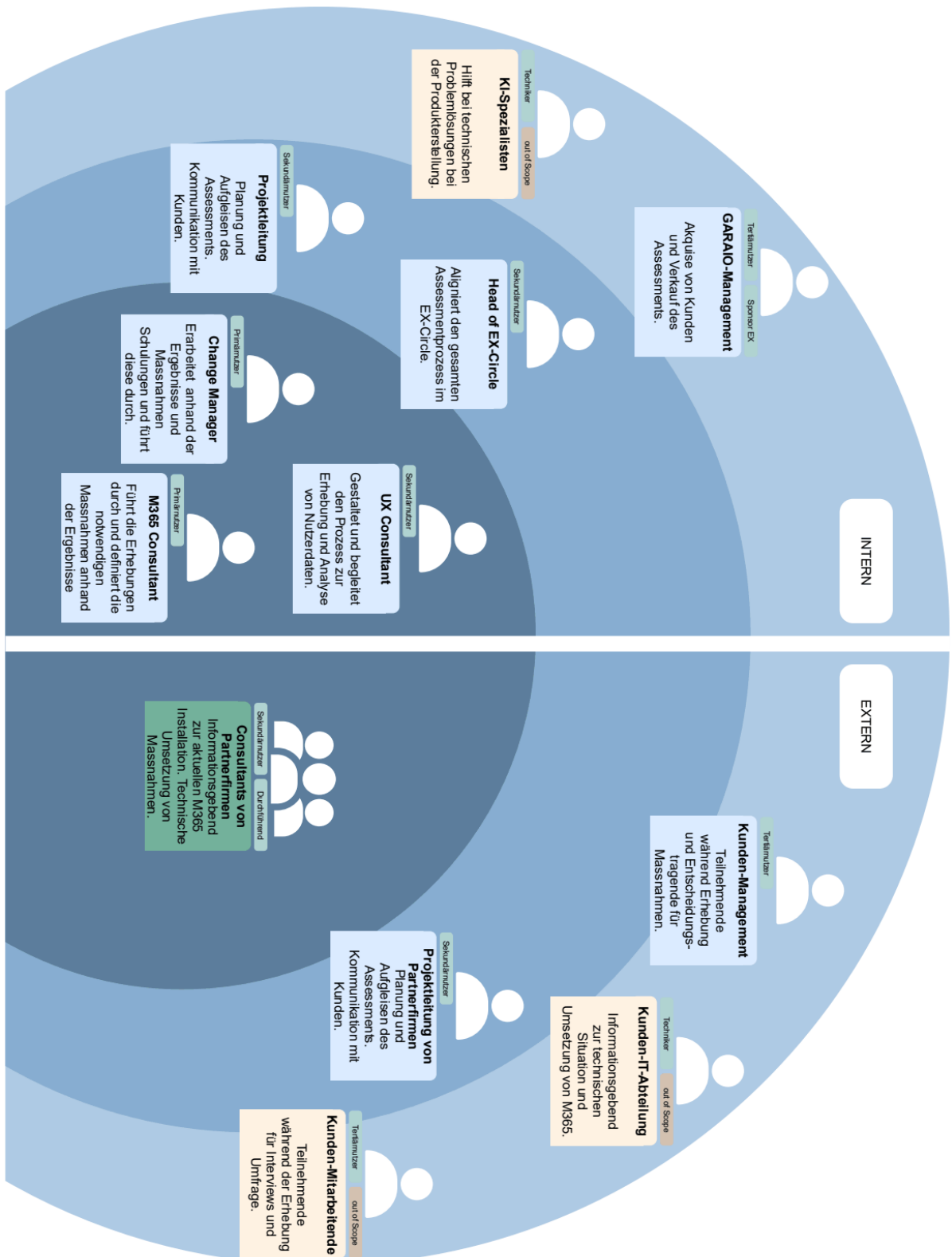
BOURNE, L. UND WALKER, D. H. T. 2005. Visualising and mapping stakeholder influence. Management Decision 43, 5, 649-660.

CHEN, L., ZAHARIA, M., UND ZOU, J. 2023. How Is ChatGPT's Behavior Changing over Time? Stanford University und UC Berkeley. arXiv:2307.09009.

- CUI, J. UND ALEXANDER, R. 2026. Same Prompt, Different Outcomes: Evaluating the Reproducibility of Data Analysis by LLMs. Working Paper, University of Toronto.
- DEUTSCHES INSTITUT FÜR NORMUNG (DIN). 2020. DIN EN ISO 9241-210:2020-03 Ergonomie der Mensch-System-Interaktion. Teil 210: Menschzentrierte Gestaltung interaktiver Systeme. Beuth Verlag, Berlin.
- FEDOSOV, A., KERN, D., UND BARTH, C.-M. 2026. LLM-powered Workflow to Support Qualitative Data Analysis: Lessons Learned from an Applied Research Project. FHNW University of Applied Sciences and Arts Northwestern Switzerland und Universität Zürich.
- GAETE CRUZ, M., ERSOY, A., CZISCHKE, D., UND VAN BUEREN, E. 2022. Towards a framework for urban landscape co-design: Linking the participation ladder and the design cycle. *CoDesign* 19, 3, 233-252.
- GARRETT, J. J. 2011. Die Elemente der User Experience: Anwender-zentriertes (Web)Design. Addison Wesley, München. Deutsche Ausgabe, Pearson Education Deutschland.
- HONG, S. UND PARK, W. 2025. Developing user-centered system design guidelines for explainable AI: a systematic literature review. *Artificial Intelligence Review* 58, 386.
- HUANG, L., YU, W., MA, W., ZHONG, W., FENG, Z., WANG, H., CHEN, Q., PENG, W., FENG, X., QIN, B., UND LIU, T. 2024. A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions. *ACM Transactions on Information Systems*.
- KALAI, A. T., NACHUM, O., VEMPALA, S. S., UND ZHANG, E. 2025. Why Language Models Hallucinate. OpenAI und Georgia Tech.
- LEW, G. UND SCHUMACHER, R. M. JR. 2020. AI and UX: Why Artificial Intelligence Needs User Experience. Apress, Berkeley, CA, USA.
- LU, C., LU, C., LANGE, R. T., FOERSTER, J., CLUNE, J., UND HA, D. 2024b. The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery. arXiv Preprint arXiv:2408.06292. Sakana AI, FLAIR/Oxford, University of British Columbia und Vector Institute. Hinweis: arXiv-Preprint, nicht peer-reviewt.
- LU, Y., YANG, Y., ZHAO, Q., ZHANG, C., UND LI, T. J.-J. 2024a. AI Assistance for UX: A Literature Review Through Human-Centered AI. University of Notre Dame, Cornell University, University of Washington und Georgia Institute of Technology.
- MATTER, P., WIEDEMANN, A., LEIKERT-BÖHM, N., UND PLEGER, L. E. 2024. The Reverse Turing Test in the Double Chinese Room: Observing GenAI Use in Practical Settings. In *Proceedings of DIGIT 2024*, AIS Electronic Library.
- MAYRING, P. 2022. Qualitative Inhaltsanalyse. Grundlagen und Techniken. 13., überarbeitete Auflage, Beltz Verlag, Weinheim und Basel.
- MESSERI, L. UND CROCKETT, M. J. 2024. Artificial intelligence and illusions of understanding in scientific research. *Nature* 627, 49-58.

- MOSQUEIRA-REY, E., HERNANDEZ-PEREIRA, E., ALONSO-RIOS, D., BOBES-BASCARAN, J., UND FERNANDEZ-LEAL, A. 2023. Human-in-the-loop machine learning: a state of the art. *Artificial Intelligence Review* 56, 3005-3054.
- PAGE, M. J., MCKENZIE, J. E., BOSSUYT, P. M., BOUTRON, I., HOFFMANN, T. C., MULROW, C. D., ET AL. 2021. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 372, n71.
- REN, S., XIE, C., JIAN, P., REN, Z., LENG, C., UND ZHANG, J. 2025. Towards Scientific Intelligence: A Survey of LLM-based Scientific Agents. *arXiv Preprint arXiv:2503.24047*. Chinese Academy of Sciences. Hinweis: arXiv-Preprint, nicht peer-reviewed.
- SANDERS, E. B.-N. UND STAPPERS, P. J. 2008. Co-creation and the new landscapes of design. *CoDesign* 4, 1, 5-18.
- SCHMIDGALL, S., SU, Y., WANG, Z., SUN, X., WU, J., YU, X., LIU, J., LIU, Z., UND BARSOUM, E. 2025. Agent Laboratory: Using LLM Agents as Research Assistants. *arXiv Preprint arXiv:2501.04227*. AMD, Johns Hopkins University und ETH Zurich. Hinweis: arXiv-Preprint, nicht peer-reviewed.
- SHNEIDERMAN, B. 2020. Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-Centered AI Systems. *ACM Transactions on Interactive Intelligent Systems* 10, 4, Article 26.
- SPINUZZI, C. 2005. The Methodology of Participatory Design. *Technical Communication* 52, 2, 163-174.
- STIGE, A., ZAMANI, E. D., MIKALEF, P., UND ZHU, Y. 2024. Artificial Intelligence (AI) and User Experience (UX) Design: A Systematic Literature Review and Future Research Agenda. *Information Technology & People*.
- VAN DIS, E. A. M., BOLLEN, J., ZUIDEMA, W., VAN ROOIJ, R., UND BOCKTING, C. L. 2023. ChatGPT: Five priorities for research. *Nature* 614, 224-226.
- VOM BROCKE, J., SIMONS, A., NIEHAVES, B., RIEMER, K., PLATTFAUT, R., UND CLEVEN, A. 2009. Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In *Proceedings of the 17th European Conference on Information Systems (ECIS 2009)*, Verona, Italien.
- WANG, H., FU, T., DU, Y., ET AL. 2023. Scientific discovery in the age of artificial intelligence. *Nature* 620, 47-60.
- WEBSTER, J. UND WATSON, R. T. 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly* 26, 2, S. xiii-xxiii.

Anhang C: Stakeholder-Analyse-Map



Anhang D: Risikoanalyse

Risikoanalyse - Tabelle 1													
ID	Gruppe (nur für Übersicht)	Risikotyp	Kurzbezeichnung	Ursache	Eutrittswahrscheinlichkeit	Auswirkung	Gesamtschätzung	Frühindikatoren	Gegenmassnahmen / Prävention	Reaktion bei Eintritt	Status	Restrisiko	Bemerkung
R04	a	Projektrisiko	Unklare Abgrenzung zwischen Assessment und Analyseprozess	Fachlicher Bezugsrahmen ist das Assessment, Untersuchungsgegenstand aber nur der Analyseprozess. Breiter Projektumfang trotz 300h-Rahmen, Iteration, Prototyping, Testing und schriftliche Arbeit konkurrieren um Zeit	mittel	mittel	mittel	Diskussionen driften in Erhebung, Präsentation oder Betriebsfragen ab	Beim schriftlichen Teil der Arbeit dem separierten Scoping für die Analyse innerhalb des Assessments	Nicht zum Gegenstand gehörende Punkte explizit ausklammern via eigenem Scoping	Beobachten	Besonders relevant für Kapitelstruktur und Evaluation.	Die Trennschärfe ist noch nicht klar genug. Darauf muss geachtet werden.
R01	Aufwand	Projektrisiko	Unterschätzung des Gesamtaufwands	Meilensteine verschieben sich, Dokumentation bleibt zurück, Evaluation und akademisches konkurrieren um Zeit	hoch	hoch	hoch	Meilensteine verschieben sich, Dokumentation bleibt zurück, Evaluation und akademisches konkurrieren um Zeit	Stakeholderanalyse durchführen um bestmögliche Probanden zu identifizieren und frühzeitige Rekrutierung.	Teilweise eingetreten - Konsequenzen wurden gezogen in anderen Risiken (R14, R02)	Beobachten	Besonders kritisch, da der Grundlegende Aufwand und Komplexität erhalten bleibt, es wird nur zeitlich geglättet.	Teilweise eingetreten. Zugegebenmassen habe ich es nicht wirklich unterschätzt. Es wurde genauso zu viel, wie ich
R13	c	Evaluationsrisiko	Stichprobe für finales User-Testing zu klein oder zu homogen	Rekrutierung aus engem internen Umfeld, wenige Rollen vertreten	hoch	mittel	hoch	Wenige Testfälle, ähnliche Profile, eingeschränkte Varianz	Einbeziehen neuer Mitarbeiter oder Projektleiter Mitarbeiter welche noch keine Assessments durchführten	Stakeholderanalyse durchführen um bestmögliche Probanden zu identifizieren und frühzeitige Rekrutierung.	beobachten	Trotz der Möglichkeiten das grundlegende Risiko durch gezieltes Rekrutieren zu vermeiden, kann es einfach sein, dass Wunschkandidaten	Ausreichende Nutzende
R05	d	Projektrisiko	Verzögerung durch akute Gesundheitsprobleme und Komplikationen bei Therapie	Es können im Zeitraum der Arbeit Therapiewechsel für die automunierkrankung des Autors stattfinden, welche eingeschränktes	mittel	hoch	mittel	Therapiewechsel wird angekündigt, Symptome kommen schlechend.	Buffer im Projektplan einbauen und Absprache mit Ärzten zur Reduktion der Symptomenrisiken.	Rechtzeitige Kommunikation mit Betreuung und Studienleitung falls termingerichteter Abschluss gefährdet ist.	Nicht eingetreten	Zu viel Verzögerung wegen eingeschränkter Arbeitsfähigkeit.	Bleib aus. Medikamentenwechsel wäre nötig gewesen und Symptome keimen auf, aber nicht beeinträchtigend. Es
R14	Engmaschiges einbinden	Evaluationsrisiko	Iterationen bleiben aus Zeitgründen unzureichend	HCD verlangt Rückmeldungen, Überarbeitung und erneute Prüfung, Zeitdruck verkürzt	mittel	hoch	mittel	Nur ein Testzyklus, Korrekturen bleiben oberflächlich	Minimale Iterationsschleifen fest einplanen, frühe Prototypen einsetzen.	Eingetreten - Ein geplanter Iterationsschritt muss ausgelassen werden. Testing findet mit Concierge-Service	Eingetreten	ISO 9241-210 fordert Zeit und Ressourcen für Iteration, Feedback und Evaluation.	Eine frühe Prototypisierung und eine Form von mehrfacher Evaluation sind eingeleitet es ist nur die
R02	Engmaschiges einbinden	Projektrisiko	Verzögerung durch Einzelarbeit	Keine Verteilung von Analyse-, Konzeptions- und Evaluationsaufgaben, reduzierte Perspektivenbreite.	hoch	hoch	hoch	Engpässe bei Parallelaufgaben, langsame Iteration, fehlendes Korrektiv	Höher Frequenziertere Zwischenvalidierungen mit Praxispartnern und stärkere Einbindung von Stakeholder bei nicht akademischen	Teilweise eingetreten - Prototyp bauen ausserhalb eigentlicher Masterarbeit Zeit hat zu viel Zeit beansprucht. Extra ferialen eingeleitet zum	abgesichert	Gleichzeitig methodisches und organisatorisches Risiko wegen aktiver Einbindung der Stakeholder in potenziell Vitale gestalterische Aspekte.	Etwas so chaotisch wie erwartet.
R03	Engmaschiges einbinden	Projektrisiko	Eingeschränkter Zugang zu Testpersonen	Abhängigkeit von M365-Spezialisten und weiteren Beteiligten für User-Testings und Feedback.	mittel	hoch	hoch	Schwierige Terminfindung, geringe Rücklaufquote, kurzfristige Absagen	Verfügbarkeit über Personalplanung sicherstellen. Garantierte Zeiten auf das Projekt für den Klare Human-in-the-loop, Logik, Reviewpunkte und Arbeitsaufwand pro Teilschritt messen. Kontrollaufwand explizit mitrechnen. Effektiven Usability-Test zur vollständigen Durchführung	abgesichert	Direkt relevant für Nutzbarkeitsbewertung.	Das Interesse war gross und es gab auch welche, die Zeit hatten. Es kam sogar zu einem Abbruch und es Aufgabenverteilung scheint vom Konzept her sinnvoll und	
R06	Engmaschiges einbinden	Prozessrisiko	Ungünstige Aufgabenverteilung zwischen Mensch und KI	Automatisierung wird zu weit oder an der falschen Stelle	mittel	hoch	hoch	Fachpersonen müssen viel nacharbeiten, Unsicherheit beim Evaluieren der Prozessschritte. M365-Spezialisten empfinden den Prozess als anstrengend.	Kritische Schritte wieder stärker manuell absichern	beobachten	ISO 9241-210 verlangt bewusste		
R09	Engmaschiges einbinden	Usability-Risiko	Fehlende tatsächliche Effizienz- oder Zufriedenheitssteigerung	KI spart Einzelaufwand, erzeugt aber Mehraufwand für Vorbereitung, Kontrolle und Korrektur	mittel	hoch	hoch	Arbeitsaufwand pro Teilschritt messen. Kontrollaufwand explizit mitrechnen. Effektiven Usability-Test zur vollständigen Durchführung	Eine Form von Engmaschiges Evaluation oder Feedbackschleifen um Teilschritte zu untersuchen. Dieses Risiko betrifft einen Punkt, der von Anfang an zu	abgesichert	Es kann sein, dass aus nicht klar identifizierbaren Gründen, die erwünschte Solizität nicht erreicht wird. Auch kann der neue Prozess als umständlich und dadurch	Ziel wurde auch im Auge des Testpersonals erreicht.	
R11	Engmaschiges einbinden	Usability-Risiko	Geringe Akzeptanz bei M365-Spezialisten	Prozess passt nicht zu realen Arbeitslogiken, ist zu komplex oder wirkt nicht vertrauenswürdig	gering	hoch	gering	Skeptische Rückmeldungen, hoher Erklärungsbedarf, Umgehungsverhalten	Engmaschiges einbeziehen der M365-Spezialisten in die Ausarbeitung des neuen Analyseprozesses.	Nicht eingetreten	Es könnte sein, dass technische Limitationen oder noch unbekannte Faktoren den Analyseprozess derart umgestalten, dass er sich von der tatsächlichen	Die Testpersonen und die partizipativen Probanden zeigen grosse Akzeptanz und Interesse.	
R12	Engmaschiges einbinden	Usability-Risiko	Zu hoher Schulungs- oder Einführungsbedarf	Fachpersonen sind nicht primär Research-Spezialisten. Zusätzlicher Prozess darf ihre Hauptarbeit nicht erschweren	gering	mittel	mittel	Testpersonen benötigen viel Anleitung, Fehler häufen sich zu Beginn	KI Unterstützung auf notwendige Schritte reduzieren und Prozess an Kernfähigkeiten ausrichten.	Nicht eingetreten	Der Prozess selbst kann einfach zu anfang für Fehler sein und zu viel verlangen um diese zu erkennen oder	Gerade mal eine Stunde Schulung und ein oder zwei Fragen beantwortet während der Durchführung, haben	
R07	Engmaschiges einbinden	Prozessrisiko	Zu geringe Nachvollziehbarkeit der KI-Auswertung	LLM-basierte Verarbeitung verdichtet Daten und trifft Entscheidungen, ohne dass Ableitungsschritte ausreichend transparent sind	hoch	hoch	hoch	Ergebnisse wirken plausibel, lassen sich aber schwer begründen	Eine Form von Nachvollziehbarkeit ist ein früh identifiziertes Risiko, welches von Anfang an zu den Grundanforderungen an den Prozess zählen. Daher	Teilweise eingetreten - Die Umfrage ist zu kompliziert um nachvollzogen zu werden und muss angepasst werden	Trotz Dokumentation und Review bleibt die vollständige Erklärbarkeit KI-gestützter Verdichtung begrenzt. Eine abschliessende fachliche Prüfung durch Menschen	Technisch gesehen konnten die Testpersonen nachvollziehen, auswerten und Beantworten der Anpassungen. Doch wegen des Umfangs und nicht	
R08	f	Prozessrisiko	Unzureichende Ergebnisverlässlichkeit	Schwankende inhaltlich Ergebnisse und zu stark variierende Interpretation, bei wiederholter Analyse identischer Datensätze	mittel	hoch	hoch	Kappa-Wert bleibt unter Zielwert, starke Streuung zwischen Durchläufen, inhaltliche und interpretative Unterschiede sind bei	Kleinsteigende Analyseschritte, manuell Quellen und Ergebnis abgleichen und Aufbau der Logik in kurzen Iterationen.	Teilweise eingetreten - Repetitive Evidenz wegen zu hoher Komplexität. Es könnte aber gegen gearbeitet und abgewendet werden	Trotz Kontrollen und Absicherungen, bleibt die Chance auf Ausreisser, welche falsche oder unstimmmige Ergebnisse	Die Verlässlichkeit wurde durch KI-Spezialisten geprüft und bestätigt.	
R10	f	Prozessrisiko	Zu starke Abhängigkeit von M365-Copilot-Funktionalitäten	Funktionsumfang, Grenzen oder Bedienlogik der verfügbaren Tools passen nicht zu den Anforderungen oder erfüllen diese nicht in ausreichendem Masse.	mittel	mittel	mittel	Workarounds häufen sich, manuelle Zwischenschritte nehmen zu	Frühe Tool-Erkundung, Alternativen, Toolgrenzen ausloten und dokumentieren. Absprachen mit internen Experten. Diese Exploration muss in	Eingetreten - Verfügbare Copilot Studio-Umgebung ist nicht optimal für den Bau des Prototypen.	Es könnte keine zufriedenstellende Lösung im Rahmen der Arbeit gefunden werden, welche alle Anforderungen erfüllt	Dieser Punkt ist für das eigentliche Projekt relativ fatal doch im Kontext der Masterarbeit, für welche hier Priorisiert ist, ist es im besten Falle Umständlich oder	
R15	Quelle	Qualitätsrisiko	Unzureichende Quellenkritik im KI-Teil	Forschungsfeld ist noch sehr neu und dynamisch. Qualität und Haltbarkeit von Quellen variieren stark und viele Quellen sind daher potenziell	hoch	mittel	hoch	Widersprüchliche Literatur, unsichere Begriffsverwendung, schwache Begründungen	Primärquellen bevorzugen, Publikationsqualität verstärken, Aussagen und Rückschlüsse vorsichtig formulieren.	Nicht eingetreten	Dieses Risiko lässt sich in diesem Sinne nicht beheben, nur damit umgehen. Es wurde bereits in der Themenbeschreibung als	ich habe getan was ich konnte und denke, ich habe das ausreichend erarbeitet.	
R16	h	Projektrisiko	Verzerrung durch bei Datenerhebung durch Doppelrolle im Unternehmenskontext.	Nähe zum Assessment, eigene Beteiligung am bestehenden Verfahren und zugleich Untersuchung desselben	hoch	mittel	mittel	Annahmen bleiben ungeprüft, kritische Punkte werden zu spät erkannt, grundlegende Fragen werden nicht gestellt oder Antworten nicht	Besondere Vorsicht beim Methodischen vorgehen. Kritische und Explorative Reflexion, klare Kommunikation der	offen	Kein Ausschlusskriterium, aber reflektionsbedürftig. Vorwurf von Fachlehre, wenn nicht klar genug erklärt.	Zwar bin ich überzeugt, dass ich ihn vermeiden konnte, aber diesen Punkt stelle ich prinzipiell nicht auf Grün. Denn die Natur dieses Risikos	
R17	i	Projektrisiko	Blinde Flecken in der Risikoanalyse	Da es sich um eine Einzelarbeit handelt, besteht die chance, dass möglich Risiken zu einseitig betrachtet werden und erweiterte Sichtweisen fehlen.	mittel	mittel	mittel	Experimentelle Nutzung von generativen KI-Systemen zum gemeinsamen Ergänzender Risiken Unverbindliche Besprechung mit Stakeholdern um Einschätzungen zu Risiken einzuholen.	Experimentelle Nutzung von generativen KI-Systemen zum gemeinsamen Ergänzender Risiken Unverbindliche Besprechung mit Stakeholdern und Betreuung, sowie achtsamkeit während dem Projekt	beobachten	Es kann ein unvorhergesehenes Risiko eintreten.	Es ist soweit nichts Relevantes eingetreten, was nicht schon ansonsten war	

Anhang E: Gesprächsleitfaden Stakeholder-Interview Consultant

Gesprächsleitfaden- Stakeholder Interview - Consultant

- 1. Du bist Teil des EX-Circles. Erzähle mir kurz, welche Rollen du hast und was du machst.**
 - a. Welches sind deine häufigsten Tätigkeiten und Aufgaben? – Beispiele?
 - b. Welche Bedeutung hat das EX Assessment für deine Arbeit?
 - i. Alternativ auch Ergebnisse aus dem Assessment?
- 2. Welche Rolle spielt das Assessment NACH dem Assessment?**
 - a. Wie sind die Endergebnisse wichtig?
 - b. Wie sind die Rohdaten/ Zusammenfassungen wichtig?
- 3. Wie ist es für dich, EX Assessments durchzuführen?**

Schilder mir bitte den Prozess, aus deiner Erinnerung.

 - a. Welche Erfahrungen wurden bis jetzt mit dem Assessment gemacht?
 - i. Was war besonders negativ?
 - ii. Was war besonders positiv?
 - b. Wie nimmst du das EX Assessment als Werkzeug für dich wahr?
 - i. Auch im Abgleich mit deiner gesamten Arbeitserfahrung im Fach.
 - c. Gibt es konkrete Erwartungen, welche nicht erfüllt wurden?
- 4. Das EX Assessment soll nun zu einem beschleunigten Light Assessment weiterentwickelt werden. Dabei sollen auch M365 eigene KI Lösungen eingesetzt werden. Hinsichtlich deiner Arbeit, im Abgleich mit deinen bisherigen Erfahrungen:**
 - a. Was sind deine Gedanken, wenn du von so einem «Light Assessment» hörst?
 - b. Was sind deine Gedanken zum KI-Einsatz für ein solches Assessment?
 - i. Nutzt du KI-Lösungen persönlich im Arbeitsalltag und wie?
 - ii. Was sind deine Gedanken, zu M365 KI-Lösungen?
- 5. Als jemand, der potenziell ein solches Light Assessment durchführen oder mit den Ergebnissen Arbeiten muss.**
 - a. Was sind Killerkriterien, welche das Light Assessment hinsichtlich der Durchführung erfüllen muss?
 - b. Gibt es aus deiner Erfahrung mit KI- und M365-Systemen konkrete Ansätze, welche du nicht sehen möchtest?
- 6. M365 erlaubt es, die gleichen Aufgaben auf viele verschiedene Arten zu lösen. Aus deiner persönlichen Sicht, deinen Präferenzen und deiner Erfahrung: Schilder mir einen idealen Ablauf eines Light Assessments, so wie vorher den aktuellen Prozess. (So grob oder detailliert, wie du möchtest.)**
- 7. Ging das Assessment oder Folgearbeiten wegen falsch Ergebnisse schon mal komplett schief?**
- 8. Gibt es noch etwas, dass du erwähnen möchtest?**

Anhang F: Gesprächsleitfaden Stakeholder-Interview Leitung EX Circle

Gesprächsleitfaden- Stakeholder Interview - Leitung EX Circle

- 1. Erzähl mir kurz von den Dienstleistungen des EX Circles.**
 - a. Du bist der Circle Lead des EX-Circles. Was sind deine Aufgaben als solcher?
 - b. Wo ist das EX Assessment in diesen Dienstleistungen verortet?/
Welche Bedeutung haben Assessments?
- 2. Was waren die Gedanken und Gründe, welche zum EX Assessment geführt haben?**
 - a. Welche Erfahrungen wurden bis jetzt mit dem Assessment gemacht?
 - i. Erfolge/Positives
 - ii. Fehlschläge/Negatives
- 3. Für dich in der Rolle als EX Circle Lead, welche sind die grössten Probleme mit dem Assessment im Moment, die du mitbekommst?**
 - a. Gibt es konkrete Erwartungen, welche nicht erfüllt wurden?
- 4. Das EX Assessment soll nun zu einem beschleunigten Light Assessment weiterentwickelt werden. Welche Bedeutung hat das Light Assessment für die Zukünftigen Pläne und Strategien des EX Circles?**
 - a. Welche Aufgaben hat das Light Assessment für den EX Circle zu erfüllen?
 - b. Was sind, von Seiten Strategie die Erwartungen an das Light Assessment?
 - c. Sind von Seiten Strategie konkrete Anforderungen definiert worden?
- 5. Eine konkrete Absicht für das Light Assessment ist die Verwendung von KI-Systemen im Prozess. Welche Bedeutung hat die Einbindung von KI-Systemen in Firmeneigene Prozesse für den EX Circle und GARAIO?**
- 6. Was sind deine persönlichen Erfahrungen beim Ausführen des bisherigen EX Assessments?**
 - a. Sowohl vor, während, als auch nach dem Assessment.
 - i. Erfolge/Positives
 - ii. Fehlschläge/Negatives
- 7. Als jemand, der potenziell ein Light Assessment durchführen oder mit den Ergebnissen Arbeiten muss. Was sind deine persönlichen Erwartungen oder Hoffnungen?**
 - a. Auch im Vergleich mit den bisherigen Erfahrungen.
- 8. Sowohl aus deiner persönlichen Sicht, als auch als EX Circle Lead. Welches sind Killerkriterien für das Light Assessment, hinsichtlich der Durchführung?**
- 9. Gibt es noch etwas, dass du erwähnen möchtest?**

Anhang G: Kodierleitfaden Hauptkategorien

Kodierleitfaden Hauptkategorien - Hauptkategorien

Kodierleitfaden Hauptkategorien (A2–L)

Deduktive Zuordnung des Materials zu acht Hauptkategorien HK2–HK9. Massgeblich ist die Definition, nicht der Kategorienname. Die Anlage ist nominal. HK1 wird nicht kodiert (Fallbeschreibung ausserhalb des Kategoriensystems). Status: durch den Menschen freigegeben (Details siehe Blatt «Kontext & Protokoll»).

Kat.	Titel	Definition	Ankerbeispiele (mit Belegstelle)	Kodierregeln
HK2	Bedeutung und Funktion des Assessments	Textstellen über das Assessment selbst, auch wertende, die seinen Stellenwert, seine Funktion und vor allem seine Auswirkungen betreffen, ausserhalb der eigentlichen Nutzung und Durchführung (z. B. Wirkung bei der Kundschaft, Standortbestimmung, Vertrauensaufbau, Bedeutung für Folgeprojekte, Rolle von Ergebnissen und Rohdaten).	Person02 3:52 — "Das Assessment ist enorm wichtig, wenn wir mit Kunden noch früh in Kontakt sind, den Kunden noch nicht gut kennen." Person01 0:56 — "Diese Messung findet eben auf der Ebene «Menschen» statt. Wir wollen wissen, was wirklich angekommen ist und was gelebt wird."	Abgrenzung zu HK5: HK2 erfasst Aussagen zum Assessment, auch wertende, zu Stellenwert, Funktion und Auswirkungen ausserhalb der konkreten Nutzung und Durchführung. HK5 erfasst Erfahrungswerte aus dem konkreten Einsatz und der Durchführung. Verbindet eine Stelle eine Auswirkung mit einer Einsatzerfahrung, Mehrfachkodierung [HK2 HK5].
HK3	Bestehender Analyseprozess: Ablauf und Werkzeuge (IST)	Textstellen, die den heutigen Ablauf des Assessments und insbesondere der Auswertung beschreiben: Arbeitsschritte, Reihenfolge, eingesetzte Werkzeuge (Excel, PowerPoint, Bericht), Schnittstellen und manuelle Tätigkeiten. Neutral-beschreibend (Ist-Zustand).	Person02 10:05 — "Ich führe ein Interview, fülle es in Excel ab, werte es noch aus, übertrage das ... auf PowerPoint, auf Word." Person01 32:06 — "Am Anfang haben wir alles manuell gemacht und mussten die Fragen manuell auswerten, man hat wenig Erfahrungsschatz gehabt, das ist der Klassiker."	Abgrenzung zu HK4: Reine Ablaufbeschreibung → HK3; benennt die Stelle den Ablauf zugleich als Problem (zu aufwendig, ineffizient), zusätzlich HK4 ([HK3 HK4]). Abgrenzung zu HK6: Ist-Beschreibung → HK3, Soll-/Wunschbeschreibung → HK6.
HK4	Schwachstellen und Pain Points	Textstellen, die Probleme, hohen Zeit- und Arbeitsaufwand, Ineffizienzen, manuelle Doppelarbeit, nicht erfüllte Erwartungen oder Belastungen im bestehenden Prozess benennen.	Person02 7:38 — "Diese Auswertung ist wahnsinnig zeitaufwendig. [...] das ist sicher nochmal die gleiche Zeit, wenn nicht das Doppelte oder mehr." Person01 15:13 — "Das, was ich mitbekommen habe, ist am Anfang sicher gewesen, dass das ein Riesenaufwand ist, das auszuwerten."	Abgrenzung zu HK9: Prozessbezogener Zeit-/Arbeitsaufwand → HK4; wirtschaftliche Rahmung (Preis, Verrechenbarkeit, Verkaufbarkeit, Marge) → HK9. Verknüpft die Stelle Aufwand ausdrücklich mit Kosten oder Verkaufbarkeit, Mehrfachkodierung [HK4 HK9].
HK5	Positive Erfahrungen und Stärken	Textstellen mit Fokus auf Erfahrungswerte aus dem konkreten Einsatz und der Durchführung des Assessments: positiv bewertete Aspekte und Stärken, die sich in der praktischen Anwendung zeigen (z. B. eingespielte Durchführung der Interviews, Nutzen der Standardisierung im Arbeiten, Framework als Orientierung in der Anwendung).	Person03 12:12 — "Von den Interviews: Das klappt unterdessen sehr gut. [...] das ist für mich jetzt kein grosser Stress mehr, die durchzuführen." Person03 41:43 — "Das als Modell beziehungsweise Framework hat mir auch geholfen ... das alles ein bisschen einzuordnen."	Abgrenzung zu HK2 siehe dort. Negative Einsatzerfahrungen gehören zu HK4, nicht zu HK5. Eine Stärke, die zugleich eine Anforderung an das Light Assessment begründet ("das muss erhalten bleiben"), zusätzlich HK6.
HK6	Anforderungen, Killerkriterien und Soll-Vorstellungen	Textstellen, die Bedürfnisse, Erwartungen, Wünsche, zwingende Kriterien (Killerkriterien) oder ideale Ablaufvorstellungen für das künftige Light Assessment ausdrücken. Soll-bezogen (was es leisten soll).	Person01 33:45 — "Es muss klipp und klar, gut strukturiert und ein transparentes Vorgehen sein, dass ich das wirklich idealerweise mit wenigen Ressourcen machen kann. [...] Idealerweise möchte ich das jemandem geben können und sagen: So, führ es durch." Person03 33:14 — "Ich habe da eine relativ gute Light-Assessment-Grundlage, mache einen Knopfdruck und habe dann gerade in meinem Projekt die relevanten Dokumente, die aktuellsten Dokumente schon drin [...] alles, was ich brauche."	Abgrenzung zu HK3 siehe dort. Bezieht sich der Wunsch auf die Aufgabenteilung Mensch-KI (z. B. "KI macht die Auswertung, ich validiere"), zusätzlich HK8 ([HK6 HK8]). Die spätere Trennung in Anforderungen/Killerkriterien und idealer Soll-Ablauf bleibt für Schritt 3 (Subkategorien) offen.
HK7	KI-Einsatz: Bedeutung, Haltung und Nutzung	Textstellen zu Bedeutung, Chancen, Skepsis und allgemeiner Einschätzung von KI sowie zur persönlichen oder firmenseitigen Nutzung von KI, Copilot und M365-KI im Arbeitsalltag.	Person01 23:35 — "Das Thema ist aktuell so ein wenig ein Hype-Thema, das Microsoft extrem auf der Roadmap hat." Person03 23:01 — "Wir haben ja im Geschäft Copilot jetzt auch die Lizenz dazu, und ich brauche das wirklich viel [...] viel für Meetings, für Zusammenfassungen, Aufgaben erstellen, das ist meiner Meinung nach der grösste Benefit."	Abgrenzung zu HK8: Allgemeine Haltung, Bedeutung oder Nutzung von KI → HK7; konkrete Aussagen zur Aufgabenteilung Mensch/KI oder zur Verlässlichkeit/Validierung/Halluzination → HK8.

Kodierleitfaden_Hauptkategorien - Hauptkategorien

Kodierleitfaden Hauptkategorien (A2–L)

HK8	Aufgabenteilung Mensch-KI, fachliche Verantwortung und Ergebnisverlässlichkeit	Textstellen dazu, welche Aufgaben automatisiert werden sollen oder können und welche beim Fachpersonal verbleiben müssen, sowie zu Validierung, Nachvollziehbarkeit, Halluzinationsrisiko und Verlässlichkeit der Ergebnisse.	Person02 21:26 — "Man muss einer KI natürlich auch immer ein bisschen entgegenreten [...] ich muss weiterhin die Fachperson bleiben und sicherstellen, dass das, was die KI ausspuckt, auch richtig ist, und das ein bisschen validieren." Person03 36:04 — "Zwischen den Zeilen lesen, Tonfälle, die Augen verdrehen, all diese Dinge, da braucht es uns immer noch."	Abgrenzung zu HK7 siehe dort. Im Zweifel, ob eine Stelle allgemeine KI-Haltung (HK7) oder konkrete Aufgabenteilung/Verlässlichkeit (HK8) ist, entscheidet, ob sie eine Grenze des KI-Einsatzes oder eine menschliche Kontrollaufgabe benennt, dann HK8.
HK9	Geschäftliche und wirtschaftliche Rahmenbedingungen	Textstellen zu Preis, Aufwand und Stunden im wirtschaftlichen Sinn, Wirtschaftlichkeit, Verrechenbarkeit, Skalierbarkeit und Verkaufbarkeit des Assessments als Rahmenbedingung.	Person01 18:21 — "Unser Assessment, das verkauft man etwa als 18000 Franken oder so [...] das ist so in einem Rahmen von 15000 bis 20000 Franken, in dem das noch verkaufbar ist [...] es darf eigentlich nichts teurer werden." Person02 17:01 — "Wir haben sehr viele Stunden aufgewendet, die zum Teil nicht verrechnet werden konnten, weil wir das Assessment meistens pauschal verkaufen, und diese Pauschalen werden dann oft überschossen."	Abgrenzung zu HK4 siehe dort. Der Teilaspekt Aufwand/Kosten kann in Schritt 3 als Subkategorie von HK9 ausdifferenziert werden.

Kodierleitfaden_Hauptkategorien - Kontext Protokoll

Kontext und Probedurchlauf-Protokoll	
Einordnung	
Zweck	Kodierleitfaden für die deduktive Zuordnung des Materials zu den acht festgelegten Hauptkategorien HK2–HK9. Je Hauptkategorie sind eine Definition, ein bis zwei Ankerbeispiele mit Belegstelle und, wo nötig, Kodierregeln
Grundsatz	Massgeblich für die Zuordnung ist die Definition, nicht der Kategoriennamen. Die Anlage ist nominal.
Status des Leitfadens	Vorschlag aus dem Probedurchlauf. Definitionen und Kodierregeln werden durch den Menschen festgelegt. Grundlage für den vollständigen Materialdurchgang (ANW)
HK1	Wird nicht kodiert (Fallbeschreibung ausserhalb des Kategoriensystems).
Probedurchlauf-Protokoll	
Kodierter Materialteil	Erstes Interview Person01 (Adrian) vollständig, als erster Teil des Materials. Repräsentativ, da es Bedeutung, Ist-Prozess, Aufwand, Stärken, KI-Haltung, Aufgabenteilung und Wirtschaftlichkeit berührt.
Vorgehen	Zuordnung repräsentativer Fundstellen anhand der Entwurfsdefinitionen; klar zuordenbare Stellen als Ankerbeispiele vorgeschlagen; nicht eindeutige Stellen als Abgrenzungsfälle in Kodierregeln überführt.
Abgrenzungsfall → [[HK3 HK4]]	Person01 32:06 ("Am Anfang haben wir alles manuell gemacht und mussten die Fragen manuell auswerten, man hat wenig Erfahrungsschatz gehabt, das ist der Klassiker." und zugleich "war mir bewusst, dass das Aufwand braucht") → Ist-Beschreibung und Pain Point zugleich,
Abgrenzungsfall → HK2	Person01 10:17 ("Was immer sehr gut angekommen ist, ist, dass wir auf den Menschen bezogen messen. [...] Wir engineeren 70% den Menschen und 30% ist Technik." als Ansatz/Funktion und "sehr gut angekommen" als Wirkung bei der Kundschaft) → nach geschärfte Regel HK2
Abgrenzungsfall → [[HK4 HK9]]	Person01 35:21 ("Wenn ich ein Assessment mit einem super Prozess und super Fragen habe und dann einer 2 Wochen investieren muss, das auszuwerten [...] es ist weniger das coole Tool, das ich wirklich kommerziell spannend einsetzen kann, es schmälert den Nutzen.",
Abgrenzungsfall → HK7 (bzw. HK8)	Person01 23:35 (allgemeine KI-Einschätzung mit Tendenz zu "zu viel Brainwork wegnehmen") → als allgemeine Haltung HK7; klare Aufgabenteilung/Kontrolle wäre HK8. Regel HK7/HK8 aufgenommen.

Kodierleitfaden_Hauptkategorien - Kontext Protokoll

Kontext und Probedurchlauf-Protokoll	
Revisionen	Keine Kategorie musste gestrichen oder umbenannt werden. Vier Abgrenzungsregeln wurden ergänzt (HK3/HK4, HK2/HK5, HK4/HK9, HK7/HK8). Keine früheren Probe-Zuordnungen wurden dadurch falsch.
Geklärt (Mensch, 2026-02-06) – HK8	HK8 bleibt als Hauptkategorie bestehen, obwohl im Probematerial Person01 schwach belegt; sie trägt vor allem in Person02 und Person03.
Geklärt (Mensch, 2026-02-06) – HK2/HK5	Abgrenzung geschärft: HK2 = Aussagen zum Assessment, auch wertend, zu Auswirkungen ausserhalb der eigentlichen Nutzung; HK5 = Erfahrungswerte bezüglich des Einsatzes. Definitionen und Kodierregeln entsprechend
Status	Kodierleitfaden durch den Menschen freigegeben.

Herleitung Bedürfnisse und Anforderungen - Komplette Herleitung

[illegible]

[illegible]

Hochschule Bielefeld und Änderungen - Komplette Neubildung															
Sinnesabschnitt-Nr.	Beleg (Person, Stelle)	Subkategorie(n)	Hauptkategorie(n)	Sinnesabschnitt (wörtliches Zitat, Schritt 3)	Paraphrasen-Nr.	Paraphrase (Schritt 4)	Begründung der NB-Zuordnung	Direkte NB-Zuordnung	Nutzerbedürfnis-Nummer	Nutzerbedürfnis-Name	Nutzerbedürfnis-Beschreibung	Nutzerbedürfnis-Erklärung	Kernbedürfnis-Nummer	Kernbedürfnis-Name	Kernbedürfnis-Beschreibung
S021	Person03.41.43	HK2 : Framework/Model als Orientierung	HK3 Positive Erfahrungen und Stärken	Das Ziel, Modell beziehungsweise Framework hat sich wohl gehalten, jetzt auch bei der Einarbeitung bei GARA, das alles ein bisschen engeordnen [...] ich bin halt jemand, der...	P10	Das Modell, das EX-Häuschen, gibt Struktur und Orientierung, hilft beim Einordnen der eigenen Arbeit und dient in Schaltungen als Hilfenetz.	Kundenverständnis als Anschluss an Kundenstammen	Zuguteilt	NB-03	In heterogenen Daten rasch Orientierung gewinnen	Als M365-Fachperson brauche ich nach intensiven Umfragen und technischen Anforderungen, welche Themen, Muster und Strukturen mich interessieren und die weitere...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-02	Orientierung und fachliche Verdichtung	M365-Fachpersonen brauchen eine Orientierung, die heterogenen Daten in eine verständliche fachliche Ordnung bringt, zentrale Muster und Stammen nutzbar macht und den Ist...
S021	Person03.41.43	HK2 : Framework/Model als Orientierung	HK3 Positive Erfahrungen und Stärken	beziehungswiese Framework Orientierung, hilft beim Einordnen der eigenen Arbeit bei GARA, das alles ein bisschen engeordnen [...] ich bin halt jemand, der...	P10	Das Modell, das EX-Häuschen, gibt Struktur und Orientierung, hilft beim Einordnen der eigenen Arbeit und dient in Schaltungen als Hilfenetz.	Kundenverständnis als Anschluss an Kundenstammen	Zuguteilt	NB-04	Den Ist-Zustand mit Zielbild und Framework verbinden	Als M365-Fachperson brauche ich eine Orientierung, die den Ist-Zustand mit dem Zielbild und dem Assessment-Framework in Beziehung zu...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-02	Orientierung und fachliche Verdichtung	M365-Fachpersonen brauchen eine Orientierung, die den Ist-Zustand mit dem Zielbild und dem Assessment-Framework in Beziehung zu...
S021	Person03.41.43	HK2 : Framework/Model als Orientierung	HK3 Positive Erfahrungen und Stärken	beziehungswiese Framework Orientierung, hilft beim Einordnen der eigenen Arbeit bei GARA, das alles ein bisschen engeordnen [...] ich bin halt jemand, der...	P10	Das Modell, das EX-Häuschen, gibt Struktur und Orientierung, hilft beim Einordnen der eigenen Arbeit und dient in Schaltungen als Hilfenetz.	Kundenverständnis als Anschluss an Kundenstammen	Zuguteilt	NB-06	Ergebnisse fachlich vertreten können	Als M365-Fachperson brauche ich eine Orientierung, die den Ist-Zustand mit dem Zielbild und dem Assessment-Framework in Beziehung zu...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-02	Fachliche Verantwortung und Beurteilbarkeit	M365-Fachpersonen brauchen eine Orientierung, die den Ist-Zustand mit dem Zielbild und dem Assessment-Framework in Beziehung zu...
S022	Person03.41.43	HK2 : Framework/Model als Orientierung	HK3 Positive Erfahrungen und Stärken	beziehungswiese Framework Orientierung, hilft beim Einordnen der eigenen Arbeit bei GARA, das alles ein bisschen engeordnen [...] ich bin halt jemand, der...	P10	Das Modell, das EX-Häuschen, gibt Struktur und Orientierung, hilft beim Einordnen der eigenen Arbeit und dient in Schaltungen als Hilfenetz.	Kundenverständnis als Anschluss an Kundenstammen	Zuguteilt	NB-03	In heterogenen Daten rasch Orientierung gewinnen	Als M365-Fachperson brauche ich nach intensiven Umfragen und technischen Anforderungen, welche Themen, Muster und Strukturen mich interessieren und die weitere...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-02	Orientierung und fachliche Verdichtung	M365-Fachpersonen brauchen eine Orientierung, die heterogenen Daten in eine verständliche fachliche Ordnung bringt, zentrale Muster und Stammen nutzbar macht und den Ist...
S022	Person03.41.43	HK2 : Framework/Model als Orientierung	HK3 Positive Erfahrungen und Stärken	beziehungswiese Framework Orientierung, hilft beim Einordnen der eigenen Arbeit bei GARA, das alles ein bisschen engeordnen [...] ich bin halt jemand, der...	P10	Das Modell, das EX-Häuschen, gibt Struktur und Orientierung, hilft beim Einordnen der eigenen Arbeit und dient in Schaltungen als Hilfenetz.	Kundenverständnis als Anschluss an Kundenstammen	Zuguteilt	NB-04	Den Ist-Zustand mit Zielbild und Framework verbinden	Als M365-Fachperson brauche ich eine Orientierung, die den Ist-Zustand mit dem Zielbild und dem Assessment-Framework in Beziehung zu...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-02	Orientierung und fachliche Verdichtung	M365-Fachpersonen brauchen eine Orientierung, die den Ist-Zustand mit dem Zielbild und dem Assessment-Framework in Beziehung zu...
S022	Person03.41.43	HK2 : Framework/Model als Orientierung	HK3 Positive Erfahrungen und Stärken	beziehungswiese Framework Orientierung, hilft beim Einordnen der eigenen Arbeit bei GARA, das alles ein bisschen engeordnen [...] ich bin halt jemand, der...	P10	Das Modell, das EX-Häuschen, gibt Struktur und Orientierung, hilft beim Einordnen der eigenen Arbeit und dient in Schaltungen als Hilfenetz.	Kundenverständnis als Anschluss an Kundenstammen	Zuguteilt	NB-06	Ergebnisse fachlich vertreten können	Als M365-Fachperson brauche ich eine Orientierung, die den Ist-Zustand mit dem Zielbild und dem Assessment-Framework in Beziehung zu...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-02	Fachliche Verantwortung und Beurteilbarkeit	M365-Fachpersonen brauchen eine Orientierung, die den Ist-Zustand mit dem Zielbild und dem Assessment-Framework in Beziehung zu...
S023	Person01.23.35	HK7 : Einschätzung von Mehrwert und Nutzen	HK7 KI Einmal: Bedienung, Nutzung und Nutzen	Ich sehe den Prozess mehrwertig, wenn ich in der KI, wo das man die Prozess unterstützt werden [...] Labor das ist nicht gerade in unserer...	P11	Der Prozess mehrwertig, wenn ich in der KI, wo das man die Prozess unterstützt werden [...] Labor das ist nicht gerade in unserer...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Berücksichtigt - kein Bedürfnis							
S024	Person02.45.7	HK3 : Auftragsanbahnung und Setup	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge	Der Einstieg erfolgt über eine Eintragsform und einen Vorbesprechungsprozess, der dann natürlich einen Auftragsprozess mit sich bringt...	P12	Der Einstieg erfolgt über eine Eintragsform und einen Vorbesprechungsprozess, der dann natürlich einen Auftragsprozess mit sich bringt...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Berücksichtigt - kein Bedürfnis							
S025	Person03.15.41	HK9 : Wirtschaftlichkeit und Verarbeitbarkeit des Aufwands	HK9 Geschäftliche und wirtschaftliche Rahmenbedingungen	Der Aufwand ist, was ich machen muss, wir sind ja in der Wirtschaft, der Aufwand ist, was ich machen muss, wir sind ja in der Wirtschaft...	P13	Der Aufwand ist, was ich machen muss, wir sind ja in der Wirtschaft, der Aufwand ist, was ich machen muss, wir sind ja in der Wirtschaft...	Einschätzung von Mehrwert und Grenzen, Kontext, stützt nicht NB-07	Zuteilung - kein direktes Nutzer-Bedürfnis	NB-01	Assessmentsdaten in eine tragfähige Einschätzung überführen	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S025	Person03.15.41	HK9 : Wirtschaftlichkeit und Verarbeitbarkeit des Aufwands	HK9 Geschäftliche und wirtschaftliche Rahmenbedingungen	Der Aufwand ist, was ich machen muss, wir sind ja in der Wirtschaft, der Aufwand ist, was ich machen muss, wir sind ja in der Wirtschaft...	P13	Der Aufwand ist, was ich machen muss, wir sind ja in der Wirtschaft, der Aufwand ist, was ich machen muss, wir sind ja in der Wirtschaft...	Einschätzung von Mehrwert und Grenzen, Kontext, stützt nicht NB-07	Zuteilung - kein direktes Nutzer-Bedürfnis	NB-12	Das Assessment auch in begrenzten Projektphasen einsetzen können	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S026	Person02.17.01	HK6 : Effizienz und Zeitersparnis (implizente Nutzer-Effizienz)	HK6 Änderungen, Klientenfragen und Soll-Vorstellungen	Meine Idee war immer: Ich will es streamline haben, dann soll das Ganze sehr schnell sein, weil ich möchte...	P14	Der Prozess soll in der eigenen Verwendung schneller gehen und Zeit sparen, sodass die Auswertung rascher verfügbar ist...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Zuguteilt	NB-01	Assessmentsdaten in eine tragfähige Einschätzung überführen	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S026	Person02.17.01	HK6 : Effizienz und Zeitersparnis (implizente Nutzer-Effizienz)	HK6 Änderungen, Klientenfragen und Soll-Vorstellungen	Meine Idee war immer: Ich will es streamline haben, dann soll das Ganze sehr schnell sein, weil ich möchte...	P14	Der Prozess soll in der eigenen Verwendung schneller gehen und Zeit sparen, sodass die Auswertung rascher verfügbar ist...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Zuguteilt	NB-02	Die eigene Aufmerksamkeits auf fachlich wertvolle Arbeit richten	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S027	Person02.17.01	HK6 : Effizienz und Zeitersparnis (implizente Nutzer-Effizienz)	HK6 Änderungen, Klientenfragen und Soll-Vorstellungen	Meine Idee war immer: Ich will es streamline haben, dann soll das Ganze sehr schnell sein, weil ich möchte...	P14	Der Prozess soll in der eigenen Verwendung schneller gehen und Zeit sparen, sodass die Auswertung rascher verfügbar ist...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Zuguteilt	NB-01	Assessmentsdaten in eine tragfähige Einschätzung überführen	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S027	Person02.17.01	HK6 : Effizienz und Zeitersparnis (implizente Nutzer-Effizienz)	HK6 Änderungen, Klientenfragen und Soll-Vorstellungen	Meine Idee war immer: Ich will es streamline haben, dann soll das Ganze sehr schnell sein, weil ich möchte...	P14	Der Prozess soll in der eigenen Verwendung schneller gehen und Zeit sparen, sodass die Auswertung rascher verfügbar ist...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Zuguteilt	NB-02	Die eigene Aufmerksamkeits auf fachlich wertvolle Arbeit richten	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S027	Person02.17.01	HK6 : Effizienz und Zeitersparnis (implizente Nutzer-Effizienz)	HK6 Änderungen, Klientenfragen und Soll-Vorstellungen	Meine Idee war immer: Ich will es streamline haben, dann soll das Ganze sehr schnell sein, weil ich möchte...	P14	Der Prozess soll in der eigenen Verwendung schneller gehen und Zeit sparen, sodass die Auswertung rascher verfügbar ist...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Zuguteilt	NB-12	Das Assessment auch in begrenzten Projektphasen einsetzen können	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S028	Person01.15.13	HK6 : Effizienz und Zeitersparnis (implizente Nutzer-Effizienz)	HK6 Änderungen, Klientenfragen und Soll-Vorstellungen	Meine Idee war immer: Ich will es streamline haben, dann soll das Ganze sehr schnell sein, weil ich möchte...	P14	Der Prozess soll in der eigenen Verwendung schneller gehen und Zeit sparen, sodass die Auswertung rascher verfügbar ist...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Zuguteilt	NB-01	Assessmentsdaten in eine tragfähige Einschätzung überführen	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S028	Person01.15.13	HK6 : Effizienz und Zeitersparnis (implizente Nutzer-Effizienz)	HK6 Änderungen, Klientenfragen und Soll-Vorstellungen	Meine Idee war immer: Ich will es streamline haben, dann soll das Ganze sehr schnell sein, weil ich möchte...	P14	Der Prozess soll in der eigenen Verwendung schneller gehen und Zeit sparen, sodass die Auswertung rascher verfügbar ist...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Zuguteilt	NB-02	Die eigene Aufmerksamkeits auf fachlich wertvolle Arbeit richten	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S029	Person03.25.42	HK6 : Effizienz und Zeitersparnis (implizente Nutzer-Effizienz)	HK6 Änderungen, Klientenfragen und Soll-Vorstellungen	Meine Idee war immer: Ich will es streamline haben, dann soll das Ganze sehr schnell sein, weil ich möchte...	P14	Der Prozess soll in der eigenen Verwendung schneller gehen und Zeit sparen, sodass die Auswertung rascher verfügbar ist...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Zuguteilt	NB-12	Das Assessment auch in begrenzten Projektphasen einsetzen können	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S029	Person03.25.42	HK6 : Effizienz und Zeitersparnis (implizente Nutzer-Effizienz)	HK6 Änderungen, Klientenfragen und Soll-Vorstellungen	Meine Idee war immer: Ich will es streamline haben, dann soll das Ganze sehr schnell sein, weil ich möchte...	P14	Der Prozess soll in der eigenen Verwendung schneller gehen und Zeit sparen, sodass die Auswertung rascher verfügbar ist...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Zuguteilt	NB-01	Assessmentsdaten in eine tragfähige Einschätzung überführen	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S029	Person03.25.42	HK6 : Effizienz und Zeitersparnis (implizente Nutzer-Effizienz)	HK6 Änderungen, Klientenfragen und Soll-Vorstellungen	Meine Idee war immer: Ich will es streamline haben, dann soll das Ganze sehr schnell sein, weil ich möchte...	P14	Der Prozess soll in der eigenen Verwendung schneller gehen und Zeit sparen, sodass die Auswertung rascher verfügbar ist...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Zuguteilt	NB-02	Die eigene Aufmerksamkeits auf fachlich wertvolle Arbeit richten	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S029	Person03.25.42	HK6 : Effizienz und Zeitersparnis (implizente Nutzer-Effizienz)	HK6 Änderungen, Klientenfragen und Soll-Vorstellungen	Meine Idee war immer: Ich will es streamline haben, dann soll das Ganze sehr schnell sein, weil ich möchte...	P14	Der Prozess soll in der eigenen Verwendung schneller gehen und Zeit sparen, sodass die Auswertung rascher verfügbar ist...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Zuguteilt	NB-12	Das Assessment auch in begrenzten Projektphasen einsetzen können	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...
S030	Person03.31.14	HK6 : Effizienz und Zeitersparnis (implizente Nutzer-Effizienz)	HK6 Änderungen, Klientenfragen und Soll-Vorstellungen	Meine Idee war immer: Ich will es streamline haben, dann soll das Ganze sehr schnell sein, weil ich möchte...	P14	Der Prozess soll in der eigenen Verwendung schneller gehen und Zeit sparen, sodass die Auswertung rascher verfügbar ist...	Effizienz und Zeitersparnis, also tragfähige Analyse mit vertretbarem Aufwand	Zuteilung - kein direktes Nutzer-Bedürfnis		Assessmentsdaten in eine tragfähige Einschätzung überführen	Als M365-Fachperson brauche ich eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...	Dieses Bedürfnis ist stärker bedient, da ich mich in der heterogenen Daten in eine verständliche fachliche Ordnung bringe, zentrale Muster und Stammen nutzbar macht und den Ist...	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, die heterogenen Daten in eine fachlich tragfähige Einschätzung zu überführen...

[illegible]

[illegible]

[illegible]

Hilfsbedürfnisse und Anforderungen - Komplette Herleitung															
Sinnesabschnitt-Nr.	Beleg (Person, Stelle)	Subkategorie(n)	Hauptkategorie(n)	Sinnesabschnitt (wörtliches Zitat, Schritt 3)	Paraphrasen-Nr.	Paraphrase (Schritt 4)	Begründung der NB-Zuordnung	Direkte NB-Zuordnung	Nutzerbedürfnis-Nummer	Nutzerbedürfnis-Name	Nutzerbedürfnis-Schilderung	Nutzerbedürfnis-Erklärung	Kernbedürfnis-Nummer	Kernbedürfnis-Name	Kernbedürfnis-Schilderung
5058	Person02.13.06		HK7.3 Grundhaltung: positiv mit Skepsis und Entschönerung Haltung und Nutzung	HK7 KI Einsatz: Bedienung, mit Skepsis und Entschönerung Haltung und Nutzung		Also, generell ja, sehr gut, dass das die Idee ist, dass es gemacht werden soll.									
5059	Person03.21.01		HK7.3 Grundhaltung: positiv mit Skepsis und Entschönerung Haltung und Nutzung	HK7 KI Einsatz: Bedienung, mit Skepsis und Entschönerung Haltung und Nutzung		Ich bin gut eingestellt, positiv mit Skepsis, die man haben muss.									
5060	Person01.23.35		HK7.3 Grundhaltung: positiv mit Skepsis und Entschönerung Haltung und Nutzung	HK7 KI Einsatz: Bedienung, mit Skepsis und Entschönerung Haltung und Nutzung		Alle wollen es gebrauchen, alle wollen es nutzen und trotzdem ist das noch ein zugleiches Entschöneren, das ist jetzt wirklich Entschöneren, das ich gut finde.									
5061	Person01.13.04		HK4.3 Interviewführung	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Wir haben einstündige Analysegespräche, aber auch mit verschiedenen Personen und managements, aber auch die, die ich interviewe.									
5062	Person03.13.58		HK4.3 Interviewführung	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Angenehm und wenn man sich nervt, und es ist ein bisschen, die Angst, ihnen die Angst zu nehmen.									
5063	Person02.31.50		HK3.2 Teilnehmerauswahl und Durchwahl	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Aus der Erfahrung heraus ist es wichtig, dass man einen guten Teilnehmerauswahl hat.									
5064	Person02.31.50		HK3.2 Teilnehmerauswahl und Durchwahl	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Aus der Erfahrung heraus ist es wichtig, dass man einen guten Teilnehmerauswahl hat.									
5064	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5064	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel.									
5065	Person03.12.12		HK3.4 Manuelle Auswertung und Werkzeugtafel	HK3 Bestehender Analyseprozess: Ablauf und Werkzeuge		Die Auswertung erfolgt manuell, entlastet eine Werkzeugtafel, die manuell entlastet eine Werkzeugtafel									

[illegible]

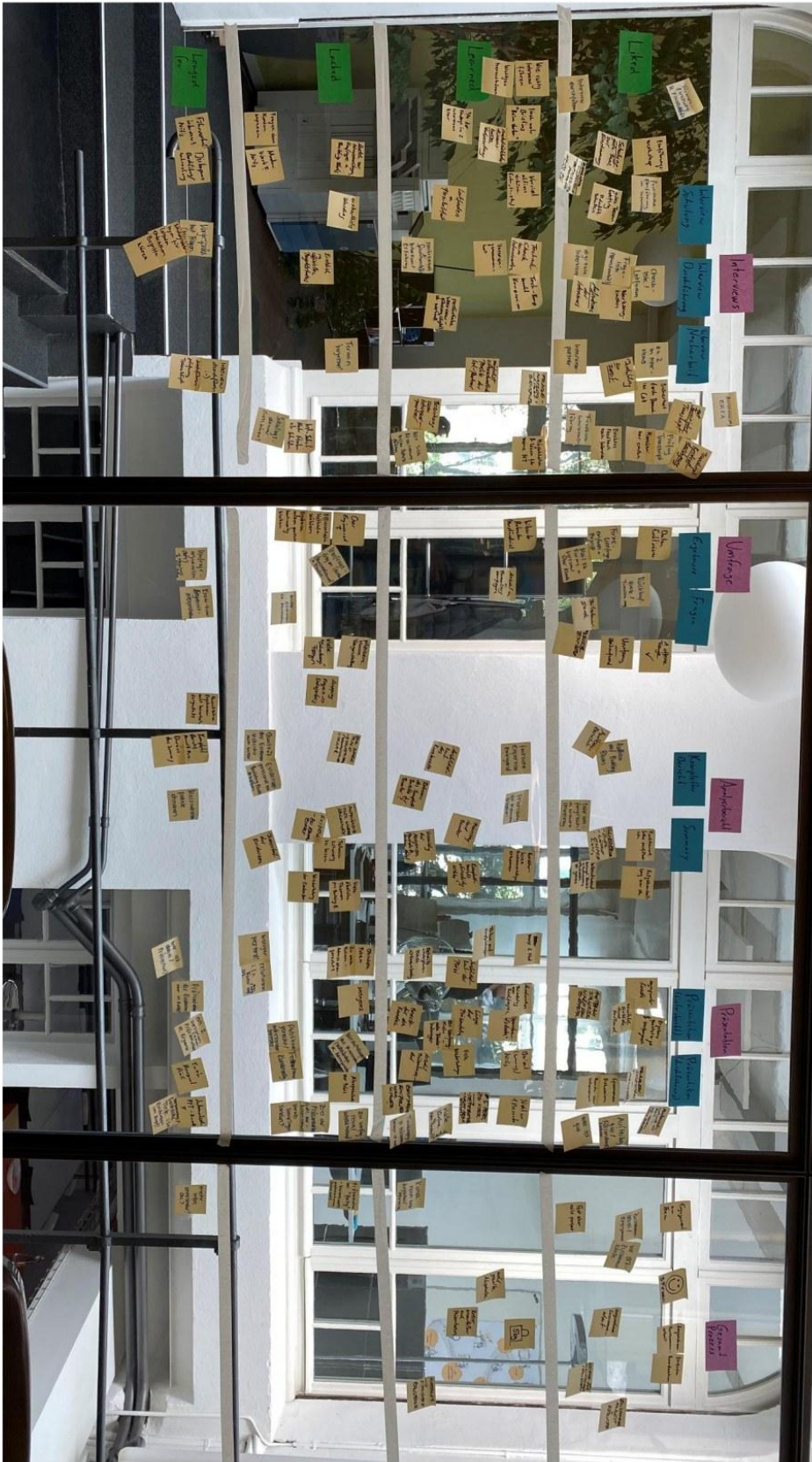
[illegible]

[illegible]

Herleitung Bedürfnisse und Anforderungen - Komplette Herleitung															
Sinnesabschnitt-Nr.	Beleg (Person, Stelle)	Subkategorie(n)	Hauptkategorie(n)	Sinnesabschnitt (wörtliches Zitat, Schritt 3)	Paraphrasen-Nr.	Paraphrase (Schritt 4)	Begründung der NB-Zuordnung	Direkte NB-Zuordnung	Nutzerbedürfnis-Nummer	Nutzerbedürfnis-Name	Nutzerbedürfnis-Schilderung	Nutzerbedürfnis-Erklärung	Kernbedürfnis-Nummer	Kernbedürfnis-Name	Kernbedürfnis-Schilderung
5095	Person02 15.15	HK7-2 Persönliche und betriebliche KI-Nutzung	HK7 KI-Einsatz: Bedeutung, Haltung und Nutzung	Was am häufigsten zum Einsatz kommt, ist eine Videoaufzeichnung und damit auch das Transkript. [...] Dann die Copilot-Appl-Lösung, also der Copilot, die KI in den	P42	KI und Copilot werden zunehmend genutzt, etwa für Transkripte, Zusammenfassungen, den Office-Copilot und Agents.	Zeigt tendenz zur Nutzung gemäss Annahmen	Indirekt Berücksichtigt							
5096	Person03 23.01	HK7-2 Persönliche und betriebliche KI-Nutzung	HK7 KI-Einsatz: Bedeutung, Haltung und Nutzung	Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen. Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen.	P42	KI und Copilot werden zunehmend genutzt, etwa für Transkripte, Zusammenfassungen, den Office-Copilot und Agents.	Zeigt tendenz zur Nutzung gemäss Annahmen	Indirekt Berücksichtigt							
5097	Person02 22.43	HK3-3 Validierung und fachliche Kontrolle	HK3 Aufgabenteilung Mensch KI, fachliche Verantwortung und Ergebnisverlässlichkeit	Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen. Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen.	P41	KI-Ergebnisse müssen validiert werden, die Fachperson bleibt verantwortlich und ein Review bleibt nötig.	Validierung und fachliche Kontrolle (in NB-06 genannt), daneben NB-08 (Unsicherheit klären)	Zugeteilt	NB-06	Ergebnisse fachlich vertreten können	Als M365-Fachperson brauche ich eine Auswertung, deren Aussagen ich prüfen, einordnen und gegenüber der Kundenschaft fachlich vertreten können. Als M365-Fachperson brauche ich eine Arbeitsweise, die meine fachliche Rolle stärkt und mir schenkt. Ergebnisse aus der Kundenschaft fachlich vertreten können.	KB-03	Fachliche Verantwortung und Beurteilbarkeit	M365-Fachpersonen brauchen Analyseergebnisse, die sie prüfen, einordnen, klären und gegenüber der Kundenschaft fachlich vertreten können.	
5097	Person02 22.43	HK3-3 Validierung und fachliche Kontrolle	HK3 Aufgabenteilung Mensch KI, fachliche Verantwortung und Ergebnisverlässlichkeit	Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen. Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen.	P41	KI-Ergebnisse müssen validiert werden, die Fachperson bleibt verantwortlich und ein Review bleibt nötig.	Validierung und fachliche Kontrolle (in NB-06 genannt), daneben NB-08 (Unsicherheit klären)	Zugeteilt	NB-07	Die eigene Fachrolle im Analyseprozess erhalten	Als M365-Fachperson brauche ich eine Arbeitsweise, die meine fachliche Rolle stärkt und mir schenkt. Ergebnisse aus der Kundenschaft fachlich vertreten. Als M365-Fachperson brauche ich eine Möglichkeit, zweifelhafte, unvollständige oder widersprüchliche Analyseergebnisse zu klären, bevor ich sie weiterverwende. Als M365-Fachperson brauche ich eine Auswertung, deren Aussagen ich prüfen, einordnen und gegenüber der Kundenschaft fachlich vertreten können.	KB-03	Fachliche Verantwortung und Beurteilbarkeit	M365-Fachpersonen brauchen Analyseergebnisse, die sie prüfen, einordnen, klären und gegenüber der Kundenschaft fachlich vertreten können.	
5097	Person02 22.43	HK3-3 Validierung und fachliche Kontrolle	HK3 Aufgabenteilung Mensch KI, fachliche Verantwortung und Ergebnisverlässlichkeit	Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen. Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen.	P41	KI-Ergebnisse müssen validiert werden, die Fachperson bleibt verantwortlich und ein Review bleibt nötig.	Validierung und fachliche Kontrolle (in NB-06 genannt), daneben NB-08 (Unsicherheit klären)	Zugeteilt	NB-08	Unsicherheit vor der Weitergabe klären	Als M365-Fachperson brauche ich eine Möglichkeit, zweifelhafte, unvollständige oder widersprüchliche Analyseergebnisse zu klären, bevor ich sie weiterverwende. Als M365-Fachperson brauche ich eine Auswertung, deren Aussagen ich prüfen, einordnen und gegenüber der Kundenschaft fachlich vertreten können.	KB-03	Fachliche Verantwortung und Beurteilbarkeit	M365-Fachpersonen brauchen Analyseergebnisse, die sie prüfen, einordnen, klären und gegenüber der Kundenschaft fachlich vertreten können.	
5098	Person03 21.25	HK3-3 Validierung und fachliche Kontrolle	HK3 Aufgabenteilung Mensch KI, fachliche Verantwortung und Ergebnisverlässlichkeit	Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen. Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen.	P41	KI-Ergebnisse müssen validiert werden, die Fachperson bleibt verantwortlich und ein Review bleibt nötig.	Unsicherheit bei Governance und Anwendung, Kontext, kein direktes Bedürfnis	Zuteilung - kein direktes Nutzerbedürfnis	NB-06	Ergebnisse fachlich vertreten können	Als M365-Fachperson brauche ich eine Auswertung, deren Aussagen ich prüfen, einordnen und gegenüber der Kundenschaft fachlich vertreten können.	KB-03	Fachliche Verantwortung und Beurteilbarkeit	M365-Fachpersonen brauchen Analyseergebnisse, die sie prüfen, einordnen, klären und gegenüber der Kundenschaft fachlich vertreten können.	
5098	Person03 21.25	HK3-3 Validierung und fachliche Kontrolle	HK3 Aufgabenteilung Mensch KI, fachliche Verantwortung und Ergebnisverlässlichkeit	Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen. Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen.	P41	KI-Ergebnisse müssen validiert werden, die Fachperson bleibt verantwortlich und ein Review bleibt nötig.	Unsicherheit bei Governance und Anwendung, Kontext, kein direktes Bedürfnis	Zuteilung - kein direktes Nutzerbedürfnis	NB-07	Die eigene Fachrolle im Analyseprozess erhalten	Als M365-Fachperson brauche ich eine Arbeitsweise, die meine fachliche Rolle stärkt und mir schenkt. Ergebnisse aus der Kundenschaft fachlich vertreten. Als M365-Fachperson brauche ich eine Möglichkeit, zweifelhafte, unvollständige oder widersprüchliche Analyseergebnisse zu klären, bevor ich sie weiterverwende. Als M365-Fachperson brauche ich eine Auswertung, deren Aussagen ich prüfen, einordnen und gegenüber der Kundenschaft fachlich vertreten können.	KB-03	Fachliche Verantwortung und Beurteilbarkeit	M365-Fachpersonen brauchen Analyseergebnisse, die sie prüfen, einordnen, klären und gegenüber der Kundenschaft fachlich vertreten können.	
5098	Person03 21.25	HK3-3 Validierung und fachliche Kontrolle	HK3 Aufgabenteilung Mensch KI, fachliche Verantwortung und Ergebnisverlässlichkeit	Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen. Man muss einer KI natürlich auch immer ein bisschen entgegengetreten [...] ich muss weiterhin die Fachperson sein und sicherstellen.	P41	KI-Ergebnisse müssen validiert werden, die Fachperson bleibt verantwortlich und ein Review bleibt nötig.	Unsicherheit bei Governance und Anwendung, Kontext, kein direktes Bedürfnis	Zuteilung - kein direktes Nutzerbedürfnis	NB-08	Unsicherheit vor der Weitergabe klären	Als M365-Fachperson brauche ich eine Möglichkeit, zweifelhafte, unvollständige oder widersprüchliche Analyseergebnisse zu klären, bevor ich sie weiterverwende. Als M365-Fachperson brauche ich eine Auswertung, deren Aussagen ich prüfen, einordnen und gegenüber der Kundenschaft fachlich vertreten können.	KB-03	Fachliche Verantwortung und Beurteilbarkeit	M365-Fachpersonen brauchen Analyseergebnisse, die sie prüfen, einordnen, klären und gegenüber der Kundenschaft fachlich vertreten können.	
5099	Person01 23.35	HK7-4 Unsicherheit bei Governance und Anwendung	HK7 KI-Einsatz: Bedeutung, Haltung und Nutzung	Was ist so eine Governance? Da geht es ein bisschen um ethische Geschichten, vor allem aber um die richtige Anwendung. Viele Leute haben wirklich wenig Ahnung, was ist so eine Governance?	P43	Rund um KI bestehen Unsicherheiten bei Governance, Ethik und richtiger Anwendung, vielen fehlt der sichere Umgang.	Hinweis auf Nicht-Präzisierung	Indirekt Berücksichtigt							
5100	Person02 7.38	HK5-1 Standardisierter und bewährter Prozessstand	HK5 Positive Erfahrungen und Stärken	Es ist schon sehr gut, standardisiert, wir haben standardisierte Fragebögen, standardisierte Interview-Fragebögen, standardisierte PowerPoint-Folien [...] das ist auch positiv.	P44	Die Standardisierung mit Fragebögen und Folien sowie der erreichte Automatisierungsgrad erleichtern die Arbeit, der Prozess wird als gut und bewährt erlebt.	Aufgaben der KI als Lösungsraum und Aufgabenteilung, dient NB-02 und NB-03	kein direktes Nutzerbedürfnis	NB-01	Assessmentsdaten in eine tragfähige Einschätzung überführen	Als M365-Fachperson brauche ich eine Möglichkeit, erhobene Assessmentsdaten in eine fachlich tragfähige Einschätzung zu überführen, ohne dass die Auswertung den Projektkontext dominiert. Als M365-Fachperson brauche ich eine Arbeitsweise, bei der meine Aufmerksamkeits auf Interpretation, Kundenverständnis und das eigene Fachwissen dominiert. Als M365-Fachperson brauche ich eine Möglichkeit, erhobene Assessmentsdaten in eine fachlich tragfähige Einschätzung zu überführen, ohne dass die Auswertung den Projektkontext dominiert.	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, Assessmentsdaten so auszuwerten, dass eine fachlich belastbare Einschätzung entsteht, ohne dass der Analyseaufwand den Projektkontext dominiert.	
5100	Person02 7.38	HK5-1 Standardisierter und bewährter Prozessstand	HK5 Positive Erfahrungen und Stärken	Es ist schon sehr gut, standardisiert, wir haben standardisierte Fragebögen, standardisierte Interview-Fragebögen, standardisierte PowerPoint-Folien [...] das ist auch positiv.	P44	Die Standardisierung mit Fragebögen und Folien sowie der erreichte Automatisierungsgrad erleichtern die Arbeit, der Prozess wird als gut und bewährt erlebt.	Aufgaben der KI als Lösungsraum und Aufgabenteilung, dient NB-02 und NB-03	kein direktes Nutzerbedürfnis	NB-02	Die eigene Aufmerksamkeits auf fachlich wertvolle Arbeit richten	Als M365-Fachperson brauche ich eine Arbeitsweise, bei der meine Aufmerksamkeits auf Interpretation, Kundenverständnis und das eigene Fachwissen dominiert. Als M365-Fachperson brauche ich eine Möglichkeit, erhobene Assessmentsdaten in eine fachlich tragfähige Einschätzung zu überführen, ohne dass die Auswertung den Projektkontext dominiert.	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, Assessmentsdaten so auszuwerten, dass eine fachlich belastbare Einschätzung entsteht, ohne dass der Analyseaufwand den Projektkontext dominiert.	
5101	Person03 32.06	HK5-1 Standardisierter und bewährter Prozessstand	HK5 Positive Erfahrungen und Stärken	Es ist schon sehr gut, standardisiert, wir haben standardisierte Fragebögen, standardisierte Interview-Fragebögen, standardisierte PowerPoint-Folien [...] das ist auch positiv.	P44	Die Standardisierung mit Fragebögen und Folien sowie der erreichte Automatisierungsgrad erleichtern die Arbeit, der Prozess wird als gut und bewährt erlebt.	Aufgaben der KI als Lösungsraum und Aufgabenteilung, dient NB-02 und NB-03	kein direktes Nutzerbedürfnis	NB-01	Assessmentsdaten in eine tragfähige Einschätzung überführen	Als M365-Fachperson brauche ich eine Möglichkeit, erhobene Assessmentsdaten in eine fachlich tragfähige Einschätzung zu überführen, ohne dass die Auswertung den Projektkontext dominiert. Als M365-Fachperson brauche ich eine Arbeitsweise, bei der meine Aufmerksamkeits auf Interpretation, Kundenverständnis und das eigene Fachwissen dominiert. Als M365-Fachperson brauche ich eine Möglichkeit, erhobene Assessmentsdaten in eine fachlich tragfähige Einschätzung zu überführen, ohne dass die Auswertung den Projektkontext dominiert.	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, Assessmentsdaten so auszuwerten, dass eine fachlich belastbare Einschätzung entsteht, ohne dass der Analyseaufwand den Projektkontext dominiert.	
5101	Person03 32.06	HK5-1 Standardisierter und bewährter Prozessstand	HK5 Positive Erfahrungen und Stärken	Es ist schon sehr gut, standardisiert, wir haben standardisierte Fragebögen, standardisierte Interview-Fragebögen, standardisierte PowerPoint-Folien [...] das ist auch positiv.	P44	Die Standardisierung mit Fragebögen und Folien sowie der erreichte Automatisierungsgrad erleichtern die Arbeit, der Prozess wird als gut und bewährt erlebt.	Aufgaben der KI als Lösungsraum und Aufgabenteilung, dient NB-02 und NB-03	kein direktes Nutzerbedürfnis	NB-02	Die eigene Aufmerksamkeits auf fachlich wertvolle Arbeit richten	Als M365-Fachperson brauche ich eine Arbeitsweise, bei der meine Aufmerksamkeits auf Interpretation, Kundenverständnis und das eigene Fachwissen dominiert. Als M365-Fachperson brauche ich eine Möglichkeit, erhobene Assessmentsdaten in eine fachlich tragfähige Einschätzung zu überführen, ohne dass die Auswertung den Projektkontext dominiert.	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, Assessmentsdaten so auszuwerten, dass eine fachlich belastbare Einschätzung entsteht, ohne dass der Analyseaufwand den Projektkontext dominiert.	
5102	Person01 29.07	HK6-5 Zielbild und Qualitätserhalt	HK6 Anforderungen, Kriterien und Soll-Vorstellungen	Ein Zielbild ist immer, dass man ein ganz klares Zielbild hat. [...] Ich finde, es ist wichtig, dass man ein ganz klares Zielbild hat. [...] Ich finde, es ist wichtig, dass man ein ganz klares Zielbild hat. [...] Ich finde, es ist wichtig, dass man ein ganz klares Zielbild hat.	P16	Zielbild und Qualitätserhalt (in NB-04 als Qualitätsanforderung), daneben NB-05 (Ergebnisqualität)	Zugeteilt	NB-04	Den Ist-Zustand mit Zielbild und Framework verbinden	Als M365-Fachperson brauche ich eine Auswertung, die mir hilft, den Ist-Zustand zu verstehen und zu verbessern. Als M365-Fachperson brauche ich eine Arbeitsweise, bei der meine Aufmerksamkeits auf Interpretation, Kundenverständnis und das eigene Fachwissen dominiert. Als M365-Fachperson brauche ich eine Möglichkeit, erhobene Assessmentsdaten in eine fachlich tragfähige Einschätzung zu überführen, ohne dass die Auswertung den Projektkontext dominiert.	KB-02	Orientierung und fachliche Verdichtung	M365-Fachpersonen brauchen eine Möglichkeit, Assessmentsdaten so auszuwerten, dass eine fachlich belastbare Einschätzung entsteht, ohne dass der Analyseaufwand den Projektkontext dominiert.		
5102	Person01 29.07	HK6-5 Zielbild und Qualitätserhalt	HK6 Anforderungen, Kriterien und Soll-Vorstellungen	Ein Zielbild ist immer, dass man ein ganz klares Zielbild hat. [...] Ich finde, es ist wichtig, dass man ein ganz klares Zielbild hat. [...] Ich finde, es ist wichtig, dass man ein ganz klares Zielbild hat. [...] Ich finde, es ist wichtig, dass man ein ganz klares Zielbild hat.	P16	Zielbild und Qualitätserhalt (in NB-04 als Qualitätsanforderung), daneben NB-05 (Ergebnisqualität)	Zugeteilt	NB-05	Aus den Daten begründete Entwicklungsschritte ableiten	Als M365-Fachperson brauche ich eine Auswertung, die mir hilft, den Ist-Zustand zu verstehen und zu verbessern. Als M365-Fachperson brauche ich eine Arbeitsweise, bei der meine Aufmerksamkeits auf Interpretation, Kundenverständnis und das eigene Fachwissen dominiert. Als M365-Fachperson brauche ich eine Möglichkeit, erhobene Assessmentsdaten in eine fachlich tragfähige Einschätzung zu überführen, ohne dass die Auswertung den Projektkontext dominiert.	KB-05	Handlungsorientierung und Ergebnisfähigkeit der Ergebnisse	M365-Fachpersonen brauchen eine Möglichkeit, Assessmentsdaten so auszuwerten, dass eine fachlich belastbare Einschätzung entsteht, ohne dass der Analyseaufwand den Projektkontext dominiert.		
5102	Person01 29.07	HK6-5 Zielbild und Qualitätserhalt	HK6 Anforderungen, Kriterien und Soll-Vorstellungen	Ein Zielbild ist immer, dass man ein ganz klares Zielbild hat. [...] Ich finde, es ist wichtig, dass man ein ganz klares Zielbild hat. [...] Ich finde, es ist wichtig, dass man ein ganz klares Zielbild hat. [...] Ich finde, es ist wichtig, dass man ein ganz klares Zielbild hat.	P16	Zielbild und Qualitätserhalt (in NB-04 als Qualitätsanforderung), daneben NB-05 (Ergebnisqualität)	Zugeteilt	NB-12	Das Assessment auch in begrenzten Projektphasen einsetzen können	Als M365-Fachperson brauche ich eine Auswertung, die mir hilft, den Ist-Zustand zu verstehen und zu verbessern. Als M365-Fachperson brauche ich eine Arbeitsweise, bei der meine Aufmerksamkeits auf Interpretation, Kundenverständnis und das eigene Fachwissen dominiert. Als M365-Fachperson brauche ich eine Möglichkeit, erhobene Assessmentsdaten in eine fachlich tragfähige Einschätzung zu überführen, ohne dass die Auswertung den Projektkontext dominiert.	KB-01	Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, Assessmentsdaten so auszuwerten, dass eine fachlich belastbare Einschätzung entsteht, ohne dass der Analyseaufwand den Projektkontext dominiert.		

Herleitung Bedürfnisse und Anforderungen - Herleitung Anforderungen					
Kernbedürfnis	Kernbedürfnisschilderung	Anforderungsnummer	Anforderungsname	Beschreibung	Prüfbarkeit
KB-01: Tragfähige Analyse mit vertretbarem Aufwand	M365-Fachpersonen brauchen eine Möglichkeit, Assessmentdaten so auszuwerten, dass eine fachlich belastbare Einschätzung entsteht, ohne dass der Analyseaufwand den Projektrahmen unverhältnismässig belastet.	NA1.1	Durchgängige Datenverarbeitung	M365-Spezialisten müssen die Daten aus Interviews, Umfrage und technischen Abklärungen in einem durchgängigen Arbeitsablauf weiterverarbeiten können, ohne Inhalte manuell zwischen nicht verbundenen Einzelaussagen zu übertragen.	Vergleich der Anzahl manueller Übertragungsschritte zwischen Ist-Prozess und Soll-Prozess.
		NA1.2	Delegation der Fleissarbeit	M365-Spezialisten müssen wiederkehrende Verarbeitungsschritte, namentlich Aufbereitung, Strukturierung und Kernaussagen-Extraktion, an die KI-Interaktion übertragen können.	Nachweis im Prozessprototyp, welche Verarbeitungsschritte ohne manuelle Ausführung durch die Fachperson erfolgen.
		NA1.3	Einhaltung der Aufwandsgrenze	M365-Spezialisten müssen den gesamten Analyseprozess eines Assessments innerhalb der unternehmensseitigen Zielgrösse von maximal rund zwanzig Arbeitsstunden durchführen können.	Effizienz-Evaluation gemäss Teilzielen, also Vergleich der dokumentierten Arbeitsstunden vergangener Assessments mit dem im Testing erhobenen Aufwand.
		NA4.3	Kalkulierbarer Zeitbedarf	M365-Spezialisten müssen den Zeitbedarf eines Analysedurchlaufs vor Beginn auf Basis des definierten	Vergleich von geschätztem und tatsächlich erhobenen Aufwand im Testing.
KB-02: Orientierung und fachliche Verdichtung	M365-Fachpersonen brauchen eine Auswertung, die heterogene Daten in eine verständliche fachliche Ordnung bringt, zentrale Muster und Spannungen sichtbar macht und den Ist-Zustand mit Zielbild und Framework in Beziehung setzt.	NA3.1	Zielbild- und Reifegradbezug	M365-Spezialisten müssen Analyseergebnisse erzeugen können, die den Bezug zum Zielbild der Kundenorganisation und zum Reifegradmodell des Assessments explizit herstellen.	Fachliche Beurteilung der Ergebnisartefakte im Testing daraufhin, ob Zielbild- und Reifegradbezug vorhanden sind.
		NA3.3	Steuerbare Analysetiefe	M365-Spezialisten müssen die fachliche Tiefe und Breite der Auswertung fallbezogen anpassen können, sodass die Verschlinkung gesteuert erfolgt und nicht pauschal zulasten der Aussagekraft geht.	Nachweis im Prozessprototyp, dass der Detaillierungsgrad je Assessment angepasst werden kann.
KB-03: Fachliche Verantwortung und Beurteilbarkeit	M365-Fachpersonen brauchen Analyseergebnisse, die sie prüfen, einordnen, klären und gegenüber der Kundschaft fachlich vertreten können.	NA2.1	Prüfung und Freigabe an Eingriffspunkten	M365-Spezialisten müssen jedes KI-generierte Zwischenergebnis an definierten Eingriffspunkten prüfen, korrigieren und freigeben können, bevor es weiterverarbeitet wird.	Nachweis der Eingriffspunkte im dokumentierten Prozessablauf sowie Beobachtung im Usability-Testing, ob Korrekturen tatsächlich vorgenommen werden können.
		NA2.2	Rückführbarkeit auf das Ausgangsmaterial	M365-Spezialisten müssen jede KI-generierte Aussage auf die zugrunde liegende Stelle im Ausgangsmaterial zurückführen können.	Belegbarkeitsprüfung im Rahmen der Evaluation der Ergebnisverlässlichkeit, Anteil der rückführbaren Aussagen.
		NA2.3	Nachvollziehbare Aufgabenteilung	M365-Spezialisten müssen für jedes Ergebnis nachvollziehen können, welche Arbeitsschritte die KI ausgeführt hat und welche Entscheide bei der Fachperson liegen.	Vorliegen eines vollständigen Protokolls je Analysedurchlauf.
KB-04: Verlässlichkeit und ausreichende Perspektivenbreite	M365-Fachpersonen brauchen eine Analyse, die über unterschiedliche Datengrundlagen hinweg zu konsistenten Einschätzungen führt und dabei genügend relevante Perspektiven berücksichtigt.	NA3.2	Kontrolle der Verdichtung	M365-Spezialisten müssen überprüfen können, ob die verdichteten Ergebnisse das zugrunde liegende Material hinreichend abbilden und keine relevanten Perspektiven verloren gegangen sind.	Vorhandensein einer Repräsentationsprüfung im Soll-Prozess und deren Anwendung im Testing.
KB-05: Handlungsorientierung und Anschlussfähigkeit der Ergebnisse	M365-Fachpersonen brauchen Analyseergebnisse, aus denen sich begründete Reifegradeinschätzungen, Handlungsfelder und nächste Entwicklungsschritte ableiten lassen und die sich in Kundengesprächen, Entscheidungsprozessen und Massnahmenplanung sinnvoll weiterverwenden lassen.	NA5.1	Anschlussfähiges Summary	M365-Spezialisten müssen am Ende des Analyseprozesses ein finales Summary erzeugen können, das ohne umfangreiche manuelle Nachbearbeitung als Lieferobjekt oder als Grundlage für die Abschlusspräsentation dient.	Fachliche Beurteilung des erzeugten Summaries im Testing, insbesondere des Umfangs der erforderlichen Nachbearbeitung.
		NA5.2	Ableitbare Handlungsfelder	M365-Spezialisten müssen aus den Analyseergebnissen Handlungsfelder und mögliche Massnahmen für die Anschlussarbeit ableiten können.	Beurteilung im Testing, ob die Ergebnisartefakte Handlungsfelder und Massnahmen erkennbar machen.
Ausserhalb der fünf Kernbedürfnisse (Prozessdimension)	Diese Anforderungen bilden eine eigenständige Prozess- und Nutzbarkeitsdimension, die in den fünf empirisch hergeleiteten Kernbedürfnissen nicht enthalten ist.	NA4.1	Definierter Prozessablauf	M365-Spezialisten müssen den Analyseprozess entlang eines definierten, schrittweisen Ablaufs mit klaren Verantwortlichkeiten und Übergaben zwischen Fachperson und KI durchführen können.	Vorliegen des dokumentierten Prozessablaufs und Beobachtung der Schrittfolge im Usability-Testing.
		NA4.2	Delegierbarkeit	M365-Spezialisten müssen den Analyseprozess nach einer kurzen Einführung selbstständig anwenden können, ohne auf die Unterstützung der prozessgestaltenden Person angewiesen zu sein.	Usability-Testing, in dem Teilnehmende den Prozess ohne Eingreifen der prozessgestaltenden Person durchführen.

Anhang I: Participatory-Design-Workshop: Foto der Workshopwand



Anhang J: Workshop-Ergebnisse Reflexionsworkshop

workshop_ergebnisse - ROH Thematisiert Frage												
Input	Phase	L	Ebene Thema	Überthema	Überthema-L	Überthema-Phase	Überthema-Komplett	Nr	Thema für die nächste Iteration	Teilprozess/Phase	Bezug B1-B5 und NA	Stützende 4L Bedunde
Technik-Check bei Interviewten	1 Interviews	Learned	Adaptive, unvoreingenommene Interviewführung bei heterogenen Voraussetzungen	Heterogene Interviewpartner bewusst und passend einbeziehen	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Adaptive Interviewführung und passende Auswahl der Interviewten	RT-06	Die Auswahl der Interviewpartner ist bewusst auf möglichst heterogene Profile auszurichten, die Rolle der Führungsebene in der Stichprobe ist zu klären, und was dem Interviewer ist	Erhebung / Interviews	B3, NA3.2	Learned: Auswahl passender Interviewpartner, möglichst unterschiedliche Profile, Technik-Check bei Interviewten, Technik-Kompetenz benötigt
Technik-Kompetenz bei Interviewten benötigt manchmal Kompromisse	1 Interviews	Learned	Adaptive, unvoreingenommene Interviewführung bei heterogenen Voraussetzungen	Heterogene Interviewpartner bewusst und passend einbeziehen	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Adaptive Interviewführung und passende Auswahl der Interviewten					
Unvereinbar einmütig	1 Interviews	Learned	Adaptive, unvoreingenommene Interviewführung bei heterogenen Voraussetzungen	Heterogene Interviewpartner bewusst und passend einbeziehen	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Adaptive Interviewführung und passende Auswahl der Interviewten					
Interviewpartner	1 Interviews	Learned	Auswahl passender Interviewpartner ist entscheidend	Heterogene Interviewpartner bewusst und passend einbeziehen	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Adaptive Interviewführung und passende Auswahl der Interviewten					
möglichst unterschiedliche Profile der Interviewten	1 Interviews	Learned	Heterogene Interviewprofile bewusst einbeziehen	Heterogene Interviewpartner bewusst und passend einbeziehen	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Adaptive Interviewführung und passende Auswahl der Interviewten					
Wie sich Menschen beim Interview fühlen	1 Interviews	Learned	Interviewqualität hängt von Beziehung, Empathie und Natürlichkeit ab	Interviewqualität entsteht stark über Beziehung und Gesprächsführung	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Adaptive Interviewführung und passende Auswahl der Interviewten					
Beziehung zum Interviewpartner	1 Interviews	Learned	Interviewqualität hängt von Beziehung, Empathie und Natürlichkeit ab	Interviewqualität entsteht stark über Beziehung und Gesprächsführung	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Adaptive Interviewführung und passende Auswahl der Interviewten					
Persönliche Interviewführung (digital) ist wertvoll	1 Interviews	Learned	Interviewqualität hängt von Beziehung, Empathie und Natürlichkeit ab	Interviewqualität entsteht stark über Beziehung und Gesprächsführung	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Adaptive Interviewführung und passende Auswahl der Interviewten					
Natürlichkeit & Spontanität Intuition/ Erfahrung	1 Interviews	Learned	Interviewqualität hängt von Beziehung, Empathie und Natürlichkeit ab	Interviewqualität entsteht stark über Beziehung und Gesprächsführung	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Adaptive Interviewführung und passende Auswahl der Interviewten					
Aufnahme der Interviews	1 Interviews	Liked	Interviewaufzeichnung unterstützt Nacharbeit und Auswertung	Nachbearbeitung und Transkriptarbeit sind gut handhabbar	Durchführung und Handhabung des Vorgehens funktionieren gut	Nachbearbeitung und Verwertung der Interviews sind zentral	Nachbearbeitung, Transkriptarbeit und Verwertung	RT-07	Der fachliche Kontrolldurchgang am Transkript (Prüfung der Worttreue vor der Transkriptprüfung sind enger zu koppeln, sodass möglichst die interviewende Person	Transkript	NA2.1, NA2.2, B2	Liked: gut bereinigtes Transkript, Transkript-Kontrolldurchgang, Prüfung Transkript, Korrektur war
Durchführen / Prüfung Transkript	1 Interviews	Longed for	Interviewdurchführung enger mit Transkriptprüfung koppeln	Interviewdurchführung und Nacharbeit enger verzahnen	Interviewvorbereitung, Schulung und Reflexion weiter ausbauen	Nachbearbeitung und Verwertung der Interviews sind zentral	Nachbearbeitung, Transkriptarbeit und Verwertung	RT-08	Die Anzahl der Interviews ist an der inhaltlichen Sättigung auszurichten, da bereits wenige Interviews einen Großteil der zentralen	Transkript	NA2.3, B2	Longed for: Durchführung/Prüfung Transkript koppeln. Learned: aktives Zuhören sichtbar im
Möglichkeiten & Ideen für kleine MT	1 Interviews	Liked	Interviews liefern konkrete Ansatzpunkte auch für kleine Mandate/Massnahmen	Interviews schaffen verwertbare Ansatzpunkte über den Einzelfall hinaus	Ergebnisse stiften erkennbaren Nutzen	Nachbearbeitung und Verwertung der Interviews sind zentral	Nachbearbeitung, Transkriptarbeit und Verwertung	RT-09	Die Anzahl der Interviews ist an der inhaltlichen Sättigung auszurichten, da bereits wenige Interviews einen Großteil der zentralen	Erhebung / Analyse	B3, B1, NA3.3	Learned: 3/4 der Findings in 2 Interviews klar. Hinweis: Grauzone, gegen Repräsentativität (B3) abzuwägen.
Gut bereinigtes Transkript	1 Interviews	Liked	Transkriptprüfung und -bereinigung werden positiv erlebt	Nachbearbeitung und Transkriptarbeit sind gut handhabbar	Durchführung und Handhabung des Vorgehens funktionieren gut	Nachbearbeitung und Verwertung der Interviews sind zentral	Nachbearbeitung, Transkriptarbeit und Verwertung					
Transkript-Kontrolldurchgang	1 Interviews	Liked	Transkriptprüfung und -bereinigung werden positiv erlebt	Nachbearbeitung und Transkriptarbeit sind gut handhabbar	Durchführung und Handhabung des Vorgehens funktionieren gut	Nachbearbeitung und Verwertung der Interviews sind zentral	Nachbearbeitung, Transkriptarbeit und Verwertung					
Prüfung Transkript	1 Interviews	Liked	Transkriptprüfung und -bereinigung werden positiv erlebt	Nachbearbeitung und Transkriptarbeit sind gut handhabbar	Durchführung und Handhabung des Vorgehens funktionieren gut	Nachbearbeitung und Verwertung der Interviews sind zentral	Nachbearbeitung, Transkriptarbeit und Verwertung					
Korrektur war einfach	1 Interviews	Liked	Transkriptprüfung und -bereinigung werden positiv erlebt	Nachbearbeitung und Transkriptarbeit sind gut handhabbar	Durchführung und Handhabung des Vorgehens funktionieren gut	Nachbearbeitung und Verwertung der Interviews sind zentral	Nachbearbeitung, Transkriptarbeit und Verwertung					
3/4 der Findings in 2 Interviews klar	1 Interviews	Learned	Wenige Interviews liefern bereits viele zentrale Findings	Wesentliche Findings lassen sich früh fokussiert herausarbeiten	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Nachbearbeitung und Verwertung der Interviews sind zentral	Nachbearbeitung, Transkriptarbeit und Verwertung					
Wichtiges heraushearsen	1 Interviews	Learned	Zentrale Aussagen im Interview aktiv herausarbeiten	Wesentliche Findings lassen sich früh fokussiert herausarbeiten	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Nachbearbeitung und Verwertung der Interviews sind zentral	Nachbearbeitung, Transkriptarbeit und Verwertung					
Terminfindung & Flexibilität	1 Interviews	Liked	Flexible Interviewplanung und Terminabstimmung	Organisation und Ablauf der Interviews	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewdurchführung und Zusammenarbeit funktionieren grundsätzlich gut	Organisation, Planung und Ablauf des Vorgehens	RT-01	Der niederschwellige Umfragezugang über Microsoft Forms mit Versand über die Kundenorganisation	Erhebung / Umfrage	B4, B3, NA1.1	Liked: Forms einfacher Zugriff, Mail zu Forms über Kunde versandt. Rücklaufquote/Timeline.
Führungsebene Interview nötig?	1 Interviews	Lacked	Rolle der Führungsebene in der Interviewstichprobe klären	Rollen und Stichprobe der Interviewpartner klären	Organisation, Schulung und Teilnehmer-Setup der Interviews	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Organisation, Planung und Ablauf des Vorgehens	RT-02	Die Terminverbindlichkeit der Interviews ist aktiv sicherzustellen (Einlassungen vermeiden)	Erhebung / Interviews	B4, NA4.3	Lacked: Termin vergessen (No-show). Liked: Terminabstimmung und
Schulungsdauer evtl. verkürzen	1 Interviews	Lacked	Schulung kompakter halten	Interviewerschulung praxisnäher und kompakter gestalten	Organisation, Schulung und Teilnehmer-Setup der Interviews	Interviewvorbereitung und Schulung weiterentwickeln	Organisation, Planung und Ablauf des Vorgehens					
Interview Schulung mehr Proben als Schildern	1 Interviews	Lacked	Schulung praxisnäher und übungsorientierter gestalten	Interviewerschulung praxisnäher und kompakter gestalten	Organisation, Schulung und Teilnehmer-Setup der Interviews	Interviewvorbereitung und Schulung weiterentwickeln	Organisation, Planung und Ablauf des Vorgehens					
Termin "Verlassen" (No show)	1 Interviews	Lacked	Terminverbindlichkeit und No-shows besser absichern	Interviewtermine verbindlicher organisieren	Organisation, Schulung und Teilnehmer-Setup der Interviews	Interviewplanung verbindlicher organisieren	Organisation, Planung und Ablauf des Vorgehens					

workshop_ergebnisse - ROH Thematisiert Frage												
Input	Phase	L	Ebene Thema	Überthema	Überthema-L	Überthema-Phase	Überthema-komplett	Nr	Thema für die nächste Iteration	Teilprozess/Phase	Bezug B1-B5 und NA	Stützende 4L Bedunde
Schulungsunterlagen für Leitfaden einbeziehen	1 Interviews	Longed for	Leitfaden in Schulungsunterlagen integrieren	Interviewschulung und Unterlagen weiter ausbauen	Interviewvorbereitung, Schulung und Reflexion weiter ausbauen	Interviewvorbereitung und Schulung weiterentwickeln	Präsentation, Moderation und adressatengerechte Kommunikation	RT-17	Die Präsentation muss verdichtet, fokussiert und redundanzarm sein, mit klarer Ergebnisdarstellung	Präsentation	B5, NA5.1, NA3.2	Learned: max. 2 Std., Fokus auf Empfehlung, viele Wiederholungen, Länge/Ausführlichkeit, zu Lacked: publikums-/adressatengerecht, weniger wertend, relativieren als Momentaufnahme der
Interview durchführen :)	1 Interviews	Longed for	Mehr praktische Interviewübungen ermöglichen	Interviewschulung und Unterlagen weiter ausbauen	Interviewvorbereitung, Schulung und Reflexion weiter ausbauen	Interviewvorbereitung und Schulung weiterentwickeln	Präsentation, Moderation und adressatengerechte Kommunikation	RT-18	Muss adressatengerecht und wertneutral sein und die Ergebnisse als Momentaufnahme rahmen, Der Interviewleitfaden ist flexibel und beispielbasiert zu gestalten und begrifflich zu vereinheitlichen, sodass Interview- und Umfragefragen müssen konsistent auf die Building Blocks, die Massnahmen und die angestrebten	Präsentation	B5, NA5.1, B3	Learned: Fragekatalog Light/Beispiele, Leitfaden und Flexibilität, Lacked: Fraealiste hemmt Match/Zuweisung Interviewfragen zu Umfrage und Building Blocks, Lacked: Umfragen checken (Massnahmen und Blocks), Lacked: präziserer klarerer Fragenkatalog, Anzahl ok und Formulierung, viele Schulungsfragen reduzieren, Learned: irritierte Probanden
Begriffe klären	1 Interviews	Longed for	Begriffe vor Interviews klären	Interviews besser vorbereiten und rahmen	Interviewvorbereitung, Schulung und Reflexion weiter ausbauen	Interviewvorbereitung und Schulung weiterentwickeln	Schulung, Leitfaden und methodische Rahmung	RT-03		Erhebung / Interviews	B3, NA3.3	
Noch mehr Briefing "Dass Keine Vorbereitung nötig"	1 Interviews	Learned	Erwartungsmanagement: Interviewe sollen ohne Vorbereitung kommen	Vorbereitung und Gesprächsverhalten beeinflussen die Datenqualität	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Schulung, Leitfaden und methodische Rahmung	RT-04		Erhebung / Analyse (Codierung)	NA3.1, B3, Codierung (T-3)	
Nachdrücklicher Hinweis: Keine Vorbereitung	1 Interviews	Learned	Erwartungsmanagement: Interviewe sollen ohne Vorbereitung kommen	Vorbereitung und Gesprächsverhalten beeinflussen die Datenqualität	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Schulung, Leitfaden und methodische Rahmung	RT-05	Der Umfragefragenkatalog ist präziser, verständlicher und auf das Wesentliche fokussiert zu gestalten, mit reduzierter Komplexität und	Erhebung / Umfrage	B3, NA3.2, B5	
Fragen dem Kunden anpassen	1 Interviews	Lacked	Fragen stärker kundenspezifisch zuschneiden	Fragenkataloge und Begrifflichkeit der Interviews überarbeiten	Fragenkataloge und Rahmung der Erhebung müssen geschärft werden	Leitfaden, Fragenlogik und Begriffe der Interviews schärfen	Schulung, Leitfaden und methodische Rahmung					
Einheitliches Wording	1 Interviews	Lacked	Fragen und Begriffe konsistent formulieren	Fragenkataloge und Begrifflichkeit der Interviews überarbeiten	Fragenkataloge und Rahmung der Erhebung müssen geschärft werden	Leitfaden, Fragenlogik und Begriffe der Interviews schärfen	Schulung, Leitfaden und methodische Rahmung					
Modern Work Frage nötig?	1 Interviews	Lacked	Fragenkatalog inhaltlich fokussieren	Fragenkataloge und Begrifflichkeit der Interviews überarbeiten	Fragenkataloge und Rahmung der Erhebung müssen geschärft werden	Leitfaden, Fragenlogik und Begriffe der Interviews schärfen	Schulung, Leitfaden und methodische Rahmung					
Vorsicht, aktives Zuhören (mhm, ja, etc.) sichtbar in Transkript.	1 Interviews	Learned	Gesprächsverhalten beeinflusst die Qualität des Transkripts	Vorbereitung und Gesprächsverhalten beeinflussen die Datenqualität	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Interviewführung muss adaptiv, bewusst und beziehungsorientiert sein	Schulung, Leitfaden und methodische Rahmung					
Interview durchführen	1 Interviews	Liked	Interviewdurchführung als praktische Erfahrung ist wertvoll	Schulung und praktischer Einstieg in die Interviewführung	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewvorbereitung und Schulung weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Match, Zuweisung von Interviewfragen zu Umfrage und building blocks	1 Interviews	Learned	Interviewfragen besser mit Umfrage und Building Blocks abgleichen	Fragenlogik und Leitfaden der Interviews schärfen	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Leitfaden, Fragenlogik und Begriffe der Interviews schärfen	Schulung, Leitfaden und methodische Rahmung					
Einführungs Workshop	1 Interviews	Liked	Interview-Schulung und Einstieg geben Sicherheit	Schulung und praktischer Einstieg in die Interviewführung	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewvorbereitung und Schulung weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Fachliche Einführung in Interviews	1 Interviews	Liked	Interview-Schulung und Einstieg geben Sicherheit	Schulung und praktischer Einstieg in die Interviewführung	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewvorbereitung und Schulung weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Schulung: Gute Basis zur Einstimmung	1 Interviews	Liked	Interview-Schulung und Einstieg geben Sicherheit	Schulung und praktischer Einstieg in die Interviewführung	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewvorbereitung und Schulung weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Gutes Team briefing, Didaktik, Verhalten	1 Interviews	Liked	Interview-Schulung und Einstieg geben Sicherheit	Schulung und praktischer Einstieg in die Interviewführung	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewvorbereitung und Schulung weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Schulung Hilfreich	1 Interviews	Liked	Interview-Schulung und Einstieg geben Sicherheit	Schulung und praktischer Einstieg in die Interviewführung	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewvorbereitung und Schulung weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Vorgespräch mit Fragen (min.)	1 Interviews	Longed for	Kurzes Vorgespräch vor Interviews einplanen	Interviews besser vorbereiten und rahmen	Interviewvorbereitung, Schulung und Reflexion weiter ausbauen	Interviewvorbereitung und Schulung weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Fragekatalog: Light! Beispiele	1 Interviews	Learned	Leichter, beispielbasierter Fragekatalog unterstützt Interviews	Fragenlogik und Leitfaden der Interviews schärfen	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Leitfaden, Fragenlogik und Begriffe der Interviews schärfen	Schulung, Leitfaden und methodische Rahmung					
Leitfaden & Flexibilität	1 Interviews	Learned	Leitfaden flexibel statt starr nutzen	Fragenlogik und Leitfaden der Interviews schärfen	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Leitfaden, Fragenlogik und Begriffe der Interviews schärfen	Schulung, Leitfaden und methodische Rahmung					
Einblick in offiziellen Projektstatus	1 Interviews	Lacked	Mehr Projektkontext für die Interviewführung nötig	Mehr Kontext und Rahmung für die Interviewführung	Fragenkataloge und Rahmung der Erhebung müssen geschärft werden	Interviewvorbereitung und Schulung weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Schulung etwas zu lange	1 Interviews	Liked	Schulung kompakter halten	Schulung und praktischer Einstieg in die Interviewführung	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewvorbereitung und Schulung weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Diskrepanz Durchführung und Wahrnehmung	1 Interviews	Longed for	Selbst- und Fremdwahrnehmung der Interviewdurchführung abgleichen	Interviewpraxis stärker reflektieren	Interviewvorbereitung, Schulung und Reflexion weiter ausbauen	Austausch und Reflexion unterstützen die Interviewpraxis	Schulung, Leitfaden und methodische Rahmung					
Frageliste (Spontanität)	1 Interviews	Lacked	Zu starre Frageliste hemmt Spontanität	Fragenkataloge und Begrifflichkeit der Interviews überarbeiten	Fragenkataloge und Rahmung der Erhebung müssen geschärft werden	Leitfaden, Fragenlogik und Begriffe der Interviews schärfen	Schulung, Leitfaden und methodische Rahmung					
Austausch/ ERFA	1 Interviews	Liked	Austausch und Erfahrungslernen sind wertvoll	Austausch und Feedback fördern die Interviewpraxis	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Austausch und Reflexion unterstützen die Interviewpraxis	Zusammenarbeit, Austausch und Rollenklarheit	RT-23	Rollen, Verantwortlichkeiten und der interne Informationsfluss müssen über den gesamten Prozess	Gesamtprozess	NA2.3, NA4.1, B4	Learned: mehr gezielte Absprache, Präsentation bei fertig, Verantwortlichkeiten wahrnehmen. Lacked: mehr

workshop_ergebnisse - ROH Thematisiert Frage												
Input	Phase	L	Ebene Thema	Überthema	Überthema-L	Überthema-Phase	Überthema-komplett	Nr	Thema für die nächste Iteration	Teilprozess/Phase	Bezug B1-B5 und NA	Stützende 4L Bedunde
Sicherheit durch Dominik im Call	1 Interviews	Liked	Co-Moderation gibt Sicherheit im Interview	Struktur und Co-Moderation unterstützen die Interviewdurchführung	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewdurchführung und Zusammenarbeit funktionieren grundsätzlich gut	Zusammenarbeit, Austausch und Rollenklarheit	RT-24	Bewährte Arbeitsformen sind als Prozessmerkmale beibehalten, namentlich die Interviewdurchführung	Interviews / Gesamtprozess	B2, B4, NA4.2	Liked: zu zweit in Interviews, Sicherheit durch Co-Moderation, direktes Feedback nach Interviews,
Digitale Interviews	1 Interviews	Liked	Digitale Interviewdurchführung ist praktikabel	Digitale Durchführung der Interviews funktioniert gut	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewdurchführung und Zusammenarbeit funktionieren grundsätzlich gut	Zusammenarbeit, Austausch und Rollenklarheit					
Werkzeuge	1 Interviews	Liked	Digitale Interviewwerkzeuge unterstützen die Durchführung	Digitale Durchführung der Interviews funktioniert gut	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewdurchführung und Zusammenarbeit funktionieren grundsätzlich gut	Zusammenarbeit, Austausch und Rollenklarheit					
Direktes Feedback nach Interviews	1 Interviews	Liked	Direktes Feedback zur Interviewführung ist wertvoll	Austausch und Feedback fördern die Interviewpraxis	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Austausch und Reflexion unterstützen die Interviewpraxis	Zusammenarbeit, Austausch und Rollenklarheit					
Feedback Interviewführung	1 Interviews	Liked	Direktes Feedback zur Interviewführung ist wertvoll	Austausch und Feedback fördern die Interviewpraxis	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Austausch und Reflexion unterstützen die Interviewpraxis	Zusammenarbeit, Austausch und Rollenklarheit					
Zu zweit in Interviews	1 Interviews	Liked	Interviewdurchführung zu zweit ist hilfreich	Struktur und Co-Moderation unterstützen die Interviewdurchführung	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewdurchführung und Zusammenarbeit funktionieren grundsätzlich gut	Zusammenarbeit, Austausch und Rollenklarheit					
Durchführung zu zweit	1 Interviews	Liked	Interviewdurchführung zu zweit ist hilfreich	Struktur und Co-Moderation unterstützen die Interviewdurchführung	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewdurchführung und Zusammenarbeit funktionieren grundsätzlich gut	Zusammenarbeit, Austausch und Rollenklarheit					
Checkliste/ Leitfaden	1 Interviews	Liked	Leitfaden unterstützt die Interviewdurchführung	Struktur und Co-Moderation unterstützen die Interviewdurchführung	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewdurchführung und Zusammenarbeit funktionieren grundsätzlich gut	Zusammenarbeit, Austausch und Rollenklarheit					
Wie richtig Interviews führen	1 Interviews	Liked	Leitfaden unterstützt die Interviewdurchführung	Struktur und Co-Moderation unterstützen die Interviewdurchführung	Durchführung und Handhabung des Vorgehens funktionieren gut	Interviewdurchführung und Zusammenarbeit funktionieren grundsätzlich gut	Zusammenarbeit, Austausch und Rollenklarheit					
Know-how Ergebnis-Interpretation	2 Umfrage	Longed for	Kompetenz in der Interpretation von Umfrageergebnissen aufbauen	Umfrageergebnisse besser interpretierbar und lesbar machen	Ergebnisse vor der Verdichtung besser prüfen und verständlich machen	Evidenzbasis und Ergebnisdarstellung der Umfrage stärken	Evidenzbasis, Ergebnisqualität und Verständlichkeit	RT-10	Die Evidenzführung in der Analyse muss verständlich, divers und redundanzarm sein und die Evidenzen klar den Building Blocks	Analyse	NA2.2, NA3.1, NA3.2, B2, B3	Lacked: Evidenzen zu kompliziert zitiert, Wiederholung der Evidenzen, viele ähnliche Themen pro Building Block,
Mehr Evidenzen zu Massnahmen	2 Umfrage	Lacked	Mehr Evidenz für Massnahmen erzeugen	Evidenzbasis der Umfrage für Massnahmen stärken	Ergebnisse und Evidenzen müssen verständlicher und belastbarer aufbereitet werden	Evidenzbasis und Ergebnisdarstellung der Umfrage stärken	Evidenzbasis, Ergebnisqualität und Verständlichkeit	RT-11	Quantitative Umfrageergebnisse (Kennzahlen wie NPS, qualitative Freitextantworten) sind	Analyse	B3, B5, NA5.1, B2	Lacked: Umfrageergebnisse (NPS) schwer zu verstehen, Longed for: Know-how Ergebnis-Interpretation, qualitative Ergebnisse nicht
Ergebnisse müssen qualitativ hochwertig bleiben	2 Umfrage	Lacked	Qualitative Evidenzbasis der Umfrage sichern	Evidenzbasis der Umfrage für Massnahmen stärken	Ergebnisse und Evidenzen müssen verständlicher und belastbarer aufbereitet werden	Evidenzbasis und Ergebnisdarstellung der Umfrage stärken	Evidenzbasis, Ergebnisqualität und Verständlichkeit					
Qualitative Ergebnisse nicht lesbar dargestellt	2 Umfrage	Longed for	Qualitative Umfrageergebnisse besser lesbar darstellen	Umfrageergebnisse besser interpretierbar und lesbar machen	Ergebnisse vor der Verdichtung besser prüfen und verständlich machen	Evidenzbasis und Ergebnisdarstellung der Umfrage stärken	Evidenzbasis, Ergebnisqualität und Verständlichkeit					
Umfrageergebnisse (NPS) Schwierig zu verstehen	2 Umfrage	Lacked	Umfrageergebnisse und Kennzahlen verständlicher darstellen	Umfrageergebnisse und Kennzahlen verständlicher aufbereiten	Ergebnisse und Evidenzen müssen verständlicher und belastbarer aufbereitet werden	Evidenzbasis und Ergebnisdarstellung der Umfrage stärken	Evidenzbasis, Ergebnisqualität und Verständlichkeit					
Forms- Umfrage einfacher Zugriff	2 Umfrage	Liked	Niedrigschwelliger Zugang über Forms ist hilfreich	Organisation, Versand und Zugang der Umfrage funktionieren gut	Durchführung und Handhabung des Vorgehens funktionieren gut	Organisation und Zugang der Umfrage funktionieren grundsätzlich gut	Organisation, Planung und Ablauf des Vorgehens					
Rücklaufquote/ Timeline	2 Umfrage	Liked	Teilnahme und Timing der Umfrage funktionieren gut	Organisation, Versand und Zugang der Umfrage funktionieren gut	Durchführung und Handhabung des Vorgehens funktionieren gut	Organisation und Zugang der Umfrage funktionieren grundsätzlich gut	Organisation, Planung und Ablauf des Vorgehens					
Teilnahmequote	2 Umfrage	Liked	Teilnahme und Timing der Umfrage funktionieren gut	Organisation, Versand und Zugang der Umfrage funktionieren gut	Durchführung und Handhabung des Vorgehens funktionieren gut	Organisation und Zugang der Umfrage funktionieren grundsätzlich gut	Organisation, Planung und Ablauf des Vorgehens					
Mail zu Forms versandt über Kunde	2 Umfrage	Liked	Versand über den Kunden erhöht Zugänglichkeit und Teilnahme	Organisation, Versand und Zugang der Umfrage funktionieren gut	Durchführung und Handhabung des Vorgehens funktionieren gut	Organisation und Zugang der Umfrage funktionieren grundsätzlich gut	Organisation, Planung und Ablauf des Vorgehens					
Viele "Schulungs" Fragen	2 Umfrage	Lacked	Anteil an Schulungsfragen im Katalog reduzieren	Fragenkatalog und Mapping der Umfrage schärfen	Fragenkataloge und Rahmung der Erhebung müssen geschärft werden	Fragenkatalog und Methodik der Umfrage weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Effizientere Methode wählen	2 Umfrage	Learned	Erhebungsmethode effizienter wählen	Umfragegestaltung und Methode beeinflussen Verständlichkeit und Effizienz	Erhebung entsteht durch bewusste Gestaltung	Fragenkatalog und Methodik der Umfrage weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Umfragen checken (Massnahmen und Blocks)	2 Umfrage	Lacked	Fragen besser auf Massnahmen und Building Blocks abstimmen	Fragenkatalog und Mapping der Umfrage schärfen	Fragenkataloge und Rahmung der Erhebung müssen geschärft werden	Fragenkatalog und Methodik der Umfrage weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Mapping Fragen -> Endergebnisse	2 Umfrage	Lacked	Fragen klarer auf Endergebnisse mappen	Fragenkatalog und Mapping der Umfrage schärfen	Fragenkataloge und Rahmung der Erhebung müssen geschärft werden	Fragenkatalog und Methodik der Umfrage weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Anzahl Ok + Formulierung von Fragen	2 Umfrage	Lacked	Fragenkatalog präziser und klarer formulieren	Fragenkatalog und Mapping der Umfrage schärfen	Fragenkataloge und Rahmung der Erhebung müssen geschärft werden	Fragenkatalog und Methodik der Umfrage weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					
Präziserer, klarerer Fragenkatalog	2 Umfrage	Lacked	Fragenkatalog präziser und klarer formulieren	Fragenkatalog und Mapping der Umfrage schärfen	Fragenkataloge und Rahmung der Erhebung müssen geschärft werden	Fragenkatalog und Methodik der Umfrage weiterentwickeln	Schulung, Leitfaden und methodische Rahmung					

workshop_ergebnisse - ROH Thematisiert Frage												
Input	Phase	L	Ebene Thema	Überthema	Überthema-L	Überthema-Phase	Überthema-komplett	Nr	Thema für die nächste Iteration	Teilprozess/Phase	Bezug B1-B5 und NA	Stützende 4L Bedunde
Irritierte Probanden (Verständnis)	2 Umfrage	Learned	Fragenverständnis bei Teilnehmenden verbessern	Umfragegestaltung und Methode beeinflussen Verständlichkeit und Effizienz	Methodische Qualität der Erhebung entsteht durch bewusste Gestaltung	Fragenkatalog und Methodik der Umfrage weiterentwickeln	Schulung, Leitfäden und methodische Rahmung					
Over Engineered?	2 Umfrage	Liked	Komplexität der Umfrage kritisch prüfen	Umfrageumfang und Frageformat sind grundsätzlich stimmig	Durchführung und Handhabung des Vorgehens funktionieren gut	Fragenkatalog und Methodik der Umfrage weiterentwickeln	Schulung, Leitfäden und methodische Rahmung					
Eine Offene Frage	2 Umfrage	Liked	Offene Fragen liefern wertvollen Zusatzkontext	Umfrageumfang und Frageformat sind grundsätzlich stimmig	Durchführung und Handhabung des Vorgehens funktionieren gut	Fragenkatalog und Methodik der Umfrage weiterentwickeln	Schulung, Leitfäden und methodische Rahmung					
Freies Textfeld	2 Umfrage	Liked	Offene Fragen liefern wertvollen Zusatzkontext	Umfrageumfang und Frageformat sind grundsätzlich stimmig	Durchführung und Handhabung des Vorgehens funktionieren gut	Fragenkatalog und Methodik der Umfrage weiterentwickeln	Schulung, Leitfäden und methodische Rahmung					
Umfang und Zeitaufwand	2 Umfrage	Liked	Umfrageumfang ist grundsätzlich passend	Umfrageumfang und Frageformat sind grundsätzlich stimmig	Durchführung und Handhabung des Vorgehens funktionieren gut	Fragenkatalog und Methodik der Umfrage weiterentwickeln	Schulung, Leitfäden und methodische Rahmung					
Teilweise schwierig zu lesen	3 Analysebericht	Lacked	Bericht ist in Teilen schwer lesbar	Lesbarkeit und Verständlichkeit des Berichts verbessern	Ergebnisse und Evidenzen müssen verständlicher und belastbarer aufbereitet werden	Lesbarkeit, Orientierung und Kohärenz des Berichts verbessern	Evidenzbasis, Ergebnisqualität und Verständlichkeit					
Aufwändig und viele Wiederholungen	3 Analysebericht	Lacked	Bericht wirkt repetitiv und dadurch aufwändig	Lesbarkeit und Verständlichkeit des Berichts verbessern	Ergebnisse und Evidenzen müssen verständlicher und belastbarer aufbereitet werden	Lesbarkeit, Orientierung und Kohärenz des Berichts verbessern	Evidenzbasis, Ergebnisqualität und Verständlichkeit					
Viele ähnliche Themen pro Building Block	3 Analysebericht	Lacked	Bericht wirkt repetitiv und dadurch aufwändig	Lesbarkeit und Verständlichkeit des Berichts verbessern	Ergebnisse und Evidenzen müssen verständlicher und belastbarer aufbereitet werden	Lesbarkeit, Orientierung und Kohärenz des Berichts verbessern	Evidenzbasis, Ergebnisqualität und Verständlichkeit					
Wiederholung der Evidenzen	3 Analysebericht	Lacked	Bericht wirkt repetitiv und dadurch aufwändig	Lesbarkeit und Verständlichkeit des Berichts verbessern	Ergebnisse und Evidenzen müssen verständlicher und belastbarer aufbereitet werden	Lesbarkeit, Orientierung und Kohärenz des Berichts verbessern	Evidenzbasis, Ergebnisqualität und Verständlichkeit					
Diversität der Evidenzen	3 Analysebericht	Lacked	Evidenzbasis diverser aufbauen	Evidenzbasis qualitativ stärken und sauber zuordnen	Ergebnisse und Evidenzen müssen verständlicher und belastbarer aufbereitet werden	Evidenzbasis und Zuordnung der Berichtsinhalte stärken	Evidenzbasis, Ergebnisqualität und Verständlichkeit					
Evidenzen passend zu Building Blocks fraglich	3 Analysebericht	Lacked	Evidenzen klarer den Building Blocks zuordnen	Evidenzbasis qualitativ stärken und sauber zuordnen	Ergebnisse und Evidenzen müssen verständlicher und belastbarer aufbereitet werden	Evidenzbasis und Zuordnung der Berichtsinhalte stärken	Evidenzbasis, Ergebnisqualität und Verständlichkeit					
Evidenzen zu kompliziert zitiert.	3 Analysebericht	Lacked	Evidenzführung verständlicher und einfacher gestalten	Lesbarkeit und Verständlichkeit des Berichts verbessern	Ergebnisse und Evidenzen müssen verständlicher und belastbarer aufbereitet werden	Lesbarkeit, Orientierung und Kohärenz des Berichts verbessern	Evidenzbasis, Ergebnisqualität und Verständlichkeit					
Qualität der Evidenzen vielleicht noch besser	3 Analysebericht	Lacked	Evidenzqualität weiter verbessern	Evidenzbasis qualitativ stärken und sauber zuordnen	Ergebnisse und Evidenzen müssen verständlicher und belastbarer aufbereitet werden	Evidenzbasis und Zuordnung der Berichtsinhalte stärken	Evidenzbasis, Ergebnisqualität und Verständlichkeit					
Wer ist dabei/ Präsentiert?	3 Analysebericht	Lacked	Zuständigkeiten für die Präsentation klar regeln	Schnittstelle zwischen Bericht und Präsentation klar regeln	Präsentation und Übergabe müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Rollen und Übergänge in der Berichterstellung klären	Präsentation, Moderation und adressatengerechte Kommunikation	RT-19	Vor der Präsentation müssen Ziel, Rollen, Anzahl Präsentierende und kundenseitige Teilnehmerschaft	Präsentation / Gesamtprozess	NA4.1, B4	Lacked: Absprache vor Präsi, Ziel der Präsentation vorbereiten, Teilnehmer kundenseitig, Engagement vom Team, Learned: Anzahl
Aufbau Bericht	3 Analysebericht	Liked	Berichtsstruktur entlang der Building Blocks ist nachvollziehbar	Struktur und Logik des Berichts sind grundsätzlich tragfähig	Struktur und Qualität der Ergebnisartefakte sind grundsätzlich tragfähig	Struktur, Logik und Bewertungsmodell des Berichts weiterentwickeln	Struktur, Lesbarkeit und Kohärenz der Ergebnisartefakte	RT-14	Die bewährte Berichtsstruktur entlang der Building Blocks mit Zielen, Lücken und Empfehlungen	Analyse	NA3.1, NA5.1, B5	Liked: Aufbau Bericht, Aufbau auf Building Blocks, Gliederung in Ziele, Lücken, Empfehlungen, Learned: Vermischung der Messgrößen (Durchschnitt, Median, Mittelwert).
Aufbau auf building blocks	3 Analysebericht	Liked	Berichtsstruktur entlang der Building Blocks ist nachvollziehbar	Struktur und Logik des Berichts sind grundsätzlich tragfähig	Struktur und Qualität der Ergebnisartefakte sind grundsätzlich tragfähig	Struktur, Logik und Bewertungsmodell des Berichts weiterentwickeln	Struktur, Lesbarkeit und Kohärenz der Ergebnisartefakte	RT-15	Die Kennzahlenlogik ist zu vereinfachen, sodass Messgrößen konsistent verwendet werden und	Analyse	B2, Ergebnisverlässlichkeit	
Gliederung in Ziele, Lücken, Empfehlungen	3 Analysebericht	Liked	Berichtsstruktur entlang der Building Blocks ist nachvollziehbar	Struktur und Logik des Berichts sind grundsätzlich tragfähig	Struktur und Qualität der Ergebnisartefakte sind grundsätzlich tragfähig	Struktur, Logik und Bewertungsmodell des Berichts weiterentwickeln	Struktur, Lesbarkeit und Kohärenz der Ergebnisartefakte	RT-16	Die Building-Block-Struktur ist fachlich feiner zu gliedern, etwa die Aufteilung des Blocks Digital Workplace in	Analyse (Codierung)	NA3.1, Codierung (T-3)	Learned: Block Digital Workplace schon unterteilen in Communicate und Collaborate (ggf. Automate).
Block Digital Workplace schon unterteilen in Communicate und Collaborate (ggf. Automate)	3 Analysebericht	Learned	Building Blocks feiner und fachlich klarer strukturieren	Struktur und Bewertungslogik des Berichts schärfen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Struktur, Logik und Bewertungsmodell des Berichts weiterentwickeln	Struktur, Lesbarkeit und Kohärenz der Ergebnisartefakte					
Vermischung der Messgrößen (Durchschnitt, Median, Mittelwert) Korrekturen lesen und Zusammenhänge erkennen.	3 Analysebericht	Learned	Kennzahlenlogik und Messgrößen verteilichen	Struktur und Bewertungslogik des Berichts schärfen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Struktur, Logik und Bewertungsmodell des Berichts weiterentwickeln	Struktur, Lesbarkeit und Kohärenz der Ergebnisartefakte					
Sehr umfangreich/ Es braucht Orientierung	3 Analysebericht	Longed for	Mehr Orientierung im umfangreichen Bericht schaffen	Mehr Orientierung und Führung im Bericht schaffen	Verdichtung besser prüfen und verständlich machen	Struktur, Logik und Bewertungsmodell des Berichts verbessern	Struktur, Lesbarkeit und Kohärenz der Ergebnisartefakte					

workshop_ergebnisse - ROH Thematisiert Frage												
Input	Phase	L	Ebene Thema	Überthema	Überthema-L	Überthema-Phase	Überthema-komplett	Nr	Thema für die nächste Iteration	Teilprozess/Phase	Bezug B1-B5 und NA	Stützende 4L Bedunde
Funktioniert, sehr ausführlich	3 Analysebericht	Liked	Bericht ist fundiert, aber sehr ausführlich	Berichtqualität ist hoch, der Umfang jedoch gross	Struktur und Qualität der Ergebniskarteifakte sind grundsätzlich tragfähig	Berichtsumfang und Priorisierung fachlich steuern	Verdichtung, Priorisierung und fachliche Steuerung	RT-12	Die Priorisierung der Massnahmen muss als bewusster, manueller Fachschritt erfolgen, aus dem anschliessend das Summary abgeleitet wird, da diese Bewertung menschliches Fachurteil erfordert und nicht an die KI delegierbar ist.	Analyse	B2, NA2.1, B5, NA5.2	Longed for: Komplettbericht manuelle Priorisierung, daraus das Summary, Massnahmenpakete checken. Learned: Priorisierung der Massnahmen fraglich, Priorisierung durch Berater (menschliches Urteil), Anschluss an Vibe-Research-Regel R2 (Bedeutungszuschreibung bei der forschenden Person).
Zeitaufwand kompletten Bericht Summary ist gross.	3 Analysebericht	Liked	Komplettbericht und Summary verursachen hohen Aufwand	Berichtqualität ist hoch, der Umfang jedoch gross	Struktur und Qualität der Ergebniskarteifakte sind grundsätzlich tragfähig	Berichtsumfang und Priorisierung fachlich steuern	Verdichtung, Priorisierung und fachliche Steuerung	RT-13	Rolle, Tiefe und Notwendigkeit des Komplettberichts sind zu klären, und der Aufwand für	Analyse	B1, NA1.3, B5	Liked: funktioniert, sehr ausführlich, Zeitaufwand Komplettbericht und Summary ist gross, M365-
Ausführlichkeit des Kurzberichts	3 Analysebericht	Lerned	Kurzbericht braucht eine passende Tiefe	Berichtsumfang und Priorisierung brauchen bewusste fachliche Steuerung	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Berichtsumfang und Priorisierung fachlich steuern	Verdichtung, Priorisierung und fachliche Steuerung					
Fachliche Expertise zwingend (M365)	3 Analysebericht	Liked	M365-Fachwissen ist für den Bericht unverzichtbar	Fachliche Expertise bleibt zentral für den Bericht	Fachliche Kompetenz ist ein zentraler Qualitätsfaktor	Berichtsumfang und Priorisierung fachlich steuern	Verdichtung, Priorisierung und fachliche Steuerung					
Massnahmenpakete Checken	3 Analysebericht	Longed for	Massnahmenpakete fachlich prüfen	Vor Verdichtung braucht es menschliche Prüfung und Priorisierung	Ergebnisse vor der Verdichtung besser prüfen und verständlich machen	Berichtsumfang und Priorisierung fachlich steuern	Verdichtung, Priorisierung und fachliche Steuerung					
Priorisierung der Massnahmen fraglich	3 Analysebericht	Lerned	Priorisierung der Massnahmen braucht menschliches Urteil	Berichtsumfang und Priorisierung brauchen bewusste fachliche Steuerung	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Berichtsumfang und Priorisierung fachlich steuern	Verdichtung, Priorisierung und fachliche Steuerung					
Priorisierung durch Berater	3 Analysebericht	Lerned	Priorisierung der Massnahmen braucht menschliches Urteil	Berichtsumfang und Priorisierung brauchen bewusste fachliche Steuerung	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Berichtsumfang und Priorisierung fachlich steuern	Verdichtung, Priorisierung und fachliche Steuerung					
Relevanz des Komplettberichts für Berater?	3 Analysebericht	Lerned	Rolle und Nutzen des Komplettberichts sind unklar	Berichtsumfang und Priorisierung brauchen bewusste fachliche Steuerung	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Berichtsumfang und Priorisierung fachlich steuern	Verdichtung, Priorisierung und fachliche Steuerung					
Komplettbericht "Einschätzung: unklar"	3 Analysebericht	Lerned	Rolle und Nutzen des Komplettberichts sind unklar	Berichtsumfang und Priorisierung brauchen bewusste fachliche Steuerung	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Berichtsumfang und Priorisierung fachlich steuern	Verdichtung, Priorisierung und fachliche Steuerung					
Komplettbericht Manuelle Priorisierung. Daraus dann das Summary.	3 Analysebericht	Longed for	Summary erst nach manueller Priorisierung ableiten	Vor Verdichtung braucht es menschliche Prüfung und Priorisierung	Ergebnisse vor der Verdichtung besser prüfen und verständlich machen	Berichtsumfang und Priorisierung fachlich steuern	Verdichtung, Priorisierung und fachliche Steuerung					
Aufgabenteilung war ok	3 Analysebericht	Liked	Aufgabenteilung in der Berichtserstellung funktioniert	Arbeitsteilung in der Berichtserstellung funktioniert	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Rollen und Übergänge in der Berichtserstellung klären	Zusammenarbeit, Austausch und Rollenklarheit					
zusätzliche Einblicke durch Austausch	4 Präsentation	Liked	Diskussion und Austausch schaffen zusätzlichen Mehrwert	Interaktion und Diskussion mit dem Kunden sind wertvoll	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Interaktion, Moderation und Engagement in der Präsentation stärken	Kundeninteraktion, Engagement und Diskussion	RT-21	Der Diskussions- und Austauschraum in der Präsentation ist bewusst einzuplanen und zu	Präsentation	B5, NA5.2	Liked: zusätzliche Einblicke durch Austausch, Austausch wertvoll, Kunde war interessiert, Fragen.
Austausch wertvoll => ggf. Abbrechen am Ende Diskussions. Spannende Diskussion	4 Präsentation	Liked	Diskussion und Austausch schaffen zusätzlichen Mehrwert	Interaktion und Diskussion mit dem Kunden sind wertvoll	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Interaktion, Moderation und Engagement in der Präsentation stärken	Kundeninteraktion, Engagement und Diskussion	RT-22	Die Kundenbegleitung und -information über den Prozess hinaus ist	Gesamtprozess / Präsentation	B5	Longed for: engagierter Kunde, Learned: Kunde noch mehr begleiten, informieren
Fragen Diskussion und Atmosphäre	4 Präsentation	Liked	Fragen, Diskussion und Atmosphäre sind positiv	Interaktion und Diskussion mit dem Kunden sind wertvoll	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Interaktion, Moderation und Engagement in der Präsentation stärken	Kundeninteraktion, Engagement und Diskussion					
Kunde war interessiert	4 Präsentation	Liked	Kundeninteresse und Beteiligung sind hoch	Interaktion und Diskussion mit dem Kunden sind wertvoll	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Interaktion, Moderation und Engagement in der Präsentation stärken	Kundeninteraktion, Engagement und Diskussion					
engagierter Kunde	4 Präsentation	Longed for	Mehr Kundenengagement in der Präsentation fördern	Interaktion und Engagement in der Präsentation weiter stärken	Präsentations- und Evidenzdarstellung weiterentwickeln	Interaktion, Moderation und Engagement in der Präsentation stärken	Kundeninteraktion, Engagement und Diskussion					
Moderation Bündelung von Rückfragen	4 Präsentation	Lerned	Rückfragen aktiv moderieren und bündeln	Diskussion und Rückfragen sollten aktiv moderiert werden	Moderation, Abstimmung und Verantwortlichkeiten brauchen klare Führung	Interaktion, Moderation und Engagement in der Präsentation stärken	Kundeninteraktion, Engagement und Diskussion					
Bericht für den Kunden	4 Präsentation	Lerned	Schriftlicher Bericht für den Kunden ist wichtig	Schriftliche Unterlagen bleiben ein wichtiger Teil der Kommunikation	Schriftliche Unterlagen bleiben ein wichtiger Teil der Kommunikation	Evidenzen und begleitende Unterlagen der Präsentation gezielt einsetzen	Kundeninteraktion, Engagement und Diskussion					
Automatische PPT- Sinnvoll?	4 Präsentation	Longed for	Automatisierung der Präsentation prüfen	Präsentationsformat und Erstellung weiterentwickeln	Präsentations- und Evidenzdarstellung weiterentwickeln	Format, Aufbau und Rollen der Präsentation weiterentwickeln	Präsentation, Moderation und adressatengerechte Kommunikation	RT-20	Ein anschlussfähiges Summary ist als Begleitdokument zur Präsentation zu erhalten.	Präsentation	B5, NA5.1, Scope-Entscheid T-6	Liked: Bericht (Summary) doch nötig, Aufbau/Struktur/Folien Darstellung nachvollziehbar
"Opportu nitylätsko sten"	4 Präsentation	Lerned	Business-Kontext und Opportunitätskosten mitdenken	Präsentation stärker auf Entscheidungen und Business-Relevanz ausrichten	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation adressatengerecht und entscheidungsorientiert ausrichten	Präsentation, Moderation und adressatengerechte Kommunikation					
Vorab unterlagen senden?	4 Präsentation	Lacked	Einsatz von Vorabunterlagen gezielt prüfen	Präsentation adressatengerecht aufbauen	Präsentation und Übergabe müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Präsentation adressatengerecht und entscheidungsorientiert ausrichten	Präsentation, Moderation und adressatengerechte Kommunikation					

workshop_ergebnisse - ROH Thematisiert Frage												
Input	Phase	L	Ebene Thema	Überthema	Überthema-L	Überthema-Phase	Überthema-komplett	Nr	Thema für die nächste Iteration	Teilprozess/Phase	Bezug B1-B5 und NA	Stützende 4L Bedunde
Fokus auf Empfehlung	4 Präsentation	Learned	Empfehlungen stärker ins Zentrum der Präsentation stellen	Präsentation stärker auf Entscheidungen und Business-Relevanz ausrichten	Verdichtung, Priorisierung und Strukturierung müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Präsentation adressatengerecht und entscheidungsorientiert ausrichten	Präsentation, Moderation und adressatengerechte Kommunikation					
relativieren i.s.v. Momentaufnahme/Foto der Organisation	4 Präsentation	Lacked	Ergebnisse als Momentaufnahme stärker relativieren	Aussagen in der Präsentation ausgewogener formulieren	Präsentation und Übergabe müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Präsentation adressatengerecht und entscheidungsorientiert ausrichten	Präsentation, Moderation und adressatengerechte Kommunikation					
Präsentation der Evidenzen nur in Gaps-Teil	4 Präsentation	Longed for	Evidenzen selektiv und gezielt im Gaps-Teil zeigen	Evidenzen gezielt und fachlich sicher präsentieren	Präsentations- und Evidenzdarstellung weiterentwickeln	Evidenzen und begleitende Unterlagen der Präsentation gezielt einsetzen	Präsentation, Moderation und adressatengerechte Kommunikation					
Aufteilung gut/ Foliendarstellung, Layout	4 Präsentation	Liked	Foliendarstellung und Aufteilung funktionieren gut	Präsentationsform und Rollenaufteilung funktionieren insgesamt gut	Zusammenarbeit und Rollenaufteilung tragen sichtbar zur Qualität bei	Format, Aufbau und Rollen der Präsentation weiterentwickeln	Präsentation, Moderation und adressatengerechte Kommunikation					
Teilnehmer Kundenseitig	4 Präsentation	Lacked	Kundenseitige Teilnehmende bewusster wählen	Vorlauf, Abstimmung und Rahmung der Präsentation verbessern	Präsentation und Übergabe müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Nicht nur ein Presenter	4 Präsentation	Liked	Mehrere Präsentierende können sinnvoll sein	Präsentationsform und Rollenaufteilung funktionieren insgesamt gut	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Format, Aufbau und Rollen der Präsentation weiterentwickeln	Präsentation, Moderation und adressatengerechte Kommunikation					
"Was ist gut" erwähnt	4 Präsentation	Liked	Positive Aspekte in der Präsentation sichtbar machen	Präsentationsform und Rollenaufteilung funktionieren insgesamt gut	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Format, Aufbau und Rollen der Präsentation weiterentwickeln	Präsentation, Moderation und adressatengerechte Kommunikation					
Max 2 Std.	4 Präsentation	Learned	Präsentation kürzer, fokussierter und weniger repetitiv gestalten	Präsentation stärker verdichten und intern klarer abstimmen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Ausführlichkeit der	4 Präsentation	Learned	Präsentation kürzer, fokussierter und weniger repetitiv gestalten	Präsentation stärker verdichten und intern klarer abstimmen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Präsentation Länge der	4 Präsentation	Learned	Präsentation kürzer, fokussierter und weniger repetitiv gestalten	Präsentation stärker verdichten und intern klarer abstimmen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Präsentation Viele Wiederholungen	4 Präsentation	Learned	Präsentation kürzer, fokussierter und weniger repetitiv gestalten	Präsentation stärker verdichten und intern klarer abstimmen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Timekeeping in Präsentation (Weniger ist mehr)	4 Präsentation	Learned	Präsentation kürzer, fokussierter und weniger repetitiv gestalten	Präsentation stärker verdichten und intern klarer abstimmen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Zeitmanagement Einhalten war schwer	4 Präsentation	Learned	Präsentation kürzer, fokussierter und weniger repetitiv gestalten	Präsentation stärker verdichten und intern klarer abstimmen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Zu Umfangreich viele Wiederholungen	4 Präsentation	Learned	Präsentation kürzer, fokussierter und weniger repetitiv gestalten	Präsentation stärker verdichten und intern klarer abstimmen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Detail => Abgeben Präsi => Kurz und Bündig	4 Präsentation	Learned	Präsentation kürzer, fokussierter und weniger repetitiv gestalten	Präsentation stärker verdichten und intern klarer abstimmen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Zu umfangreich/ Wiederholend	4 Präsentation	Learned	Präsentation kürzer, fokussierter und weniger repetitiv gestalten	Präsentation stärker verdichten und intern klarer abstimmen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Wiederholungen der Massnahmen (Vollständigkeit vs. Adressgerecht)	4 Präsentation	Lacked	Präsentation stringenter und adressatengerechter ausrichten	Präsentation adressatengerechter aufbauen	Präsentation und Übergabe müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Präsentation adressatengerecht und entscheidungsorientiert ausrichten	Präsentation, Moderation und adressatengerechte Kommunikation					
Stringenz	4 Präsentation	Lacked	Präsentation stringenter und adressatengerechter ausrichten	Präsentation adressatengerechter aufbauen	Präsentation und Übergabe müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Präsentation adressatengerecht und entscheidungsorientiert ausrichten	Präsentation, Moderation und adressatengerechte Kommunikation					
Obsolete Folien, Zu viele Folien - adressatengerecht?	4 Präsentation	Lacked	Präsentation stringenter und adressatengerechter ausrichten	Präsentation adressatengerechter aufbauen	Präsentation und Übergabe müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Präsentation adressatengerecht und entscheidungsorientiert ausrichten	Präsentation, Moderation und adressatengerechte Kommunikation					
Publikum gerecht/ adressatengerecht	4 Präsentation	Lacked	Präsentation stringenter und adressatengerechter ausrichten	Präsentation adressatengerechter aufbauen	Präsentation und Übergabe müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Präsentation adressatengerecht und entscheidungsorientiert ausrichten	Präsentation, Moderation und adressatengerechte Kommunikation					
Ergebnis qualifiziert & Umfangreich	4 Präsentation	Liked	Präsentationsinhalt ist fundiert und umfangreich	Inhalt und Aufbau der Präsentation sind grundsätzlich tragfähig	Struktur und Qualität der Ergebnisartefakte sind grundsätzlich tragfähig	Format, Aufbau und Rollen der Präsentation weiterentwickeln	Präsentation, Moderation und adressatengerechte Kommunikation					
Aufbau struktur Nachvollziehbar	4 Präsentation	Liked	Präsentationsstruktur und Aufbau sind nachvollziehbar	Inhalt und Aufbau der Präsentation sind grundsätzlich tragfähig	Struktur und Qualität der Ergebnisartefakte sind grundsätzlich tragfähig	Format, Aufbau und Rollen der Präsentation weiterentwickeln	Präsentation, Moderation und adressatengerechte Kommunikation					
Präsentierende Vorbereitung/ Fachliche Tiefe i.s.v. Evidenzen im Kopf!	4 Präsentation	Longed for	Präsentierende brauchen fachliche Tiefe und Evidenzsicherheit	Evidenzen gezielt und fachlich sicher präsentieren	Präsentations- und Evidenzdarstellung weiterentwickeln	Evidenzen und begleitende Unterlagen der Präsentation gezielt einsetzen	Präsentation, Moderation und adressatengerechte Kommunikation					
Weniger wertend	4 Präsentation	Lacked	Sprache in der Präsentation stärker neutralisieren	Aussagen in der Präsentation ausgewogener formulieren	Präsentation und Übergabe müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Präsentation adressatengerecht und entscheidungsorientiert ausrichten	Präsentation, Moderation und adressatengerechte Kommunikation					
Bericht (Summary) Doch nötig	4 Präsentation	Liked	Summary als Begleitdokument bleibt sinnvoll	Inhalt und Aufbau der Präsentation sind grundsätzlich tragfähig	Struktur und Qualität der Ergebnisartefakte sind grundsätzlich tragfähig	Format, Aufbau und Rollen der Präsentation weiterentwickeln	Präsentation, Moderation und adressatengerechte Kommunikation					

workshop_ergebnisse - ROH Thematisiert Frage												
Input	Phase	L	Ebene Thema	Überthema	Überthema-L	Überthema-Phase	Überthema-komplett	Nr	Thema für die nächste Iteration	Teilprozess/Phase	Bezug B1-B5 und NA	Stützende 4L Bedunde
Absprache vor Präsi	4 Präsentation	Lacked	Teamabstimmung vor der Präsentation verbessern	Vorlauf, Abstimmung und Rahmung der Präsentation verbessern	Präsentation und Übergabe müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Engagement vom Team	4 Präsentation	Lacked	Teamabstimmung vor der Präsentation verbessern	Vorlauf, Abstimmung und Rahmung der Präsentation verbessern	Präsentation und Übergabe müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
5 Personen zu viel (heterogen, Nicht abstimmt)	4 Präsentation	Learned	Weniger Präsentierende und klarere Abstimmung	Präsentation stärker verdichten und intern klarer abstimmen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Viele Köche -> Roter Faden/ Einheitslich	4 Präsentation	Learned	Weniger Präsentierende und klarere Abstimmung	Präsentation stärker verdichten und intern klarer abstimmen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Anzahl der Präsentierenden	4 Präsentation	Learned	Weniger Präsentierende und klarere Abstimmung	Präsentation stärker verdichten und intern klarer abstimmen	Verdichtung, Priorisierung und Strukturierung müssen aktiv gesteuert werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Ziel der Präsentation (evtl. Vorgesprochen)	4 Präsentation	Lacked	Ziel und Erwartung der Präsentation vorab klären	Vorlauf, Abstimmung und Rahmung der Präsentation verbessern	Präsentation und Übergabe müssen besser vorbereitet und adressatengerecht ausgerichtet werden	Präsentation inhaltlich verdichten und intern besser abstimmen	Präsentation, Moderation und adressatengerechte Kommunikation					
Evtl. zwei Versionen für Kunden - Präsentation vs. Ergebnisse	4 Präsentation	Longed for	Zielgruppenspezifische Präsentationsvarianten entwickeln	Präsentationsformat und Erstellung weiterentwickeln	Präsentations- und Evidenzdarstellung weiterentwickeln	Format, Aufbau und Rollen der Präsentation weiterentwickeln	Präsentation, Moderation und adressatengerechte Kommunikation					
Zweite management	4 Präsentation	Longed for	Zielgruppenspezifische Präsentationsvarianten entwickeln	Präsentationsformat und Erstellung weiterentwickeln	Präsentations- und Evidenzdarstellung weiterentwickeln	Format, Aufbau und Rollen der Präsentation weiterentwickeln	Präsentation, Moderation und adressatengerechte Kommunikation					
Kunde noch mehr begleiten und Informieren	5 Gesamtprozess	Learned	Kundenkommunikation und Begleitung ausbauen	Kundenkommunikation im Gesamtprozess ausbauen	Moderation, Abstimmung und Verantwortlichkeiten brauchen klare Führung	Kundenkommunikation im Gesamtprozess ausbauen	Kundeninteraktion, Engagement und Diskussion					
Gut aber nicht perfekt	5 Gesamtprozess	Liked	Gesamtprozess funktioniert, bleibt aber ausbaufähig	Der Gesamtprozess stiftet Nutzen, bleibt aber optimierbar	Ergebnisse stiftet erkennbaren Nutzen	Gesamtprozess gezielt weiterentwickeln und standardisieren	Prozessweiterentwicklung, Standardisierung und Reflexion	RT-25	Wiederverwendbare Standardbausteine (Katalog) sind aufbauend von den	Gesamtprozess	B4, NA4.1, NA4.2	Learned: Katalog Standardbausteine. Liked: wichtige für Grundfragen für
Wichtig für Grundlagen für Entscheidungen	5 Gesamtprozess	Liked	Prozess schafft gute Entscheidungsgrundlagen	Der Gesamtprozess stiftet Nutzen, bleibt aber optimierbar	Ergebnisse stiftet erkennbaren Nutzen	Gesamtprozess gezielt weiterentwickeln und standardisieren	Prozessweiterentwicklung, Standardisierung und Reflexion	RT-26	Eine Reflexion bzw. Retrospektive ist als fester Schritt im Prozess zu verankern, unmittelbar nach	Gesamtprozess	B4, menschenzentrierte Iteration	Learned: Retrospektive unmittelbar nach der Präsentation. Longed for: Reflexion als festen Schritt
Reflexionsworkshop 23.04.2026	5 Gesamtprozess	Longed for	Reflexion als festen Schritt im Prozess verankern	Reflexion als festen Schritt im Prozess verankern	Reflexion und Standardisierung sollten im Prozess verankert werden	Reflexion als festen Prozessschritt verankern	Prozessweiterentwicklung, Standardisierung und Reflexion					
Retro unmittelbar nach Präsentieren	5 Gesamtprozess	Learned	Retrospektive unmittelbar nach der Präsentation durchführen	Reflexion zeitnah im Prozess verankern	Reflexion und Standardisierung sollten im Prozess verankert werden	Reflexion als festen Prozessschritt verankern	Prozessweiterentwicklung, Standardisierung und Reflexion					
Katalog: Standard Bausteine	5 Gesamtprozess	Learned	Standardbausteine für den Prozess aufbauen	Prozess stärker standardisieren	Reflexion und Standardisierung sollten im Prozess verankert werden	Gesamtprozess gezielt weiterentwickeln und standardisieren	Prozessweiterentwicklung, Standardisierung und Reflexion					
5 Kg	5 Gesamtprozess	Liked	Unklarer Einzelpunkt / bitte im Original prüfen	Unklarer Einzelpunkt / bitte im Original prüfen	Unklar / bitte im Original prüfen	Unklarer Einzelpunkt / bitte im Original prüfen	Unklar / bitte im Original prüfen					
Mehr infos untereinander? Ok?	5 Gesamtprozess	Lacked	Interne Abstimmung und Informationsfluss stärken	Interne Abstimmung und Informationsfluss stärken	Interne Abstimmung und Informationsfluss stärken	Interne Abstimmung, Rollen und Informationsfluss verbessern	Zusammenarbeit, Austausch und Rollenklarheit					
Mehr gezielte Absprache	5 Gesamtprozess	Learned	Interne Abstimmung und Informationsfluss stärken	Interne Abstimmung und Verantwortlichkeiten klarer führen	Moderation, Abstimmung und Verantwortlichkeiten brauchen klare Führung	Interne Abstimmung, Rollen und Informationsfluss verbessern	Zusammenarbeit, Austausch und Rollenklarheit					
Einfache Koordination	5 Gesamtprozess	Liked	Koordination im Gesamtprozess funktioniert gut	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit, Austausch und Rollenklarheit					
Präsentation bei "fertig" - Verantwortlichkeiten wahrnehmen	5 Gesamtprozess	Learned	Verantwortlichkeiten bis zur Präsentation klar wahrnehmen	Interne Abstimmung und Verantwortlichkeiten klarer führen	Moderation, Abstimmung und Verantwortlichkeiten brauchen klare Führung	Interne Abstimmung, Rollen und Informationsfluss verbessern	Zusammenarbeit, Austausch und Rollenklarheit					
Vor Ort- Austausch künfte	5 Gesamtprozess	Liked	Vor-Ort-Austausch stärkt die Zusammenarbeit	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit, Austausch und Rollenklarheit					
Angenehme Zusammenarbeit	5 Gesamtprozess	Liked	Zusammenarbeit im Gesamtprozess ist angenehm	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit, Austausch und Rollenklarheit					
Angenehme Zusammenarbeit	5 Gesamtprozess	Liked	Zusammenarbeit im Gesamtprozess ist angenehm	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit, Austausch und Rollenklarheit					
Zusammenarbeit und Resultate	5 Gesamtprozess	Liked	Zusammenarbeit trägt zu guten Resultaten bei	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit, Austausch und Rollenklarheit					
: 3 Firmen	5 Gesamtprozess	Liked	Zusammenarbeit über mehrere Organisationen ist grundsätzlich tragfähig	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit, Austausch und Rollenklarheit					
Zusammenarbeit / Engagement	5 Gesamtprozess	Liked	Zusammenarbeit und Engagement im Gesamtprozess sind gut	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit und Austausch tragen sichtbar zur Qualität bei	Zusammenarbeit und Koordination im Gesamtprozess funktionieren gut	Zusammenarbeit, Austausch und Rollenklarheit					

workshop_ergebnisse - RT und RL Übersicht

Nr.	Überthema-komplett	Thema für die nächste Iteration	Teilprozess / Phase	Bezug B1-B5 / NA	Stützende 4L-Befunde
RT-01	Organisation, Planung und Ablauf des Vorgehens	Der niederschwellige Umfragezugang über Microsoft Forms mit Versand über die Kundenorganisation hat sich bewährt und ist als Verteilweg beizubehalten, da er Rücklauf und Teilnahme sichert.	Erhebung / Umfrage	B4, B3, NA1.1	Liked: Forms einfacher Zugriff, Mail zu Forms über Kunde versandt, Rücklaufquote/Timeline, Teilnahmequote, Terminfindung und Flexibilität.
RT-02	Organisation, Planung und Ablauf des Vorgehens	Die Terminverbindlichkeit der Interviews ist aktiv abzusichern (Erinnerungen, klare Zusagen), um No-shows und Terminausfälle	Erhebung / Interviews	B4, NA4.3	Lacked: Termin vergessen (No-show). Liked: Terminfindung und Flexibilität (beibehalten).
RT-03	Schulung, Leitfäden und methodische Rahmung	Der Interviewleitfaden ist flexibel und beispielbasiert zu gestalten und begrifflich zu vereinheitlichen, sodass er Orientierung gibt.	Erhebung / Interviews	B3, NA3.3	Learned: Fragekatalog Light/Beispiele, Leitfaden und Flexibilität. Lacked: Frageliste hemmt Spontaneität, einheitliches Wording, Fragen dem Kunden anpassen / Lerned für Beispielfragen. Unnötigkeit mit Fragen.
RT-04	Schulung, Leitfäden und methodische Rahmung	Interview- und Umfragefragen müssen konsistent auf die Building Blocks, die Massnahmen und die angestrebten Endergebnisse abgebildet sein (Mapping), damit die nachgelagerte Analyse nachvollziehbar und verständlich ist.	Erhebung / Analyse (Codierung)	NA3.1, B3, Codierung (T-3)	Learned: Match/Zuweisung Interviewfragen zu Umfrage und Building Blocks. Lacked: Umfragen checken (Massnahmen und Blocks), Mapping Fragen zu Endergebnissen, Fragen besser auf Massnahmen und Building Blocks abstimmen.
RT-05	Schulung, Leitfäden und methodische Rahmung	Der Umfragefragenkatalog ist präziser, verständlicher und auf das Wesentliche fokussiert zu gestalten, mit reduzierter Komplexität und weniger reinen Schulungsfragen, damit Teilnehmende die Ergebnisse leichter verstehen.	Erhebung / Umfrage	B3, NA3.2, B5	Lacked: präziserer klarerer Fragenkatalog, Anzahl ok und Formulierung, viele Schulungsfragen reduzieren. Learned: irritierte Probanden (Verständnis), effizientere Methode wählen. Liked: Umfang/Zeitaufwand passend, offene/freie Textfelder wertvoll (beibehalten), Komplexität kritisch.
RT-06	Adaptive Interviewführung und passende Auswahl der Interviewten	Die Auswahl der Interviewpartner ist bewusst auf möglichst heterogene Profile auszurichten, die Rolle der Führungsebene in der Stichprobe ist zu klären, und vor dem Interview ist ein kurzer Technik-Check vorzusehen.	Erhebung / Interviews	B3, NA3.2	Learned: Auswahl passender Interviewpartner, möglichst unterschiedliche Profile, Technik-Check bei Interviews, Technik-Kompetenz benötigt Kompromisse. Lacked (aus Cluster Organisation): Führungsebene Interview nötig? Übrige Befunde des Clusters (Beziehung, Empathie, Natürlichkeit, Unvoreingenommenheit) siehe Learnings RL-01.
RT-07	Nachbearbeitung, Transkriptarbeit und Verwertung	Der fachliche Kontrolldurchgang am Transkript (Prüfung der Worttreue vor der sprachlichen Glättung, Bereinigung nach definierten Regeln) hat sich bewährt und ist als verbindlicher Schritt im Transkriptprozess zu nehmen.	Transkript	NA2.1, NA2.2, B2	Liked: gut bereinigtes Transkript, Transkript-Kontrolldurchgang, Prüfung Transkript, Korrektur war einfach, Aufnahme der Interviews.
RT-08	Nachbearbeitung, Transkriptarbeit und Verwertung	Durchführung und Transkriptprüfung sind enger zu koppeln, sodass möglichst die interviewende Person das zugehörige	Transkript	NA2.3, B2	Longed for: Durchführung/Prüfung Transkript koppeln. Learned: aktives Zuhören sichtbar im Transkript, Wichtiges heraushören.
RT-09	Nachbearbeitung, Transkriptarbeit und Verwertung	Die Anzahl der Interviews ist an der inhaltlichen Sättigung auszurichten, da bereits wenige Interviews einen Grossteil der zentralen Befunde liefern, was die angestrebte Verschlinkung des Lohs-Ansatzes erleichtert. Die Evidenzführung in der Analyse muss verständlich, divers und redundanzarm sein und die Evidenzen klar den Building Blocks zuordnen, sodass Befunde nachvollziehbar belegt und nicht durch Mehrdeutigkeiten und Widersprüche verunsichert werden.	Erhebung / Analyse	B3, B1, NA3.3	Learned: 3/4 der Findings in 2 Interviews klar. Hinweis: Grauzone, gegen Repräsentativität (B3) abzuwägen.
RT-10	Evidenzbasis, Ergebnisqualität und Verständlichkeit	Die Evidenzführung in der Analyse muss verständlich, divers und redundanzarm sein und die Evidenzen klar den Building Blocks zuordnen, sodass Befunde nachvollziehbar belegt und nicht durch Mehrdeutigkeiten und Widersprüche verunsichert werden.	Analyse	NA2.2, NA3.1, NA3.2, B2, B3	Lacked: Evidenzen zu kompliziert zitiert, Wiederholung der Evidenzen, viele ähnliche Themen pro Building Block, Diversität der Evidenzen, Evidenzqualität besser, Evidenzen passend zu Building Blocks fraglich, mehr Evidenzen zu Massnahmen, Ergebnisse qualitativ hochwertig halten.
RT-11	Evidenzbasis, Ergebnisqualität und Verständlichkeit	Quantitative Umfrageergebnisse (Kennzahlen wie NPS, qualitative Freitextantworten) sind fachlich einzuordnen und verständlich aufzubereiten, damit sie ohne Fehlinterpretation herangezogen werden können.	Analyse	B3, B5, NA5.1, B2	Lacked: Umfrageergebnisse (NPS) schwer zu verstehen. Longed for: Know-how Ergebnis-Interpretation, qualitative Ergebnisse nicht lesend dargestellt. Bezug Kapitel 7, T-2: quantitative Auswertungen ohne fachliche Einordnung begünstigen Fehlinterpretationen.
RT-12	Verdichtung, Priorisierung und fachliche Steuerung	Die Priorisierung der Massnahmen muss als bewusster, manueller Fachschritt erfolgen, aus dem anschliessend das Summary abgeleitet wird, da diese Bewertung menschliches	Analyse	B2, NA2.1, B5, NA5.2	Longed for: Kompletbericht manuelle Priorisierung, daraus das Summary, Massnahmenpakete checken. Learned: Priorisierung der Massnahmen fraglich, Priorisierung durch Berater (menschliches Urteil), Anschluss an Vibe-Research-Regel R2 (Bedeutungszuschreibung bei der forschenden
RT-13	Verdichtung, Priorisierung und fachliche Steuerung	Rolle, Tiefe und Notwendigkeit des Kompletberichtes sind zu klären, und der Aufwand für Kompletbericht und Summary ist zu reduzieren, etwa indem der Kompletbericht als Zwischenartefakt	Analyse	B1, NA1.3, B5	Liked: funktioniert, sehr ausführlich, Zeitaufwand Kompletbericht und Summary ist gross, M365-Fachwissen zwingend (bestätigt B2). Learned: Ausführlichkeit des Kurzberichts, Relevanz/Einschätzung des Kompletberichtes unklar.
RT-14	Struktur, Lesbarkeit und Kohärenz der Ergebnisartefakte	Die bewährte Berichtsstruktur entlang der Building Blocks mit Zielen, Lücken und Empfehlungen ist zu erhalten und um Orientierungshilfen zu ergänzen (Navigation, Lesehilfe), damit der	Analyse	NA3.1, NA5.1, B5	Liked: Aufbau Bericht, Aufbau auf Building Blocks, Gliederung in Ziele, Lücken, Empfehlungen. Longed for: sehr umfangreich, es braucht Orientierung.
RT-15	Struktur, Lesbarkeit und Kohärenz der Ergebnisartefakte	Die Kennzahlenlogik ist zu vereinheitlichen, sodass Messgrössen konsistent verwendet werden und keine Vermischung von	Analyse	B2, Ergebnisverlässlichkeit	Learned: Vermischung der Messgrössen (Durchschnitt, Median, Mittelwert).
RT-16	Struktur, Lesbarkeit und Kohärenz der Ergebnisartefakte	Die Building-Block-Struktur ist fachlich feiner zu gliedern, etwa die Aufteilung des Blocks Digital Workplace in Communicate, Collaborate und gegebenenfalls Automate, was direkt in die	Analyse (Codierung)	NA3.1, Codierung (T-3)	Learned: Block Digital Workplace schon unterteilen in Communicate und Collaborate (ggf. Automate).
RT-17	Präsentation, Moderation und adressatengerechte Kommunikation	Die Präsentation muss verdichtet, fokussiert und redundanzarm sein, mit klarer Empfangsorientierung und eingehaltenen Zeitrahmen, sodass Vollständigkeit nicht zulasten der Wirkung	Präsentation	B5, NA5.1, NA3.2	Learned: max. 2 Std., Fokus auf Empfehlung, viele Wiederholungen, Länge/Ausführlichkeit, zu umfangreich, Timekeeping (weniger ist mehr), Detail abgeben, kurz und bündig. Lacked: Stringenz, obsoleter/zu viele Folien.
RT-18	Präsentation, Moderation und adressatengerechte Kommunikation	Die Ergebnisdarstellung muss adressatengerecht und wertneutral sein und die Ergebnisse als Momentaufnahme rahmen, bei Bedarf mit zielgruppenspezifischen Varianten (Kunde vs. Management) und einheitlichen Präsentationsregeln im Gesamte	Präsentation	B5, NA5.1, B3	Lacked: publikums-adressatengerecht, weniger wertend, relativieren als Momentaufnahme der Organisation. Longed for: zwei Versionen Kunde vs. Management, Präsentation der Evidenzen nur im Gaps-Teil.
RT-19	Präsentation, Moderation und adressatengerechte Kommunikation	Vor der Präsentation müssen Ziel, Rollen, Anzahl Präsentierende und kundenseitige Teilnehmerschaft abgestimmt sein, damit ein roter Faden entsteht und mehrere Präsentierende abgestimmt	Präsentation / Gesamtprozess	NA4.1, B4	Lacked: Absprache vor Präsi, Ziel der Präsentation vorbesprechen, Teilnehmer kundenseitig, Engagement vom Team. Learned: Anzahl der Präsentierenden, 5 Personen zu viel, viele Köche, roter Faden/vereinheitlich. Liked: nicht nur ein Presenter (mehrere sinnvoll, aber abgestimmt).
RT-20	Präsentation, Moderation und adressatengerechte Kommunikation	Ein anschliessendes Summary ist als Begleitdokument zur Präsentation zu erhalten, und die Befüllung der Präsentationsunterlagen bleibt bewusst manuell (keine	Präsentation	B5, NA5.1, Scope-Entscheid T-6	Liked: Bericht (Summary) doch nötig, Aufbau/Struktur/Folienanstellung nachvollziehbar (beibehalten). Longed for: automatische PPT, Sinnvoll? (kritisch geprüft und verworfen).
RT-21	Kundeninteraktion, Engagement und Diskussion	Der Diskussions- und Austauschraum in der Präsentation ist bewusst einzuplanen und zu moderieren (Rückfragen bündeln), da er nachweislich Mehrwert schafft, notfalls zulasten des	Präsentation	B5, NA5.2	Liked: zusätzliche Einblicke durch Austausch, Austausch wertvoll, Kunde war interessiert, Fragen, Diskussion und Atmosphäre. Learned: spannende Diskussion, Moderation/Bündelung von Rückfragen.
RT-22	Kundeninteraktion, Engagement und Diskussion	Die Kundenbegleitung und -information über den Prozess hinweg ist auszubauen, um Engagement und Anschlussfähigkeit der	Gesamtprozess / Präsentation	B5	Longed for: engagierter Kunde. Learned: Kunde noch mehr begleiten/informieren, schriftlicher Bericht für den Kunden wichtig.
RT-23	Zusammenarbeit, Austausch und Rollenklarheit	Rollen, Verantwortlichkeiten und der interne Informationsfluss müssen über den gesamten Prozess klar definiert sein, insbesondere die durchgängige Wahrnehmung der	Gesamtprozess	NA2.3, NA4.1, B4	Learned: mehr gezielte Absprache, Präsentation bei fertig, Verantwortlichkeiten wahrnehmen. Lacked: mehr Infos untereinander.
RT-24	Zusammenarbeit, Austausch und Rollenklarheit	Bewährte Arbeitsformen sind als Prozessmerkmale beizubehalten, namentlich die Interviewdurchführung zu zweit mit Co-Moderation, direkte Feedbackschleifen nach Interviews und	Interviews / Gesamtprozess	B2, B4, NA4.2	Liked: zu zweit in Interviews, Sicherheit durch Co-Moderation, direktes Feedback nach Interviews, Checklisten/Leitfäden, digitale Interviews und Werkzeuge, Aufgabenteilung war ok.

workshop_ergebnisse - RT und RL Übersicht

Nr.	Überthema-komplett	Thema für die nächste Iteration	Teilprozess / Phase	Bezug B1-B5 / NA	Stützende 4L-Befunde
RT-25	Prozessweiterentwicklung, Standardisierung und Reflexion	Wiederverwendbare Standardbausteine (Katalog) sind aufzubauen, um den Prozess planbarer und delegierbarer zu	Gesamtprozess	B4, NA4.1, NA4.2	Learned: Katalog Standardbausteine. Liked: wichtig für Grundlagen für Entscheidungen.
RT-26	Prozessweiterentwicklung, Standardisierung und Reflexion	Eine Reflexion bzw. Retrospektive ist als fester Schritt im Prozess zu verankern, unmittelbar nach der Präsentation, um die iterative	Gesamtprozess	B4, menschenzentrierte Iteration	Learned: Retrospektive unmittelbar nach der Präsentation. Longed for: Reflexion als festen Schritt im Prozess verankern (Reflexionsworkshop 23.04.2026).
RL-01	Learnings (übergreifend, prozessunabhängig)	Interviewqualität hängt wesentlich von Beziehung, Empathie, Natürlichkeit und Unvoreingenommenheit ab, also von Kompetenzen, die sich nicht in den Prozess hineingestalten lassen, sondern durch Erleben und Schulung entstehen.	Interviews	Interviewkompetenz (nicht prozessgestaltbar)	Learned: wie sich Menschen beim Interview fühlen, Beziehung zum Interviewpartner, persönliche Interviewführung (digital) wertvoll, Natürlichkeit und Spontanität/intuition/erfahrung, Unvoreingenommenheit, adaptive Führung bei heterogenen Voraussetzungen.
RL-02	Learnings (übergreifend, prozessunabhängig)	Einführungsworkshop und Interviewschulung haben sich bewährt und geben Sicherheit, sollten aber kompakter und übungsorientierter angelegt werden. Hinweis: Wird Onlineinterviewschulung als Teil der Didaktik verwendet, werden	Interviews	Schulung/Onboarding (Grauzone zu Prozess)	Liked: Einführungsworkshop, Schulung gute Basis/hilfreich, gutes Teambriefing (Didaktik, Verhalten). Lacked: Schulung mehr Proben als Schildern, Schulungsdauer verkürzen. Learned: Erwartungsmanagement (Interviewte ohne Vorbereitung).
RL-03	Learnings (übergreifend, prozessunabhängig)	Präsentierende benötigen fachliche Tiefe und Evidenzsicherheit sowie ein Bewusstsein für den Geschäftskontext (z. B. Opportunitätskosten). Dies ist eine Frage der Kompetenz und Unabhängigkeit des Betrachters, nicht der Prozessneutralität.	Präsentation	Präsentationskompetenz/Mindset	Longed for: Präsentation Vorbereitung, fachliche Tiefe (Evidenzen im Kopf). Learned: Opportunitätskosten, Business-Kontext mitdenken.
RL-04	Learnings (übergreifend, prozessunabhängig)	Die Zusammenarbeit über drei Firmen hinweg war angenehm, gut koordiniert und durch Vor-Ort-Austausch und ERFA wertvoll, was die grundsätzliche Tragfähigkeit eines firmenübergreifenden Vorgehens bestätigt.	Gesamtprozess / Interviews	Zusammen Arbeitsklima (bestätigend)	Liked: 3 Firmen, angenehme Zusammenarbeit, einfache Koordination, Vor-Ort-Zusammenkünfte, Austausch/ERFA, Zusammenarbeit und Engagement.
RL-05	Learnings (übergreifend, prozessunabhängig)	Die Interviews lieferten Ansatzpunkte und Ideen auch für kleinere Mandate und Massnahmen, was den Geschäftswert des	Interviews	Geschäftschance (nicht prozessgestaltend)	Liked: Möglichkeiten und Ideen für kleine MT.
RL-06	Learnings (übergreifend, prozessunabhängig)	Selbst- und Fremdwahrnehmung der Interviewdurchführung können auseinanderfallen, was als Erfahrungswert für die	Interviews	Reflexion/Kompetenz	Longed for: Diskrepanz Durchführung und Wahrnehmung.
RL-07	Learnings (übergreifend, prozessunabhängig)	Die Kombination des Reflexionsworkshops mit einer Retrospektive hat sich bewährt und liess sich auf einen Nachmittag begrenzen, was als methodischer Erfahrungswert für weitere Reflexionsworkshops dient.	Gesamtprozess	Workshop-Methodik	Bezug zum Reflexionsworkshop selbst (Kapitel 7, Lernreflexion).
RL-08	Learnings (übergreifend, prozessunabhängig)	Der Einzelpunkt 5 Kg (Liked, Gesamtprozess) ist inhaltlich unklar und im Originalmaterial zu prüfen, bevor er zugeordnet wird.	Gesamtprozess	Klärung offen (grau)	Liked: 5 Kg (unklar / bitte im Original prüfen).

workshop_ergebnisse - RT Verdichtet					
Nr.	Kurzbezeichnung	Verdichtetes Thema (ausführlich)	Verdichtetes Thema (konsolidiert)	Gruppierte RT	Bezug B1–B5 / NA
VT-1	Erhebung verlässlich, fokussiert und passend zusammensetzen	Die Datenerhebung muss organisatorisch verlässlich, instrumentell fokussiert und in der Stichprobe bewusst zusammengesetzt sein. Der bewährte niederschwellige Umfragezugang und die flexible Terminplanung sind zu erhalten, Termine verbindlicher abzusichern, Leitfaden und Fragenkatalog flexibel, verständlich und auf das.	Die Datenerhebung muss verlässlich organisiert und fachlich fokussiert sein, mit verbindlich abgesicherten Interviewterminen, dem bewährten niederschwelligen Umfragezugang, einem verständlichen und auf das	RT-01, RT-02, RT-03, RT-05, RT-06, RT-09	B3, B4 (NA3.2, NA3.3, NA4.3)
VT-2	Building-Block- und Codierungsgerüst kohärent und trennscharf halten	Das Building-Block- und Codierungsgerüst, das den Analyseprozess strukturiert, muss fachlich trennscharf und kohärent sein. Erhebungsinstrumente sind konsistent auf Building Blocks, Massnahmen und Endergebnisse abzubilden, und die Bausteine sind dort feiner zu gliedern, wo sie inhaltlich überladen sind (etwa Digital Workplace in der Aufbereitung der Transkripte muss inhaltstreu und kontrolliert bleiben. Der fachliche Kontrolldurchgang mit Prüfung der Worttreue vor der sprachlichen Glättung ist als verbindlicher Schritt zu erhalten und möglichst eng an die interviewende Person zu koppeln.	Das Building-Block- und Codierungsgerüst muss fachlich trennscharf und kohärent sein, indem überladene Bausteine feiner gegliedert werden, etwa Digital Workplace in Communicate, Collaborate und Automate, und Interview- wie	RT-04, RT-16	NA3.1, B3
VT-3	Transkripte inhaltstreu und kontrolliert aufbereiten	Die Transkripte müssen inhaltstreu aufbereitet werden, mit einem verbindlichen fachlichen Kontrolldurchgang, der die Worttreue vor der sprachlichen Glättung prüft und möglichst von der interviewenden Person durchgeführt wird.	Die Transkripte müssen inhaltstreu aufbereitet werden, mit einem verbindlichen fachlichen Kontrolldurchgang, der die Worttreue vor der sprachlichen Glättung prüft und möglichst von der interviewenden Person durchgeführt wird.	RT-07, RT-08	B2 (NA2.1, NA2.2, NA2.3)
VT-4	Analyseergebnisse belegbar, verständlich und fachlich gesteuert verdichten	Die Analyse muss belegbare, verständliche und fachlich gesteuerte Ergebnisse erzeugen. Die Evidenzführung ist verständlich, divers und redundanzarm an die Building Blocks zu binden, quantitative Ergebnisse sind fachlich einzuordnen, die Priorisierung bleibt menschliches Fachurteil als Grundlage des Summaries, der Aufwand für Komplettbericht und	Die Analyse muss belegbare und verständliche Ergebnisse erzeugen, indem Evidenzen redundanzarm den Building Blocks zugeordnet und quantitative Kennzahlen fachlich eingeordnet werden, wobei die Priorisierung als manueller	RT-10, RT-11, RT-12, RT-13, RT-14, RT-15	B2, B3, B5, B1 (NA2.1, NA2.2, NA3.1, NA3.2, NA5.1, NA5.2, NA1.3)
VT-5	Ergebnisse verdichtet, adressatengerecht und dialogorientiert präsentieren	Die Präsentation muss die Ergebnisse verdichtet, adressatengerecht und dialogorientiert vermitteln. Sie ist auf Empfehlungen zu fokussieren und im Zeitrahmen zu halten, wertneutral als Momentaufnahme zu rahmen, in Ziel, Rollen und Beteiligten vorab abzustimmen, durch ein anschlussfähiges Summary zu begleiten (bei bewusst manueller	Die Präsentation muss die Ergebnisse fokussiert und empfehlungsorientiert im vorgegebenen Zeitrahmen vermitteln, wertneutral als Momentaufnahme gerahmt, in Ziel und Rollen vorab abgestimmt, durch ein anschlussfähiges	RT-17, RT-18, RT-19, RT-20, RT-21, RT-22	B5, B3, B4 (NA5.1, NA5.2, NA3.2, NA4.1)
VT-6	Prozess durch klare Rollen, bewährte Arbeitsformen und Reflexion steuern	Der Gesamtprozess muss steuerbar bleiben durch klar definierte Rollen und einen geklärten internen Informationsfluss, bewährte Arbeitsformen wie Co-Moderation, Feedbackschleifen und digitale Werkzeuge beizubehalten, wiederverwendbare	Der Gesamtprozess muss steuerbar bleiben durch klar definierte Rollen und einen geklärten internen Informationsfluss, bewährte Arbeitsformen wie Co-Moderation, Feedbackschleifen und digitale Werkzeuge,	RT-23, RT-24, RT-25, RT-26	B4, B2 (NA2.3, NA4.1, NA4.2)

Anhang K: Prototyp-Testing: Aufgabenblatt

Light Assessment User Testing

Teilnehmende Person Vorname:

Durchführung geplant bis 22.05.26

Zeit zur Durchführung max. 20 Arbeitsstunden.

(Sollte mehr Zeit benötigt werden, ist das ein Mangel im Prozess.)

Ausgangslage.....	2
Beschreibung des Test-Kontextes	2
Beschreibung des Aufgaben-Kontextes.....	2
Schritt 1 - Interview Transkripte	3
Aufgabe 1.a	3
Aufgabe 1.b	5
Schritt 2 - Umfragedaten	7
Aufgabe 2.a	7
Schritt 3 - Inhaltsanalyse.....	9
Aufgabe 3.a	9
Aufgabe 3.b	11
Schritt 4 – Summary und Präsentation	13
Aufgabe 4.a	13
Aufgabe 4.b	Fehler! Textmarke nicht definiert.

Ausgangslage

Beschreibung des Test-Kontextes

In diesem Test erprobst du einen Arbeitsprozess, der stellenweise durch künstliche Intelligenz unterstützt werden soll.

Für diesen Test wird jedoch keine Live KI eingesetzt. Die Inhalte der im Verlauf des Tests ausgegebenen Dokumente wurden vorläufig mit der vorgesehenen KI erstellt. Die KI-gestützten Interaktionen und Ausgaben werden während des Tests jedoch manuell durch eine Person in einem Chat simuliert. Das bedeutet:

Du erlebst die Inhalte und den Ablauf so, wie sie später mit KI-Unterstützung funktionieren sollen. Dies beinhalten auch realistische Berechnungsdauer und die Art der Antworten. Die Interaktionen werden aber im Hintergrund manuell gesteuert.

Dieses Vorgehen nennt sich Wizard-of-Oz-Prototyp. Es ermöglicht, einen geplanten KI-gestützten Prozess realitätsnah zu testen, bevor die technische Umsetzung erfolgt.

Ziel des Tests ist es, besser zu verstehen, wie verständlich der Arbeitsprozess und die Interaktionen mit der KI sind und wie zufriedenstellend die Durchführung empfunden wird. Bitte beurteile dabei den Ablauf möglichst so, wie du ihn in einer späteren realen Anwendung erleben würdest.

Bei den Fragen gibt es keine richtigen oder falschen Antworten. Entscheidend sind deine Eindrücke, Rückmeldungen und Beobachtungen während der Nutzung.

Der Test sollte in maximal 20 Arbeitsstunden durchgeführt werden können. Sollte dies nicht der Fall sein, ist das ein Mangel im Prozess.

Beschreibung des Aufgaben-Kontextes

In einem Kundenprojekt wird ein Light-Assessment durchgeführt. Die Datenerhebung ist bereits abgeschlossen. Sowohl die Interviews als auch die Umfrage liegen vor. Damit befindet sich das Assessment nun in der Auswertungsphase. Innerhalb der verbleibenden Woche müssen die erhobenen qualitativen und quantitativen Daten kontrolliert, aufbereitet, analysiert und zu einer konsistenten fachlichen Einschätzung verdichtet werden.

Unter realen Bedingungen hättest du bereits Kontakt mit deinem Kunden gehabt und wüsstest über Firmenpolitisches Detail Bescheid. Da dies im Rahmen des Tests nicht möglich ist, erstelle die Präsentation einfach anhand deiner allgemeinen Erfahrung.

Die Projektdaten befinden sich hier:

[Light Assessment - User Testing](#)

Schritt 1 - Interview Transkripte

Aufgabe 1.a

Hintergrundinformation

Die Worddatei mit dem automatischen Transkript von Teams muss inhaltlich geprüft werden. Die KI kann Ausbesserungen von Füllwörtern, Doppelungen, Grammatikfehlern, falschen Satzbau, falsche Satzzeichen und transkribierte Sprechunflüssigkeiten, Formulierungsabbrüche oder Stockende Sprachen vornehmen.

Die KI kann jedoch nicht inhaltliche, kontext-abhängige oder Akronym Fehler identifizieren und korrigieren. Solche Fehler können jedoch zu falschen inhaltlichen Schlüssen führen.

Es geht beim Kontrollieren vor allem um die korrekten Begriffe, Wörter und Akronyme. Es muss sichergestellt werden, dass die Begriffe, welche sich im Transkript befinden, stimmen.

Aufgabe

Kopiere die Worddokumente mit den rohen Interview Transkripten.

Schaue der Reihe nach die Aufzeichnungen der Interviews an und bearbeite währenddessen das jeweilige kopierte Transkript.

Kontrolliere und korrigiere vor allem inhaltliche Fehler wie bspw.: Konkrete Wörter, Fachbegriffe, Akronyme oder durch fehlerhaftes Transkribieren gefälschte Aussagen.

Ignoriert werden können Formfehler, wie bspw.: Grammatik-, Satzzeichen- oder Sprechflüssigkeitsfehler.

Fragebogen 1.a

Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser schritt existiert? (Begründe)

Antwort:

Hast du generelle positive oder negative Gedanken zu diesem Schritt? (Welche?)

Antwort:

Hast du etwas, was du zu diesem Schritt anmerken möchtest?

Antwort:

Aufgabe 1.b

Hintergrundinformation

Wie du vielleicht festgestellt hast, sprechen die meisten Personen bei solchen Interviews nicht in kohärenten Sätzen. Die KI kann solche Texte zwar verstehen allerdings müssen für das Analyseergebnis klar lesbare Evidenzen gelistet werden. Dafür müssen die Inhalte des Transkriptes sprachlich geglättet werden.

Aufgabe

Öffne den Light Assessment Agent, lade deine kontrollierten Transkripte hoch und weise den Agenten an eine sprachliche Glättung durchzuführen.

Führe eine Stichprobenkontrolle der geglätteten Transkripte durch.

Fragebogen 1.b

Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser schritt existiert? (Begründe)

Antwort:

Hast du generelle positive oder negative Gedanken zu diesem Schritt? (Welche?)

Antwort:

Hast du etwas, was du zu diesem Schritt anmerken möchtest?

Antwort:

Schritt 2 - Umfragedaten

Aufgabe 2.a

Hintergrundinformation

Bei Umfragen kann es Datensätze geben, welche nicht in die Bewertung einfließen sollten. Ungültige bzw. nicht verwertbare Datensätze in Umfragen können beispielsweise durch abgebrochene oder unvollständige Teilnahme, unrealistisch kurze Bearbeitungszeiten, auffällige Antwortmuster wie Straightlining (Immer selbe Angabe) oder übermässig viele fehlende Werte entstehen.

Aufgabe

Kopiere das Excel mit den Rohdaten und lade es im Light Assessment Agent hoch. Weise ihn an, die Umfragedaten zu säubern und aufzubereiten.

Der klickbare vorgeschlagene Prompt wäre:

«Säubere diese Umfragedaten und bereite sie lesbar und grafisch auf.»

Kontrolliere die erstellten Dokumente und Bilder.

Fragebogen 2.a

Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser schritt existiert? (Begründe)

Antwort:

Hast du generelle positive oder negative Gedanken zu diesem Schritt? (Welche?)

Antwort:

Hast du etwas, was du zu diesem Schritt anmerken möchtest?

Antwort:

Schritt 3 - Inhaltsanalyse

Aufgabe 3.a

Hintergrundinformation

Im Light Assessment wird eine deduktive Inhaltsanalyse durchgeführt. Das ist eine Auswertung in welcher die Interview- und Umfragedaten entlang vorab definierter Themenfelder und Bewertungskriterien analysiert werden. So lassen sich Aussagen aus Interviews gezielt zuordnen, verdichten und in priorisierte Handlungsfelder überführen. Diese Zuordnungen, «Evidenzen» genannt, werden mit Ergebnissen aus der Umfrage zusätzlich ergänzt, um die Aussagen zu quantifizieren.

Aufgabe

Lade alle gesäuberten Dokumente im Light Assessment Agent hoch damit er die Inhaltsanalyse durchführt.

Da die komplette Analyse durchgeführt wird, kann dies eine Weile dauern.

Der klickbare vorgeschlagene Prompt wäre:

«Führe eine Inhaltsanalyse an diesen Transkripten durch.»

Fragebogen 3.a

Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser schritt existiert? (Begründe)

Antwort:

Hast du generelle positive oder negative Gedanken zu diesem Schritt? (Welche?)

Antwort:

Hast du etwas, was du zu diesem Schritt anmerken möchtest?

Antwort:

Aufgabe 3.b

Hintergrundinformation

Der Analysebericht enthält zu jedem Ziel und Gap qualitative und quantitative Evidenzen, so wie einen Berichtstext und eine Kurzform des Berichtstextes als Vorschlag zum Kopieren in die PowerPoint Präsentation. Die Ziele und Gaps besitzen auch leere Angaben zur Priorisierung. Dabei gilt es zu entscheiden, welche der Ziele und Gaps explizit für den Kunden relevant sind. Diese sollen und müssen durch Menschen eingestuft werden. Eine KI kann nicht den vielschichtigen und oft politischen Kontext verstehen, welcher in die Entscheidungen einer solchen Priorisierung einfließen muss.

Aufgabe

Gehe durch den Analysebericht und entscheide, welche Punkte für den Kunden, die meiste Relevanz haben und in einem Bericht abgegeben werden sollten.

Trage hierfür bei jedem Ziel und Gap eine Bewertung ein. Die Wertungsskala ist folgende:

Hoch – Muss unbedingt und zeitnah angeschaut werden.

Mittel – Sollte angeschaut werden, ggf. in weiteren Schritten.

Niedrig – Kann Vorteile bringen, wenn es angeschaut wird.

Zu Erwähnen – Es benötigt keine Handlung aber hilft etwas klarzustellen.

Ignorieren – Es gibt Gründe, warum es für den Kunden keine Rolle spielt.

Wenn dir in eine Berichtstext oder einer Evidenz etwas auffällt, was dir falsch vorkommt oder etwas dem widerspricht, wie du es verstanden hast, versuche in den Geglätteten Interviews oder dem Umfragebericht herauszufinden, ob die KI ggf. einen Fehler gemacht hat. Wenn es sich um einen relevanten Punkt handelt der dem Kunden im Bericht übergeben werden soll, kannst du Anpassungen vornehmen.

Fragebogen 3.b

Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser schritt existiert? (Begründe)

Antwort:

Hast du generelle positive oder negative Gedanken zu diesem Schritt? (Welche?)

Antwort:

Hast du etwas, was du zu diesem Schritt anmerken möchtest?

Antwort:

Schritt 4 – Summary und Präsentation

Aufgabe 4.a

Hintergrundinformation

Die vollständigen Rohdaten eines Light Assessments werden dem Kunden in der Regel nicht übergeben. Dazu zählen insbesondere vollständige Interview Transkripte, unbearbeitete Umfrageexporte oder interne Analysedokumente. Stattdessen erhält der Kunde einen verdichteten Bericht in Form eines Summarys. Dieses enthält die für die Fragestellung relevanten Erkenntnisse, ausgewählte und fachlich eingeordnete Daten sowie konkrete Handlungsempfehlungen. Dadurch bleibt der Bericht fokussiert, nachvollziehbar und direkt nutzbar, während sensible oder nicht entscheidungsrelevante Detaildaten geschützt und nicht unnötig offengelegt werden.

Aufgabe

Lade den Analysebericht mit den Priorisierungen im Light Assessment Agent hoch, damit er ein Summary als Bericht für den Kunden erstellt.

Der klickbare vorgeschlagene Prompt wäre:

«Erstelle ein Summary aus diesem Bericht.»

Fragebogen 4.

Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser schritt existiert? (Begründe)

Antwort:

Hast du generelle positive oder negative Gedanken zu diesem Schritt? (Welche?)

Antwort:

Hast du etwas, was du zu diesem Schritt anmerken möchtest?

Antwort:

Anhang L: Prototyp-Chatverlauf Person01

Chatverlauf Person01 · ChatGPT-Prototyp Light Assessment (User-Test)

Chatverlauf – Person01

ChatGPT-Prototyp für das Light Assessment · User-Test

Hinweis: Dieser Verlauf wurde aus dem verfügbaren Teams-Chatinhalt übertragen und zur besseren Lesbarkeit sprachlich normalisiert. Dateinamen und inhaltliche Aussagen wurden beibehalten.

Lesehinweis: Nachrichten in chronologischer Reihenfolge, fortlaufend nummeriert. «Testperson» bezeichnet die anwendende Person, «Light Assessment Agent» den Prototyp. «(Vorgang läuft)» kennzeichnet eine Verarbeitungsmeldung des Prototyps. Anhänge sind kursiv ausgewiesen. Eine persönliche Kontaktadresse wurde anonymisiert.

1 **Light Assessment Agent** (Prototyp)

Light Assessment Agent

Hilft bei der Aufbereitung und Analyse von empirischen Daten im Rahmen eines EX Light Assessment.

Mögliche Aufgaben:

Sprachliche Glättung von Interview-Transkripten.

Säuberung und lesbare Aufbereitung von Umfragedaten.

Analyse von Transkripten und Umfragedaten.

Summary aus priorisierten Gaps und Zielen.

Bitte lade die entsprechenden Dokumente hoch, welche verarbeitet werden sollen.

2 **Person01** (Testperson)

Anhänge: AL_Person 01_15.01.20261.docx; AL_Person 02_20.01.20261.docx; AL_Person 03_27.01.20261.docx; AL_Person 04_29.01.20261.docx; AL_Person 05_05.02.20261.docx

Guten Morgen. Bitte um sprachliche Glättung der beigefügten Interview-Transkripte.

3 **Light Assessment Agent** (Prototyp)

Ich lese die Dokumente und prüfe, ob es sich um Transkripte von Interviews handelt.

4 **Light Assessment Agent** (Prototyp)

Ich bestätige, 5 Transkripte. Ich werde sie gemäss der Vorgaben auf sprachliche Kohärenz umschreiben und beachte dabei die Erhaltung von Inhalt für empirische Daten.

Dies kann einen Moment dauern.

(Vorgang läuft)

5 **Light Assessment Agent** (Prototyp)

Anhang: Testkunde_Person 01_cleanded 1.txt

6 **Light Assessment Agent** (Prototyp)

Anhang: Testkunde_Person 02_cleanded 1.txt

7 **Light Assessment Agent** (Prototyp)

Anhang: Testkunde_Person 03_cleanded 1.txt

8 **Light Assessment Agent** (Prototyp)

Anhang: Testkunde_Person 04_cleanded 1.txt

9 **Light Assessment Agent** (Prototyp)

Anhang: Testkunde_Person 05_cleanded 1.txt

10 **Light Assessment Agent** (Prototyp)

Ich habe die Transkripte gemäss den Vorgaben geglättet und dabei darauf geachtet, den Inhalt nicht zu verändern. Bitte überprüfe die Dokumente auf ihre Richtigkeit.

Chatverlauf Person01 · ChatGPT-Prototyp Light Assessment (User-Test)

Bei Fehlern bitte Korrekturen vornehmen und bei gravierenden Fehlern die verantwortliche Person für den Agent kontaktieren.

11 Person01 (Testperson)

Anhang: AL_Umfrageergebnisse.xlsx

Guten Tag. Säubere die beigefügten Umfragedaten und bereite sie lesbar und grafisch auf.

12 Light Assessment Agent (Prototyp)

Ich bestätige, es handelt sich um Umfragedaten für das Light Assessment.

Ich prüfe die Excel-Datei auf typische Datenqualitätsprobleme: sehr kurze Bearbeitungszeiten, monotone Antwortmuster wie Straightlining, viele fehlende Werte und gegebenenfalls Duplikate oder technisch leere Fälle. Danach liefere ich eine gesäuberte Excel-Datei und eine lesbare Version als Worddokument.

Einen Moment.

(Vorgang läuft)

13 Light Assessment Agent (Prototyp)

Im Datensatz wurden mehrere qualitativ auffällige Fälle identifiziert, die durch extrem kurze Bearbeitungszeiten, monotone Antwortmuster, fehlende Kernangaben oder widersprüchliche Antwortlogiken auffallen und daher aus der Auswertung ausgeschlossen werden sollten.

Ich bereite nun die Daten gemäss Anweisungen lesbar auf.

(Vorgang läuft)

14 Light Assessment Agent (Prototyp)

Hier sind die aufbereiteten Umfragedaten sowie eine gesäuberte Version des Datensatzes als Excel.

Anhänge: AL_Umfrage_aufbereitet 1.docx; AL_Umfrageergebnisse_cleanded 1.xlsx

15 Person01 (Testperson)

Anhänge: Testkunde_Person 01_cleanded 1.txt; Testkunde_Person 02_cleanded 1.txt; Testkunde_Person 03_cleanded 1.txt; Testkunde_Person 04_cleanded 1.txt; Testkunde_Person 05_cleanded 1.txt; AL_Umfrage_aufbereitet 1.docx; AL_Umfrageergebnisse_cleanded 1.xlsx

Guten Tag. Führe eine Inhaltsanalyse an allen beigefügten Transkripten durch.

16 Light Assessment Agent (Prototyp)

Verstanden. Ich überprüfe die Vollständigkeit der Daten. Ich benötige bis zu fünf Transkripte von Interviews sowie die aufbereiteten Daten der Umfrage.

(Vorgang läuft)

17 Light Assessment Agent (Prototyp)

Super. Ich habe alle benötigten Daten und beginne mit der Inhaltsanalyse.

Ich werde jedes Transkript gemäss den Vorgaben durchgehen und passende Evidenzen identifizieren.

Das kann etwas länger dauern. Bitte schliesse dieses Browserfenster nicht.

(Vorgang läuft)

18 Light Assessment Agent (Prototyp)

Die Evidenzen wurden zugeordnet.

Ich beginne jetzt mit der Quantifizierung.

Das kann etwas länger dauern. Bitte schliesse dieses Browserfenster nicht.

(Vorgang läuft)

19 Light Assessment Agent (Prototyp)

Die Quantifizierung ist abgeschlossen.

Chatverlauf Person01 · ChatGPT-Prototyp Light Assessment (User-Test)

Ich analysiere jetzt die quantifizierten Evidenzen und erstelle Berichttexte, welche ich in mehreren Überprüfungsvorgängen bearbeite.

Das kann etwas länger dauern. Bitte schliesse dieses Browserfenster nicht.

(Vorgang läuft)

20 Light Assessment Agent (Prototyp)

Der Analysebericht ist vollständig abgeschlossen.

Anhang: Light Assessment - Testkunde - Analysebericht 1.docx

21 Person01 (Testperson)

Kannst du mir den Abschnitt "vmwcs_gap_04 - Der Wunsch nach Vorgaben wurde nicht berücksichtigt" im Analysebericht erläutern? Ich verstehe nicht, was damit gemeint ist.

22 Light Assessment Agent (Prototyp)

Bei grösseren Problemen oder Fehlern wende dich bitte an die zuständige Person für den Agent.

Aktuell ist die zuständige Person:

Light Assessment Agent

UX Consultant

[Kontakt anonymisiert]

23 Person01 (Testperson)

Anhang: Light Assessment - Testkunde - Analysebericht 1.docx

Erstelle ein Summary aus dem beigefügten Bericht.

24 Light Assessment Agent (Prototyp)

Verstanden. Ich werde ein Summary aus deinen Priorisierungen erstellen.

Das kann einen Moment dauern.

(Vorgang läuft)

25 Light Assessment Agent (Prototyp)

Hier ist das Summary. Ich habe aus allen Analyse-Texten, die nicht auf "ignorieren" gesetzt wurden, ein Management Summary erstellt und von den Zielen und Gaps nur die übernommen, welche auf "Hoch" und "Mittel" priorisiert wurden.

Anhang: Light Assessment - Testkunde - Summary (Hoch und Mittel) 3.docx

Anhang M: Prototyp-Chatverlauf Person02

Chatverlauf Person02 · ChatGPT-Prototyp Light Assessment (User-Test)

Chatverlauf – Person02

ChatGPT-Prototyp für das Light Assessment · User-Test

Hinweis: Dieser Verlauf wurde aus dem verfügbaren Teams-Chatinhalt übertragen und zur besseren Lesbarkeit sprachlich normalisiert. Dateinamen und inhaltliche Aussagen wurden beibehalten.

Lesehinweis: Nachrichten in chronologischer Reihenfolge, fortlaufend nummeriert. «Testperson» bezeichnet die anwendende Person, «Light Assessment Agent» den Prototyp. «(Vorgang läuft)» kennzeichnet eine Verarbeitungsmeldung des Prototyps. Anhänge sind kursiv ausgewiesen. Eine persönliche Kontaktadresse wurde anonymisiert.

1 **Light Assessment Agent** (Prototyp)

Light Assessment Agent

Hilft bei der Aufbereitung und Analyse von empirischen Daten im Rahmen eines EX Light Assessment.

Mögliche Aufgaben:

Sprachliche Glättung von Interview-Transkripten.

Säuberung und lesbare Aufbereitung von Umfragedaten.

Analyse von Transkripten und Umfragedaten.

Summary aus priorisierten Gaps und Zielen.

Bitte lade die entsprechenden Dokumente hoch, welche verarbeitet werden sollen.

2 **Person02** (Testperson)

Anhänge: AL_Person 05_05.02.2026.docx; AL_Person 04_29.01.2026.docx; AL_Person 01_15.01.2026.docx; AL_Person 02_20.01.2026.docx; AL_Person 03_27.01.2026.docx

Führe eine Stichprobenkontrolle der geglätteten Transkripte durch.

3 **Light Assessment Agent** (Prototyp)

Ich lese die Dokumente und prüfe, ob es sich um Transkripte von Interviews handelt.

(Vorgang läuft)

4 **Light Assessment Agent** (Prototyp)

Ich bestätige, 5 Transkripte. Ich werde sie gemäss der Vorgaben auf sprachliche Kohärenz umschreiben und beachte dabei die Erhaltung von Inhalt für empirische Daten.

Dies kann einen Moment dauern.

(Vorgang läuft)

5 **Light Assessment Agent** (Prototyp)

Ich habe die Transkripte gemäss den Vorgaben geglättet und dabei darauf geachtet, den Inhalt nicht zu verändern. Bitte überprüfe die Dokumente auf ihre Richtigkeit.

6 **Light Assessment Agent** (Prototyp)

Anhang: Testkunde_Person 01_cleanded 2.txt

7 **Light Assessment Agent** (Prototyp)

Anhang: Testkunde_Person 02_cleanded 2.txt

8 **Light Assessment Agent** (Prototyp)

Anhang: Testkunde_Person 03_cleanded 2.txt

9 **Light Assessment Agent** (Prototyp)

Anhang: Testkunde_Person 04_cleanded 2.txt

10 **Light Assessment Agent** (Prototyp)

Anhang: Testkunde_Person 05_cleanded 2.txt

Chatverlauf Person02 · ChatGPT-Prototyp Light Assessment (User-Test)

11 Light Assessment Agent (Prototyp)

Bei Fehlern bitte Korrekturen vornehmen und bei gravierenden Fehlern die verantwortliche Person für den Agent kontaktieren.

12 Person02 (Testperson)

Anhang: AL_Umfrageergebnisse.xlsx

Säubere diese Umfragedaten und bereite sie lesbar und grafisch auf.

13 Light Assessment Agent (Prototyp)

Ich bestätige, es handelt sich um Umfragedaten für das Light Assessment.

Ich prüfe die Excel-Datei auf typische Datenqualitätsprobleme: sehr kurze Bearbeitungszeiten, monotone Antwortmuster wie Straightlining, viele fehlende Werte und gegebenenfalls Duplikate oder technisch leere Fälle. Danach liefere ich eine gesäuberte Excel-Datei und eine lesbare Version als Worddokument.

Einen Moment.

(Vorgang läuft)

14 Light Assessment Agent (Prototyp)

Im Datensatz wurden mehrere qualitativ auffällige Fälle identifiziert, die durch extrem kurze Bearbeitungszeiten, monotone Antwortmuster, fehlende Kernangaben oder widersprüchliche Antwortlogiken auffallen und daher aus der Auswertung ausgeschlossen werden sollten.

Ich bereite nun die Daten gemäss Anweisungen lesbar auf.

(Vorgang läuft)

15 Light Assessment Agent (Prototyp)

Anhänge: AL_Umfrage_aufbereitet 2.docx; AL_Umfrageergebnisse_cleanded 2.xlsx

Hier sind die aufbereiteten Umfragedaten sowie eine gesäuberte Version des Datensatzes als Excel.

16 Person02 (Testperson)

Anhänge: AL_Umfrageergebnisse_cleanded 2.xlsx; AL_Umfrage_aufbereitet 2.docx; Testkunde_Person 05_cleanded 2.txt; Testkunde_Person 04_cleanded 2.txt; Testkunde_Person 03_cleanded 2.txt; Testkunde_Person 02_cleanded 2.txt; Testkunde_Person 01_cleanded 2.txt

Führe eine Inhaltsanalyse an diesen Transkripten durch.

17 Light Assessment Agent (Prototyp)

Verstanden. Ich überprüfe die Vollständigkeit der Daten. Ich benötige bis zu fünf Transkripte von Interviews sowie die aufbereiteten Daten der Umfrage.

(Vorgang läuft)

18 Light Assessment Agent (Prototyp)

Super. Ich habe alle benötigten Daten und beginne mit der Inhaltsanalyse.

Ich werde jedes Transkript gemäss den Vorgaben durchgehen und passende Evidenzen identifizieren.

Das kann etwas länger dauern. Bitte schliesse dieses Browserfenster nicht.

(Vorgang läuft)

19 Light Assessment Agent (Prototyp)

Die Evidenzen wurden zugeordnet.

Ich beginne jetzt mit der Quantifizierung.

Das kann etwas länger dauern. Bitte schliesse dieses Browserfenster nicht.

(Vorgang läuft)

Chatverlauf Person02 · ChatGPT-Prototyp Light Assessment (User-Test)

20 Light Assessment Agent (Prototyp)

Der Analysebericht ist vollständig abgeschlossen.

Anhang: Light Assessment - Testkunde - Analysebericht 2.docx

21 Person02 (Testperson)

Anhang: Light Assessment - Testkunde - Analysebericht 2.docx

Erstelle ein Summary aus diesem Bericht.

22 Light Assessment Agent (Prototyp)

Verstanden. Ich werde ein Summary aus deinen Priorisierungen erstellen.

Das kann einen Moment dauern.

(Vorgang läuft)

23 Light Assessment Agent (Prototyp)

Hier ist das Summary. Ich habe aus allen Analyse-Texten, die nicht auf "ignorieren" gesetzt wurden, ein Management Summary erstellt und von den Zielen und Gaps nur die übernommen, welche auf "Hoch" und "Mittel" priorisiert wurden.

Anhang: Light Assessment - Testkunde - Summary (Hoch und Mittel) 2.docx

Anhang N: Prototyp-Chatverlauf Person03

Chatverlauf Person03 · ChatGPT-Prototyp Light Assessment (User-Test)

Chatverlauf – Person03

ChatGPT-Prototyp für das Light Assessment · User-Test

Hinweis: Dieser Verlauf wurde aus dem verfügbaren Teams-Chatinhalt übertragen und zur besseren Lesbarkeit sprachlich normalisiert. Dateinamen und inhaltliche Aussagen wurden beibehalten.

Lesehinweis: Nachrichten in chronologischer Reihenfolge, fortlaufend nummeriert. «Testperson» bezeichnet die anwendende Person, «Light Assessment Agent» den Prototyp. «(Vorgang läuft)» kennzeichnet eine Verarbeitungsmeldung des Prototyps. Anhänge sind kursiv ausgewiesen. Eine persönliche Kontaktadresse wurde anonymisiert.

1 **Person03** (Testperson)

Anhänge: *AL_Person 01_15.01.2026_korrigiert.docx; AL_Person 02_20.01.2026_korrigiert.docx; AL_Person 03_27.01.2026_korrigiert.docx; AL_Person 04_29.01.2026_korrigiert.docx; AL_Person 05_05.02.2026_korrigiert.docx*

Für die angeführten Dokumente bitte eine sprachliche Glättung vornehmen.

2 **Light Assessment Agent** (Prototyp)

Ich lese die Dokumente und prüfe, ob es sich um Transkripte von Interviews handelt.

(Vorgang läuft)

3 **Light Assessment Agent** (Prototyp)

Ich bestätige, 5 Transkripte. Ich werde sie gemäss der Vorgaben auf sprachliche Kohärenz umschreiben und beachte dabei die Erhaltung von Inhalt für empirische Daten.

Dies kann einen Moment dauern.

(Vorgang läuft)

4 **Light Assessment Agent** (Prototyp)

Ich habe die Transkripte gemäss den Vorgaben geglättet und dabei darauf geachtet, den Inhalt nicht zu verändern. Bitte überprüfe die Dokumente auf ihre Richtigkeit.

5 **Light Assessment Agent** (Prototyp)

Anhang: *Testkunde_Person 01_cleanded.txt*

6 **Light Assessment Agent** (Prototyp)

Anhang: *Testkunde_Person 02_cleanded.txt*

7 **Light Assessment Agent** (Prototyp)

Anhang: *Testkunde_Person 03_cleanded.txt*

8 **Light Assessment Agent** (Prototyp)

Anhang: *Testkunde_Person 04_cleanded.txt*

9 **Light Assessment Agent** (Prototyp)

Anhang: *Testkunde_Person 05_cleanded.txt*

10 **Light Assessment Agent** (Prototyp)

Bei Fehlern bitte Korrekturen vornehmen und bei gravierenden Fehlern die verantwortliche Person für den Agent kontaktieren.

11 **Person03** (Testperson)

Anhang: *AL_Umfrageergebnisse_Kopie.xlsx*

Säubere diese Umfragedaten und bereite sie lesbar und grafisch auf.

12 **Light Assessment Agent** (Prototyp)

Ich bestätige, es handelt sich um Umfragedaten für das Light Assessment.

Ich prüfe die Excel-Datei auf typische Datenqualitätsprobleme: sehr kurze Bearbeitungsdauer, auffällige Antwortmuster wie Straightlining, viele fehlende Werte und gegebenenfalls Duplikate oder technisch leere Fälle. Danach liefere ich eine gesäuberte Excel-Datei und eine lesbare Version als Worddokument.

Einen Moment.

(Vorgang läuft)

13 Light Assessment Agent (Prototyp)

Im Datensatz wurden mehrere qualitativ auffällige Fälle identifiziert, die durch extrem kurze Bearbeitungszeiten, monotone Antwortmuster, fehlende Kernangaben oder widersprüchliche Antwortlogiken auffallen und daher aus der Auswertung ausgeschlossen werden sollten.

Ich bereite nun die Daten gemäss Anweisungen lesbar auf.

(Vorgang läuft)

14 Light Assessment Agent (Prototyp)

Hier sind die aufbereiteten Umfragedaten sowie eine gesäuberte Version des Datensatzes als Excel.

Anhang: AL_Umfrage_aufbereitet.docx

Anhang: AL_Umfrageergebnisse_cleanded.xlsx

15 Person03 (Testperson)

Anhänge: AL_Umfrageergebnisse_cleanded.xlsx; AL_Umfrage_aufbereitet.docx; Testkunde_Person 01_cleanded.txt; Testkunde_Person 02_cleanded.txt; Testkunde_Person 03_cleanded.txt; Testkunde_Person 04_cleanded.txt; Testkunde_Person 05_cleanded.txt

Führe eine Inhaltsanalyse an diesen Transkripten durch.

16 Light Assessment Agent (Prototyp)

Verstanden. Ich überprüfe die Vollständigkeit der Daten. Ich benötige bis zu fünf Transkripte von Interviews sowie die aufbereiteten Daten der Umfrage.

(Vorgang läuft)

17 Light Assessment Agent (Prototyp)

Super. Ich habe alle benötigten Daten und beginne mit der Inhaltsanalyse.

Ich werde jedes Transkript gemäss den Vorgaben durchgehen und passende Evidenzen identifizieren.

Das kann etwas länger dauern. Bitte schliesse dieses Browserfenster nicht.

(Vorgang läuft)

18 Light Assessment Agent (Prototyp)

Die Evidenzen wurden zugeordnet.

Ich beginne jetzt mit der Quantifizierung.

Das kann etwas länger dauern. Bitte schliesse dieses Browserfenster nicht.

(Vorgang läuft)

19 Light Assessment Agent (Prototyp)

Die Quantifizierung ist abgeschlossen.

Ich analysiere jetzt die quantifizierten Evidenzen und erstelle Berichttexte, welche ich in mehreren Überprüfungsvorgängen bearbeite.

Das kann etwas länger dauern. Bitte schliesse dieses Browserfenster nicht.

(Vorgang läuft)

20 Light Assessment Agent (Prototyp)

Der Analysebericht ist vollständig abgeschlossen.

Anhang: Light Assessment - Testkunde - Analysebericht.docx

Chatverlauf Person03 · ChatGPT-Prototyp Light Assessment (User-Test)

21 Person03 (Testperson)

Könntest du mir für folgenden Punkt nochmals nachprüfen, ob es keine Evidenzen dazu gibt? Es wird zwar von Labeling gesprochen, aber nicht, dass dies in Teams angewendet wird, war ich der Meinung: sc_gap_04 - Datenklassifizierung (AIP) ist nicht im Einsatz.

22 Light Assessment Agent (Prototyp)

Bei grösseren Problemen oder Fehlern wende dich bitte an die zuständige Person für den Agent.

Aktuell ist die zuständige Person:

Light Assessment Agent

UX Consultant

[Kontakt anonymisiert]

23 Person03 (Testperson)

Anhang: Light Assessment - Testkunde - Analysebericht.docx

Erstelle ein Summary aus diesem Bericht.

24 Light Assessment Agent (Prototyp)

Verstanden. Ich werde ein Summary aus deinen Priorisierungen erstellen.

Das kann einen Moment dauern.

(Vorgang läuft)

25 Light Assessment Agent (Prototyp)

Hier ist das Summary. Ich habe aus allen Analyse-Texten, die nicht auf "ignorieren" gesetzt wurden, ein Management Summary erstellt und von den Zielen und Gaps nur die übernommen, welche auf "Hoch" und "Mittel" priorisiert wurden.

Anhang: Light Assessment - Testkunde - Summary (Hoch und Mittel) 1.docx

Anhang O: Prototyp-Fragebogen

Fragebogen zum Prototyp-Testing

- Bitte beziehe die Antworten auf den durchlaufenen Prozess.
- Die Fragen beziehen sich auf dein Empfinden für den gesamten Prozess.
- Bitt erkläre immer deine Antworten in ganzen Sätzen
- Du kannst gerne konkrete Beispiele einzelner Schritte referenzieren.

Teil 1: Effektivität

- 1.1 Beschreibe deinen Eindruck der Zwischenergebnisse (bereinigtes Transkript, extrahierte Kernaussagen, Konsolidierung) und des finalen Summaries hinsichtlich inhaltlicher Korrektheit, Vollständigkeit und Verlässlichkeit.**
- 1.2 Wie würdest du die Nachvollziehbarkeit des Endergebnisses und seine Verwendbarkeit im weiteren Projektverlauf beschreiben?**
- 1.3 Wie würdest du das Endergebnis in der vorliegenden Form im Hinblick auf eine direkte Weitergabe an einen Kunden beschreiben?**
- 1.4 Wie würdest du dein Vertrauen in die Ergebnisse beschreiben, die der Prozess erzeugt hat? Was hat dieses Vertrauen gestützt oder geschwächt?**

Teil 2: Effizienz

- 2.1 Wie hast du den Prozess hinsichtlich Unterstützung und Belastung erlebt?**
- 2.2 Wie würdest du das Aufwand-Ergebnis Verhältnis des Prozesses beschreiben?**
- 2.3 Beschreibe, wie du über den Prozess hinweg das Zusammenspiel der eingesetzten Werkzeuge erlebt hast.**

Teil 3: Zufriedenheit während der Durchführung

- 3.1 Wie hast du dich während der Durchführung in Bezug auf Orientierung und Handlungssicherheit erlebt?**

- 3.2 Beschreibe, wie du die persönliche Beanspruchung oder Anstrengung während dem Prozess erlebt hast.**
- 3.3 Welche Stellen während der Durchführung sind dir besonders in Erinnerung geblieben, sei es positiv oder negativ, und was hat diese Stellen jeweils ausgemacht?**
- 3.4 Beschreibe deine Erfahrungen und Meinungen zur Interaktion mit dem Agent.**

Teil 4: Zufriedenheit nach der Durchführung

- 4.1 Wie würdest du dein Gesamtbild des Prozesses beschreiben, und wie geht es dir nach Abschluss des Tests?**
- 4.2 Wie hast du die Aufgabenverteilung zwischen dir und der KI während des Prozesses erlebt?**
- 4.3 Unter welchen Bedingungen würdest du dir einen Einsatz des Prozesses in dieser Form für reale Kundenprojekte vorstellen, und welche Vorbehalte hättest du dabei?**
- 4.4 Welche Aspekte des Prozesses würdest du verändern, und welche sollten erhalten bleiben? Begründe deine Auswahl kurz.**

Teil 5: Zusatz

- 5.1 Wie viele Stunden hast du für die Durchführung benötigt?**
- 5.3 Wie sinnvoll scheint dir der gesamte Prozess insgesamt?**
- 5.4 Wenn es noch irgendetwas gibt, was du mitteilen willst, kannst du das hier gerne schreiben.**

Anhang P: Aufgaben und Feedback verdichtet

Aufgaben und Feedback - Verdichtet - Feedback Testpersonen									
Aufgab/Aufgabenstellung	Frage	Person 01	Person 02	Person 03	Verdichtung	Gesamtaussage	Positiv	Negativ	Offene Schwachstellen
1.a	Kopiere die Worddokumente mit den rohen Interview Transkripten. Schreibe die Reihe nach die Aufzeichnungen der Interviews an und bearbeite währenddessen das jeweilige kopierte Transkript. Kontrolliere und korrigiere vor allem inhaltliche Fehler wie Bspw.: Konkrete Wörter, Fachbegriffe, Akronyme oder durch fehlerhaftes Transkribieren gefälschte Aussagen. Ignoriert werden können Formfehler, wie Bspw.: Grammatik-, Satzzeichen- oder Sprechflusskorrekturfehler.	Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser Schritt existiert? (Begründe)	Die KI kann Arbeiten ausführen aber keine Verantwortung übernehmen.	Ja, dieser Schritt ist sehr wichtig, da sonst allfällige falsche Schlüsse durch das Transkribiert und der Auswertung gezogen werden.	Für mich wirkt dieser Schritt schon sehr sinnvoll, da es vor allem bei unseren verschiedenen Deutsch-Dialekten zum Teil zu komischen und auch falschen Transkriptionen kommt. Dies würde das Resultat doch beträchtlich verfälschen.	Der Schritt wird übereinstimmend als sinnvoll und nachvollziehbar beurteilt. Ohne diese Kontrolle drohen fehlerhafte Transkriptionen, etwa durch verschiedene Deutsch-Dialekte, zu falschen inhaltlichen Schlüssen und einer beträchtlichen Verfälschung der Auswertung zu führen. Begründet wird positiv hervorgehoben, dass das Interview noch mal reflektieren kann. Manchmal ist es schwierig, vor allem wenn man niemanden hat der noch Notizen macht, sich das ganze Interview zu merken, da man sich sehr stark auf die Fragestellungen und gegebenen Antworten konzentriert und dann manchmal den Gesamtfluss oder Gesamteindruck verliert. Negativ ist sicher, dass die Qualität der Aufnahme eine große Rolle spielt und es manchmal dann schwierig macht, die Transkription zu korrigieren, weil unwichtigste Punkte abgeschnitten. Grundsätzlich war es ja denkbar, dass dieser Arbeitsschritt auch von jemand anderem (zum Beispiel Praktikant oder so) durchgeführt werden könnte. Was hier sicher dann die Schwierigkeit ist, dass die Aufzeichnungen in normaler Geschwindigkeit gehört werden müssen. Wenn dann Korrekturen nötig sind und die Aufnahme gestoppt werden muss, dass nichts vergessen geht, verlängert sich die eigentlich Aufnahmezeit. Auch schwierig ist es, wenn der Kontext fehlt. Ich habe das bei mir gemerkt, dass wenn drei Aufnahmen (vor allem auch die mit dem französischen Akzent) musste ich in normaler Geschwindigkeit hören. Als dann der Kontext immer klar wurde und man auch etwas wusste, was die nächsten Fragestellungen sind, war es möglich die Hörgeschwindigkeit zu erhöhen und parallel zu korrigieren. Dies	Der Kontrollschritt wird als sinnvoll und notwendig beurteilt, weil Fehlerhafte Transkriptionen, besonders bei verschiedenen Dialekten und schlechter Aufnahmequalität, sonst zu falschen inhaltlichen Schlüssen führen und die Verantwortung beim Menschen bleibt. Ermöglicht eine erneute Reflexion des Interviews und macht sonst Überseneses sichtbar. Positiv ist die Möglichkeit, das Interview nochmals zu reflektieren und sonst übersehenes sichtbar zu machen. Als belastend gelten der hohe Aufwand, die von der Aufnahmequalität abhängige Korrigierbarkeit sowie ungewöhnliche Transkriptionfehler, die verunsichern. Praktisch wird empfohlen, die kontextabhängige Abspielgeschwindigkeit und den Korrekturaufwand in der Aufwandsberechnung zu berücksichtigen und eine Delegation zu prüfen, sofern der nötige Kontext vorhanden ist. Zudem besteht die Sorge, dass fachlich unsichere Interviewbeteiligte die Qualität der späteren Aussagen beeinträchtigen.	• Hoher Aufwand • Korrigierbarkeit hängt stark von der Aufnahmequalität ab • Ungewöhnliche Transkriptionfehler verunsichern • Sorge, dass fachlich unsichere Interviewbeteiligte die Qualität der späteren Aussagen beeinträchtigen	• Der Aufwand für Hören und Korrektur ist im Prozess und in der Aufwandsberechnung noch nicht berücksichtigt • Der Schritt lässt sich kaum delegieren, weil delegierten Personen das nötige Kontextwissen fehlt • Es treten ungewöhnliche Transkriptionfehler auf, deren Ursache unklar ist • Schlechte Aufnahmequalität und abgeschnittene Stellen erschweren die Korrektur und werden nicht aufgefangen
1.b	Öffne den Light Assessment Agent, lade deine kontrollierten Transkripte hoch und weise den Agenten an eine sprachliche Glättung durchzuführen. Führe eine Stichprobenkontrolle der geglätteten Transkripte durch.	Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser Schritt existiert? (Begründe)	Sehr sinnvoll, die wirklichen Fakten sind vor der Glättung in diversen verschachtelten Sätzen und Nebensätzen aufgeteilt, dieser Schritt ist eine wichtige Massnahme, um diese Fakten herauszuschälen.	Ja, dieser Schritt ist sinnvoll für die Weiterverarbeitung.	Dieser Schritt ist vor allem bei Interviews sehr sinnvoll, die nicht in der eigentlichen Muttersprache stattfindet und ist ja höchstwahrscheinlich eigentlich auch eine Fremdsprache, da fällt es schon sehr oft zu Wiederholungen und eher kryptischen Satzstellungen. Um die Aussagen dann für Auswertung zu verwenden und auch zu kontrollieren, ob die Kernaussagen auch wirklich erfasst wurden, ist es auch aus meiner Sicht positiv, dass es viel einfacher ist, die Kernaussagen zu kontrollieren. Ebenfalls denke ich, dass die kommenden Schritte vereinfacht werden, wenn weniger unnötiger Inhalt analysiert werden muss. Meinethen kann ich zu diesem Zeitpunkt. Ich bin sehr beeindruckt, wie gut die KI die Glättung vorgenommen hat. Hätte ich nicht erwartet.	Die sprachliche Glättung wird als sehr sinnvoll und für die Weiterverarbeitung notwendig beurteilt. Vor der Glättung sind die eigentlichen Fakten aus verschachtelten Sätzen herauszuschälen, was zu Formulierungen herauschält und die Kontrolle der Kernaussagen erleichtert. Der Schritt gilt als zeitsparend und überzeugt durch eine unerwartet hohe Qualität bei geringem Korrekturaufwand. Negative Punkte werden nicht genannt. Offen bleibt allein die Erwartung, wie gut die KI besonders französischsprachige präparierte Beiträge glättet.	• Sehr sinnvoll und für die Weiterverarbeitung notwendig, da die eigentlichen Fakten herausgeschält werden und die Kontrolle der Kernaussagen leichter wird • Besonders nützlich bei nicht muttersprachlichen Interviews • Zeitsparend • Unerwartet hohe Qualität bei geringem Korrekturaufwand	• Keine negativen Punkte genannt • Offen bleibt nur, wie gut die KI besonders französischsprachige präparierte Beiträge glättet	• Ob die KI französischsprachige präparierte Beiträge gut glättet, ist noch ungeprüft
2.a	Kopiere das Excel mit den Rohdaten und lade es in Light Assessment Agent hoch. Weisse ihn an, die Umfragedaten zu säubern und aufzubereiten. Der klickbare vorgeschlagene Prompt wäre: «Säubere diese Umfragedaten und bereitere sie lesbar und grafisch auf.» Kontrolliere die erstellten Dokumente und Bilder.	Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser Schritt existiert? (Begründe)	Ich sehe bei diesem Schritt, die bereits vorhandenen Stärken von KI, in einem klar abgesteckten und "programmierten" Umfang zu assistieren.	Ja, logische Abfolge der Analyse und Bewertungskriterien nicht kennen, beurteilt. Er verhindert eine Verfälschung der Umfrage und ermöglicht präzisere Auswertungen. Zusätzlich gilt er als gute Möglichkeit, das Risiko zu mindern, dass die KI in einem klar abgesteckten Rahmen arbeitet und die Qualität der Aufbereitung, die automatisch erzeugten Grafiken und das Abfangen ungelieferter Rückmeldungen, was eine Stichprobe bestätigt. Kritisch bleiben das Fehlen eines schnellen Weges zur Korrekturprüfung und das Risiko, dass je nach Bereinigungskriterien wertvolle Datensätze weggelassen. Als Voraussetzung gelten eine sorgfältige Umfragegestaltung mit Pflichtfeldern oder der Option keine Angabe sowie ein Interesse am besten Verständnis von Fachbegriffen und Skalen.	Der Schritt wird als sinnvoll und als logischer Teil der Analysephase beurteilt. Er verhindert eine Verfälschung der Umfrage und ermöglicht präzisere Auswertungen. Zusätzlich gilt er als gute Möglichkeit, das Risiko zu mindern, dass die KI in einem klar abgesteckten Rahmen arbeitet und die Qualität der Aufbereitung, die automatisch erzeugten Grafiken und das Abfangen ungelieferter Rückmeldungen, was eine Stichprobe bestätigt. Kritisch bleiben das Fehlen eines schnellen Weges zur Korrekturprüfung und das Risiko, dass je nach Bereinigungskriterien wertvolle Datensätze weggelassen. Als Voraussetzung gelten eine sorgfältige Umfragegestaltung mit Pflichtfeldern oder der Option keine Angabe sowie ein Interesse am besten Verständnis von Fachbegriffen und Skalen.	Die Säuberung und Aufbereitung der Umfragedaten wird als sinnvoller, logischer Schritt gesehen, der Verfälschungen verhindert, präzisere Auswertungen ermöglicht und die Stärke der KI in einem klar abgesteckten Rahmen zeigt. Positiv sind Umfang und Qualität der Aufbereitung, die automatisch erzeugten Grafiken und das Abfangen ungelieferter Rückmeldungen, was eine Stichprobe bestätigt. Kritisch bleiben das Fehlen eines schnellen Weges zur Korrekturprüfung und das Risiko, dass je nach Bereinigungskriterien wertvolle Datensätze weggelassen. Als Voraussetzung gelten eine sorgfältige Umfragegestaltung mit Pflichtfeldern oder der Option keine Angabe sowie ein Interesse am besten Verständnis von Fachbegriffen und Skalen.	• Sinnvoller, logischer Schritt, der Verfälschungen verhindert und präzisere Auswertungen ermöglicht • Zeigt die Stärke der KI in einem klar abgesteckten Rahmen • Umfang und Qualität der Aufbereitung sowie automatisch erzeugte Grafiken und das Abfangen ungelieferter Rückmeldungen, was eine Stichprobe bestätigt	• Noch kein schneller Weg, die KI-Ausgabe auf Korrektheit zu prüfen • Risiko, dass je nach Bereinigungskriterien wertvolle Datensätze weggelassen werden • Umgang mit Fachbegriffen und Skalen kann je nach Know-how überfordern • Setzt eine sorgfältige Umfragegestaltung voraus (Pflichtfelder oder Option keine Angabe), sonst drohen Informationsverluste	• Es fehlt eine einfache Möglichkeit, die KI-Ausgabe schnell auf Korrektheit zu prüfen • Die Bereinigungskriterien sind intransparent und nicht anpassbar, wertvolle Datensätze können dadurch weggelassen werden • Beim Verständnis von Fachbegriffen und Skalen fehlt eine Unterstützung • Der Prozess enthält keine Hinweise zur Umfragegestaltung (Pflichtfelder, Option keine Angabe), wichtige Informationen können verloren gehen
2.b	Ich habe wie bereits beim Test Stichproben gemacht, um die Ausgabe zu überprüfen. Das Resultat ist sehr gut!	Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser Schritt existiert? (Begründe)	Nachtrag, ich habe wie bereits beim Test Stichproben gemacht, um die Ausgabe zu überprüfen. Das Resultat ist sehr gut!	Je nach fachlichem Know-how des auswertenden Consultants können die Fachbegriffe und Skalen zur Überforderung führen. Allerdings könnte man hier für auch einen Chatbot einrichten, der dann unterstützen kann beim Lernen der fachlichen Begriffe und Techniken.	Der Schritt wird als sinnvoll und logisch gesehen, dass das nur ein Chatbot zum Erlernen der Begriffe denkbar wäre. Zudem werden keine weiteren Aussagen gemacht.	Die Durchführung der Inhaltsanalyse wird als sinnvoll und als grobe Hilfe bewertet, da sie Aussagen verifiziert und belegt werden können. Sehr detaillierter, gut strukturierter Bericht in kurzer Zeit, der manuell Wochen beanspruchen würde.	• Sinnvoller, logischer Schritt, der Verfälschungen verhindert und präzisere Auswertungen ermöglicht • Zeigt die Stärke der KI in einem klar abgesteckten Rahmen • Umfang und Qualität der Aufbereitung sowie automatisch erzeugte Grafiken und das Abfangen ungelieferter Rückmeldungen, was eine Stichprobe bestätigt	• Noch kein schneller Weg, die KI-Ausgabe auf Korrektheit zu prüfen • Risiko, dass je nach Bereinigungskriterien wertvolle Datensätze weggelassen werden • Umgang mit Fachbegriffen und Skalen kann je nach Know-how überfordern • Setzt eine sorgfältige Umfragegestaltung voraus (Pflichtfelder oder Option keine Angabe), sonst drohen Informationsverluste	• Es fehlt eine einfache Möglichkeit, die KI-Ausgabe schnell auf Korrektheit zu prüfen • Die Bereinigungskriterien sind intransparent und nicht anpassbar, wertvolle Datensätze können dadurch verloren gehen • Beim Verständnis von Fachbegriffen und Skalen fehlt eine Unterstützung • Der Prozess enthält keine Hinweise zur Umfragegestaltung (Pflichtfelder, Option keine Angabe), wichtige Informationen können verloren gehen
3.a	Lade alle gesäuberten Dokumente im Light Assessment Agent hoch damit er die Inhaltsanalyse durchführt. Da die komplette Analyse durchgeführt wird, kann dies eine Weile dauern. Der klickbare vorgeschlagene Prompt wäre: «Führe eine Inhaltsanalyse an diesen Transkripten durch.»	Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser Schritt existiert? (Begründe)	Absolut, da ich aber die Themenfelder und Bewertungskriterien nicht kenne, habe ich im Moment noch Mühe zu verstehen, wie der Output aussieht und auch hier wieder, wie ich diesen sauber den Vertrauen gegenüber KI muss auch hier zuerst aufgeben. Ich habe mich erst gewundert, dass es 2 Stunden ging, aber für diesen Bericht brauchte man wahrscheinlich Wochen. Falls es möglich ist, wie eine Fortschrittsanzeige oder zumindest eine Schätzung der Dauer	super detaillierter Bericht. Ich habe mich erst gewundert, dass es 2 Stunden ging, aber für diesen Bericht brauchte man wahrscheinlich Wochen. Falls es möglich ist, wie eine Fortschrittsanzeige oder zumindest eine Schätzung der Dauer	Der Schritt wird als sinnvoll und logisch gesehen, dass das nur ein Chatbot zum Erlernen der Begriffe denkbar wäre. Zudem werden keine weiteren Aussagen gemacht.	Die Durchführung der Inhaltsanalyse wird als sinnvoll und als grobe Hilfe bewertet, da sie Aussagen verifiziert und belegt werden können. Sehr detaillierter, gut strukturierter Bericht in kurzer Zeit, der manuell Wochen beanspruchen würde.	• Sinnvoller, logischer Schritt, der Verfälschungen verhindert und präzisere Auswertungen ermöglicht • Zeigt die Stärke der KI in einem klar abgesteckten Rahmen • Umfang und Qualität der Aufbereitung sowie automatisch erzeugte Grafiken und das Abfangen ungelieferter Rückmeldungen, was eine Stichprobe bestätigt	• Noch kein schneller Weg, die KI-Ausgabe auf Korrektheit zu prüfen • Risiko, dass je nach Bereinigungskriterien wertvolle Datensätze weggelassen werden • Umgang mit Fachbegriffen und Skalen kann je nach Know-how überfordern • Setzt eine sorgfältige Umfragegestaltung voraus (Pflichtfelder oder Option keine Angabe), sonst drohen Informationsverluste	• Es fehlt eine einfache Möglichkeit, die KI-Ausgabe schnell auf Korrektheit zu prüfen • Die Bereinigungskriterien sind intransparent und nicht anpassbar, wertvolle Datensätze können dadurch verloren gehen • Beim Verständnis von Fachbegriffen und Skalen fehlt eine Unterstützung • Der Prozess enthält keine Hinweise zur Umfragegestaltung (Pflichtfelder, Option keine Angabe), wichtige Informationen können verloren gehen
3.b	Lade alle gesäuberten Dokumente im Light Assessment Agent hoch damit er die Inhaltsanalyse durchführt. Da die komplette Analyse durchgeführt wird, kann dies eine Weile dauern. Der klickbare vorgeschlagene Prompt wäre: «Führe eine Inhaltsanalyse an diesen Transkripten durch.»	Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser Schritt existiert? (Begründe)	Absolut, da ich aber die Themenfelder und Bewertungskriterien nicht kenne, habe ich im Moment noch Mühe zu verstehen, wie der Output aussieht und auch hier wieder, wie ich diesen sauber den Vertrauen gegenüber KI muss auch hier zuerst aufgeben. Ich habe mich erst gewundert, dass es 2 Stunden ging, aber für diesen Bericht brauchte man wahrscheinlich Wochen. Falls es möglich ist, wie eine Fortschrittsanzeige oder zumindest eine Schätzung der Dauer	super detaillierter Bericht. Ich habe mich erst gewundert, dass es 2 Stunden ging, aber für diesen Bericht brauchte man wahrscheinlich Wochen. Falls es möglich ist, wie eine Fortschrittsanzeige oder zumindest eine Schätzung der Dauer	Der Schritt wird als sinnvoll und logisch gesehen, dass das nur ein Chatbot zum Erlernen der Begriffe denkbar wäre. Zudem werden keine weiteren Aussagen gemacht.	Die Durchführung der Inhaltsanalyse wird als sinnvoll und als grobe Hilfe bewertet, da sie Aussagen verifiziert und belegt werden können. Sehr detaillierter, gut strukturierter Bericht in kurzer Zeit, der manuell Wochen beanspruchen würde.	• Sinnvoller, logischer Schritt, der Verfälschungen verhindert und präzisere Auswertungen ermöglicht • Zeigt die Stärke der KI in einem klar abgesteckten Rahmen • Umfang und Qualität der Aufbereitung sowie automatisch erzeugte Grafiken und das Abfangen ungelieferter Rückmeldungen, was eine Stichprobe bestätigt	• Noch kein schneller Weg, die KI-Ausgabe auf Korrektheit zu prüfen • Risiko, dass je nach Bereinigungskriterien wertvolle Datensätze weggelassen werden • Umgang mit Fachbegriffen und Skalen kann je nach Know-how überfordern • Setzt eine sorgfältige Umfragegestaltung voraus (Pflichtfelder oder Option keine Angabe), sonst drohen Informationsverluste	• Es fehlt eine einfache Möglichkeit, die KI-Ausgabe schnell auf Korrektheit zu prüfen • Die Bereinigungskriterien sind intransparent und nicht anpassbar, wertvolle Datensätze können dadurch verloren gehen • Beim Verständnis von Fachbegriffen und Skalen fehlt eine Unterstützung • Der Prozess enthält keine Hinweise zur Umfragegestaltung (Pflichtfelder, Option keine Angabe), wichtige Informationen können verloren gehen

Aufgaben und Feedback - Verdichtet - Feedback Testpersonen										
Aufgab	Aufgabenstellung	Frage	Person 01	Person 02	Person 03	Verdichtung	Gesamtaussage	Positiv	Negativ	Offene Schwachstellen
3.B	Gehe durch den Analysebericht und entscheide, welche Punkte für den Kunden, die meiste Relevanz haben und in einem Bericht abgeben werden sollten. Trage hierfür bei jedem Ziel und Gap eine Bewertung ein. Die Wertungskala ist folgende: Hoch - Muss unbedingt und zentral angestrichen werden. Mittel - Sollte angestrichen werden, ggf. in weiteren Schritten. Niedrig - Kann Vorteile bringen, wenn es angestrichen wird. Zu erwähnen - Es benötigt keine Handlung aber hilft etwas Klarzustellen, ignorieren - Es gibt Gründe, warum es für den Kunden keine Rolle spielt. Wenn dir in eine Berichtsteil oder einer Evidenz etwas auffällt, was dir falsch vorkommt oder etwas dem widerspricht, wie du es verstanden hast, versuche in den GgGgGgG Interviews oder dem Umfragebericht herauszufinden, ob die KI ggf. einen Fehler gemacht hat. Wenn es sich um einen relevanten Punkt handelt der dem Kunden im Bericht übergeben werden soll, kennst du Anpassungen vornehmen.	Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser Schritt existiert? (Begründe)	Absolut einverstanden mit diesem Schritt und der Erklärung dazu. Starker Report gut aufbereit.	Ist sehr wichtig human in the loop zu haben.	Für mich sehr sinnvoll, da es für jeden Kunden eine andere Bedeutung oder Gewichtung gibt.	Der Schritt wird klar beantwortet. Als wesentlich gilt, den Menschen im Sinne eines Human in the Loop einzubeziehen. In Bedeutung und Gewichtung der Punkte je Kunde unterschiedlich. Mehrere kritische Punkte werden genannt. Erstens ist bei fehlenden Evidenzen nicht immer aus dem Gap-Titel um die „subjektive“ Meinung von Einzelpersonen. Daher kann es immer noch passieren, dass Bewertungen gesamtunternehmerlich gesehen nicht korrekt vorgenommen werden. Zweitens gilt der Schritt für eine Anwendung in größerem Massstab als noch zu aufwendig. Vorgeschlagen wird aus dem ausführlichen Bericht einen zusätzlichen, abgespeckten Bericht zu erzeugen, dessen Aufwand sich nach der	Die menschliche Priorisierung wird klar beantwortet, da die Gewichtung je Kunde unterschiedlich ist und ein Mensch im Sinne eines Human in the Loop nötig bleibt, gestützt auf einen als stark beurteilten Bericht. Zugleich wird der Schritt als sehr komplex und für eine Anwendung in größerem Massstab als zu aufwendig erlebt, weshalb eine stärkere Unterteilung und ein zusätzlicher, abgespeckter Bericht je Kundengröße vorgeschlagen werden. Die Bewertung wird durch unklare Vorgaben erschwert, etwa fehlende Evidenzen und schwer verständliche Gap-Titel, eine offene Anleitung zum Ausfüllen sowie inhaltliche Zusammenhänge und Redundanzen zwischen Zielen, Gaps und Empfehlungen. Der Schritt wird als sehr komplex erlebt, weshalb viele Punkte zu berücksichtigen und an die jeweilige Kundenumgebung anzupassen, weshalb eine zusätzliche Unterteilung gewünscht wird. Zudem ist schwer abzuschätzen, wie viele Punkte auf hoch oder mittel gesetzt werden müssen, um zugleich guten Umfang und klare Botschaft zu erreichen. Mehrere Verständnislücken betreffen die Herleitung des Berichts: Unklar ist, weshalb Ziele ohne jegliche Evidenz erscheinen und wofür die Gaps stammen, und Empfehlungen, Ziele und Gaps sind teilweise inhaltlich gleich, aber mehrfach unterschiedlich formuliert.	• Menschliche Priorisierung wird klar beantwortet, da die Gewichtung je Kunde unterschiedlich ist und ein Mensch im Sinne eines Human in the Loop nötig bleibt • Der zugrunde liegende Bericht gilt als stark	• Sehr komplex und für eine Anwendung in größerem Massstab zu aufwendig • Bewertung durch unklare Vorgaben erschwert fehlende Evidenzen, schwer verständliche Gap-Titel, offene Anleitung zum Ausfüllen • Inhaltliche Zusammenhänge und Redundanzen zwischen Zielen, Gaps und Empfehlungen • Herleitung evidenzloser Ziele und die Herkunft der Gaps sind nicht nachvollziehbar • Priorisierung bleibt subjektive Einschätzung Einzelner und kann gesamtunternehmerlich unzutreffend ausfallen	• Die Anleitung ist unklar, welche Fehler auszufüllen sind (alle Einschätzungen jeder nur die Empfehlungstabelle) • Gap-Titel sind ohne Evidenz nicht verständlich • Es gibt keine abgespeckten Berichtsvarianten, der Schritt ist für eine Anwendung in größerem Massstab zu aufwendig • Die Herleitung evidenzloser Ziele und die Herkunft der Gaps sind nicht nachvollziehbar • Ziele, Gaps und Empfehlungen sind teilweise redundant • Es fehlt Orientierung, wie viele Punkte sinnvoll auf hoch oder mittel zu setzen sind • Ein Gap liefert teils keine Treffer, obwohl er laut Gesprächen relevant ist (Beispiel Teams-Labeling)
4.A	Lade den Analysebericht mit den Priorisierungen im Light Assessment Agent hoch, damit er ein Summary als Bericht für den Kunden erstellt. Der klickbare vorgeschlagene Prompt wäre: «Erstelle ein Summary aus diesem Bericht»	Wie sinnvoll wirkt dieser Schritt auf dich und ist es nachvollziehbar, warum dieser Schritt existiert? (Begründe)	Wie bereits erwähnt kann nur ein Mensch die für den Kunden relevanten Punkte aus dem Bericht entnehmen und diese so zusammenführen, dass sie für den Kunden und das Unterfangen von Vorteil sind. Dieser Schritt ist sehr komplex ein direkter Kundenbezug ist in „z“ zwingend nötig. Ich kann den Output nicht abschätzen und stehe diesem Schritt deshalb kritisch gegenüber. Ich werde unten die Anmerkung beifügen, sobald ich das Resultat sehe. Ich vermute, dass ich durch falsche Bewertung/en einen unklaren oder nicht zielführenden Output bekomme. Nach dem ich nun den Output gelesen habe, bin ich beeindruckt von der Zeit, die ich in diesem Ablauf gespart habe und der hohen Qualität der Management Summary.	Ja, der logische Abschluss der gesamten Vorarbeit und am Ende des Lieferobjekt für den Kunden.	Dieser Schritt ist sehr sinnvoll. Es braucht für das Management eine prägnante Übersicht über die betroffenen Themenbereiche.	Der Schritt wird als sinnvoller und logischer Abschluss der gesamten Vorarbeit beurteilt, der mit dem kundenrechtlichen Bericht das eigentliche Lieferobjekt darstellt. Der Bericht wird zugleich die Komplexität, der Direkter Kundenbezug Sache des Menschen bleiben. Vor Sicht des Ergebnisses besteht eine abwartende, eher kritische Haltung wegen des schwer abschätzbaren Outputs, nach Schlichtung überzeugen jedoch die hohe Qualität der Management Summary, die geparte Zeit und die grafische Aufbereitung, für ja/Nein-Fragen wird ein Kuchendiagramm angeregt. Als wichtiger Punkt wird der Datenschutz genannt, da trotz anonymisierter Kürzel Rückschlüsse auf Kollegen oder Abteilungen möglich bleiben und unklar ist, ob die KI oder ein vorgelagerter Korrekturschritt solche Bezüge entfernen müsste.	Die Erstellung des Summaries wird als sinnvoller und logischer Abschluss der Vorarbeit beurteilt, der mit dem kundenrechtlichen Bericht das eigentliche Lieferobjekt darstellt. Der Bericht wird zugleich die Komplexität, der Direkter Kundenbezug Sache des Menschen bleiben. Vor Sicht des Ergebnisses besteht eine abwartende, eher kritische Haltung wegen des schwer abschätzbaren Outputs, nach Schlichtung überzeugen jedoch die hohe Qualität der Management Summary, die geparte Zeit und die grafische Aufbereitung, für ja/Nein-Fragen wird ein Kuchendiagramm angeregt. Als wichtiger Punkt wird der Datenschutz genannt, da trotz anonymisierter Kürzel Rückschlüsse auf Kollegen oder Abteilungen möglich bleiben und unklar ist, ob die KI oder ein vorgelagerter Korrekturschritt solche Bezüge entfernen müsste.	• Sinnvoller, logischer Abschluss, der das eigentliche Lieferobjekt und eine prägnante Management-Übersicht erzeugt • Auswahl der relevanten Punkte und direkter Kundenbezug bleiben sinnvollerweise Sache des Menschen • Nach Schlichtung überzeugen die hohe Qualität der Management Summary, die geparte Zeit und die grafische Aufbereitung	• Vor Sicht des Ergebnisses abwartende, eher kritische Haltung wegen des schwer abschätzbaren Outputs und der Sorge, dass fehlerhafte Bewertungen zu einem unklaren Ergebnis führen • Datenschutz: trotz anonymisierter Kürzel sind Rückschlüsse auf Kollegen oder Abteilungen möglich, und es ist unklar, ob die KI oder ein vorgelagerter Korrekturschritt solche Bezüge entfernen müsste	• Für ja/Nein-Fragen werden Balkendiagramme statt anschaulicherer Kuchendiagramme verwendet • Der Datenschutz ist nicht sicher gestellt, da Bezüge auf Arbeitskollegen oder Abteilungen im Bericht Rückschlüsse erlauben

Anhang Q: Fragebogen Abschluss verdichtet

Fragebogen Abschluss Verdichtet - Fragebogen Abschluss							
Teil	Frage	Person 01	Person 02	Person 03	Verdichtung	Positiv	Negativ
Teil 1: Effektivität	Frage 1.1: Beschreibe deinen Eindruck der Zwischenergebnisse (Bereinigtes Transkript, extrahierte Kernausagen, Konsolidierung) und des finalen Summary hinsichtlich inhaltlicher Korrektheit, Vollständigkeit und Verständlichkeit.	Generell, absolut Top, mit einigen Punkten, die ich jetzt noch genauer anschauen müsste. Der Grund hierfür könnte mein Argwohn gegenüber Automatismen sein (im Sinne von: Traue nie einer Präsidi die du nicht selber...).	Die Ergebnisse sind auf einem hohen Niveau, aber nicht perfekt. Im Kosten-Nutzen-Verhältnis aber sehr sinnvoll und auch ohne Zweifel weiterverwendbar.	Ich war ehrlich gesagt doch etwas überrascht, wie gut die KI die Bereinigungen und Extrahierungen machen konnte. Aus meiner Sicht hat inhaltlich nichts gefehlt und es mussten nur kleinere Anpassungen gemacht werden, die darauf zurück zu führen sind, dass die Interviewpartner nicht muttersprachlich Hochdeutsch sind.	Die Zwischenergebnisse und das Summary sind inhaltlich als sehr gut und auf hohem Niveau beurteilt, wenn auch nicht durchgehend perfekt. Inhaltlich fehlte praktisch nichts, nötig waren nur kleinere Anpassungen, etwa wegen nicht muttersprachlich Hochdeutsch sprechender Interviewpartner. Die Ergebnisse gelten als zweifelslos weiterverwendbar und im Kosten-Nutzen-Verhältnis als sehr sinnvoll. Einzelne Punkte sollen noch genauer geprüft werden, was auch mit der Nachvollziehbarkeit wird unterschiedlich beurteilt. Bezogen auf Ablauf, Umsetzung und Resultat sowie dank der Quellenangaben gilt sie als hoch bis vollständig gegeben, ebenso die Verwendbarkeit. Einschränkung wird angemerkt, dass der Bericht allein sehr lang und für Laien nicht vollständig nachvollziehbar ist und sich erst mit einer ergänzenden Präsentation gut weiterverwenden lässt. Als Voraussetzungen werden klare	• Inhaltlich sehr gut und auf hohem Niveau, praktisch nichts fehlte • Nur kleinere Anpassungen nötig, bedingt durch nicht muttersprachliches Hochdeutsch • Zweifellos weiterverwendbar und im Kosten-Nutzen-Verhältnis sehr sinnvoll	• Nicht durchgehend perfekt, einzelne Punkte sollten noch genauer geprüft werden • Grundsätzliche Skepsis gegenüber Automatismen bleibt
	Frage 1.2: Wie würdest du die Nachvollziehbarkeit des Endergebnisses und seine Verwendbarkeit im weiteren Projektverlauf beschreiben?	Ich würde sie als hoch einschätzen, bezogen auf Ablauf Umsetzung und Resultat. Zeitersparnis sollte deshalb auch separat genannt werden. Es bedarf allerdings auch klare Rahmenbedingungen z.B. Kunden und Projektbezug muss sichergestellt sein.	Durch die Quellenangaben hat das Endergebnis eine maximale Nachvollziehbarkeit, dementsprechend ist auch die Verwendbarkeit in meinen Augen vollständig gegeben.	Der Bericht als solches ist aus meiner Sicht immer noch sehr lang und für einen Laien nicht zu 100% nachvollziehbar. Ich gehe aber davon aus, dass die Präsentation da deutlich Klarheit geschaffen hätte. Nur den Bericht alleine finde ich nicht passend genug für die Weiterverwendung.	• Hohe bis vollständige Nachvollziehbarkeit dank Ablauf, Umsetzung, Resultat und Quellenangaben • Verwendbarkeit grundsätzlich gegeben (sichergestellter Kunden- und Projektbezug)	• Bericht allein sehr lang und für Laien nicht vollständig nachvollziehbar • Gut weiterverwendbar erst mit ergänzender Präsentation • Setzt klare Rahmenbedingungen voraus (sichergestellter Kunden- und Projektbezug)	• Der Bericht allein ist sehr lang und für Laien nicht vollständig nachvollziehbar • Ohne ergänzende Präsentation ist die Weiterverwendung eingeschränkt
	Frage 1.3: Wie würdest du das Endergebnis in der vorliegenden Form im Hinblick auf eine direkte Weitergabe an einen Kunden beschreiben?	Der Fall würde schonungslos analysiert, ausgewertet und zusammengefasst. Der Inhalt müsste nun noch "verdaubar" gemacht werden. Die Präsentation sollte demzufolge unbedingt mit "Fun facts" angereichert werden. Diese Postiven Aspekte gab es im vorliegenden Fall. Ich habe diese aber als irrelevant abgehan (personal learning).	Es ist zu umfangreich und zu wenig selbsterklärend, sowie nicht ausreichend visuell aufbereitet. Ich würde mir noch ein paar sinnvoll eingesetzte Grafiken mit Highlights wünschen z.B. mit einem Kreisdiagramm, das zeigt wie viele Personen oder Prozent Mängel oder Handlungsbedarf sehen. Je nachdem wer auf Kundenseite der Ansprechpartner ist, kann es den Gegenüber fast schon "erschlagen". Es fehlen Deckblatt und GARAIO-Netics Design und Logo im gesamten Dokument. Allenfalls könnte man auch mit Platzhaltern arbeiten, sodass auf dem Deckblatt dann steht: Eine Analyse für (Kundenname) von GARAIO low, Netics.	Den Bericht alleine würde ich nicht an den Kunden weitergeben. Dies müsste zwingend in Form einer Präsentation passieren. Der Bericht ist eindeutig zu lang und doch noch etwas kompliziert aufgebaut. Es bräuchte aus meiner Sicht zu viel Erklärung. Ebenfalls müsste zu 100% sichergestellt sein, dass der Bericht und nach die Präsentation anonymisiert ist und aus keiner Aussage Rückschlüsse auf Personen gemacht werden könnten.	Für eine direkte Weitergabe an Kunden wird das Endergebnis in der vorliegenden Form als noch nicht geeignet beurteilt. Übereinstimmend gilt der Bericht als zu lang, zu wenig selbsterklärend und erklärungsbedürftig, weshalb die Weitergabe zwingend über eine Präsentation erfolgen sollte. Inhaltlich ist der Bericht sehr gut und vollständig, müsste aber verständlicher und visuell ansprechender aufbereitet werden, etwa mit gezielten Grafiken und Highlights wie einem Kreisdiagramm zum Anteil festgestellter Mängel. Handlungsbefehle sowie mit positiven, aufbauenden Elementen. Es fehlen zudem ein Deckblatt sowie das GARAIO-beziehungswise Netics-Design und -Logo, denkbar mit Platzhaltern für den Kundenamen. Als zwingende Bedingung	• Inhaltlich schonungslos und vollständig analysiert	• In vorliegender Form für die direkte Kundenweitergabe noch nicht geeignet • Zu lang, zu wenig selbsterklärend und erklärungsbedürftig • Zu wenig visuell aufbereitet, es fehlen Grafiken und Highlights wie ein Kreisdiagramm • Es fehlen Deckblatt sowie GARAIO-beziehungswise Netics-Design und -Logo • Die vollständige Anonymisierung muss noch sichergestellt werden
	Frage 1.4: Wie würdest du dein Vertrauen in die Ergebnisse beschreiben, die der Prozess erzeugt hat? Was hat dieses Vertrauen erzeugt oder geschwächt?	Ganz klar gestützt, aber ebenso klar die Grenzen aufgezeigt.	Mein Vertrauen in KI würde ich generell als eher mittelmäßig, aber vorhanden beschreiben. Die Ergebnisse sind für mich zufriedenstellend, insbesondere im Verhältnis zum Aufwand. Wichtig ist, dass der Mensch (Consultant) weiterhin die finalen Entscheidungen und Einschätzungen trifft. Viele Menschen tendieren dazu, einer KI übermäßig zu vertrauen, da sie als "intelligent" wahrgenommen wird und allenfalls sogar als neutral, da es ja eine Maschine ist (Bias).	Grundsätzlich habe ich sehr großes Vertrauen in KI, aber natürlich auch gewisse Vorbehalte. Die Ergebnisse sind, wie gut die Ergebnisse sind. Trotzdem ist es aber so, dass auch immer noch der Mensch mitbestimmt. Die KI kann nicht verhindern, dass bei den manuellen Korrekturen oder bei der Einschätzung der Kriterien nicht Fehler passieren. Die KI ist nur so gut, wie wir sie auch beliefern. Aber ich würde das auf jeden Fall einsetzen in den Projekten.	Das Vertrauen in die erzeugten Ergebnisse wird grundsätzlich als gestützt und teils als sehr gross beschrieben, getragen durch die sichtbare Ergebnisqualität und das gute Verhältnis zum Aufwand. Zugleich werden die Grenzen gesehen. Betont wird, dass der Mensch die finalen Entscheidungen und Einschätzungen treffen muss, da die KI Fehler bei manuellen Korrekturen oder bei der Kriterienbewertung nicht verhindern kann und nur so gut ist wie ihre Datengrundlage. Eine eher kritische Grundhaltung gegenüber KI bleibt bestehen, verbunden mit dem Hinweis,	• Vertrauen grundsätzlich gestützt, teils sehr groß • Getragen durch sichtbare gute Ergebnisqualität und gutes Verhältnis zum Aufwand	• Grenzen der KI klar sichtbar, finale Entscheidungen müssen noch Menschen bleiben • KI kann Fehler bei manuellen Korrekturen oder bei der Kriterienbewertung nicht verhindern und ist nur so gut wie ihre Datengrundlage • Eher kritische Grundhaltung bleibt, Gefahr übermäßigen Vertrauens in die KI
	Frage 2.1: Wie hast du den Prozess hinsichtlich Unterstützung und Belastung erlebt?	Sehr gut ausbalanciert. Einzig bei Aufgabe 3b war ich wegen der der Cap und Ziel-Titel (zu wenig klare Definition) unsicher. Der Bericht habe ich bereits eine Aussage gemacht.	Die Unterstützung des Agents war sehr zufriedenstellend. Es ist eine klare Zeitersparnis erkennbar. Zur Belastung habe ich bereits eine Aussage gemacht.	Ich muss ehrlich sagen, dass ich die Belastung vor allem zu Beginn doch als sehr hoch empfand. Dies ist aber darauf zurückzuführen, dass ich überhaupt nicht im Thema war und auch nicht an den Interviews beteiligt war. Ich musste zuerst selbst in das Thema hinein-wachsen und konnte daher die ersten Interviews auch nicht mit höherer Geschwindigkeit hören. Anschließend ist dann die Belastung natürlich deutlich kleiner geworden und die KI war eine grosse Unterstützung. Ich hätte das Gute ist auch noch, dass man in der Zeit wo die KI arbeitet andere Tasks erledigen kann. Das heisst, die Effizienz steigt deutlich.	Unterstützung und Belastung werden insgesamt als gut ausbalanciert erlebt, mit klar erkennbarer Zeitersparnis durch den Agenten. Die Belastung war vor allem zu Beginn hoch, was auf fehlendes Themen- und Kontextwissen zurückgeführt wird, etwa wenn die ersten Interviews noch nicht beschleunigt gehört werden konnten. Mit zunehmender Unterstützung erlebt, auch weil sich während ihrer Rechenzeit andere Aufgaben erledigen lassen, was die Effizienz spürbar erhöht. Als einzelne unsichere Stelle wird Aufgabe 3b wegen der zu wenig klar definierten Cap- und Ziel-Titel genannt.	• Unterstützung und Belastung gut ausbalanciert, klare Zeitersparnis • Nach Einarbeitung grosse Unterstützung durch die KI, parallele Aufgaben erhöhen die Effizienz	• Belastung vor allem zu Beginn hoch, solange Themen- und Kontextwissen • Aufgabe 3b wegen zu wenig klar definierter Cap- und Ziel-Titel unsicher
Teil 2: Effizienz	Frage 2.2: Wie würdest du das Aufwand-Ergebnis-Verhältnis des Prozesses beschreiben?	Sehr gut, allerdings (so nehme ich an) nur bei Personen im Projekt resp. Kundenbezug haben.	Das Verhältnis ist meiner Meinung nach sehr gut.	Im Test würde ich es als eher gut überbewertet sehr gut beurteilen. Das Projekt gewesen wäre und ich schon viel für den Prozess und Projekt und Kundenbezug gilt. Im Test ohne eigenes Projekt fällt die Einschätzung etwas zurückhaltender aus und würde bei besserem Kundenwissen als sehr gut bewertet. Als Verbesserung wird angeregt, Interviewtranskripte einzeln eingeben zu können, um die Nachbearbeitung zutrah und möglichst nah am Gespräch vorzunehmen.	• Aufwand-Ergebnis-Verhältnis überbewertet sehr gut	• Volle Wirkung vor allem für Personen im Projekt und Kundenbezug • Interviewtranskripte einzeln eingeben werden, was eine zeitnahe Nachbearbeitung erschwerte	• Die volle Wirkung zeigt sich vor allem bei Personen im Projekt und Kundenbezug • Interviewtranskripte einzeln eingeben werden, was eine zeitnahe Nachbearbeitung erschwerte
	Frage 2.3: Beschreibe, wie du über den Prozess hinweg das Zusammenspiel der eingesetzten Werkzeuge erlebt hast.	Absolut Zielführend, einfach und klar.	Sehr sinnvoll. Wie bereits erwähnt, habe ich großen Interesse an dem eingesetzten Agenten, sodass man hier ein besseres Bild bekommt, zu was der Agent in der Lage ist. Welche Interaktionen können allenfalls noch über den aktuellen Prozess hinaus eingesetzt werden? Welche Hilfestellungen kann der Agent direkt selbst liefern?	Ich empfand das Zusammenspiel der Werkzeuge als sehr logisch und konsistent. Die Rückmeldungen der KI waren als gut verständlich, sodass zu jedem Zeitpunkt, was der genaue Schritt ist. Am aktuellen Prozess hinaus eingesetzt werden? Welche Hilfestellungen kann der Agent direkt selbst liefern?	Das Zusammenspiel der Werkzeuge wird als zielführend, logisch, konsistent und klar erlebt. Die Rückmeldungen der KI waren als gut verständlich, sodass zu jedem aktuellen Schritt jederzeit klar war. Darüber hinaus besteht Interesse daran, die Fähigkeiten des Agenten besser kennenzulernen, etwa welche Interaktionen über den aktuellen Prozess hinaus möglich sind und welche Hilfestellungen er selbst liefern kann.	• Zusammenspiel der Werkzeuge zielführend, logisch, konsistent und klar • Rückmeldungen der KI gut verständlich, der aktuelle Schritt war jederzeit klar	• Keine negativen Punkte genannt
							• Die Fähigkeiten und der mögliche Funktionsumfang des Agenten sind noch nicht vollständig transparent
						Gesamtaussage	
						Bezogen auf die Effektivität, verstanden als das vollständige Erreichen des angestrebten Arbeitsergebnisses, zeichnen die Rückmeldungen ein überwiegend hohes, aber nicht vollständig abstrichfreies Bild. Die inhaltlichen Zwischenergebnisse und das Summary werden als weitgehend korrekt, vollständig und verlässlich beschrieben, mit nur kleineren nötigen Anpassungen und hoher Nachvollziehbarkeit durch Quellenangaben. Abstriche zeigen sich vor allem beim fertigen Lieferprodukt für den Kunden, da der Bericht in der vorliegenden Form als zu lang, zu wenig aufbereitet und nicht ohne ergänzende Präsentation direkt weitervergeben beschrieben wird und eine vollständige Anonymisierung sowie ein kundengerechtes Layout als noch offen bekannt werden. Das Vertrauen in die Ergebnisse wird als gestützt geschildert, bleibt aber an die menschliche Kontrolle und die Qualität der Datengrundlage gebunden.	

Fragebogen Abschluss Verdichtet - Fragebogen Abschluss							
	Frage	Person 01	Person 02	Person 03	Verdichtet	Positiv	Negativ
Teil 3: Zufriedenheit während der Durchführung	Frage 3.1: Wie hast du dich während der Durchführung in Bezug auf Orientierung und Handlungssicherheit erlebt?	Besser als erwartet, bestätigt durch den jeweiligen Output bis auf Aufgabe 3b.	Sehr zufriedenstellend.	Ich habe mich sehr gut geführt und sicher gefühlt. Allenfalls könnte es helfen, wenn bei den einzelnen Schritten der Ki eine der durchschnittlichen Bearbeitungszeit angegeben würde, nur so als Orientierung.	Orientierung und Handlungssicherheit werden als sehr zufriedenstellend und besser als erwartet erlebt, bestätigt durch den jeweiligen Output und das Gefühl, gut geführt zu sein. Eine Ausnahme bildet Aufgabe 3.b. Als hilfreiche Ergänzung wird angeregt, bei den Ki-Schritten eine durchschnittliche Bearbeitungszeit als Orientierung anzugeben.	• Orientierung und Handlungssicherheit sehr zufriedenstellend und besser als erwartet, Gefühl gut geführt zu sein • Ausnahme bildet Aufgabe 3.b	• Handlungssicherheit eingeschränkt • Bei den Ki-Schritten fehlt eine Angabe der durchschnittlichen Bearbeitungszeit zur Orientierung
	Frage 3.2: Beschreibe, wie du die persönliche Beanspruchung oder Anstrengung während dem Prozess erlebt hast.	Sehr gut - ich kann diese Arbeiten im stop-go-Verfahren behandeln. Was in meinem Arbeitsumfeld sehr relevant ist, da ich selbst einige Arbeiten mit ähnlichem Design (qualitative Inhaltsanalyse) verfasst habe und es somit für mich eine willkommene Abwechslung war. Die inhaltliche Tiefe hat mir sehr gut gefallen und hätte sogar noch tiefer sein können.	Aufgrund anderer dringender Kundentermine habe ich leider die Wochenfrist nicht einhalten können. Die Anstrengung war nicht sehr groß, eher eine angenehme Arbeit, da ich selbst einige Arbeiten mit ähnlichem Design (qualitative Inhaltsanalyse) verfasst habe und es somit für mich eine willkommene Abwechslung war. Die inhaltliche Tiefe hat mir sehr gut gefallen und hätte sogar noch tiefer sein können.	Das Hören der Interviews und Korrektur der Transkripte fand ich zu Beginn sehr anstrengend, da ich nicht im Thema drin war. Auch war teilweise die Aufnahmequalität nicht sehr gut, was mich nicht einfach machte, das Gesprochene mit den Transkripten zu vergleichen. Mit der Zeit viel es etwas einfacher, da man schon ungefähr wusste, welche Frage als nächstes kommt und man so schon den Kontext hatte. Die weiteren Prozessschritte empfand ich als nicht sehr anstrengend und auch die Belastung war in einem mehr als vertretbaren Rahmen.	Die Beanspruchung wird insgesamt als gut handhabbar und eher angenehm beschrieben, auch weil sich die Arbeit im Stop-and-go-Verfahren erledigen lässt. Als besonders anstrengend gilt der Anfang, vor allem das Hören und Korrigieren der Transkripte bei fehlendem Themenkontext und teils schlechter Aufnahmequalität, was den Abgleich von Gesprochenem und Transkript erschweren kann. Mit zunehmendem Kontext und gegen Ende sinkt die Beanspruchung deutlich und bleibt in einem gut vertretbaren Rahmen. Punktuell erhöhte Beanspruchung bei Schritt 3.b.	• Beanspruchung insgesamt gut handhabbar und eher angenehm, auch dank Stop-and-go-Arbeitweise • Gegen Ende deutlich gesunken und in gut vertretbarem Rahmen • Inhaltliche Tiefe positiv hervorgehoben	• Anfang besonders anstrengend, vor allem Hören und Korrigieren der Transkripte bei fehlendem Themenkontext • Teils schlechte Aufnahmequalität erschweren den Abgleich von Gesprochenem und Transkript • Punktuell erhöhte Beanspruchung bei Schritt 3.b
	Frage 3.3: Welche Stellen während der Durchführung und der besonders in Erinnerung geblieben, sei es positiv oder negativ, und was hat diese Stellen jeweils ausgemacht?	Schritt 3b, weil es einige Unsicherheiten (meinerseits) gab die mich gehindert haben nur auf die Aufgabe zu fokussieren. Mein Bezug zu einem ähnlichen Projekt, mit dem ich die vorliegende Situation automatisch verglichen habe. Abgrenzung ist auch hier wichtig.	Einen 136-seitigen Bericht auszufüllen. Ich war zunächst überrascht, ob der Seitenanzahl und der Ausführlichkeit des Berichts.	Besonders positiv ist mir die Glättung der Transkripte in Erinnerung geblieben. Ich fand es sehr faszinierend, wie gut dies durch die KI gemacht wurde. Egal ob die Wortwiederholungen waren, es kam anschließend ein konkreter und im Kontext geübener Satz heraus. Das fand ich doch sehr beeindruckend.	Die Interaktion mit dem Agenten wird als sehr gut, klar und angenehm beurteilt. Besonders positiv ist mir die Glättung der Transkripte in Erinnerung geblieben. Ich fand es sehr faszinierend, wie gut dies durch die KI gemacht wurde. Egal ob die Wortwiederholungen waren, es kam anschließend ein konkreter und im Kontext geübener Satz heraus. Das fand ich doch sehr beeindruckend.	• Sprachliche Glättung der Transkripte besonders beeindruckend, korrekte und kontextgenaue Sätze trotz Akzent und Wortwiederholungen	• Schritt 3.b durch eigene Unsicherheiten und den Vergleich mit einem ähnlichen Projekt erschwert • Der 136-seitige Bericht wirkt überraschend umfangreich
	Frage 3.4: Beschreibe deine Erfahrungen und Meinungen zur Interaktion mit dem Agent.	Logisch und klar.	Angenehm, aber ich hätte gerne einen echten Agenten gehabt, um die Interaktion und seine Fähigkeiten noch besser einschätzen zu können.	Ich fand die Interaktion mit dem Agenten als sehr gut. Besonders gut fand ich, dass der Agent sehr flexibel war und auf meine Fragen und Anforderungen eingegangen ist. Die Interaktion war sehr angenehm und ich habe mich sehr wohl gefühlt. Ich finde die Interaktion mit dem Agenten als sehr gut. Besonders gut fand ich, dass der Agent sehr flexibel war und auf meine Fragen und Anforderungen eingegangen ist. Die Interaktion war sehr angenehm und ich habe mich sehr wohl gefühlt.	Die Interaktion mit dem Agenten wird als sehr gut, klar und angenehm beurteilt. Besonders positiv ist mir die Glättung der Transkripte in Erinnerung geblieben. Ich fand es sehr faszinierend, wie gut dies durch die KI gemacht wurde. Egal ob die Wortwiederholungen waren, es kam anschließend ein konkreter und im Kontext geübener Satz heraus. Das fand ich doch sehr beeindruckend.	• Interaktion logisch, klar und angenehm • Kurze Zusammenfassungen zu hochgeladenen Dokumenten und nächsten Schritt erlauben Kontrolle, ob alles erkannt wurde	• Es fehlt eine sichtbare Bearbeitungsanzeige wie eine Sanduhr oder ein drehendes Symbol, was Unsicherheit auslösen kann, ob der Agent noch arbeitet
Teil 4: Zufriedenheit nach der Durchführung	Frage 4.1: Wie würdest du dein Gesamtbild des Prozesses beschreiben, und wie geht es dir nach Abschluss des Tests?	Positives Gesamtbild, durch nachvollziehbare Schritte und einen logischen Ablauf, bis hin zur kundengerechten „Aufbereitung“ des Assessments. Sollte als Beispiel genommen werden, um die rein technischen Probleme unserer Kunden den Griff zu kriegen.	Sehr positiv, das kann bei korrektem Einsatz wirklich einen Mehrwert für uns und die Kunden bieten. Wichtig ist mir, dass die Methoden verstanden werden und somit korrekt eingesetzt werden. Hier hatte ich auch intern Schulungen für den richtigen Weg. Das Assessment muss erklärt werden und auch der Einsatz von KI muss besser oder überhaupt. Insbesondere intern geschult werden.	Ich finde den Prozess wie auch die Umsetzung mit der KI sehr sinnvoll. Im Arbeitsalltag kann mir so sehr viel Arbeit abgenommen werden und ich kann mich auf andere Tasks konzentrieren. Ich bin sehr zuversichtlich, dass dies ein viel eingesetzter Arbeitsschritt gilt und mit stetiger Erfahrung auch noch weitere Verbesserungen oder Optimierungen mit einfließen können.	Das Gesamtbild fällt durchweg positiv aus. Hervorgehoben werden die nachvollziehbaren Schritte, der logische Ablauf bis zur kundengerechten Aufbereitung und der erwartete Mehrwert für Unternehmen und Kunden bei korrektem Einsatz. Bis hin zur Idee, den Prozess als Standardbeispiel für die Lösung technischer Kundenprobleme zu nutzen. Betont wird, dass die zugrunde liegenden Methoden und der KI-Einsatz verstanden und korrekt angewendet werden müssen, wofür interne Schulungen als wesentlicher Weg gelten.	• Durchweg positives Gesamtbild, nachvollziehbare Schritte und logischer Ablauf bis zur kundengerechten Aufbereitung • Erwarteter Mehrwert für Unternehmen und Kunden bei korrektem Einsatz	• Voraussetzung ist, dass Methoden und der KI-Einsatz verstanden und korrekt angewendet werden, wofür interne Schulungen nötig sind
	Frage 4.2: Wie hast du die Aufgabenverteilung zwischen dir und der KI während des Prozesses erlebt?	So wie sie m.e. sein sollten!	So wie es sein soll. Die Fleissarbeit macht die KI und die Leitplanken sowie Entscheidungen trifft der Mensch.	Die Aufgabenverteilung habe ich als sehr sinnvoll erlebt. Es gibt einfache Dinge, die kann die KI übernehmen.	Die Aufgabenverteilung zwischen Mensch und KI wird einheitlich als richtig und sinnvoll erlebt. Die KI übernimmt die Fleissarbeit, während Leitplanken und Entscheidungen beim Menschen bleiben, da es Aufgaben gibt, welche die KI nicht übernehmen kann.	• Aufgabenverteilung einheitlich als richtig und sinnvoll erlebt, die KI übernimmt die Fleissarbeit, Entscheidungen bleiben beim Menschen	• Keine negativen Punkte genannt
	Frage 4.3: Unter welchen Bedingungen würdest du dir einen Einsatz des Prozesses in dieser Form für reale Kundenprojekte vorstellen, und welche Vorbehalte hättest du dabei?	Wir schaffen damit dringend benötigte Transparenz, ich für meinen Teil werde empfehlen, dass die Sales und Account Management Teams den Einsatz dieses Assessments als Standard-Lieferobjekt in das Projektskoping aufnehmen.	Ich würde mich sehr freuen, wenn ich es bei einem Kunden verwenden könnte. Die Bedenken zur Anwendung habe ich 4.1 bereits geschlüsselt.	Grundsätzlich kann ich mir den Einsatz bei jedem Projekt vorstellen, bei dem ein Assessment durchgeführt wird. Es muss natürlich mit dem Kunden geklärt werden, ob die Informationen durch KI analysiert werden dürfen aufgrund von Datenschutzbestimmungen beim Kunden.	Ein Einsatz in realen Kundenprojekten wird grundsätzlich befürwortet und für jedes Projekt mit Assessment für vorstellbar gehalten, da der Prozess dringend benötigte Transparenz schafft. Angeregt wird, das Assessment als Standard-Lieferobjekt ins Projektskoping aufzunehmen. Als Bedingungen und Vorbehalte gelten das nötige Verständnis der Methoden samt interner Schulung sowie die datenschutzrechtliche Klärung.	• Einsatz grundsätzlich befürwortet und für jedes Assessment-Projekt vorstellbar, da es dringend benötigte Transparenz schafft • Anregung, das Assessment als Standard-Lieferobjekt ins Projektskoping aufzunehmen	• Vorbehalt: nötiges Verständnis der Methoden samt interner Schulung • Datenschutzrechtliche Klärung mit dem Kunden erforderlich, ob dessen Informationen durch KI analysiert werden dürfen
	Frage 4.4: Welche Aspekte des Prozesses würdest du verändern, und welche sollten erhalten bleiben? Begründe deine Auswahl kurz.	Ich würde genauso starten mit Ausnahme von Schritt 3b (bereits begründet).	Keine wesentlichen Änderungen.	Im Moment würde ich den Prozess so belassen. Ich würde aber empfehlen, wenn der Prozess breitgefächert eingesetzt wird, nach einer gewissen Zeit nochmals eine Evaluation zu machen. Unter realen Bedingungen kommen dann vielleicht doch noch Dinge zum Vorschein, die jetzt im Test nicht bedacht wurden.	Im Moment würde ich den Prozess so belassen. Ich würde aber empfehlen, wenn der Prozess breitgefächert eingesetzt wird, nach einer gewissen Zeit nochmals eine Evaluation zu machen. Unter realen Bedingungen kommen dann vielleicht doch noch Dinge zum Vorschein, die jetzt im Test nicht bedacht wurden.	• Prozess soll im Wesentlichen erhalten bleiben, kaum Änderungswünsche	• Einzige Ausnahme ist Schritt 3.b
Teil 5: Zusatz	Frage 5.1: Wie viele Stunden hast du für die Durchführung benötigt?	Ohne Präz bis ich auf 1,3h gekommen.	Ca. 25 Stunden, wahrscheinlich ist es in 20 Stunden möglich, aber ich konnte es leider nicht möglichst am Stück bearbeiten und so musste ich mich immer erst wieder einfinden.	(Wegen Dokumentenfehler erst später nochmal gefragt.) Insgesamt 23 Stunden aber davon 2-3 Stunden für Onboarding und Fragebogen.	Der genannte Zeitaufwand liegt zwischen rund 13 Stunden ohne Präsentation und etwa 23 bis 25 Stunden. Die höheren Werte werden damit relativiert, dass die Bearbeitung nicht am Stück erfolgte und ein wiederholtes Einarbeiten nötig war beziehungsweise dass davon 2 bis 3 Stunden auf Onboarding und Fragebogen entfielen. Eine Bearbeitung in etwa 20	• Eine Bearbeitung in etwa 20 Stunden wird als realistisch eingeschätzt, ohne Präsentation deutlich weniger (rund 13 Stunden)	• Einzelne lagen bei rund 23 bis 25 Stunden, unter anderem weil nicht am Stück gearbeitet wurde und Zeit für Onboarding und Fragebogen anfiel
							• Keine prozessbedingte Schwachstelle, die teils höhere Stundenzahl (23 bis 25) ist beizubehalten, etwa weil nicht am Stück gearbeitet wurde und Onboarding anfiel
Der Zusatzteil bildet ergänzende Angaben ohne eigenes H2D-Überschema. Der genannte Zeitaufwand liegt zwischen circa 13 Stunden ohne Präsentation und etwa 23 bis 25 Stunden, wobei höhere Werte mit teils bedingten Gründen wie unterbreiteter Bearbeitung und Onboarding erklärt werden und eine Bearbeitung in rund 20 Stunden als							

Fragebogen Abschluss Verdichtet - Fragebogen Abschluss								
Teil		Person 01	Person 02	Person 03	Verdichtung	Positiv	Negativ	Offene Schwachstellen
	Frage 5.2: Wie sinnvoll scheint dir der gesamte Prozess insgesamt?	Gemessen am Resultat verstehe ich das als neuen Standard bei Projekten.	Sehr sinnvoll.	(Wegen Dokumentenfehler erst später mündlich gefragt.) Die Meinung ist positiv am Resultat wird er sogar als möglicher neuer Standard für Projekte verstanden.	Der Gesamtprozess wird als sehr sinnvoll und als Erleichterung beurteilt. Gemessen am Resultat wird er sogar als möglicher neuer Standard für Projekte verstanden.	Gesamtprozess als sehr sinnvoll und als Erleichterung beurteilt, gemessen am Resultat als möglicher neuer Standard.	Keine negativen Punkte genannt	Keine offene Schwachstelle genannt
	Frage 5.3: Wenn es noch irgendwas gibt, was du mitteilen willst, kannst du das hier gerne schreiben.	Danke das ich hier mitwirken durfte.	Mir gefällt die Idee sehr gut und auch die Umsetzung ist sehr durchdacht. Top Arbeit!	Vielen Dank dass ich Teil dieses Tests sein durfte. Es war sehr spannend für mich zu sehen, was alles möglich ist mit KI, wenn man es richtig und gezielt einsetzt. Ich bin sehr sehr durchdacht und gelange zu KI, bin ich sehr überzeugt, dass wir das noch immer viel zu wenig nutzen in unseren Arbeitsalltag und wir unsere Arbeit noch viel mehr erleichtern könnten.	Die abschließenden Rückmeldungen sind durchweg positiv und dankbar für die Teilnahme. Idee und Umsetzung werden nun als richtig und gezielt einsetzt. Ich bin sehr sehr durchdacht und gelange zu KI, bin ich sehr überzeugt, dass wir das noch immer viel zu wenig nutzen in unseren Arbeitsalltag und wir unsere Arbeit noch viel mehr erleichtern könnten.	Durchweg positive und dankbare Rückmeldungen. Idee und Umsetzung als sehr durchdacht und gelungen bezeichnet. Überzeugung, dass mit gezieltem KI Einsatz viel möglich ist und das Potenzial bislang zu wenig genutzt wird.	Keine negativen Punkte genannt	Keine offene Schwachstelle genannt