

Bachelorarbeit

Einfluss von Meetings auf Produktivität, Flow und Wohlbefinden von Softwareentwickelnden

Projektnummer: 25HS_IIT34

Eingereicht von:

Xeno Isenegger

xeno.isenegger@students.fhnw.ch

Gideon Monterosa

gideon.monterosa@students.fhnw.ch

Fachbetreuer FHNW:

Prof. Dr. Norbert Seyff

norbert.seyff@fhnw.ch

Dr. Nitish Patkar

nitish.patkar@fhnw.ch

Abstract

Meetings zählen zu den häufigsten Unterbrechungen im Arbeitsalltag von Softwareentwickelnden, doch ihre Auswirkungen auf Produktivität, Flow und Wohlbefinden werden in der Forschung selten gemeinsam betrachtet. Diese Arbeit untersucht den Einfluss von Meetings auf diese drei Dimensionen in einem zweiphasigen Mixed-Methods-Design. Eine Online-Befragung ($N = 55$) erfasst retrospektive Einschätzungen zu Meeting-Praktiken und deren wahrgenommenen Auswirkungen. Eine anschließende Feldstudie mit *Experience Sampling* ($N = 3,73$ Meeting-Bewertungen, 23 Tagesbewertungen) erhebt situative Echtzeit-Daten über zwei Arbeitswochen. Zur Datenerhebung wurde eine Webapplikation entwickelt, die automatisch synchronisierte Kalenderdaten mit strukturiertem subjektivem Feedback kombiniert und Reflexionsdashboards bereitstellt.

Die Ergebnisse zeigen, dass Meetings über unterschiedliche Mechanismen auf die drei Outcome-Konstrukte wirken. Die Meeting-Anzahl fragmentiert den Arbeitstag und beeinträchtigt primär den Flow, während die empfundene Belastung als Vermittler zwischen objektiver Meeting-Last und Wohlbefinden fungiert. Die wahrgenommene Produktivität hängt stärker von der subjektiven Belastungswahrnehmung ab als vom objektiven Ausmass der Meetings. Ein zentraler Befund ist die Dissoziation zwischen Nützlichkeit und Disruption: Meetings können als wertvoll und zugleich als störend erlebt werden, eine Differenzierung, die nur durch mehrdimensionale Echtzeit-Erhebung sichtbar wird. Der Vergleich beider Studienphasen offenbart zudem eine Wahrnehmungs-Realitäts-Diskrepanz, insbesondere beim Wohlbefinden, dessen Beeinträchtigung retrospektiv kaum wahrgenommen wird. Die Compliance-Raten von über 90 % bestätigen die Praktikabilität des entwickelten Erhebungsinstruments im Arbeitsalltag.

Keywords: Meetings, Softwareentwicklung, Produktivität, Flow, Wohlbefinden, Experience Sampling Method, Meeting-Kultur, Fragmentierung

Inhaltsverzeichnis

Abstract	2
1. Einleitung	8
1.1. Zielsetzung der Arbeit	8
1.2. Forschungsfragen	9
1.3. Wissenschaftlicher und praktischer Beitrag	9
1.4. Aufbau der Arbeit	10
2. State of the Art	11
2.1. Meeting-Last, Produktivität und Wohlbefinden	11
2.1.1. Multidimensionale Produktivitätsmessung	11
2.2. Unterbrechungen, Flow und Recovery	12
2.2.1. Flow als Voraussetzung produktiver Softwareentwicklung	12
2.2.2. Kognitive Kosten und Recovery	12
2.3. Daily Stand-ups in der Softwareentwicklung	13
2.4. Meeting-Effektivität und Wohlbefinden	13
2.5. Wahrnehmungs-Realitäts-Diskrepanz und Experience Sampling	14
2.6. Forschungslücken und Einordnung der vorliegenden Arbeit	14
3. Forschungsmodell und Forschungsdesign	16
3.1. Konzeptionelles Forschungsmodell	16
3.1.1. Outcome-Konstrukte: Produktivität, Flow und Wohlbefinden	17
3.1.2. Vom Meeting zum Tageserleben: Variablengruppen und Wirkpfade	18
3.1.3. Querschnittsperspektive: Wahrnehmung vs. Realität	19
3.2. Forschungsdesign	20
3.2.1. Gesamtdesign der Studie	20
3.2.2. Methodische Qualitätssicherung	21
4. Phase 1: Befragungsstudie	22
4.1. Fragebogendesign	22
4.2. Erhebungsverfahren	22
4.3. Stichprobenbeschreibung	23
4.4. Ergebnisse	23
4.4.1. Meeting-Last und -Struktur	24
4.4.2. Wahrgenommene Nützlichkeit	24
4.4.3. Auswirkungen auf Produktivität	26
4.4.4. Auswirkungen auf Flow	26
4.4.5. Auswirkungen auf Wohlbefinden und Arbeitszufriedenheit	27
4.4.6. Meeting-Vermeidungsverhalten	27
4.4.7. Reflexion der Meeting-Kultur und Akzeptanz von Feedback	28
4.5. Qualitative Ergänzungen	28
4.6. Zusammenfassung und Implikationen für Phase 2	29
5. Konzeptlösung	30
5.1. Zentrales Problemstatement	30
5.2. Teilprobleme	30
5.3. Stakeholder	31
5.4. Vision, Mission und Leitprinzipien	32
5.5. Produktziele	34
5.6. Einordnung und Traceability Matrix	34

5.7. Anforderungen	36
5.7.1. Funktionale Anforderungen	36
5.7.2. Nicht-funktionale Anforderungen	37
6. Systemarchitektur und Implementierung	39
6.1. Systemüberblick	39
6.1.1. Authentifizierung und Benutzerverwaltung	40
6.1.2. Onboarding und Guided Tour	41
6.2. Objektive Datengrundlage: Kalendersynchronisation	42
6.2.1. Anbindung externer Kalenderanbieter	42
6.2.2. Hintergrund-Synchronisation	43
6.2.3. Kalenderansicht und manuelle Verwaltung	43
6.3. KI-gestützte Kontextualisierung	44
6.3.1. Modellwahl und Datenschutz	44
6.3.2. Kategorisierungsstrategie	44
6.4. Erfassung der subjektiven Wahrnehmung	45
6.4.1. Push-Benachrichtigungen	45
6.4.2. Zwei komplementäre Feedback-Mechanismen	45
6.4.3. Feedback-Inbox als zentrale Sammelstelle	48
6.4.4. Flexible Datenspeicherung	48
6.5. Visualisierung und Reflexion: Dashboards	48
6.5.1. Structure Dashboard	49
6.5.2. Impact Dashboard	51
6.6. Technische Architektur und Betrieb	53
6.6.1. Containerisierung und Netzwerk	53
6.6.2. SSL und Verschlüsselung	53
6.6.3. Frontend	53
6.6.4. REST-API-Schnittstelle	53
6.6.5. CI/CD-Pipeline	54
6.7. Datenmodell	54
7. Validierung	55
7.1. Usability-Test	55
7.1.1. Zielsetzung	55
7.1.2. Methodik und Setup	55
7.1.3. Test-Szenarien und Aufgaben	56
7.1.4. Ergebnisse	57
7.1.5. Abgeleitete Massnahmen	58
7.2. Anforderungvalidierung	58
7.2.1. Funktionale Anforderungen	59
7.2.2. Nicht-funktionale Anforderungen	61
7.2.3. Erreichung der Produktziele	62
8. Phase 2: Feldstudie	63
8.1. Methodik	63
8.2. Stichprobe und Durchführung	63
8.3. Ergebnisse auf Meeting-Ebene	64
8.3.1. Meeting Struktur	64
8.3.2. Einfluss der Meeting-Dauer auf das Erleben	65

8.3.3. Einfluss der Tageszeit auf das Erleben	66
8.3.4. Meeting-Typ-Profile	66
8.3.5. Qualitative Kontextualisierung über Tags	67
8.4. Ergebnisse auf Tages-Ebene	68
8.4.1. Individuelle Profile	69
8.5. Zusammenfassung	69
9. Diskussion, Fazit und Ausblick	70
9.1. Einordnung der Ergebnisse	70
9.1.1. Integration dreier Forschungsstränge	70
9.1.2. Meeting-Level-Analyse mit objektiven Kalendermerkmalen	70
9.1.3. Wahrnehmungs-Realitäts-Diskrepanz	70
9.1.4. Rolle der subjektiven Wahrnehmung	70
9.2. Beantwortung der Forschungsfragen	71
9.2.1. FF1: Meeting-Praktiken und retrospektive Wahrnehmung	71
9.2.2. FF2: Einfluss von Meeting-Häufigkeit, -Dauer und -Timing	72
9.2.3. FF3: Wahrnehmung und Realität	74
9.3. Praktische Implikationen	75
9.4. Die Applikation als Reflexionsinstrument	75
9.5. Limitationen	76
9.5.1. Stichprobengrösse in Phase 2	76
9.5.2. Unvollständige Erfassung von Unterbrechungen	76
9.5.3. Single-Item-Messung der Outcome-Konstrukte	76
9.5.4. Selbstselektion und Stichprobenhomogenität	76
9.5.5. Beobachtungszeitraum	77
9.5.6. Retrospektive Verzerrung in Phase 1	77
9.6. Fazit	77
9.7. Ausblick	78
A Eigenständigkeitserklärung	79
B Hilfsmittelverzeichnis	80
C Glossar	81
D Variablenübersicht	84
E Vollständiger Fragebogen für die Umfrage in Phase 1	86
F Marktanalyse bestehender Tools	92
G Anhang: Wireframes	93
H Feedback Instrumente	95
H.1 Meeting Feedback Instrument	95
H.2 End of Day Feedback Instrument	97
I Meeting-Kategorisierung LLM Model vergleichs Test	98
I.1 Usability-Test Demodaten	99
J Einverständniserklärung Phase 2	102
K Externe Ressourcen	104
K.1 Quellcode der Applikation	104
K.2 Anonymisierte Rohdaten	104
Literaturverzeichnis	105

Abbildungsverzeichnis

Abbildung 1	Vorschau der entwickelten Applikation: Wochenansicht mit Kalenderintegration und Meeting-Bewertungen	9
Abbildung 2	Konzeptionelles Forschungsmodell mit vier Variablengruppen, drei Outcome-Konstrukten und vier Wirkpfaden.	16
Abbildung 3	Zusammenhang zwischen genereller Nützlichkeitsbewertung und Häufigkeit der Ineffizienzwahrnehmung.	24
Abbildung 4	Häufigste negative Einflussfaktoren auf die Meeting-Erfahrung (Mehrfachauswahl, N = 55).	25
Abbildung 5	Retrospektive Wahrnehmung der Meeting-Auswirkungen auf die drei Outcome-Konstrukte.	26
Abbildung 6	Architekturübersicht: Die sechs Docker-Services und ihre Kommunikationspfade.	40
Abbildung 7	Login- und Registrierungs-Ansicht der Applikation im Vergleich.	41
Abbildung 8	Die interaktive Guided Tour erklärt die Komponenten der Applikation schrittweise.	41
Abbildung 9	Die Tour führt den Nutzer durch die Grundkonfiguration der App.	42
Abbildung 10	Wochenansicht des Kalenders mit Terminen.	43
Abbildung 11	Meeting-Feedback-Formular mit ROTI, Stimmung, Energie und Tags.	46
Abbildung 12	End-of-Day-Feedback zur Erfassung der subjektiven Tagesbewertung.	47
Abbildung 13	Die Feedback-Inbox zeigt alle noch nicht bewerteten Meetings und Arbeitstage.	48
Abbildung 14	Formel welche zur berechnung des Flow Potential genutzt wurde.	49
Abbildung 15	Das Structure Dashboard visualisiert die objektiven Merkmale des Arbeitstages.	50
Abbildung 16	Das Impact Dashboard visualisiert die subjektiven Merkmale des Arbeitstages.	52
Abbildung 17	Das Datenmodel Unified Modeling Language (UML) Notation.	54
Abbildung 18	Subjektive Bewertungen nach Meeting-Dauer (N = 73)	65
Abbildung 19	Meeting-Typ-Profil mit drei Dimensionen	66
Abbildung 20	Häufigkeit der kontextabhängigen Tags	67
Abbildung 21	Wireframe der Dashboard-Ansicht.	93
Abbildung 22	Wireframe der Kalenderansicht.	94
Abbildung 23	Wireframe der Einstellungsansicht.	94

Tabellenverzeichnis

Tabelle 1	Zuordnung der subjektiven Variablen zu den Outcome-Konstrukten auf beiden Messebenen.	19
Tabelle 2	Vergleich der beiden Studienphasen im Überblick.	20
Tabelle 3	Traceability Matrix: Zuordnung von Teilproblemen, Produktzielen, Forschungsfragen und Beitragsebene.	35
Tabelle 4	Abgeleitete Massnahmen aus dem Usability-Test.	58
Tabelle 5	Validierung der funktionalen Anforderungen (FA-01 bis FA-09).	60
Tabelle 6	Validierung der nicht-funktionalen Anforderungen (NFA-01 bis NFA-07).	62
Tabelle 7	Charakteristika der Studienteilnehmenden	63
Tabelle 8	Spearman-Korrelationen zwischen Meeting-Last und Tages-Outcomes (N = 23)	68
Tabelle 9	Eingesetzte Hilfsmittel mit Verwendungszweck und betroffenen Stellen.	80
Tabelle 10	Übersicht der erhobenen Variablen mit Konstrukt, Datenquelle und Operationalisierung.	84
Tabelle 11	Vollständiger Fragebogen der Online-Umfrage (Phase 1).	86
Tabelle 12	Vergleich bestehender Meeting- und Produktivitätstools mit der Konzeptlösung.	92
Tabelle 13	Meeting Feedback Instrument: Hauptfragen (Q1-Q4)	95
Tabelle 14	Meeting Feedback Instrument: Follow-up-Fragen (Q5-Q6)	96
Tabelle 15	End of Day Feedback Instrument	97

1. Einleitung

Softwareentwicklung erfordert lange Phasen konzentrierter Denkarbeit. Gleichzeitig verbringen Beschäftigte in der Wissensarbeit durchschnittlich über fünf Stunden pro Woche in geplanten Meetings [1], und Softwareentwickelnde verwenden rund ein Viertel ihrer Arbeitszeit für kollaborative Aktivitäten wie Meetings, Gespräche und E-Mails [2]. Meetings sind dabei nicht per se schädlich: Sie ermöglichen Abstimmung, Entscheidungsfindung und Wissensaustausch. Doch sie fragmentieren den Arbeitstag und unterbrechen die tiefen Konzentrationsphasen, die für produktive Softwareentwicklung essenziell sind [2], [3]. Je mehr Meetings an einem Tag stattfinden, desto stärker steigen Erschöpfung und empfundene Arbeitsbelastung [4]. Zugleich folgt die Wirkung von Meetings einem paradoxen Muster: Sie unterstützen produktive Beiträge bis zu einer gewissen Schwelle, ab der sie Aufmerksamkeits- und Motivationsressourcen erschöpfen [1].

Trotz dieser Befunde fehlt in vielen Teams eine datenbasierte Grundlage, um Meetings differenziert zu bewerten. Entscheidungen über Häufigkeit, Dauer oder Timing beruhen häufig auf Gewohnheiten statt auf Evidenz [5]. Hinzu kommen versteckte Kosten: Nach einem Meeting vergehen durchschnittlich rund 20 Minuten, bis Wissensarbeitende wieder produktiv arbeiten können [6]. Dieser Erholungsbedarf wird bei der Planung selten berücksichtigt [6]. Zudem bilden retrospektive Berichte über Meetingqualität eher generelle Einstellungen ab als tatsächlich Erlebtes [7]. Ohne Echtzeiterhebung bleibt diese Diskrepanz zwischen wahrgenommener und tatsächlicher Wirkung unsichtbar.

Diese Problemstellung motiviert die vorliegende Arbeit. Sie untersucht, wie Meetings auf der Ebene einzelner Meetings, auf Tagesebene und im Hinblick auf die individuelle Gesamtwahrnehmung Produktivität, Flow und Wohlbefinden von Softwareentwickelnden beeinflussen.

1.1. Zielsetzung der Arbeit

Die Arbeit verfolgt drei Ziele. Erstens soll empirisch erfasst werden, welche Meeting-Praktiken unter Softwareentwickelnden verbreitet sind und wie diese die Auswirkungen von Meetings retrospektiv einschätzen. Zweitens wird analysiert, wie Meeting-Charakteristika (z.B. Häufigkeit, Dauer, Timing) mit Produktivität, Flow-Erleben und emotionalem Zustand zusammenhängen. Drittens wird untersucht, ob retrospektive Einschätzungen über die Wirkung von Meetings mit den situativ erhobenen Effekten im Arbeitsalltag übereinstimmen.

Zur Umsetzung dieser Ziele wird ein zweiphasiges Forschungsdesign eingesetzt: Eine Online-Befragung (Phase 1) erfasst retrospektive Einschätzungen einer breiten Stichprobe. Eine Feldstudie mit *Experience Sampling Method* (ESM) (Phase 2) erhebt situative Echtzeit-Bewertungen während des Arbeitsalltags. Die Kombination beider Zugänge ermöglicht sowohl eine breite Bestandsaufnahme als auch eine granulare Analyse auf Meeting- und Tagesebene.

1.2. Forschungsfragen

Die folgenden Forschungsfragen strukturieren die empirische Untersuchung. FF1 wird durch Phase 1 (Online-Befragung) adressiert, FF2 primär durch Phase 2 (Feldstudie mit ESM) und FF3 durch den Vergleich der Ergebnisse beider Phasen.

FF1: Welche Meeting-Praktiken sind unter Softwareentwickelnden verbreitet, und wie nehmen sie die Auswirkungen von Meetings auf ihre Produktivität, ihren Flow und ihr Wohlbefinden retrospektiv wahr?

FF2: Welchen Einfluss haben Meeting-Häufigkeit, -Dauer, und -Timing auf Softwareentwickelnde?

- **FF2a:** ... auf deren wahrgenommene Produktivität?
- **FF2b:** ... auf deren Wohlbefinden?
- **FF2c:** ... auf deren Flow?

FF3: Welche Muster zeigen sich in den Echtzeit-Bewertungen von Meetings durch Softwareentwickelnde, und inwiefern spiegeln diese die in der Befragung identifizierten Meeting-Eigenschaften und Problemfelder wider?

1.3. Wissenschaftlicher und praktischer Beitrag

Auf wissenschaftlicher Ebene verbindet die Arbeit Meetingsforschung, Produktivität und Wohlbefinden in der Softwareentwicklung. Diese drei Forschungsstränge werden in der bestehenden Literatur selten integriert betrachtet (vgl. Abschnitt 2). Das zweiphasige Design ermöglicht die Analyse von Meeting-Effekten auf mehreren Ebenen: auf Meeting-Ebene, auf Tagesebene und als Gesamtwahrnehmung.

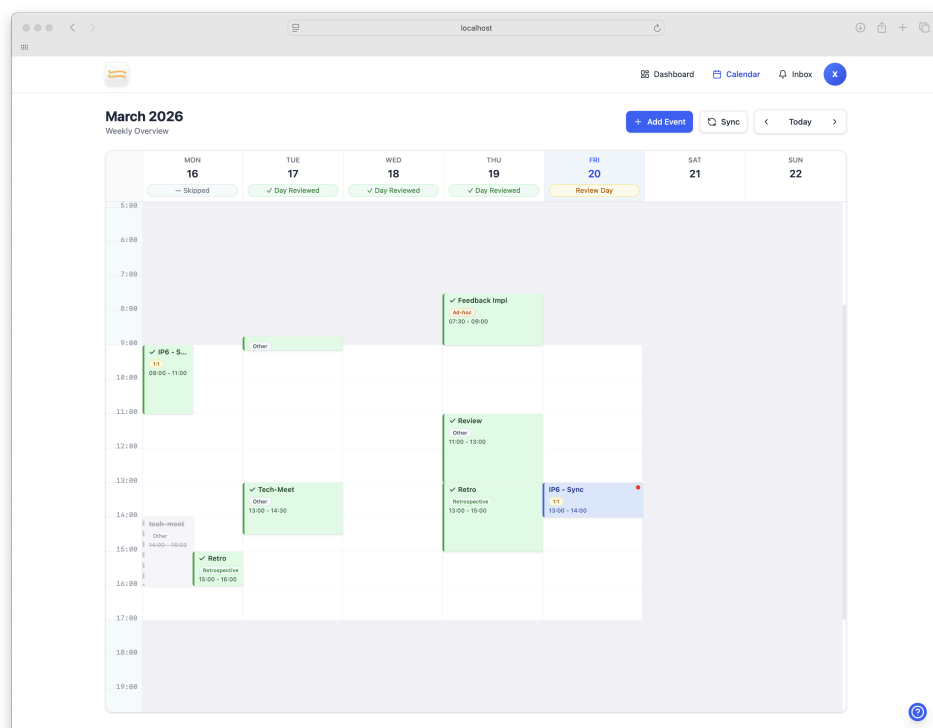


Abbildung 1: Vorschau der entwickelten Applikation: Wochenansicht mit Kalenderintegration und Meeting-Bewertungen

Auf praktischer Ebene wird eine Applikation entwickelt, die objektive Kalenderdaten mit strukturiertem subjektivem Feedback kombiniert. Sie ermöglicht Teams eine evidenzbasierte Reflexion ihrer Meeting-Kultur und adressiert damit den fehlenden Feedback-Mechanismus, den die Literatur als zentrales Hindernis für die Verbesserung von Meetings identifiziert [5]. Abbildung 1 zeigt eine Vorschau der entwickelten Applikation.

1.4. Aufbau der Arbeit

Abschnitt 2 verortet die Untersuchung in der bestehenden Literatur und leitet die Forschungslücken ab. Abschnitt 3 entwickelt darauf aufbauend das konzeptionelle Modell und beschreibt das zweiphasige Forschungsdesign. Abschnitt 4 dokumentiert die Online-Befragung mit Design, Stichprobe und Ergebnissen. Abschnitt 5 überführt die identifizierten Problemfelder in eine Konzeptlösung mit Vision, Produktzielen und Anforderungen. Abschnitt 6 beschreibt die technische Umsetzung der Applikation, die als Erhebungsinstrument für die Feldstudie dient. Abschnitt 7 dokumentiert den Usability-Test und die Anforderungvalidierung. Abschnitt 8 analysiert die Experience-Sampling-Daten auf Meeting- und Tagesebene. Abschnitt 9 ordnet die Ergebnisse in den Forschungskontext ein, beantwortet die Forschungsfragen, leitet praktische Implikationen ab und formuliert Anknüpfungspunkte für zukünftige Forschung.

2. State of the Art

Meetings gehören zu den allgegenwärtigsten Elementen moderner Wissensarbeit und zugleich zu den am kontroversesten diskutierten. Sie ermöglichen Koordination, Entscheidungsfindung und Wissensaustausch, fragmentieren jedoch den Arbeitstag und können konzentriertes Arbeiten sowie das Wohlbefinden beeinträchtigen. In der Softwareentwicklung, wo tiefe Konzentration und Flow-Zustände für die Problemlösung zentral sind, wiegen diese Effekte besonders schwer. Die vorliegende Arbeit verbindet drei Forschungsstränge, nämlich Meetingforschung, Produktivität und Wohlbefinden, die in der bestehenden Literatur selten integriert betrachtet werden. Dieses Kapitel ordnet die relevante Forschungslage ein, identifiziert Forschungslücken und begründet das methodische Vorgehen der Studie.

2.1. Meeting-Last, Produktivität und Wohlbefinden

Meetings zählen zu den zeitintensivsten wiederkehrenden Aktivitäten in der Wissensarbeit. Beschäftigte verbringen im Schnitt über fünf Stunden pro Woche in geplanten Meetings [1], wobei Softwareentwickelnde rund ein Viertel ihrer Arbeitszeit auf kollaborative Aktivitäten wie Meetings, informelle Gespräche und E-Mails verwenden [2]. Frühere empirische Arbeiten beschrieben einen nicht-linearen Zusammenhang zwischen Meeting-Belastung und Wohlbefinden: Meetings unterstützen produktive Beiträge bis zu einer gewissen Schwelle, ab der sie Aufmerksamkeits- und Motivationsressourcen erschöpfen, ein Muster, das als *Meeting Load Paradox* bezeichnet wird [1]. Ergänzende Tagebuchstudien verankern diesen Effekt auf der Tagesebene und zeigen, dass die tägliche Meeting-Anzahl Erschöpfung und subjektive Arbeitsbelastung vorhersagt [4].

Innerhalb der Softwareentwicklung hat sich gezeigt, dass Produktivitätswahrnehmungen mehrdimensional sind und Aufgabenabschluss, Flow-Zustand und Zufriedenheit umfassen [2], [3]. Meetings wurden dabei als signifikante Quelle sowohl produktiver als auch unproduktiver Zeit identifiziert, wobei ihr Wert davon abhängt, ob sie konkrete Ergebnisse liefern und die richtigen Personen involvieren [2]. Eine dreimonatige Langzeitstudie mit 25 Entwickelnden in fünf agilen Teams ergab, dass die in Meetings verbrachte Zeit kein signifikanter Prädiktor der wahrgenommenen Teamproduktivität war, während die wahrgenommene individuelle Produktivität den stärksten Prädiktor darstellte [8]. Dieser Befund deutet darauf hin, dass Meetings nicht pauschal produktivitätsmindernd wirken, sondern ihre Wirkung von kontextuellen Faktoren abhängt, ein Aspekt, den die vorliegende Studie durch die Analyse spezifischer Meeting-Eigenschaften adressiert.

2.1.1. Multidimensionale Produktivitätsmessung

Die Produktivitätsforschung in der Softwareentwicklung hat sich von einfachen Output-Metriken zu mehrdimensionalen Modellen entwickelt. Das SPACE-Framework systematisiert Produktivität in fünf Dimensionen: *Satisfaction and well-being*, *Performance*, *Activity*, *Communication and collaboration* sowie *Efficiency and flow* [9]. Das Framework empfiehlt, mindestens drei Dimensionen gleichzeitig zu erheben. Für die vorliegende Arbeit ist dies in zweierlei Hinsicht relevant: Es integriert Flow explizit als Produktivitätskomponente und erhebt Wohlbefinden als Produktivitätsdimension, was die Verknüpfung der drei Outcome-Konstrukte dieser Arbeit (Produktivität, Flow, Wohlbefinden) theoretisch unterstreicht. Ergänzend destilliert das Konzept der *Developer Experience* (DevEx) Produktivität auf drei Kernfaktoren: Feedback-Schleifen, kognitive Last und Flow-Zustand [10].

In einer Befragung von 413 professionellen Softwareentwickelnden bei Microsoft identifizierten Meyer et al. sechs distinkte Produktivitätstypen. Über die meisten Cluster hinweg bewerteten Entwickelnde Tage mit wenig oder keinen Meetings als produktiver und nannten die längste ununterbrochene Fokuszeit als besonders interessante Metrik zur Produktivitätsreflexion [3].

2.2. Unterbrechungen, Flow und Recovery

Produktive Softwareentwicklung erfordert zusammenhängende Konzentrationsphasen. Zwei Forschungsstränge sind dafür zentral: erstens die Flow-Forschung, die beschreibt, unter welchen Bedingungen vertiefte kognitive Arbeit entsteht [11], und zweitens die Unterbrechungsforschung, die zeigt, wie Kontextwechsel diesen Zustand gefährden und welche kognitiven Kosten dabei entstehen [12], [13].

2.2.1. Flow als Voraussetzung produktiver Softwareentwicklung

Csikszentmihalyis Flow-Theorie beschreibt einen Bewusstseinszustand völliger Absorption in einer Tätigkeit, der durch drei Bedingungen ermöglicht wird: klare Ziele, unmittelbares Feedback und ein Gleichgewicht zwischen Herausforderung und Fähigkeit [11]. Softwareentwicklung erfüllt diese Bedingungen in besonderer Weise, da Code kompiliert oder nicht, Tests bestehen oder scheitern, und die Aufgabenkomplexität dem Fähigkeitsniveau angepasst werden kann. Dies erklärt, warum Entwickelnde Flow-Zustände besonders häufig berichten und warum deren Unterbrechung als besonders kostspielig empfunden wird [2], [3].

2.2.2. Kognitive Kosten und Recovery

Die gravierendsten Auswirkungen von Meetings resultieren nicht aus deren Dauer allein, sondern aus den kognitiven Kosten des Kontextwechsels. Das Konzept der *Attention Residue* beschreibt, dass beim Wechsel zwischen Aufgaben ein Teil der kognitiven Aufmerksamkeit bei der vorherigen Aufgabe verbleibt und die Leistung bei der neuen Aufgabe beeinträchtigt [12]. Dieser Effekt ist besonders gravierend, wenn die vorherige Aufgabe unvollständig bleibt [12], eine Situation, die bei meetingbedingten Unterbrechungen regelmässig eintritt. Empirische Studien bestätigen, dass Unterbrechungen zwar die aufgabenbezogene Bearbeitungszeit verkürzen können, dies jedoch mit signifikant erhöhtem Stress und kognitiver Belastung einhergeht [13]. In der Softwareentwicklung sind die Wiederherstellungszeiten beträchtlich: Häufige Kontextwechsel und die resultierende Attention Residue führen dazu, dass die Rückkehr zu tiefem Fokus deutlich länger dauert als die eigentliche Unterbrechung [12], [13].

Allen et al. lieferten mit der *Meeting Recovery Scale* die erste empirische Operationalisierung des Aufwands für die Rückkehr zur produktiven Arbeit nach einem Meeting [6]. Meetingzufriedenheit und wahrgenommene Effektivität erwiesen sich als inverse Prädiktoren des Erholungsbedarfs, wobei die empfundene Relevanz des Meetings den Zusammenhang moderierte [6]. Die durchschnittliche *Meeting-to-Work Transition Time* betrug 20.22 Minuten (SD = 23.30), was die zeitliche Dimension des Post-Meeting-Erholungsbedarfs quantifiziert [6]. Positive affektive Zustände wurden zudem als förderlich für die Problemlösungsfähigkeit von Softwareentwickelnden nachgewiesen [14], was nahelegt, dass die emotionale Nachwirkung von Meetings Konsequenzen für die nachfolgende Arbeitsleistung haben kann. Zusammengenommen zeigen diese Befunde, dass Meetings Post-Meeting-Kosten verursachen, die über die im Meeting verbrachte Zeit hinausgehen. Keine der bestehenden Studien verbindet jedoch objektive Kalendermerkmale mit diesen Kosten innerhalb einer repräsentativen Arbeitswoche.

2.3. Daily Stand-ups in der Softwareentwicklung

Daily Stand-ups (DSMs) verdienen als häufigstes wiederkehrendes Meeting-Format in agilen Teams besondere Aufmerksamkeit. Eine Studie etablierte, dass DSMs Funktionen jenseits der drei Scrum-Fragen erfüllen, nämlich psychologische Sicherheit, frühe Problemerkennung und Teamkommunikation, wobei ihr Wert stark davon abhängt, ob sie selbstorganisiert oder von Vorgesetzten gesteuert werden [15]. Auf der Mikro-Interaktionsebene zeigte eine Transkriptanalyse von acht Daily Meetings, dass nur 24 % der Gesprächsinhalte die Scrum-Fragen adressierten, während problemorientierte Kommunikation mit 35 % dominierte [16].

Der wahrgenommene Wert von Stand-ups ist entlang von Seniorität und Teamgrösse polarisiert: Senior-Entwickelnde und Mitglieder grösserer Teams bewerten das Format besonders kritisch [17]. Gleichzeitig fördern Stand-ups Kommunikation und Problemlösung, überschreiten in der Praxis aber regelmässig ihre vorgesehene Dauer [18]. Neuere Forschung differenziert dieses Bild: Eine quantitative Studie lieferte den ersten empirischen Nachweis, dass DSMs die psychologische Sicherheit im Team positiv vorhersagen, welche wiederum als Mediator Verbesserungen bei arbeitsbezogener Zufriedenheit und wahrgenommener Teamleistung vermittelt [19]. Stand-ups erfüllen demnach unter geeigneten Bedingungen nicht nur koordinative, sondern auch sozial-psychologische Funktionen. Keine der bestehenden Studien untersucht jedoch, wie die Planungeigenschaften von DSMs (Timing, Dauer, Back-to-Back-Platzierung) das individuelle Erleben von Produktivität und Wohlbefinden innerhalb einer Arbeitswoche beeinflussen.

2.4. Meeting-Effektivität und Wohlbefinden

Ausserhalb des Software-Engineering-Kontexts hat die organisationspsychologische Forschung untersucht, welche Meeting-Design-Merkmale die wahrgenommene Effektivität vorhersagen. Neun Design-Charakteristika, darunter zeitliche, prozedurale und teilnehmerbezogene Dimensionen, wurden als Prädiktoren identifiziert [20]. Supervisor-Verhalten in Meetings beeinflusst die wahrgenommene organisationale Unterstützung. In diesem Zusammenhang wurden *Meeting Citizenship Behaviors* eingeführt: freiwillige Beiträge von Teilnehmenden, die über die normale Meetingteilnahme hinausgehen, etwa aktives Einbringen von Ideen oder das Fördern der Partizipation anderer [21]. Authentisches Engagement fördert die wahrgenommene Meetingeffektivität, während *Surface Acting*, also das Vortäuschen von Emotionen entgegen dem eigenen Empfinden, negativ mit Effektivität und Wohlbefinden assoziiert ist [22].

In der agilen Praxis hat sich das Instrument ROTI (Return on Time Invested) etabliert, das den wahrgenommenen Wert eines Meetings im Verhältnis zur investierten Zeit auf einer fünfstufigen Skala erfasst [23]. Ursprünglich als Closing-Aktivität für Retrospektiven konzipiert, wird ROTI in der Praxis breit für die Evaluation beliebiger Meeting-Formate eingesetzt. In der vorliegenden Studie dient es als quantitativer Kern der Meeting-Level-Bewertung.

Die Verbindung zwischen Wohlbefinden und Produktivität ist konsistent belegt: Entwickelnde mit positiveren affektiven Zuständen schneiden bei analytischen Problemlösungsaufgaben signifikant besser ab [14], und eine breite Befragung identifizierte zahlreiche Faktoren für Unzufriedenheit, wobei der grösste wahrgenommene Einfluss auf Produktivität und kognitive Leistung entfiel [24]. Theoretisch lassen sich diese Zusammenhänge über zwei Rahmenwerke verankern: Das *Job Demands-Resources*-Modell (JD-R) beschreibt, wie Arbeitsanforderungen bei chronischer Überlastung zu Burnout führen können [25], und die *Affective Events Theory* (AET) postuliert, dass Arbeitsereignisse wie Meetings als bidirektionale affektive Auslöser wirken, die sowohl positive als auch negative Emotionen hervorrufen [26].

2.5. Wahrnehmungs-Realitäts-Diskrepanz und Experience Sampling

Ein besonders relevanter Befund für die vorliegende Arbeit ist die Diskrepanz zwischen generellen Überzeugungen über Meetings und den tatsächlich erlebten Effekten. Öffentliche Aussagen über Meetings fallen systematisch negativer aus als private, anonyme Bewertungen [5]. Gleichzeitig bewerten Meeting-Organisierende den Erfolg ihrer eigenen Meetings systematisch höher als die Teilnehmenden, ein Phänomen, das als *Organizer Blindspot* beschrieben wird [5]. Da in den meisten Organisationen eine systematische, datengetriebene Feedback-Schleife für Meetings fehlt, perpetuieren sich ineffiziente Formate.

Diese Diskrepanz wird durch methodische Erkenntnisse zur retrospektiven Urteilsbildung verschärft. Retrospektive Berichte über Meetingqualität tendieren dazu, generelle Einstellungen statt tatsächlicher Erlebnisse abzubilden [7]. Die *Experience Sampling Method* (ESM) adressiert dieses Problem, indem Erfahrungen zeitnah zum Erleben erfasst werden und retrospektive Verzerrungen reduziert werden [27]. ESM gilt als methodischer Standard für die Erfassung dynamischer Prozesse in der Organisationsforschung [7].

In der Softwareentwicklungsforschung wurde ESM massgeblich durch Meyer und Fritz vorangetrieben: In einer Pionierstudie kombinierten sie Produktivitätsselbsteinschätzungen alle 90 Minuten mit automatisiertem Computer-Monitoring [2]. Dieser Ansatz ermöglichte erstmals die Korrelation von momentaner Produktivitätswahrnehmung mit objektiven Aktivitätsdaten und bildet ein methodisches Vorbild für das Design der vorliegenden Studie. Die Validität und Reliabilität der ESM-Methode wurde bereits in den 1980er-Jahren durch Csikszentmihalyi und Larson nachgewiesen [28] und ist seither in der psychologischen Forschung breit etabliert.

2.6. Forschungslücken und Einordnung der vorliegenden Arbeit

Die Analyse der bestehenden Literatur offenbart drei zentrale Forschungslücken, die die vorliegende Arbeit adressiert.

Fehlende ganzheitliche Integration der Forschungsstränge. Es existiert umfangreiche Forschung sowohl zu Meetingeffektivität und Wohlbefinden aus der Organisationspsychologie [1], [4], [6] als auch zu Unterbrechungen und Produktivität aus dem Software Engineering [2], [29]. Diese Stränge werden jedoch selten zusammengeführt: Die Software-Engineering-Forschung fokussiert hauptsächlich auf allgemeine Unterbrechungen, nicht spezifisch auf Meetings als deren häufigste strukturelle Quelle, während die Meetingstudien sich selten auf die spezifischen Anforderungen von Softwareentwickelnden beziehen. Die vorliegende Arbeit adressiert diese Lücke durch die simultane Untersuchung von Produktivität, Flow und Wohlbefinden im Kontext von Meetings bei Softwareentwickelnden (FF1, FF2).

Fehlende Meeting-Level-Analyse mit objektiven Kalendermerkmalen. Die bestehende Forschung operiert überwiegend auf der Ebene der wöchentlichen Meetingbelastung [1] oder der täglichen Meetinganzahl [4]. Es fehlt an Studien, die einzelne Meetings als Analyseeinheit verwenden und deren spezifische objektive Merkmale, insbesondere Timing und Dauer, systematisch mit subjektiven Bewertungen und Tagesergebnissen verknüpfen [2], [6], [29]. Die vorliegende Arbeit adressiert diese Lücke, indem echtzeitnahe subjektive Meetingbewertungen mit automatisch erhobenen Kalenderdaten auf Meeting- und Tagesebene kombiniert werden (FF2).

Wahrnehmungs-Realitäts-Diskrepanz nicht systematisch untersucht. Obwohl die Diskrepanz zwischen retrospektiven und echtzeitnahen Bewertungen methodisch gut dokumentiert ist [7], wurde sie für Meetings bei Softwareentwickelnden bisher nicht systematisch untersucht. Das zweiphasige Design der vorliegenden Studie, eine retrospektive Online-Befragung (Phase 1) und eine Feldstudie mit ESM (Phase 2), wurde gezielt entwickelt, um innerhalb derselben Personen Survey-Einschätzungen mit Feld-Erfahrungen zu vergleichen (FF3).

Zusammenfassend verbindet die vorliegende Arbeit drei etablierte, aber bisher getrennte Forschungstraditionen und nutzt ein methodisch etabliertes Design, um die Auswirkungen von Meetings auf Softwareentwickelnde granularer und valider zu erfassen, als dies bisherige Studien erlaubt haben.

3. Forschungsmodell und Forschungsdesign

Dieses Kapitel entwickelt das konzeptionelle Forschungsmodell, das die Forschungsfragen FF1, FF2 und FF3 operationalisierbar macht. Zunächst wird das Gesamtmodell mit seinen vier Variablengruppen und drei Outcome-Konstrukten vorgestellt. Anschliessend werden die einzelnen Bestandteile und deren Zusammenhänge im Detail beschrieben.

3.1. Konzeptionelles Forschungsmodell

Das Modell beschreibt, wie Meeting-Eigenschaften mit den drei Outcome-Konstrukten *Produktivität*, *Flow* und *Wohlbefinden* zusammenhängen. Die zentrale Annahme ist, dass objektive Meeting-Merkmale nicht direkt auf diese Outcomes wirken. Stattdessen wird ihre Wirkung durch zwei Mechanismen vermittelt: durch die subjektive Bewertung einzelner Meetings und durch die resultierende Struktur des Arbeitstages. Abbildung 2 zeigt das Gesamtmodell im Überblick.

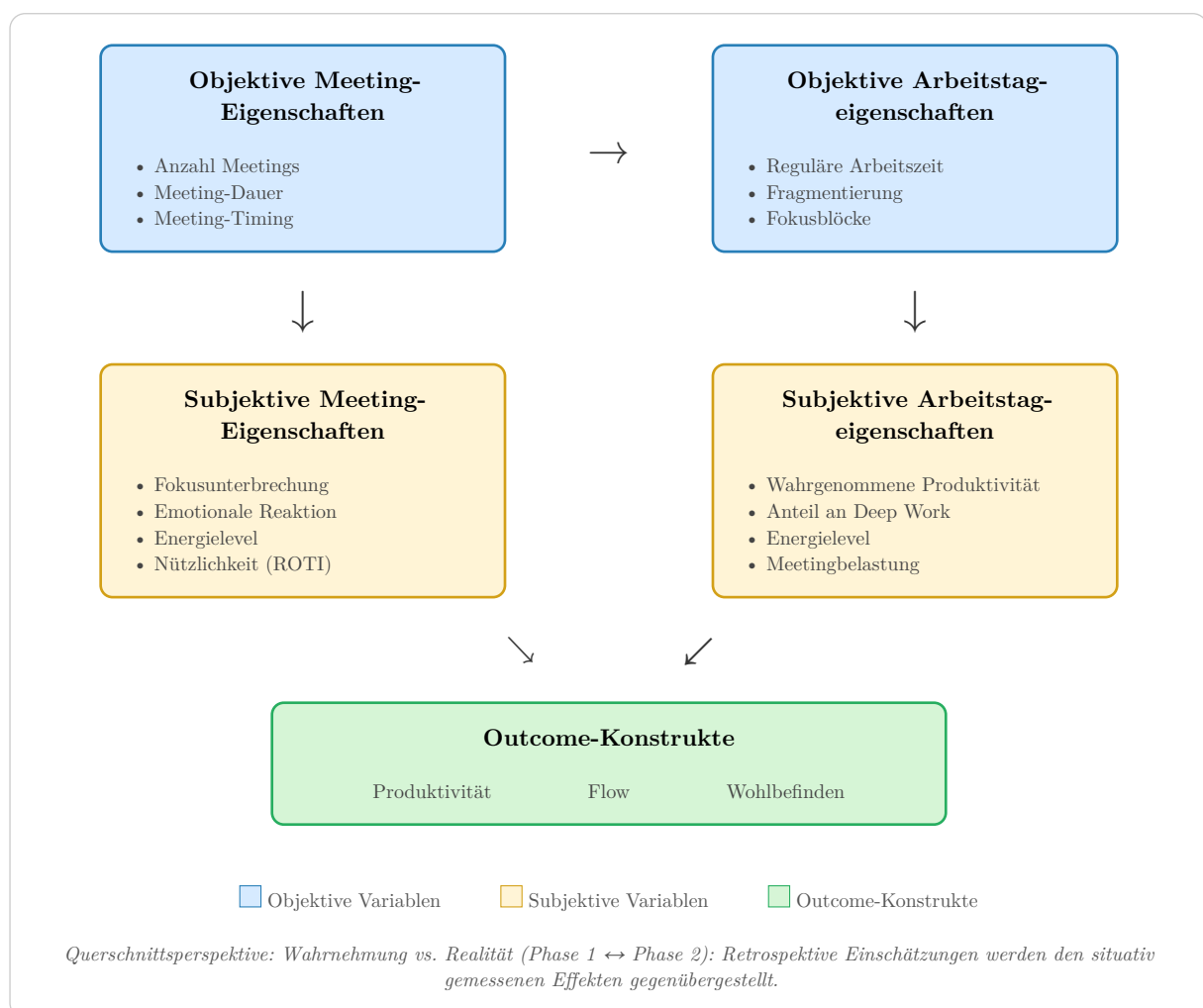


Abbildung 2: Konzeptionelles Forschungsmodell mit vier Variablengruppen, drei Outcome-Konstrukten und vier Wirkpfaden.

Das Diagramm zeigt vier Variablengruppen und drei Outcome-Konstrukte, verbunden durch vier Wirkpfade. Die obere Ebene (blau) enthält objektiv messbare Grössen: links die Eigenschaften einzelner Meetings, rechts die daraus resultierende Tagesstruktur. Die mittlere Ebene

(gelb) bildet die subjektive Wahrnehmung ab: links die Bewertung einzelner Meetings, rechts die Einschätzung des gesamten Arbeitstages. Beide subjektiven Gruppen münden in die drei Outcome-Konstrukte (grün). In den folgenden Abschnitten werden zunächst die Outcome-Konstrukte definiert, dann die vier Variablengruppen beschrieben und schliesslich die Wirkpfade erläutert, die sie verbinden.

3.1.1. Outcome-Konstrukte: Produktivität, Flow und Wohlbefinden

Die Forschungsfragen FF2a, FF2b und FF2c adressieren drei Outcome-Konstrukte. Diese beschreiben, was durch Meetings beeinflusst wird. Die Frage, wie sie konkret gemessen werden, beantwortet die Operationalisierung in den nachfolgenden Abschnitten.

Produktivität beschreibt die subjektiv wahrgenommene Fähigkeit, geplante Aufgaben effektiv umzusetzen und messbaren Fortschritt zu erzielen. Im SPACE-Framework [9] entspricht dies einer Kombination aus Performance (Aufgabenerledigung) und Satisfaction (Zufriedenheit mit dem Arbeitsergebnis). Produktivität ist eine Ergebniswahrnehmung: Wurde etwas geschafft?

Flow (Deep Focus) bezeichnet einen Zustand anhaltender, ununterbrochener Konzentration bei der Bearbeitung anspruchsvoller Aufgaben. Flow erfordert klare Ziele, unmittelbares Feedback und ein Gleichgewicht zwischen Herausforderung und Fähigkeit [11]. Im Vordergrund steht die Qualität der Arbeitsausführung: die Tiefe und Kontinuität der Konzentration. Entwickelnde können sich im Flow befinden, ohne am Tagesende messbaren Fortschritt zu erleben, und umgekehrt. Flow und Produktivität sind damit verwandt, aber trennbar.

Wohlbefinden umfasst das subjektive Wohlbefinden im Arbeitskontext. Im Mittelpunkt steht das Kontinuum zwischen Erschöpfung und Energie. Gemäss der AET wirken Meetings als affektive Ereignisse, die sowohl positive Emotionen (Motivation, Zielklärung) als auch negative Emotionen (Frustration, Kontrollverlust) auslösen können [26]. Wohlbefinden bezieht sich auf das Wie des Erlebens: Wie fühlt sich eine Person am Ende des Tages?

Die drei Konstrukte sind konzeptionell verwandt, aber analytisch trennbar: Produktivität beschreibt das wahrgenommene Arbeitsergebnis, Flow die Qualität des Arbeitsprozesses und Wohlbefinden den emotionalen und energetischen Zustand. Die beiden Studienphasen erfassen diese Konstrukte mit unterschiedlicher Messlogik. Phase 1 erhebt retrospektive Einschätzungen darüber, wie Meetings Produktivität, Flow und Wohlbefinden im Allgemeinen beeinflussen. Phase 2 differenziert die Messung auf zwei Ebenen: auf der Meeting-Ebene (situative Wirkung einzelner Meetings) und auf der Tagesebene (aggregierte Tagesbewertung). Genau dieser Unterschied zwischen retrospektiver Wahrnehmung und situativer Messung ermöglicht den in FF3 formulierten Vergleich.

3.1.2. Vom Meeting zum Tageserleben: Variablengruppen und Wirkpfade

Dieser Abschnitt beschreibt die vier Variablengruppen des Modells und die vier Wirkpfade, die sie mit den Outcome-Konstrukten verbinden. Die Darstellung folgt der Logik des Diagramms (Abbildung 2): zunächst die vier Gruppen von oben nach unten, dann die Wirkpfade als Verbindungen.

Objektive Meeting-Eigenschaften bilden den Ausgangspunkt des Modells (oben links im Diagramm). Sie umfassen kalenderbasierte Kenngrößen, die ohne subjektive Bewertung aus automatisch erhobenen Kalenderdaten abgeleitet werden: die Anzahl der Meetings pro Tag, die Dauer einzelner Meetings sowie deren zeitliche Platzierung im Tagesverlauf. Diese Variablen beschreiben, was an einem Arbeitstag strukturell stattgefunden hat.

Objektive Arbeitstageigenschaften beschreiben die strukturellen Merkmale des Arbeitstages, die sich aus der Kombination der Meeting-Eigenschaften mit den individuellen Arbeitszeiten ergeben (oben rechts im Diagramm). Dazu zählen die reguläre Arbeitszeit (Arbeitsbeginn, Arbeitsende, Arbeitstage) sowie die daraus abgeleiteten Metriken Fragmentierung (wie stark der Tag durch Meetings zerstückelt wird) und Fokusblöcke (zusammenhängende Zeiträume ohne Meetings). Während die objektiven Meeting-Eigenschaften einzelne Meetings beschreiben, bilden die objektiven Arbeitstageigenschaften deren kumulativen Effekt auf die Tagesstruktur ab.

Subjektive Meeting-Eigenschaften erfassen die individuelle Bewertung einzelner Meetings unmittelbar nach Meetingende (Mitte links im Diagramm). Sie umfassen vier Dimensionen: die wahrgenommene Fokusunterbrechung, die emotionale Reaktion, das Energielevel nach dem Meeting sowie die wahrgenommene Nützlichkeit, operationalisiert über den Return on Time Invested (ROTI). Diese Variablen bilden das subjektive Erleben auf der Mikro-Ebene einzelner Meetings ab.

Subjektive Arbeitstageigenschaften erfassen die individuelle Bewertung des gesamten Arbeitstages über eine tägliche Selbsteinschätzung am Tagesende (Mitte rechts im Diagramm). Sie umfassen vier Dimensionen: die wahrgenommene Produktivität, den Anteil an Deep Work (konzentrierte Tiefenarbeit), das Energielevel sowie die empfundene Meetingbelastung. Diese Variablen repräsentieren die aggregierte Tageswahrnehmung.

Beide subjektiven Variablengruppen münden in die drei Outcome-Konstrukte, die in Abschnitt 3.1.1 definiert wurden. Aufgrund der ESM-Anforderung an minimale Teilnehmerbelastung wird jedes Konstrukt durch einzelne Indikatoren gemessen.

Die Pfeile in Abbildung 2 repräsentieren vier postulierte Wirkpfade, die beschreiben, wie die Variablengruppen zusammenhängen.

Pfad 1: Objektive Meeting-Eigenschaften → Objektive Arbeitstageigenschaften

Die Anzahl, Dauer und zeitliche Platzierung der Meetings bestimmen, wie stark der Arbeitstag fragmentiert wird und wie viele zusammenhängende Fokusblöcke verbleiben. Ein Tag mit fünf kurzen, über den Tag verteilten Meetings erzeugt eine andere Tagesstruktur als ein Tag mit einem einzigen langen Meeting am Vormittag.

Pfad 2: Objektive Meeting-Eigenschaften → Subjektive Meeting-Eigenschaften

Die objektiven Merkmale eines Meetings beeinflussen, wie es subjektiv erlebt wird. Ein langes Meeting kann als energieraubend empfunden werden. Die zeitliche Platzierung beeinflusst, wie stark eine laufende Arbeitsphase betroffen ist.

Pfad 3: Objektive Arbeitstageigenschaften → Subjektive Arbeitstageigenschaften

Die Tagesstruktur beeinflusst die Tageswahrnehmung. Ein stark fragmentierter Tag mit wenigen Fokusblöcken wird tendenziell als weniger produktiv und erschöpfender erlebt als ein Tag mit langen, ununterbrochenen Arbeitsphasen.

Pfad 4: Subjektive Variablen → Outcome-Konstrukte

Die situativen Meeting-Bewertungen erfassen die Wirkung auf der Mikro-Ebene (einzelnes Meeting), die Tageseinschätzungen auf der Makro-Ebene (gesamter Tag). Gemeinsam ergeben sie ein differenziertes Bild der Meeting-Auswirkungen auf Produktivität, Flow und Wohlbefinden.

Die folgende Zuordnung zeigt, welche Variablen welches Konstrukt operationalisieren: Tabelle 1 zeigt, welche Variablen welches Konstrukt operationalisieren.

Konstrukt	Meeting-Ebene (Subj. Meeting-Eigenschaften)	Tages-Ebene (Subj. Arbeitstageigenschaften)
Produktivität	Nützlichkeit (ROTI)	Wahrgenommene Produktivität
Flow	Fokusunterbrechung	Anteil Deep Work
Wohlbefinden	Emotionale Reaktion, Energielevel nach Meeting	Energielevel (Tagesende), Empfundene Meetingbelastung

Tabelle 1: Zuordnung der subjektiven Variablen zu den Outcome-Konstrukten auf beiden Messebenen.

Die vollständige Variablenübersicht mit Datenquelle, Operationalisierung und Skalenart ist in Anhang D dokumentiert. Die konkreten Erhebungsinstrumente, Fragebogenitems und Skalenbezeichnungen werden in Abschnitt 4 (für Phase 1) und Abschnitt 6.4 (für Phase 2) im Detail beschrieben.

3.1.3. Querschnittsperspektive: Wahrnehmung vs. Realität

Neben den vier Wirkpfaden verankert das Modell eine methodische Querschnittsperspektive, die FF3 operationalisiert. Retrospektive Einschätzungen aus Phase 1 werden den situativ gemessenen Effekten aus Phase 2 gegenübergestellt. Dieser Vergleich durchzieht das gesamte Modell: Er betrifft sowohl die Bewertung einzelner Meetings als auch die Einschätzung ganzer Arbeitstage. Die Querschnittsperspektive ist im Diagramm (Abbildung 2) als Annotation dargestellt. Sie kontextualisiert die Befunde zu FF1 und FF2, inwiefern retrospektive Einschätzungen mit konkreten Echtzeit-Erfahrungen übereinstimmen.

3.2. Forschungsdesign

Dieses Kapitel beschreibt das methodische Vorgehen der Studie im Überblick. Die detaillierte Beschreibung der beiden Studienphasen folgt in Abschnitt 4 und Abschnitt 8; hier werden das Gesamtdesign und die Massnahmen zur methodischen Qualitätssicherung eingeführt.

3.2.1. Gesamtdesign der Studie

Die Studie verfolgt ein zweiphasiges Mixed-Methods-Design, das retrospektive Selbstberichte mit Echtzeit-Erhebungen kombiniert. Phase 1 liefert eine breite Baseline über Meeting-Praktiken und deren wahrgenommene Auswirkungen und adressiert damit FF1 sowie FF2. Phase 2 ergänzt diese retrospektiven Einschätzungen durch Echtzeit erhobene Daten und ermöglicht so den in FF3 formulierten Abgleich. Tabelle 2 fasst die beiden Phasen im Vergleich zusammen.

	Phase 1: Online-Befragung	Phase 2: Feldstudie (ESM)
Ziel	Quantitative Erfassung von Meeting-Praktiken und retrospektive Wahrnehmung der Auswirkungen	Echtzeit-Messung situativer Meeting-Effekte und Validierung der Baseline-Ergebnisse
Methode	Strukturierter Online-Fragebogen (35 Items)	Experience Sampling: situative Bewertung nach Meetings + tägliche Selbsteinschätzung + automatisch erhobene Kalenderdaten
Stichprobe	N = 55 Softwareentwickelnde	N = 3 Teilnehmende (Subset aus Phase 1)
Dauer	Einmalige Erhebung (Feb. 2026)	2 Wochen (2.–13. März 2026)
Datenquellen	Retrospektive Selbstberichte (Likert, Multiple Choice)	Objektive Kalenderdaten + subjektive Echtzeit-Bewertungen
Primäre FF	FF1, FF2 (Baseline)	FF2 (Vertiefung), FF3

Tabelle 2: Vergleich der beiden Studienphasen im Überblick.

Die detaillierte Beschreibung beider Phasen, einschliesslich Fragebogendesign, Stichprobenrekrutierung und Erhebungsinstrumenten, folgt in Abschnitt 4 (Phase 1) und Abschnitt 8 (Phase 2).

3.2.2. Methodische Qualitätssicherung

Die folgenden Abschnitte beschreiben, wie das Studiendesign Validität, Reliabilität und Bias-Reduktion adressiert, und dokumentieren die ethischen Rahmenbedingungen der Erhebung.

Validität

Die zweiphasige Anlage stärkt die interne Validität, indem retrospektive Einschätzungen (Phase 1) durch Echtzeit-Bewertungen (Phase 2) trianguliert werden. Experience Sampling reduziert retrospektive Verzerrungen [7], und die Kombination von objektiven Kalenderdaten mit subjektiven Bewertungen ermöglicht es, wahrgenommene Effekte durch strukturelle Merkmale zu kontextualisieren. Die drei Outcome-Konstrukte sind theoretisch in etablierten Rahmenwerken verankert (SPACE-Framework [9], Flow-Theorie [11], AET [26]), was die Konstruktvalidität stützt. Da jedes Konstrukt durch einzelne Indikatoren operationalisiert wird, ist diese jedoch eingeschränkt. Die Phase-1 Stichprobe umfasst Softwareentwickelnde mit unterschiedlicher Berufserfahrung, Teamgrösse und Arbeitsarrangement, was die externe Validität unterstützt. Die Generalisierbarkeit ist durch die Stichprobengrösse ($N = 55$ in Phase 1, $N = 3$ in Phase 2) und die überwiegend deutschsprachige Stichprobe eingeschränkt. Beide Einschränkungen werden im Limitationskapitel aufgegriffen.

Reliabilität

Der Fragebogen in Phase 1 verwendet standardisierte Likert-Skalen mit konsistenten Ankerbezeichnungen. Die Erhebungsinstrumente in Phase 2 verwenden identische Skalen über alle Erhebungszeitpunkte hinweg, was die Vergleichbarkeit über Tage und Personen sicherstellt. Die automatische Kalender-Synchronisation eliminiert Messfehler bei den objektiven Metriken.

Bias-Reduktion

Die Phase-1 Befragung wurde anonym durchgeführt, um soziale Erwünschtheit zu reduzieren. Die ESM-basierte Erhebung in Phase 2 minimiert den *Recall Bias* [7], da Bewertungen zeitnah zum Erleben erfolgen. Die Kombination von subjektiven und objektiven Datenquellen ermöglicht zudem eine Plausibilitätsprüfung der Selbstberichte. Die Erhebungsinstrumente wurden bewusst kurz gehalten (vier Items pro Meeting, vier Items pro Tag), um Ermüdungseffekte zu minimieren.

Ethische Aspekte

Die Studie wurde unter Einhaltung forschungsethischer Grundsätze durchgeführt. In beiden Phasen wurde eine informierte Einwilligung eingeholt; die Teilnahme war freiwillig und konnte jederzeit beendet werden. In Phase 2 verarbeitet die Applikation sensible Kalenderdaten, deren Zugriff strikt auf den jeweiligen Nutzenden beschränkt ist. Meeting-Titel und -Inhalte werden ausschliesslich für die Kategorisierung und Anzeige verwendet und nicht an Dritte weitergegeben. Die Einverständniserklärung für Phase 2 ist in Anhang J dokumentiert.

4. Phase 1: Befragungsstudie

Die Online-Befragung bildet die erste Phase des zweistufigen Forschungsdesigns und adressiert FF1: Welche Meeting-Praktiken sind verbreitet, und wie nehmen Softwareentwickelnde deren Auswirkungen auf Produktivität, Flow und Wohlbefinden wahr? Dazu erfasst die Befragung einerseits die Meeting-Landschaft (Häufigkeit, Dauer, Formate, wahrgenommene Nützlichkeit) und andererseits die retrospektive Wahrnehmung der Auswirkungen auf die drei Outcome-Konstrukte. Diese retrospektiven Einschätzungen dienen zugleich als Vergleichsgrundlage für den in FF3 formulierten Abgleich mit den Echtzeit-Daten der Feldstudie (Phase 2). Ergänzend identifiziert die Befragung problematische Muster in der Meeting-Kultur, etwa bezüglich Timing, wahrgenommener Ineffizienz und Vermeidungsverhalten.

4.1. Fragebogendesign

Der Fragebogen wurde als strukturierter Online-Fragebogen in englischer Sprache konzipiert und über Google Forms implementiert. Er umfasst thematische Sektionen mit insgesamt 35 inhaltlichen Items sowie zwei vorgelagerten Filterfragen. Der vollständige Fragebogen ist in Anhang E dokumentiert.

Die inhaltlichen Items decken folgende Themenbereiche ab: Meeting-Last und -Häufigkeit, Meeting-Charakteristika und -Formate, wahrgenommene Nützlichkeit und Ineffizienz, Auswirkungen auf Produktivität und Flow, Einfluss auf Wohlbefinden und Arbeitszufriedenheit, Meeting-Vermeidungsverhalten, Reflexion der Meeting-Kultur sowie Akzeptanz von Feedback-Mechanismen. Ergänzt werden diese durch demografische Fragen zu Rolle, Berufserfahrung, Organisationskontext und technischem Arbeitsanteil. Die Items wurden systematisch den drei Outcome-Konstrukten zugeordnet.

Der Fragebogen verwendet drei Antwortformate. Likert-Skalen mit fünf Stufen kommen für die Mehrzahl der Items zum Einsatz, wobei die Ankerbezeichnungen je nach Konstrukt variieren. Multiple-Choice-Fragen werden für nominale Variablen wie Rolle, Arbeitsarrangement und Meeting-Formate eingesetzt, teilweise mit Mehrfachauswahl. Offene Textfelder ergänzen den Fragebogen für qualitative Kommentare und Kontaktdaten.

4.2. Erhebungsverfahren

Die Befragung wurde über zwei Distributionskanäle verteilt. Der erste Kanal war eine öffentlich zugängliche Online-Befragung, die über professionelle Netzwerke, Social-Media-Plattformen und persönliche Kontakte verbreitet wurde. Der zweite Kanal waren Softwareentwicklungsteams in den Organisationen der Autoren, die gezielt zur Teilnahme eingeladen wurden. Beide Varianten verwendeten den identischen Kernfragebogen. Die Erhebung erfolgte im Zeitraum vom 1. Februar 2026 bis zum 28. Februar 2026. Die Teilnahme war freiwillig und anonym.

Zu Beginn des Fragebogens wurden zwei Filterfragen eingesetzt. Die erste Frage prüfte, ob die Person in den letzten vier Wochen an arbeitsbezogenen Meetings teilgenommen hatte. Die zweite Frage erfasste die informierte Einwilligung. Nur Personen, die beide Fragen bejahten, wurden zum Hauptteil des Fragebogens weitergeleitet. Von insgesamt 56 Einreichungen wurde ein Teilnehmender durch diesen Filter ausgeschlossen.

4.3. Stichprobenbeschreibung

Insgesamt wurden $N = 55$ gültige Antworten erhoben. Die Stichprobe zeigt eine ausgewogene Altersverteilung: 22 Befragte (40,0 %) gehören der Altersgruppe 18 bis 29 an, 22 (40,0 %) der Gruppe 30 bis 44, 8 (14,5 %) der Gruppe 45 bis 60 und 2 (3,6 %) sind über 60 Jahre alt. Eine Person machte keine Angabe. Die Geschlechterverteilung ist mit 45 männlichen (81,8 %) und 9 weiblichen (16,4 %) Teilnehmenden männlich dominiert.

Die Berufserfahrung verteilt sich über alle Karrierestufen: 17 Teilnehmende (30,9 %) haben weniger als 5 Jahre Erfahrung, 15 (27,3 %) verfügen über 5 bis 10 Jahre, 10 (18,2 %) über 11 bis 20 Jahre und 13 (23,6 %) über mehr als 20 Jahre. Hinsichtlich der Rollen sind Software Engineers und Developers mit 40 Personen (72,7 %) die grösste Gruppe, gefolgt von Team Leads und Engineering Managers mit 8 Personen (14,5 %) sowie Product Owners und Product Managers mit 3 Personen (5,5 %). Die verbleibenden 4 Personen verteilen sich auf QA Engineering, Testautomation, Administration und Forschung.

Der Grossteil (40 Personen, 72,7 %) arbeitet in IT-Abteilungen innerhalb von Nicht-IT-Organisationen. Die übrigen verteilen sich auf Softwareunternehmen und IT-Dienstleister (5, 9,1 %), Forschung und Bildung (5, 9,1 %), öffentliche Organisationen (4, 7,3 %) und ein Startup (1, 1,8 %). Hinsichtlich der Unternehmensgrösse dominieren Grossunternehmen mit 250 oder mehr Mitarbeitenden deutlich: 48 Befragte (87,3 %) arbeiten in solchen Organisationen, während 3 (5,5 %) in mittleren Unternehmen (50 bis 249 Mitarbeitende) und 4 (7,3 %) in Kleinunternehmen (10 bis 49 Mitarbeitende) tätig sind. Diese Konzentration auf grosse IT-Abteilungen innerhalb etablierter Organisationen ist bei der Interpretation der Ergebnisse zu berücksichtigen.

Der Anteil an technischer Arbeit variiert erwartungsgemäss mit der Rolle: 22 Befragte (40,0 %) verbringen 81 bis 100 % ihrer Arbeitszeit mit technischer Arbeit, 17 (30,9 %) verbringen 61 bis 80 %, und die restlichen 16 Befragten (29,1 %) liegen unter 60 %, was den Einschluss von Personen mit hybridem Rollenprofil (z. B. Team Leads) reflektiert.

Die Frage zum Arbeitsstandort (D2, vgl. Anhang E) wurde erst während der Erhebungsphase ergänzt und liegt daher nur für einen Teil der Stichprobe vor. Da die Rekrutierung primär über professionelle Netzwerke in der Deutschschweiz erfolgte, ist eine geografische Konzentration anzunehmen.

4.4. Ergebnisse

Die Ergebnisdarstellung folgt der in FF1 formulierten Struktur: Zunächst werden die verbreiteten Meeting-Praktiken beschrieben, anschliessend die retrospektiven Wahrnehmungen zu den drei Outcome-Konstrukten Produktivität, Flow und Wohlbefinden. Ergänzend werden Befunde zu Meeting-Vermeidungsverhalten, Reflexion der Meeting-Kultur und Akzeptanz von Feedback-Mechanismen berichtet, die zusätzlichen Kontext für die Interpretation liefern und das Design der Feldstudie informieren. Die Rohdaten der Befragung sind im Anhang unter Anhang K.2 verfügbar.

Für die Analyse von Zusammenhängen wird durchgängig der Spearman-Rangkorrelationskoeffizient verwendet. Dieser eignet sich für Daten wie Likert-Skalen, da er keine Normalverteilung voraussetzt und auf Rangfolgen statt auf absoluten Werten basiert [30].

4.4.1. Meeting-Last und -Struktur

Die Ergebnisse zeigen eine moderate Meeting-Last. 23 Teilnehmende (41,8 %) berichten von durchschnittlich einem Meeting pro Arbeitstag, 24 (43,6 %) von 2 bis 3 Meetings, 7 (12,7 %) von 4 bis 5 Meetings und eine Person von mehr als 5 Meetings täglich. Die tägliche Meetingzeit verteilt sich wie folgt: 9 Personen (16,4 %) verbringen weniger als 30 Minuten in Meetings, 17 (30,9 %) zwischen 30 Minuten und einer Stunde, 17 (30,9 %) zwischen 1 und 2 Stunden, 10 (18,2 %) zwischen 2 und 4 Stunden und 2 Personen (3,6 %) mehr als 4 Stunden. Die Meetingmodalität verteilt sich nahezu gleichmässig zwischen Remote-Formaten ($n = 21$, 38,2 %), hybriden Formaten ($n = 21$, 38,2 %) und reinen Vor-Ort-Meetings ($n = 13$, 23,6 %).

Die am häufigsten besuchten Meeting-Formate sind Daily Stand-ups (78,2 % der Befragten), Planning- oder Refinement-Sessions (76,4 %), Ad-hoc- oder spontane technische Syncs (69,1 %), One-on-One-Meetings (67,3 %), unternehmensweite Meetings (61,8 %) und Retrospektiven (54,5 %). Bei der Frage, welche Formate den grössten Anteil der Meetingzeit ausmachen, dominieren Planning- und Refinement-Sessions (49,1 %) vor Daily Stand-ups (41,8 %) und Ad-hoc-Syncs (34,5 %).

4.4.2. Wahrgenommene Nützlichkeit

Die wahrgenommene Nützlichkeit von Meetings fällt insgesamt moderat positiv aus ($M = 3,59$, $SD = 0,96$, $N = 54$). Knapp zwei Drittel der Befragten bewerten ihre Meetings als eher oder sehr nützlich. Rund ein Sechstel bewertet sie neutral, ein weiteres Sechstel negativ. Die hohe Standardabweichung zeigt, dass die Bewertungen innerhalb der Stichprobe stark streuen.

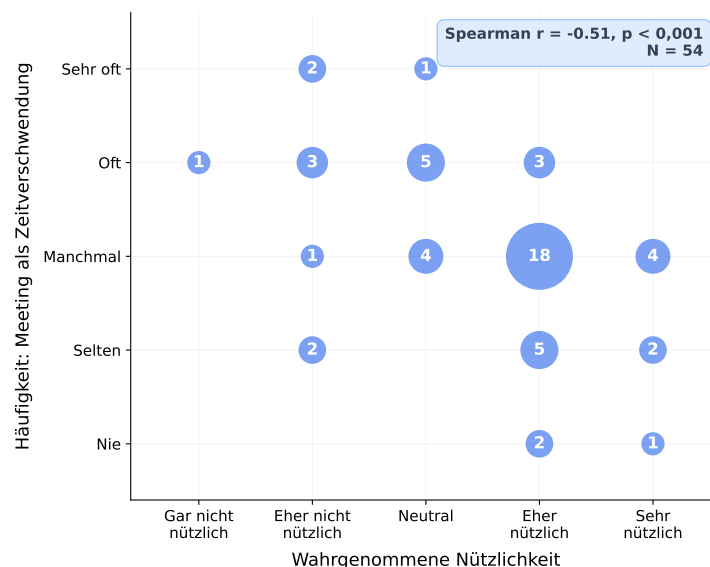


Abbildung 3: Zusammenhang zwischen genereller Nützlichkeitsbewertung und Häufigkeit der Ineffizienzwahrnehmung.

Auf einer separaten Skala wurde erhoben, wie häufig Befragte ein Meeting als Zeitverschwendung empfinden ($M = 3,05$, $SD = 0,91$). Gut ein Viertel (27,3 %) berichtet dies häufig oder sehr häufig, die Hälfte (50,9 %) gelegentlich (vgl. Abbildung 3). Beide Items korrelieren signifikant negativ (Spearman $r = -0,51$, $p < 0,001$): Wer Meetings insgesamt als nützlicher bewertet, berichtet seltener von Zeitverschwendung. Dass dennoch auch Personen mit positiver Gesamtbewertung gelegentlich Ineffizienz erleben, deckt sich mit der von Rogelberg et al.

beschriebenen Tendenz, dass generelle Meeting-Einstellungen und tatsächlich erlebte Meeting-Qualität auseinanderfallen können [5].

Die negativen Einflussfaktoren auf die Meeting-Erfahrung (Mehrfachauswahl) konkretisieren diese Streuung: Am häufigsten wird das Fehlen eines klaren Ziels genannt (45,5 %), gefolgt von übermässiger Dauer (30,9 %), irrelevanten oder zu vielen Teilnehmenden (29,1 %) und fehlenden Massnahmen oder nächsten Schritten (25,5 %). Fehlende Agenda und ungünstiges Timing nennen jeweils 12,7 % (vgl. Abbildung 4).

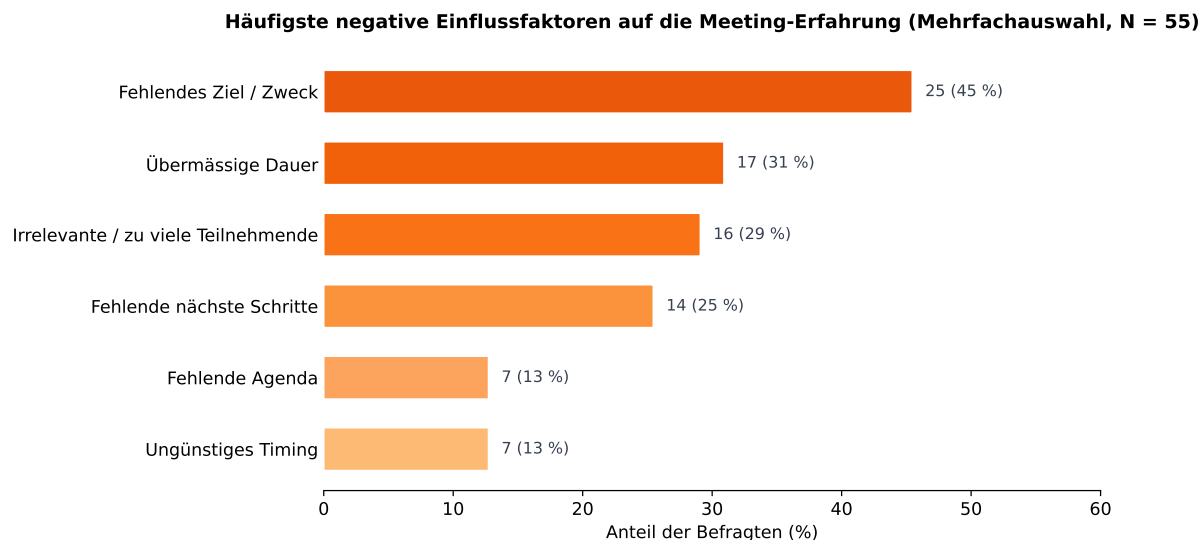


Abbildung 4: Häufigste negative Einflussfaktoren auf die Meeting-Erfahrung (Mehrfachauswahl, N = 55).

Diese Faktoren sind überwiegend prozeduraler Natur und entsprechen den von Leach et al. identifizierten Design-Charakteristika, die die Meeting-Effektivität vorhersagen [20].

Hohe Streuung der Meeting-Qualität

Zwei Drittel der Befragten bewerten Meetings generell als nützlich, doch 27,3 % erleben häufig oder sehr häufig Zeitverschwendung. Die häufigste Ursache ist ein fehlendes Ziel (45,5 %).

4.4.3. Auswirkungen auf Produktivität

Die Produktivitätswirkung von Meetings wird ambivalent bewertet. Auf die Frage, wie Meetings die Fähigkeit beeinflussen, Entwicklungsaufgaben zu erledigen ($M = 3,05$), antworten 21 Befragte (38,2 %) neutral, 16 (29,1 %) eher positiv und 15 (27,3 %) eher negativ, während jeweils 1 bis 2 Personen die Extrempositionen wählen.

Deutlich eindeutiger fällt der Vergleich zwischen Tagen mit vielen und wenigen Meetings aus (vgl. Abbildung 5, linkes Panel): 32 Befragte (60,4 %) berichten von deutlich niedrigerer Produktivität an meetingreichen Tagen, 10 (18,9 %) von leicht niedrigerer und nur 6 (11,3 %) von gleichbleibender Produktivität ($M = 1,75$, $SD = 1,16$, $N = 53$). Lediglich 5 Personen (9,4 %) geben eine höhere Produktivität an Tagen mit vielen Meetings an. **Insgesamt berichten somit 79,2 % der Befragten von niedrigerer Produktivität an meetingreichen Tagen.**

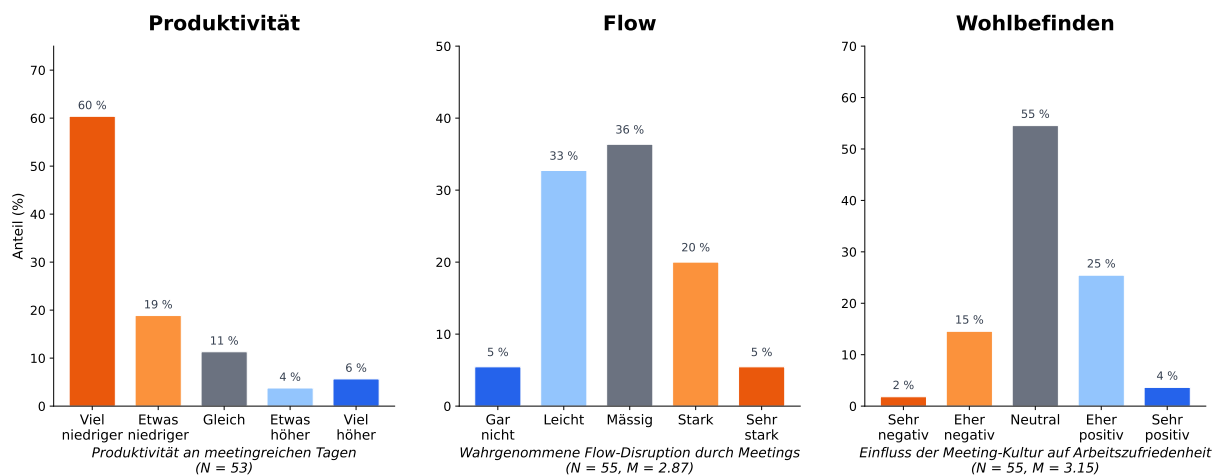


Abbildung 5: Retrospektive Wahrnehmung der Meeting-Auswirkungen auf die drei Outcome-Konstrukte.

Eine ergänzende Korrelationsanalyse zeigt einen signifikanten positiven Zusammenhang zwischen der Anzahl täglicher Meetings und der wahrgenommenen Produktivitätseinbusse (Spearman $r = 0,293$, $p = 0,033$, $N = 53$): Je mehr Meetings berichtet werden, desto stärker fällt die wahrgenommene Produktivitätseinbusse aus.

4.4.4. Auswirkungen auf Flow

Die Daten zeigen einen deutlichen Zusammenhang zwischen Meetings und der Unterbrechung von Flow-Zuständen (vgl. Abbildung 5, mittleres Panel). Die durchschnittliche wahrgenommene Flow-Disruption liegt bei $M = 2,87$ ($SD = 0,98$). 11 Befragte (20,0 %) berichten von starker und 3 (5,5 %) von sehr starker Flow-Unterbrechung, während 20 (36,4 %) eine moderate und 18 (32,7 %) eine leichte Unterbrechung angeben. Nur 3 Personen (5,5 %) erleben keine Flow-Disruption. **Zusammengenommen berichten 61,8 % der Befragten von mindestens moderater Flow-Unterbrechung.**

Besonders aufschlussreich ist die Zustimmung zur Aussage „After a meeting, I usually need at least 15 minutes before I can work productively again“ ($M = 3,54$, $SD = 1,18$). 22 Befragte (40,0 %) stimmen zu und 12 (21,8 %) stimmen stark zu. **Insgesamt berichten somit 61,8 % der Teilnehmenden eine mindestens 15-minütige Wiederherstellungszeit nach Meetings.** Dieses Ergebnis steht im Einklang mit der von Allen et al. berichteten durchschnittlichen Meeting-to-Work Transition Time von rund 20 Minuten nach virtuellen Meetings [6]. Die vorliegenden Daten bestätigen diesen Effekt spezifisch für Softwareentwickelnde.

Bestätigung des Recovery-Effekts für Softwareentwickelnde

61,8 % der Befragten berichten eine Wiederherstellungszeit von mindestens 15 Minuten nach Meetings. Dieser Befund bestätigt die in der Literatur dokumentierten Recovery-Kosten spezifisch für Softwareentwickelnde.

Hinsichtlich der Tageszeit werden Meetings am Vormittag ($n = 12$, 21,8 %) und am Nachmittag ($n = 12$, 21,8 %) als am disruptivsten empfunden. Bemerkenswert ist, dass 14 Befragte (25,5 %) angeben, die Tageszeit mache für sie keinen Unterschied. Dies deutet darauf hin, dass für einen substanziellen Anteil der Befragten Meetings unabhängig von ihrer zeitlichen Platzierung als störend empfunden werden.

4.4.5. Auswirkungen auf Wohlbefinden und Arbeitszufriedenheit

Der Einfluss der Meeting-Kultur auf die Arbeitszufriedenheit wird mehrheitlich als neutral wahrgenommen (vgl. Abbildung 5, rechtes Panel). 30 Befragte (54,5 %) bewerten den Einfluss neutral, 14 (25,5 %) eher positiv und 2 (3,6 %) sehr positiv ($M = 3,15$, $SD = 0,78$). Allerdings berichten 9 Befragte (16,4 %) von einem eher negativen oder sehr negativen Einfluss, was auf eine Teilgruppe hinweist, deren Wohlbefinden durch die Meeting-Praxis spürbar beeinträchtigt wird.

Der Zusammenhang zwischen Meeting-Häufigkeit und Arbeitszufriedenheit ist nicht signifikant (Spearman $r = 0,075$, $p = 0,589$). Dies deutet darauf hin, dass die Meeting-Last allein die Arbeitszufriedenheit nicht wesentlich beeinflusst. Vielmehr scheinen qualitative Faktoren wie die Nützlichkeit und Gestaltung der Meetings eine grössere Rolle zu spielen als die bloße Anzahl [20]. Im Vergleich zu den klaren Befunden bei Produktivität und Flow fällt die Wohlbefindenskomponente damit zurückhaltender aus. Die situative Erfassung von Stimmung und Energie in Phase 2 (vgl. Abschnitt 8) ermöglicht eine differenziertere Analyse dieses Zusammenhangs.

4.4.6. Meeting-Vermeidungsverhalten

Das Vermeidungsverhalten ist moderat ausgeprägt. Die Häufigkeit, mit der Meetingeinladungen abgelehnt werden, ist gering ($M = 2,07$): 10 Personen (18,2 %) geben an, nie Meetings abzulehnen, und 33 (60,0 %) tun dies nur selten. Noch seltener werden Meetings tatsächlich übersprungen ($M = 1,71$): 24 Befragte (43,6 %) geben an, nie ein Meeting zu überspringen, und 24 (43,6 %) tun dies selten.

Die häufigsten Gründe für das Ablehnen oder Überspringen sind fehlende Relevanz für die eigenen Aufgaben (16 Nennungen), Überlappung mit wichtigen Deadlines (9 Nennungen) und zu häufige oder zu lange Meetings (5 Nennungen). Der Bedarf an ungestörter Fokuszeit wird lediglich von 3 Personen als Grund genannt, obwohl 61,8 % Flow-Disruption berichten. Dies deutet darauf hin, dass der Schutz von Fokuszeit in der Praxis selten als legitimer Grund für das Ablehnen von Meetings akzeptiert wird.

Auffällig ist, dass weniger erfahrene Entwickelnde (weniger als 5 Jahre) Meetings seltener ablehnen ($M = 1,76$) als erfahrenere (5 bis 10 Jahre: $M = 2,27$; über 20 Jahre: $M = 2,23$). Dies könnte auf eine zunehmende Selbstbestimmung mit wachsender Berufserfahrung hindeuten.

4.4.7. Reflexion der Meeting-Kultur und Akzeptanz von Feedback

Die systematische Reflexion der Meeting-Kultur ist wenig verankert. **Systematische Meeting-Reflexion ist kaum verankert:** Nur 3 Befragte (5,5 %) berichten, dass ihr Team regelmässig über Meeting-Verbesserung diskutiert, 18 (32,7 %) berichten gelegentliche Diskussionen, 19 (34,5 %) seltene und 9 (16,4 %) gar keine.

Gleichzeitig besteht eine deutliche Offenheit für datenbasierte Verbesserung. Die Aussage, dass Teams die Erfahrungen der Teilnehmenden erfassen sollten (nicht nur Dauer und Häufigkeit), findet 61,8 % Zustimmung ($M = 3,64$). Die Aussage, dass kurzes Feedback nach Meetings helfen könnte, erzielt 65,5 % Zustimmung ($M = 3,68$). Die Bereitschaft, Meeting-Statistiken zu nutzen, ist hingegen heterogen verteilt ($M = 2,45$): 25,5 % halten solche Statistiken für nicht nützlich, was darauf hindeutet, dass einfache quantitative Metriken allein als unzureichend empfunden werden und subjektive Erfahrungsdaten ergänzend benötigt werden.

Offenheit für Veränderung bei fehlender Praxis

Nur 5,5 % der Teams diskutieren regelmässig über ihre Meeting-Kultur. Gleichzeitig befürworten 65,5 % der Befragten kurzes Feedback nach Meetings.

Für das Design der Feldstudie besonders relevant: Unmittelbar nach einem Meeting wird als am wenigsten disruptiver Feedback-Zeitpunkt bewertet ($M = 2,25$, $SD = 1,47$). Am disruptivsten wird eine Stunde nach dem Meeting empfunden ($M = 3,40$, $SD = 1,27$), vermutlich weil dieser Zeitpunkt in die Phase des Wiedereinstiegs in produktive Arbeit fällt. Das Ende des Tages ($M = 2,95$) und das Ende der Woche ($M = 3,09$) liegen dazwischen. Diese Ergebnisse bestätigen die Designentscheidung, das Meeting-Feedback in der Applikation direkt nach Meetingende auszulösen.

4.5. Qualitative Ergänzungen

Die offenen Kommentare ($n = 8$) verdichten die quantitativen Befunde und lassen sich thematisch in zwei Bereiche gliedern.

Der am häufigsten adressierte Bereich betrifft Vorbereitung, Struktur und Zeitbegrenzung. Mehrere Befragte fordern klare Agenden, strikte Zeitlimits und eine verbindliche Vorbereitung aller Teilnehmenden. Ein Befragter fasst dies pointiert zusammen: „*Clear agenda/goal. Time-box strictly (15–25 min max). Prep required.*“ Ein anderer fordert, dass bei mehr als fünf Teilnehmenden eine Begründung erforderlich sein solle und greift dabei explizit die Unterschei-

dung zwischen *Maker's Schedule* und *Manager's Schedule* auf [31]. Dieses Konzept beschreibt, dass Entwickelnde lange zusammenhängende Arbeitsblöcke benötigen, während Managende ihren Tag in kurze Intervalle unterteilen können. Es unterstreicht, warum Meetings für Softwareentwickelnde besonders kostspielig sind: Ein einzelnes Meeting kann einen halben Arbeitstag fragmentieren.

Der zweite Bereich betrifft asynchrone Alternativen und selektive Teilnahme. Mehrere Befragte argumentieren, dass Meetings nur stattfinden sollten, wenn sie echten Mehrwert bieten: „*Meetings should only cover essential topics. If everything has been discussed, the meeting can end early.*“ Ein anderer kritisiert unangekündigte Anrufe für Fragen, die asynchron beantwortet werden könnten, und verweist auf die hohe Meetingbelastung durch Scrum. Bilaterale Vorabklärungen könnten viele Diskussionen überflüssig machen.

4.6. Zusammenfassung und Implikationen für Phase 2

Die Ergebnisse der Online-Befragung adressieren FF1 und lassen sich entlang der drei Outcome-Konstrukte zusammenfassen.

Produktivität. Der deutlichste Befund betrifft die Produktivitätswirkung: 79,2 % der Befragten berichten von niedrigerer Produktivität an meetingreichen Tagen. Meeting-Häufigkeit korreliert signifikant mit der wahrgenommenen Produktivitätseinbusse ($r = 0,293$, $p = 0,033$). Gleichzeitig wird die generelle Produktivitätswirkung von Meetings ambivalent bewertet ($M = 3,05$), was auf eine differenzierte Wahrnehmung hindeutet: Meetings sind nicht pauschal schädlich, aber ihre Kumulation an einzelnen Tagen hat deutlich negative Effekte.

Flow. 61,8 % der Teilnehmenden berichten von mindestens moderater Flow-Disruption durch Meetings, und ebenfalls 61,8 % benötigen mindestens 15 Minuten Wiederherstellungszeit. Diese Befunde bestätigen die in der Literatur dokumentierte Attention Residue [6], [12] und unterstreichen, dass Meetings über ihre nominelle Dauer hinaus Arbeitszeit beanspruchen.

Wohlbefinden. Die Befunde zum Wohlbefinden sind weniger eindeutig: 54,5 % bewerten den Einfluss als neutral. Die Meeting-Häufigkeit allein erklärt die Arbeitszufriedenheit nicht signifikant. Dies legt nahe, dass qualitative Meeting-Eigenschaften (Nützlichkeit, emotionale Wirkung) eine grössere Rolle spielen als die blosse Quantität.

Übergreifend. Die Daten zeigen eine hohe situative Variabilität der Meeting-Qualität: Die generelle Nützlichkeitsbewertung fällt moderat positiv aus ($M = 3,59$), doch 27,3 % der Befragten erleben häufig oder sehr häufig Zeitverschwendung. Die dominierenden Ursachen sind prozeduraler Natur (fehlendes Ziel, übermässige Dauer, irrelevante Teilnehmende).

Für das Design der Feldstudie ergeben sich drei konkrete Implikationen: Die Bestätigung des unmittelbaren Post-Meeting-Zeitpunkts als am wenigsten disruptive Feedback-Option ($M = 2,25$ vs. $M = 3,40$ für den einstündigen Verzug); die Notwendigkeit, neben quantitativen Metriken auch subjektive Erfahrungsdimensionen (Emotion, Energie, Fokusunterbrechung) zu erfassen; und die Relevanz der Tagesstruktur (Fragmentierung, Meetingfrequenz) als zentrale Einflussvariable.

5. Konzeptlösung

Dieses Kapitel überführt die in der Literatur identifizierten Problemfelder und die empirischen Befunde der Befragungsstudie (vgl. Abschnitt 4) in eine strukturierte Konzeptlösung. Ausgehend von einem zentralen Problemstatement werden zunächst Teilprobleme und Stakeholder definiert. Darauf aufbauend werden Vision, Mission und Leitprinzipien formuliert, aus denen konkrete Produktziele abgeleitet werden. Den Abschluss bildet die Spezifikation funktionaler und nicht-funktionaler Anforderungen, die das nachfolgende Kapitel zur Systemarchitektur und Implementierung (vgl. Abschnitt 6) als Ausgangspunkt verwendet.

Die im Rahmen dieser Arbeit entwickelte Applikation erfüllt dabei eine Doppelrolle: Sie dient als Erhebungsinstrument für die Feldstudie (vgl. Abschnitt 8) und stellt zugleich einen funktionsfähigen Prototyp dar, der den in Abschnitt 1.3 formulierten Praxisbeitrag konkretisiert. Diese Doppelrolle wird in Abschnitt 5.6 explizit eingeordnet.

Die strukturierte Darstellung orientiert sich an konzeptionellen Artefakten des *Rational Unified Process* (RUP) [32] und dient dazu, eine konsistente Verbindung zwischen Problemdefinition, Zielsetzung und Anforderungsspezifikation herzustellen.

5.1. Zentrales Problemstatement

Das Problem von ineffizienten, häufigen und schlecht reflektierten Meetings betrifft Softwareentwickelnde in modernen Softwareprojekten und wirkt sich negativ auf Produktivität, Flow-Zustand, Wohlbefinden und Zufriedenheit aus. Gleichzeitig fehlt es an datenbasierten Grundlagen, um die Auswirkungen einzelner Meetings systematisch zu bewerten und die eigene Meeting-Praxis gezielt zu verbessern.

Dieses Problemstatement wird sowohl durch die bestehende Literatur (vgl. Abschnitt 2) als auch durch die Ergebnisse der Befragungsstudie gestützt (vgl. Abschnitt 4.6).

5.2. Teilprobleme

Aus dem zentralen Problemstatement (vgl. Abschnitt 5.1) lassen sich fünf Teilprobleme ableiten, die jeweils durch Literatur und eigene Befragungsdaten fundiert sind.

P1: Fragmentierung des Arbeitstages durch Meetings. Häufige und ungünstig platzierte Meetings fragmentieren den Arbeitstag und unterbrechen Flow-Zustände von Softwareentwickelnden. Tage mit wenigen oder keinen Meetings werden über verschiedene Produktivitätstypen hinweg als produktiver bewertet [3], und die Wirkung von Meetings folgt einem nicht-linearen Muster: Ab einer bestimmten Schwelle erschöpfen sie Aufmerksamkeits- und Motivationsressourcen [1]. Hinzu kommt, dass beim Wechsel zwischen Meeting und Arbeit kognitive Aufmerksamkeit bei der vorherigen Tätigkeit verbleibt (Attention Residue), was die Rückkehr zu konzentrierter Arbeit erschwert [12]. Die durchschnittliche Wiederherstellungszeit nach einem Meeting beträgt rund 20 Minuten [6]. Die Befragungsdaten bestätigen diesen Befund: 61,8 % der Teilnehmenden berichten von mindestens moderater Flow-Unterbrechung durch Meetings, und ebenfalls 61,8 % benötigen nach einem Meeting mindestens 15 Minuten, um wieder produktiv zu arbeiten (vgl. Abschnitt 4.4.4).

P2: Fehlende datenbasierte Grundlage zur Verbesserung der Meeting-Kultur. Teams verfügen über keine systematische, datengestützte Feedback-Schleife für Meetings. Ohne einen solchen Mechanismus perpetuieren sich ineffiziente Formate [5]. Die Befragungsdaten verdeutlichen diese Lücke: Nur 5,5 % der Teilnehmenden geben an, dass ihr Team regelmässig über die Verbesserung der Meeting-Kultur diskutiert, obwohl 65,5 % kurzes Meeting-Feedback als potenziell hilfreich einschätzen (vgl. Abschnitt 4.4.7).

P3: Wahrgenommene Ineffizienz ohne systematische Reflexion. Viele Meetings werden als ineffizient wahrgenommen, ohne dass diese Wahrnehmung systematisch erfasst oder zur Anpassung bestehender Formate genutzt wird. Die Befragung zeigt eine aufschlussreiche Diskrepanz: Während die Mehrheit Meetings insgesamt als eher nützlich bewertet, berichten gleichzeitig 78,2 % mindestens gelegentlich, ein Meeting als Zeitverschwendung empfunden zu haben. Die häufigsten Ursachen sind prozeduraler Natur und decken sich mit den von Leach et al. identifizierten Design-Charakteristika, die Meeting-Effektivität vorhersagen [20]: fehlendes Ziel (45,5 %), übermässige Dauer (30,9 %) und irrelevante Teilnehmende (29,1 %) (vgl. Abschnitt 4.4.2).

P4: Nicht-erfasste emotionale Auswirkungen von Meetings. Die emotionalen Auswirkungen von Meetings werden von Softwareentwickelnden nicht systematisch erfasst oder reflektiert. Das ist problematisch, weil positive affektive Zustände die Problemlösungsfähigkeit nachweislich fördern [14] und Unzufriedenheit sich negativ auf Produktivität und kognitive Leistung auswirkt [24]. Die AET [26] postuliert, dass Meetings als affektive Ereignisse wirken, die sowohl positive Emotionen (Motivation, Zielklärung) als auch negative Emotionen (Frustration, Stress) auslösen können. Die Befragung zeigt, dass 16,4 % der Teilnehmenden einen negativen Einfluss der Meeting-Kultur auf ihre individuelle Arbeitszufriedenheit berichten (vgl. Abschnitt 4.4.5).

P5: Fehlende Transparenz über die tatsächliche Meeting-Last. Softwareentwickelnde verfügen über keine transparente Übersicht über ihre tatsächliche Meeting-Last. Die Befragung ergibt ein heterogenes Bild: Während 41,8 % der Befragten ein Meeting pro Tag angeben, berichten 58,2 % von zwei oder mehr Meetings täglich (vgl. Abschnitt 4.4.1). Zugleich bewerten die Befragten den potenziellen Nutzen rein quantitativer Meeting-Statistiken (z.B. Anzahl und Dauer) nur als moderat, was darauf hindeutet, dass solche Kennzahlen allein nicht ausreichen, um die eigene Meeting-Praxis wirksam zu reflektieren (vgl. Abschnitt 4.4.7).

5.3. Stakeholder

Im Kontext dieser Arbeit wird eine primäre Stakeholder-Gruppe identifiziert.

Softwareentwickelnde sind die Hauptnutzenden der Applikation. Sie geben Feedback zu einzelnen Meetings und am Tagesende ab und nutzen das Dashboard zur Reflexion über ihre individuelle Meeting-Last und deren Auswirkungen. Ihr zentrales Interesse ist der Schutz ungestörter Arbeitsphasen und die Reduktion von Meeting-Belastung.

Forschende (Sekundäre Stakeholder) nutzen die Applikation nicht im täglichen Arbeitsalltag, sind aber auf die aggregierten und pseudonymisierten Daten angewiesen, um die Forschungsfragen (vgl. Abschnitt 1.2) auszuwerten. Ihr primäres Interesse liegt in der hohen Datenqualität, einer lückenlosen Erfassung (hohe Compliance-Raten) sowie der Möglichkeit, die erhobenen Daten für die spätere statistische Auswertung der Feldstudie (vgl. Abschnitt 8)

zuverlässig und sicher exportieren zu können. Die strengen Anforderungen an die Systemstabilität und den Datenschutz leiten sich massgeblich aus diesem Stakeholder-Interesse ab.

Die Applikation dient zusätzlich als Erhebungsinstrument für die Feldstudie (vgl. Abschnitt 8). Die daraus resultierenden Anforderungen an die Datenqualität und Zuverlässigkeit des Systems fliessen in die Anforderungen ein (vgl. Abschnitt 5.7), werden aber nicht als eigene Stakeholder-Gruppe modelliert, da sie keine eigenständigen Nutzungsanforderungen an die Benutzeroberfläche stellen.

5.4. Vision, Mission und Leitprinzipien

Ausgehend von den identifizierten Stakeholder-Bedürfnissen werden im Folgenden Vision und Mission der Konzeptlösung definiert. Die Vision beschreibt den angestrebten Zielzustand der Meeting-Kultur, die Mission den konkreten Beitrag der Applikation zu diesem Ziel.

Vision

Die Vision der Konzeptlösung ist eine Arbeitsweise in Softwareprojekten, in der Meetings gezielt, reflektiert und mit erkennbarem Mehrwert eingesetzt werden. Meetings sollen Zusammenarbeit unterstützen, ohne fokussierte Arbeitsphasen zu fragmentieren oder den Flow-Zustand, die Produktivität und das Wohlbefinden von Softwareentwickelnden negativ zu beeinflussen.

Langfristig soll eine Meeting-Kultur gefördert werden, in der der Nutzen einzelner Meetings transparent ist und deren Auswirkungen auf Arbeitserleben und Leistungsfähigkeit bewusst berücksichtigt werden.

Mission

Die Mission der Konzeptlösung besteht darin, eine strukturierte und evidenzbasierte Grundlage zur Bewertung von Meetings in Softwareentwicklungsteams zu schaffen. Durch die systematische Erfassung sowohl objektiver als auch subjektiver Meeting-Eigenschaften (vgl. Abschnitt 3.1) sollen die Auswirkungen einzelner Meetings auf individueller Ebene nachvollziehbar gemacht werden.

Bestehende Kalender-Analysetools beschränken sich auf quantitative Metriken wie die Anzahl und Dauer von Meetings. Die Konzeptlösung geht darüber hinaus, indem sie diese objektiven Daten mit strukturiertem subjektivem Feedback auf zwei Ebenen verknüpft: auf der Ebene einzelner Meetings und auf der Ebene des gesamten Arbeitstages. Diese Kombination ermöglicht es, nicht nur zu erkennen, wie viele Meetings stattfinden, sondern wie sie tatsächlich erlebt werden und welche konkreten Auswirkungen sie auf die drei Outcome-Konstrukte Produktivität, Flow und Wohlbefinden haben (vgl. Abschnitt 3.1.1).

Die Konzeptlösung zielt darauf ab, Softwareentwickelnde bei der reflektierten Analyse ihrer Meeting-Praxis zu unterstützen und fundierte Anpassungen von Meeting-Formaten zu ermöglichen, ohne dabei Kontroll- oder Überwachungsmechanismen zu etablieren.

Eine Analyse bestehender Tools bestätigt, dass keines der untersuchten Produkte alle drei Outcome-Dimensionen mit einer echtzeitbasierten, entwicklerspezifischen Erhebungsmethodik kombiniert (vgl. Anhang F).

Leitprinzipien

Die folgenden Leitprinzipien definieren die Designentscheidungen, die der Konzeptlösung zugrunde liegen. Sie leiten sich aus den Forschungsfragen (vgl. Abschnitt 1.2), der Literatur (vgl. Abschnitt 2) und den Befragungsergebnissen (vgl. Abschnitt 4.4) ab.

Fokus auf Flow und ungestörte Arbeit Meetings werden nicht isoliert betrachtet, sondern in ihrem Einfluss auf zusammenhängende Arbeitsphasen und den Flow-Zustand von Softwareentwickelnden bewertet. Dieses Prinzip leitet sich aus der empirisch dokumentierten Fragmentierungsproblematik ab (vgl. P1) und stellt sicher, dass die Datenerhebung den Arbeitstag als zusammenhängende Struktur abbildet.

Kombination aus Daten und Wahrnehmung. Objektive Kennzahlen und subjektive Einschätzungen werden gleichwertig berücksichtigt, um ein ganzheitliches Bild der Meeting-Wirkung zu erhalten. Dieser Grundsatz entspricht der Empfehlung des SPACE-Frameworks, Produktivität nicht eindimensional zu messen, sondern mehrere Dimensionen gleichzeitig zu erheben [9]. Die Konzeptlösung setzt dieses Prinzip um, indem sie die drei Outcome-Konstrukte Produktivität, Flow und Wohlbefinden parallel erfasst (vgl. Abschnitt 3.1.1).

Reflexion statt Bewertung. Die erhobenen Daten dienen der individuellen Selbstreflexion und nicht der Leistungsbeurteilung einzelner Personen oder Teams. Diese Designentscheidung folgt der Empfehlung des SPACE-Frameworks, Produktivitätsmetriken nicht zur Bewertung Einzelner einzusetzen [9], und soll die Bereitschaft der Nutzenden zur ehrlichen Dateneingabe fördern.

Transparenz als Grundlage für Reflexion. Die Konzeptlösung geht davon aus, dass eine nachvollziehbare Übersicht über Meeting-Last, Meeting-Struktur und deren Wirkung eine notwendige Voraussetzung für fundierte Anpassungen der Meeting-Praxis darstellt. Diese Annahme stützt sich auf die Befragungsergebnisse, wonach 65,5 % der Teilnehmenden kurzes Meeting-Feedback als potenziell hilfreich einschätzen (vgl. Abschnitt 4.4.7), während gleichzeitig nur 5,5 % angeben, dass ihr Team regelmässig über die Meeting-Kultur diskutiert (vgl. P2).

Minimale Disruption bei der Datenerhebung. Die Erhebungsinstrumente sind bewusst kurz gehalten und zeitlich so platziert, dass sie den Arbeitsfluss möglichst wenig beeinträchtigen. Diese Entscheidung stützt sich auf die Befragungsergebnisse, die den unmittelbaren Post-Meeting-Zeitpunkt als am wenigsten disruptiv identifizieren (vgl. Abschnitt 4.4.7).

5.5. Produktziele

Aus der Mission (vgl. Abschnitt 5.4) und den identifizierten Teilproblemen (vgl. Abschnitt 5.2) werden zwei übergeordnete Produktziele abgeleitet.

PZ1: Kontextbezogene Datenerhebung mit minimaler Störung des Arbeitsalltags.

Das System soll objektive Kalenderdaten mit subjektiven Echtzeitbewertungen zusammenführen (vgl. Abschnitt 3.1). Die Kalenderdaten werden automatisch über bestehende Kalenderanbieter synchronisiert. Subjektives Feedback wird über kurze, strukturierte Abfragen direkt nach einem Meeting sowie am Ende des konfigurierten Arbeitstages erfasst. Die Datenerhebung gilt als erfolgreich in den Arbeitsalltag integriert, wenn eine einzelne Feedback-Eingabe in unter 30 Sekunden abschliessbar ist und die Benachrichtigung unmittelbar nach Meetingende ausgelöst wird. Dieses Produktziel adressiert die Teilprobleme P2, P4 und P5.

PZ2: Aufbereitung und Visualisierung der erhobenen Daten zur Selbstreflexion.

Die erhobenen Daten sollen so aufbereitet und dargestellt werden, dass Softwareentwickelnde Muster in ihrem Arbeitsalltag erkennen und ihre Meeting-Gewohnheiten eigenständig hinterfragen können. Dazu gehört einerseits die Darstellung, wie stark der Arbeitstag durch Meetings fragmentiert wird und wo zusammenhängende Fokusphasen liegen. Andererseits werden die erfassten Bewertungen nach Meeting-Typ aufgeschlüsselt dargestellt. Das Produktziel gilt als erreicht, wenn das System mindestens ein Dashboard bereitstellt, das objektive Strukturmetriken und subjektive Bewertungsdaten gemeinsam darstellt. Ziel ist nicht, durch automatische Systemeingriffe zu steuern, sondern Transparenz zu schaffen, damit Softwareentwickelnde selbst erkennen können, wo Anpassungen sinnvoll wären. Dieses Produktziel adressiert die Teilprobleme P1, P3 und P5. Die gestalterische Konkretisierung beider Produktziele wurde vor der Implementierung in digitalen Wireframes festgehalten, die den Entwurfsprozess und die zentralen Designentscheidungen dokumentieren (vgl. Anhang G). Die Wireframes wurden gemeinsam im Team erstellt, um frühzeitig ein einheitliches Verständnis der geplanten Funktionalität zu schaffen und die anschliessende Umsetzung effizient koordinieren zu können. Dabei wurde die Applikation gezielt in einzelne Views aufgegliedert, darunter Kalenderansicht, Feedback-Eingabe, Dashboard und Einstellungen, sodass die Verantwortlichkeiten klar verteilt und die einzelnen Funktionsbereiche parallel entwickelt werden konnten. Dieser Ansatz ermöglichte es, die Implementierung zügig aufzunehmen, da jede beteiligte Person auf Basis der abgestimmten Wireframes eigenständig an ihrer jeweiligen View arbeiten konnte, ohne auf Abstimmungsschleifen warten zu müssen.

5.6. Einordnung und Traceability Matrix

Die Produktziele bedienen zwei Ebenen, die im Folgenden unterschieden werden.

Forschungsrelevante Features dienen der Datenerhebung für die Feldstudie (vgl. Abschnitt 8) und sind unmittelbar an die Forschungsfragen gebunden. Das Meeting-Feedback-Survey und das End-of-Day-Feedback operationalisieren die subjektiven Eigenschaften des Forschungsmodells (vgl. Abschnitt 3.1) und adressieren FF2 sowie FF3. Die automatische Kalendersynchronisation liefert die objektiven Eigenschaften und adressiert FF2 sowie FF3 als objektive Vergleichsgrundlage.

Infrastruktur-Features sind technisch notwendig, damit die Applikation als Erhebungsinstrument funktionieren kann. Dazu gehören die Benutzerverwaltung, die Kalenderanbindung und die Datenhaltung.

Praxisorientierte Features gehen über die unmittelbare Datenerhebung hinaus und adressieren den praktischen Beitrag der Arbeit (vgl. PZ2). Die Visualisierung der objektiven Meeting-Struktur und der subjektiven Erfahrungsdaten in Form von Dashboards ermöglicht die Reflexion der eigenen Meeting-Praxis.

Tabelle 3 ordnet die Teilprobleme (P) den Produktzielen (PZ) und Forschungsfragen (FF) zu und unterscheidet zwischen Forschungs- und Praxisbeitrag.

P	Teilproblem	PZ	FF	Forschungsbeitrag (Phase 2)	Praxisbeitrag (Individuum)
P1	Fragmentierung	PZ2	FF2	Situative Fokusphasen-Analyse	Persönliche Visualisierung von Fokusblöcken zum Schutz von Deep Work
P2	Fehlende Datenbasis	PZ1	FF2, FF3	ESM-Surveys und Kalender-Sync	Bereitstellung einer privaten Datengrundlage für Self-Tracking
P3	Ineffizienz ohne Reflexion	PZ2	FF2, FF3	Situative Nützlichkeitsbewertung (RO-TI), KI-Kategorisierung	Individuelle Effizienzanalyse nach Meeting-Typ zur Priorisierung
P4	Emotionale Auswirkungen	PZ1, PZ2	FF2	Situative Erfassung von Stimmung und Energie	Darstellung persönlicher emotionaler Trends im Arbeitsalltag
P5	Fehlende Transparenz	PZ1, PZ2	FF2	Erhebung objektiver Kalendermetriken	Transparente private Dashboard-Metriken zur Selbstreflexion

Tabelle 3: Traceability Matrix: Zuordnung von Teilproblemen, Produktzielen, Forschungsfragen und Beitragsebene.

Ergänzend erfordert PZ1 diverse Infrastruktur-Features, die die technische Voraussetzung für das gesamte System bilden. Diese werden nicht als eigenständige Anforderungen modelliert, sondern ergeben sich als technische Konsequenzen der funktionalen Anforderungen (vgl. Abschnitt 5.7.1).

5.7. Anforderungen

Die Anforderungen leiten sich aus den Produktzielen PZ1 und PZ2 und der Doppelrolle der Applikation als Erhebungsinstrument für die Feldstudie (vgl. Abschnitt 8) und als funktionsfähiger Prototyp ab. Ein Erhebungsinstrument, das den Arbeitsalltag fragmentiert, würde genau das Phänomen verstärken, das es messen soll. Ein Prototyp, der bei Ausfällen Daten verliert, würde die gesamte Feldstudie entwerten. Aus diesem Spannungsfeld ergeben sich Anforderungen, die über die einer reinen Forschungssoftware hinausgehen.

Die Anforderungen verwenden zur Kennzeichnung der Verbindlichkeit die Schlüsselwörter *muss* (verpflichtend für den Erfolg der Feldstudie oder die Kernfunktionalität des Prototyps), *soll* (wichtig für die Qualität der Erhebung oder der Nutzungserfahrung, aber nicht zwingend erforderlich) und *kann* (wünschenswert, aber für den Projekterfolg nicht erforderlich).

5.7.1. Funktionale Anforderungen

FA-01 (Muss, Herkunft: PZ1, P2, P5) Das System muss Kalenderdaten externer Anbieter (Google, Microsoft) automatisiert über OAuth 2.0 synchronisieren und lokal persistieren.

Abnahmekriterium: Kalendereinträge eines verbundenen Kontos sind innerhalb von 5 Minuten nach Erstellung im System sichtbar.

FA-02 (Muss, Herkunft: PZ1, P3) Das System muss importierte Kalendereinträge anhand ihres Titels automatisch in vordefinierte Meeting-Kategorien klassifizieren.

Abnahmekriterium: Mindestens 80 % der Kalendereinträge eines Testdatensatzes werden der korrekten Kategorie zugeordnet.

FA-03 (Muss, Herkunft: PZ1, P2, P3, P4) Das System muss nach Beendigung eines Meetings eine Push-Benachrichtigung auslösen, die zur Abgabe eines strukturierten Meeting-Feedbacks auffordert. Das Feedback umfasst die wahrgenommene Nützlichkeit (ROTI), die Fokusunterbrechung, die Stimmung und das Energielevel.

Abnahmekriterium: Nach dem im Kalender definierten Meetingende wird eine Benachrichtigung auf dem Endgerät der nutzenden Person angezeigt.

FA-04 (Muss, Herkunft: PZ1, P2, P4) Das System muss am Ende des konfigurierten Arbeitstages ein End-of-Day-Feedback auslösen, das die subjektive Tagesbewertung (wahrgenommene Produktivität, Deep-Work-Anteil, Energielevel, Meetingbelastung) erhebt.

Abnahmekriterium: Die Benachrichtigung wird eine Stunde vor konfigurierten Zeitpunkt (Arbeitszeitende) ausgelöst und das Feedback wird vollständig persistiert.

FA-05 (Muss, Herkunft: PZ2, P1, P5) Das System muss ein Dashboard bereitstellen, das die Fragmentierung des Arbeitstages, zusammenhängende Fokusblöcke und die Verteilung der Meeting-Typen Visualisiert.

Abnahmekriterium: Das Dashboard zeigt für einen ausgewählten Zeitraum mindestens die Tagesfragmentierung, die Fokusblöcke und die Verteilung nach Meeting-Kategorie an. Der initiale Entwurf dieser Ansicht ist im Dashboard-Wireframe dokumentiert (vgl. Abbildung 21).

FA-06 (Soll, Herkunft: PZ2, P3, P4, P5) Das System soll ein Dashboard bereitstellen, das die erfassten Feedback-Scores Visualisiert.

Abnahmekriterium: Die Feedback-Werte (Nützlichkeit, Stimmung, Energie) werden pro Meeting-Kategorie und im Zeitverlauf dargestellt. Die Entwurfsgrundlage für diese Ansicht bildet der Dashboard-Wireframe (vgl. Abbildung 21).

FA-07 (Muss, Herkunft: PZ1) Nutzende müssen Arbeitszeiten und Benachrichtigungspräferenzen individuell konfigurieren können.

Abnahmekriterium: Arbeitszeiten und Benachrichtigungszeitpunkte lassen sich über die Benutzeroberfläche ändern und werden bei der nächsten Benachrichtigung berücksichtigt. Die gestalterische Umsetzung dieser Konfigurationsansicht ist im Settings-Wireframe festgehalten (vgl. Abbildung 23).

FA-08 (Muss, Herkunft: PZ1) Das System muss eine Benutzerverwaltung mit Registrierung und tokenbasierter Authentifizierung bereitstellen.

Abnahmekriterium: Eine neue nutzende Person kann sich registrieren, anmelden und erhält ein gültiges Authentifizierungstoken für nachfolgende Anfragen.

FA-09 (Kann, Herkunft: PZ1, P5) Nutzende können Kalendereinträge manuell erfassen, bearbeiten und löschen.

Abnahmekriterium: Ein manuell erstellter Eintrag erscheint in der Kalenderübersicht und wird bei der Feedbackerhebung berücksichtigt. Der Entwurf der Kalenderansicht mit den vorgesehenen Interaktionsmöglichkeiten ist im Kalender-Wireframe dokumentiert (vgl. Abbildung 22).

5.7.2. Nicht-funktionale Anforderungen

NFA-01 (Muss, Qualitätsmerkmal: Benutzbarkeit) Die Feedback-Eingabe muss den Arbeitsfluss minimal unterbrechen und in unter 30 Sekunden abschliessbar sein.

Abnahmekriterium: In einem Usability-Test schliessen mindestens 80 % der Testpersonen eine Meeting-Feedback-Eingabe in unter 30 Sekunden ab.

NFA-02 (Muss, Qualitätsmerkmal: Sicherheit) Sämtliche Verbindungen müssen über HTTPS verschlüsselt sein, und OAuth-Tokens müssen verschlüsselt persistiert werden.

Abnahmekriterium: Alle API-Endpunkte sind ausschliesslich über HTTPS erreichbar, und OAuth-Tokens sind in der Datenbank nicht im Klartext gespeichert.

NFA-03 (Muss, Qualitätsmerkmal: Zuverlässigkeit) Die Hintergrundsynchonisierung muss gegen temporäre API-Ausfälle tolerant sein, um Datenverluste während der Feldstudie (vgl. Abschnitt 8) zu vermeiden.

Abnahmekriterium: Nach einem simulierten API-Ausfall von bis zu 15 Minuten synchronisiert das System die fehlenden Daten beim nächsten erfolgreichen Verbindungsaufbau vollständig nach.

NFA-04 (Soll, Qualitätsmerkmal: Zeitverhalten) Push-Benachrichtigungen sollen zeitnah nach Meetingende zugestellt werden, um die Erinnerungsgenauigkeit der situativen Bewertung sicherzustellen.

Abnahmekriterium: Mindestens 90 % der Benachrichtigungen werden innerhalb von 2 Minuten nach dem kalenderbasierten Meetingende zugestellt.

NFA-05 (Soll, Qualitätsmerkmal: Wartbarkeit) Das System soll containerisiert und über eine CI/CD-Pipeline automatisiert deploybar sein.

Abnahmekriterium: Ein Deployment lässt sich über die Pipeline ohne manuelle Eingriffe auf dem Server durchführen.

NFA-06 (Soll, Qualitätsmerkmal: Änderbarkeit) Die Datenstrukturen für Feedback sollen flexibel erweiterbar sein, ohne dass bei der Ergänzung neuer Feedback-Felder Schemamigrationen erforderlich werden.

Abnahmekriterium: Ein neues Feedback-Feld kann hinzugefügt werden, ohne dass eine Datenbankmigration notwendig ist.

NFA-07 (Muss, Qualitätsmerkmal: Sicherheit) Das System muss den Schutz sensibler Kalenderinhalte (wie Meeting-Titel und Teilnehmerdaten) bei der automatisierten Verarbeitung gewährleisten. Zur Wahrung des Datenschutzes dürfen diese Informationen für die in FA-02 definierte Kontextualisierung nicht an Systeme, Dienste oder APIs externer Drittanbieter übermittelt werden.

Abnahmekriterium: Bei der automatisierten Kategorisierung der Kalendereinträge verlassen zu keinem Zeitpunkt Kalenderdaten die Systemgrenzen der eigenen Applikationsinfrastruktur.

6. Systemarchitektur und Implementierung

Die Konzeptlösung (Abschnitt 5) definiert zwei Produktziele: eine automatisierte, kontextbezogene Datenerhebung (PZ1) und die Aufbereitung der erhobenen Daten zur Selbstreflexion (PZ2). Diese Ziele stellen spezifische technische Anforderungen, die über die einer reinen Forschungssoftware hinausgehen. PZ1 verlangt die Integration externer Kalender-APIs, eine intelligente Kontextualisierung roher Termineinträge und eine Feedback-Erfassung, die den Arbeitsfluss nicht zusätzlich fragmentiert. PZ2 erfordert Aggregationslogik und Visualisierungen, die objektive Meeting-Daten sowie subjektive Bewertungen jeweils aufbereiten und zugänglich machen. Gleichzeitig muss das System als Erhebungsinstrument für eine zweiwöchige Feldstudie zuverlässig funktionieren: Ausfälle oder Datenverluste während des Erhebungszeitraums würden die gesamte empirische Grundlage der Arbeit gefährden.

Dieses Kapitel dokumentiert, wie diese Anforderungen architektonisch und implementierungstechnisch umgesetzt wurden. Die Darstellung folgt dem logischen Datenfluss des Systems: von der automatisierten Gewinnung objektiver Kalenderdaten (Abschnitt 6.2) über die KI-gestützte Anreicherung (Abschnitt 6.3) und die Erfassung subjektiver Wahrnehmungen (Abschnitt 6.4) bis hin zur Visualisierung in Dashboards (Abschnitt 6.5). Die technischen Querschnittsaspekte wie Infrastruktur, Deployment und Datenmodell werden abschliessend in den Abschnitt 6.6 und Abschnitt 6.7 behandelt.

Der vollständige Quellcode der Applikation ist im Anhang unter Anhang K.1 verfügbar.

6.1. Systemüberblick

Bevor die folgenden Abschnitte die einzelnen Subsysteme im Detail beschreiben, gibt dieser Abschnitt einen Überblick über die Gesamtarchitektur.

Konzeptionell lässt sich das System entlang seines Datenflusses in vier fachliche Subsysteme gliedern. Die Kalendersynchronisation importiert objektive Termindaten externer Anbieter und bildet damit die quantitative Grundlage des Forschungsmodells: Anzahl, Dauer und Timing der Meetings. Die KI-gestützte Kontextualisierung reichert diese Rohdaten semantisch an, indem sie Kalendereinträge automatisch in Meeting-Kategorien klassifiziert und so eine differenzierte Analyse nach Meeting-Format ermöglicht. Die Feedback-Erfassung ergänzt die objektive Datengrundlage um die subjektive Perspektive der Nutzenden: ereigniskontingentes Meeting-Feedback und zeitkontingentes End-of-Day-Feedback. Die Dashboards schliesslich überführen beide Datenströme in Visualisierungen, die eine evidenzbasierte Reflexion der Meeting-Kultur ermöglichen.

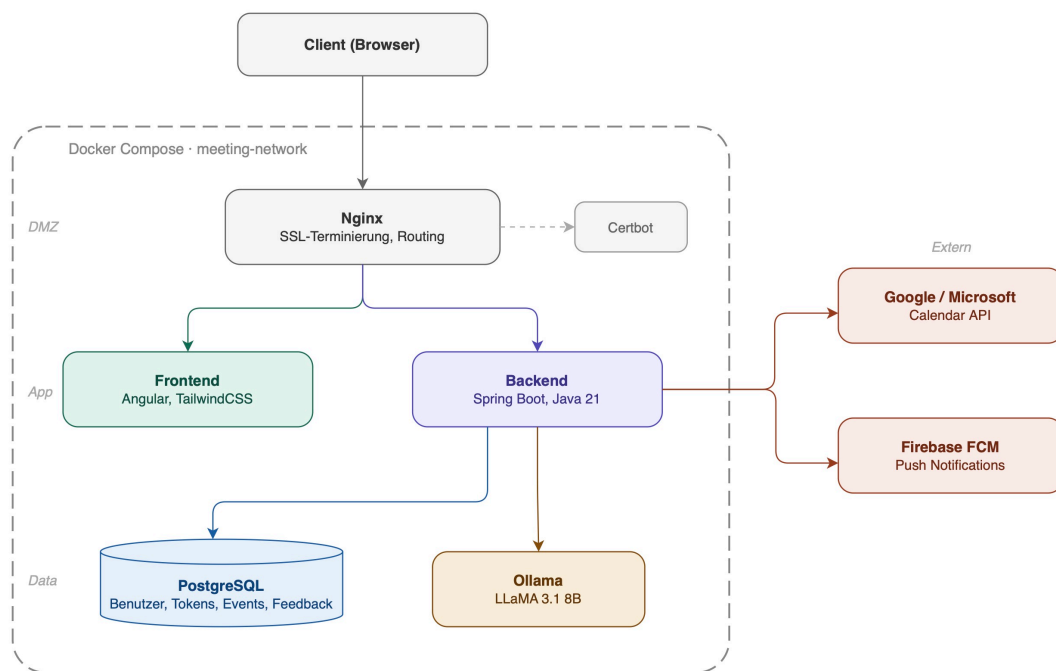


Abbildung 6: Architekturübersicht: Die sechs Docker-Services und ihre Kommunikationspfade.

Technisch ist die Applikation als Docker-basierte Architektur aus sechs Containern umgesetzt (vgl. Abbildung 6): einem Backend für Geschäftslogik und REST-API, einem Frontend für die Benutzeroberfläche, einer relationalen Datenbank, einem lokalen Sprachmodell-Container für die Meeting-Kategorisierung, einem Reverse-Proxy sowie einem Container für die automatisierte Zertifikatsverwaltung. Diese Containerisierung ermöglichte es, während der laufenden Feldstudie Anpassungen an einzelnen Komponenten auszurollen, ohne den Betrieb der übrigen zu unterbrechen (vgl. NFA-05). Die konkreten Technologieentscheidungen, das Zusammenspiel der Services und die Deployment-Pipeline werden in Abschnitt 6.6 beschrieben.

6.1.1. Authentifizierung und Benutzerverwaltung

Die Applikation setzt eine eigenständige Benutzerverwaltung um (vgl. FA-08 und Abbildung 7). Die Registrierung erfordert E-Mail, Benutzername und Passwort; Letzteres wird mit BCrypt gehasht in der Datenbank abgelegt. Nach erfolgreicher Anmeldung erhält das Frontend ein JSON Web Token (JWT), das bei jeder nachfolgenden API-Anfrage im **Authorization**-Header mitgesendet wird. Ein Angular HTTP Interceptor fügt das Token automatisch in alle ausgehenden Requests ein, sodass sich die Feature-Komponenten nicht um die Authentifizierungslogik kümmern müssen. Auf Backend-Seite validiert Spring Security das Token und extrahiert die Nutzeridentität, bevor die Anfrage an den jeweiligen Controller weitergereicht wird.

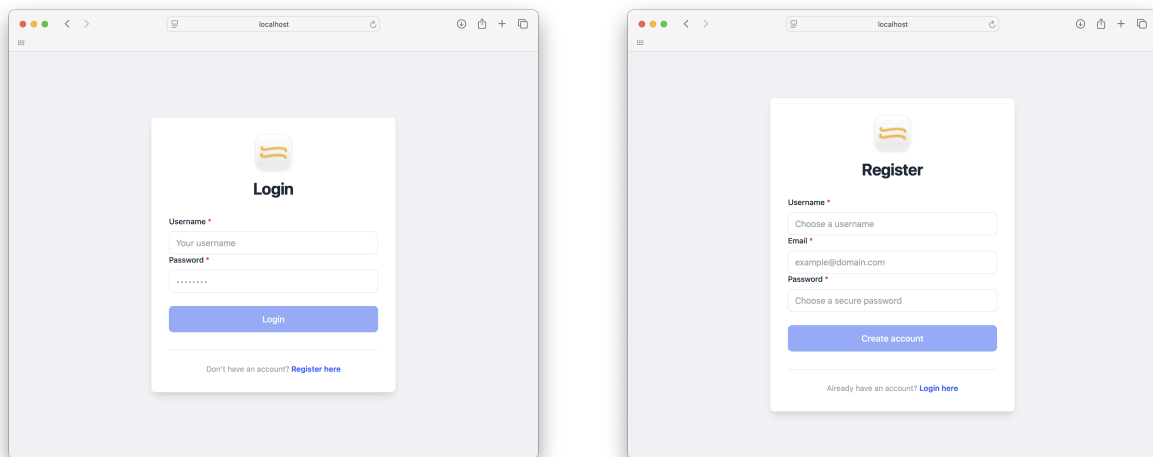


Abbildung 7: Login- und Registrierungs-Ansicht der Applikation im Vergleich.

6.1.2. Onboarding und Guided Tour

Erstmalige Nutzende werden durch einen dreistufigen Onboarding-Prozess in Form einer interaktiven Tour geführt, die mit der Bibliothek Shepherd.js implementiert ist (vgl. Abbildung 8). Diese Tour erklärt die App und die wichtigsten Konfigurationen kontextsensitiv, darunter die Kalenderverbindung, die Konfiguration der Arbeitszeiten und die Aktivierung von Push-Benachrichtigungen (vgl. Abbildung 9). Dies reduziert die Einstiegshürde für Studienteilnehmende, die das System zum ersten Mal nutzen. Über einen schwebenden Hilfe-Button kann die Tour zudem zu jedem Zeitpunkt neu gestartet werden.

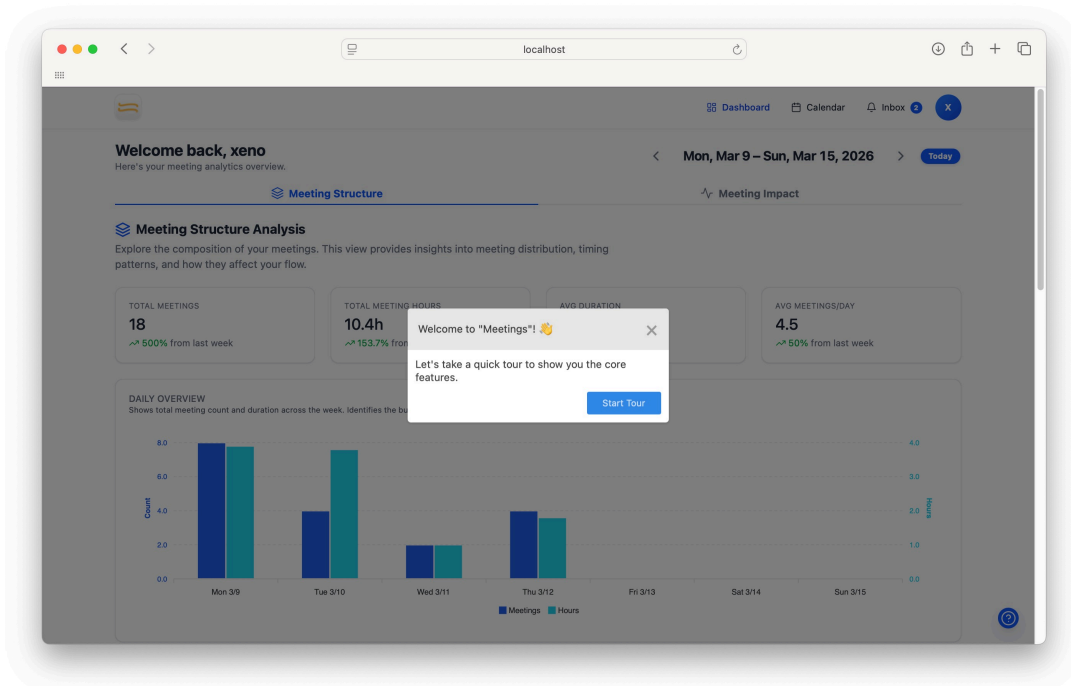


Abbildung 8: Die interaktive Guided Tour erklärt die Komponenten der Applikation schrittweise.

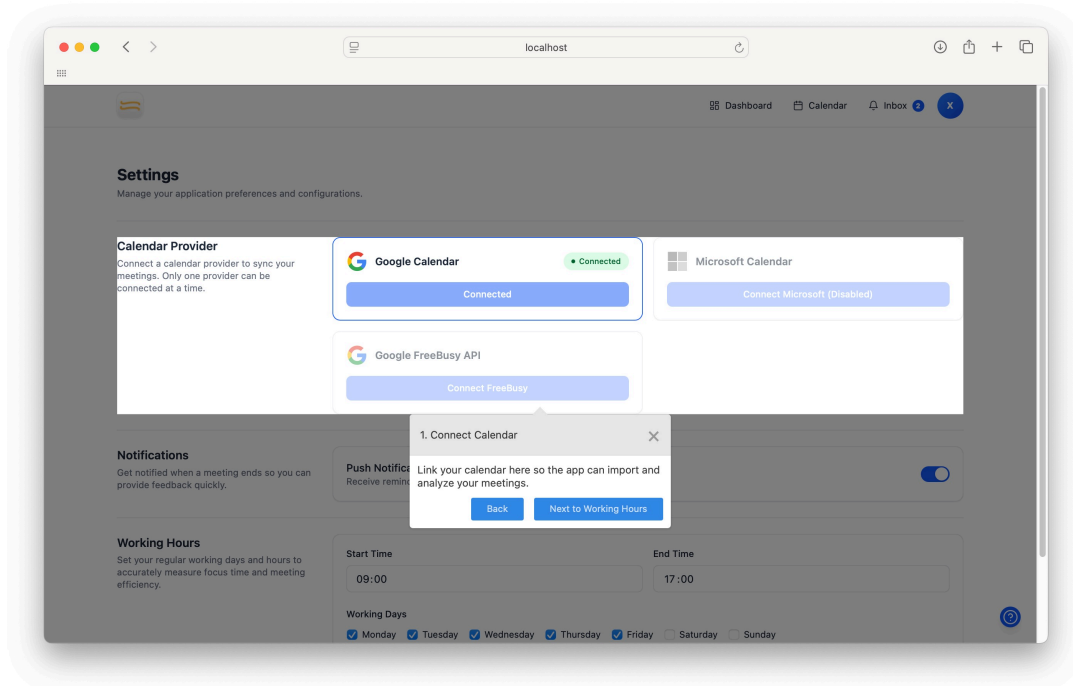


Abbildung 9: Die Tour führt den Nutzer durch die Grundkonfiguration der App.

6.2. Objektive Datengrundlage: Kalendersynchronisation

Die automatisierte Kalendersynchronisation liefert die objektiven Meeting-Eigenschaften des Forschungsmodells, also Anzahl, Dauer und Timing der Meetings. Durch die automatische Übernahme der Termine entfällt für die Probanden die manuelle Erfassung, was dem Leitprinzip der minimalen Disruption entspricht und die Datenqualität sichert (vgl. FA-01).

6.2.1. Anbindung externer Kalenderanbieter

Die Anbindung externer Kalenderanbieter erfolgt über den OAuth 2.0 Authorization Code Flow. Die Architektur ist über eine Provider-Abstraktion so konzipiert, dass mehrere Anbieter unterstützt werden können. Im Rahmen dieser Arbeit wurde die Integration mit Google Calendar vollständig umgesetzt; die geplante Anbindung von Microsoft Outlook konnte aus zeitlichen Gründen nicht realisiert werden, lässt sich jedoch ohne strukturelle Änderungen ergänzen.

Bei der Google-Anbindung stehen den Nutzenden zwei Integrationsstufen zur Verfügung. Die vollständige Integration synchronisiert Kalendereinträge einschliesslich ihrer Titel, was die automatische KI-Kategorisierung nach Meeting-Typ ermöglicht (vgl. Abschnitt 6.3). Alternativ bietet die Free/Busy-Integration eine datenschutzsensiblere Variante: Hier werden ausschliesslich Zeitfenster und Dauer der Termine übernommen, ohne deren Titel zu übertragen. Die objektiven Meeting-Eigenschaften des Forschungsmodells (Anzahl, Dauer, Timing) bleiben damit grundsätzlich erfassbar, die automatische Kategorisierung entfällt jedoch. Diese Abstufung ermöglicht es Nutzenden mit sensiblen Kalenderinhalten, das System zu verwenden, ohne projektbezogene Informationen preiszugeben.

Der konkrete Ablauf beginnt auf der Settings-Seite, wo Nutzende über einen Button den Verbindungsprozess starten und die gewünschte Integrationsstufe wählen. Das Frontend leitet an die Autorisierungs-URL von Google weiter, die das Backend über den `CalendarAuthController` generiert. Nach der Zustimmung durch die nutzende Person leitet der Anbieter an die konfi-

gurierte Redirect-URI (`/calendar-callback`) zurück. Das Backend tauscht den erhaltenen Autorisierungscode gegen ein Access- und ein Refresh-Token, die verschlüsselt in der Tabelle `user_oauth_tokens` persistiert werden (vgl. NFA-02). Eine Eindeutigkeitsbedingung auf `(user_id, provider)` stellt sicher, dass pro Anbieter nur ein aktiver Token-Satz existiert, und bildet gleichzeitig die Grundlage für die spätere Unterstützung weiterer Anbieter.

6.2.2. Hintergrund-Synchronisation

Ein Hintergrund-Job (`CalendarSyncScheduler`) synchronisiert im 10-Minuten-Intervall die Kalendereinträge aller registrierten Nutzenden über die jeweilige Anbieter-API. Die abgerufenen Events werden lokal in einer `events`-Tabelle gespeichert. Diese lokale Datenhaltung entkoppelt die Applikation von den Rate Limits der externen APIs, ermöglicht die Zuordnung applikationsspezifischer Metadaten (Meeting-Kategorie, Feedback-Status) und stellt sicher, dass die Dashboards auch bei temporärer API-Nichterreichbarkeit weiterhin Daten anzeigen können (vgl. NFA-03). Eine Eindeutigkeitsbedingung auf `(user_id, provider, external_id)` verhindert das Duplizieren bereits importierter Termine. Ergänzend können Nutzende über einen Refresh-Button jederzeit eine manuelle Synchronisation auslösen.

6.2.3. Kalenderansicht und manuelle Verwaltung

Die Kalenderansicht bietet eine interaktive Wochenübersicht der synchronisierten Termine (vgl. Abbildung 10). Nutzende können Termine betrachten, manuell erstellen, bearbeiten und löschen (vgl. FA-09). Beim Klick auf einen bestehenden Termin öffnet sich ein Popover mit den wichtigsten Optionen: Feedback geben, Kategorie korrigieren, Termin bearbeiten oder die Feedback-Aufforderung verwerfen. Letzteres ist besonders wichtig für automatisch synchronisierte Einträge, die keine eigentlichen Meetings darstellen, etwa geblockte Fokuszeiten oder ganztägige Abwesenheitseinträge.

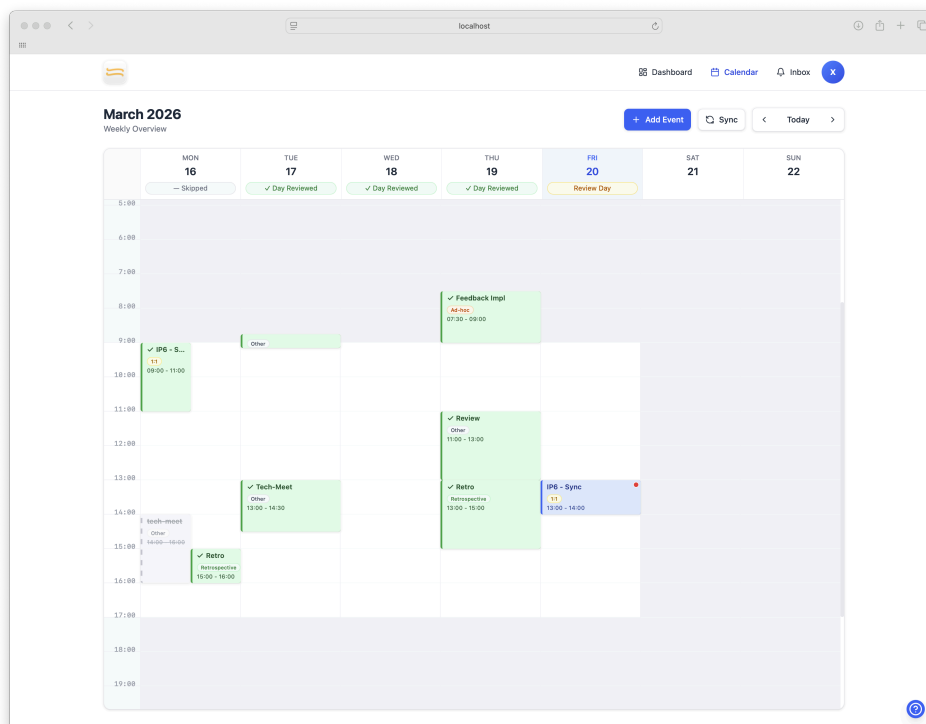


Abbildung 10: Wochenansicht des Kalenders mit Terminen.

6.3. KI-gestützte Kontextualisierung

Rohe Kalenderdaten allein unterscheiden nicht zwischen einem 15-minütigen Stand-up und einem mehrstündigen Planning-Workshop. Um die in der Konzeptlösung identifizierte Ineffizienz (vgl. P3) differenziert nach Meeting-Formaten analysieren zu können, kategorisiert das System die importierten Einträge automatisch (vgl. FA-02). Erst diese Kontextualisierung ermöglicht es den Dashboards, Muster nach Meeting-Typ aufzuzeigen, eine Voraussetzung, die reine Kalender-Analysetools typischerweise nicht erfüllen.

Die automatische Kategorisierung greift ausschliesslich bei vollständig synchronisierten Einträgen, deren Titel dem Modell als Eingabe dienen. Manuell erstellte Events werden nicht automatisch kategorisiert; stattdessen kann die Kategorie direkt im Erstellformular gesetzt werden. Einträge aus der Free/Busy-Integration, die keinen Titel enthalten, erhalten die Standardkategorie `OTHER`. Unabhängig vom Ursprung lässt sich die Kategorie jedes Eintrags jederzeit über die Kalenderansicht manuell überstimmen.

6.3.1. Modellwahl und Datenschutz

Die Kategorisierung basiert auf dem Sprachmodell LLaMA 3.1 (8B), das über einen lokalen Ollama-Container bereitgestellt wird. Die Entscheidung für ein lokales Modell war primär datenschutzgetrieben (vgl. NFA-07): Meeting-Titel enthalten häufig projektbezogene Informationen, Personennamen oder interne Begriffe, die nicht an externe Cloud-Dienste übertragen werden sollen. Um die Modelle objektiv zu vergleichen, wurde ein standardisierter Test-Prompt entwickelt (vgl. Anhang I), der 20 realistische Meeting-Titel unterschiedlicher Sprachen und Ambiguitätsstufen enthält und deren korrekte Zuordnung zu den sieben vordefinierten Kategorien prüft. Ein Modell galt als geeignet, wenn es in mehrfachen Durchläufen alle Titel konsistent und fehlerfrei klassifizierte. In einem Vergleich mehrerer kleinerer Modelle (LLaMA 3.2 1B, LLaMA 3.2 3B, Qwen2 1.5B, Gemma3 4B) konnte keines diesen Prompt erfolgreich verarbeiten und die Meeting-Titel zuverlässig den vordefinierten Kategorien zuordnen. LLaMA 3.1 (8B) war das kleinste getestete Modell, das diesen Test-Prompt bei mehrfachen Durchläufen zu 100 % erfüllte und die Kategorisierung konsistent korrekt durchführte. Da der Produktionsserver über keine GPU verfügt, läuft die Inferenz ausschliesslich auf der CPU, was den Ressourcenverbrauch zu einem zentralen Faktor macht. Das Modell belegt rund 7 GB Arbeitsspeicher und erreicht eine Inferenzgeschwindigkeit von etwa 80 Sekunden für 20 Events, was für die asynchrone Hintergrundverarbeitung ausreichend ist.

6.3.2. Kategorisierungsstrategie

Zur Kategorisierung wird ein *Large Language Model* (LLM) verwendet. Die LLM-Inferenz (Generierung von Vorhersagen oder Klassifikationen) kann bei grossen Kalenderbeständen mehrere Minuten dauern. Damit Nutzende beim ersten Öffnen nicht auf eine vollständige Kategorisierung warten müssen, kategorisiert der `MeetingCategorizationService` alle Meetings asynchron. Die unkategorisierten Einträge werden dabei vollständig im Hintergrund in Batches abgearbeitet.

Das Modell ordnet jeden Titel einer vordefinierten Kategorie zu: `STAND_UP`, `PLANNING`, `ONE_ON_ONE`, `TEAM_MEETING`, `WORKSHOP`, `AD_HOC` oder `OTHER`. Diese Kategorisierung ist die Voraussetzung dafür, dass die Dashboards Feedback-Scores und Fragmentierungsmuster nach Meeting-Typ aufschlüsseln können (vgl. FA-05, FA-06). Ergänzend können Nutzende die KI-Kategorisierung über die Kalenderansicht manuell korrigieren.

6.4. Erfassung der subjektiven Wahrnehmung

Die vorangehenden Abschnitte beschreiben, wie die objektive Datengrundlage entsteht. Für die Beantwortung der Forschungsfragen sind jedoch die subjektiven Bewertungen der Nutzenden entscheidend, da sie die subjektiven Meeting- und Arbeitstageigenschaften des Forschungsmodells operationalisieren. Dieser Abschnitt beschreibt, wie das System diese Bewertungen zeitnah und mit minimaler Disruption erfasst (vgl. FA-03, FA-04).

Die Befragungsstudie (vgl. Abschnitt 4) zeigt, dass eine Mehrheit der Befragten kurzes Post-Meeting-Feedback befürwortet und sich die Erfassung subjektiver Erfahrungen wünscht. Gleichzeitig identifiziert die Studie den unmittelbaren Post-Meeting-Zeitpunkt als am wenigsten disruptive Feedback-Option (vgl. Abschnitt 4.4.7). Die 30-Sekunden-Obergrenze für die Eingabe (vgl. NFA-01) und die prompte Benachrichtigung unmittelbar nach Meetingende (vgl. NFA-04) sind daher direkte Designentscheidungen aus den empirischen Befunden.

6.4.1. Push-Benachrichtigungen

Die situative Erfassung erfordert, dass Nutzende zum richtigen Zeitpunkt auf die Feedback-Möglichkeit hingewiesen werden (vgl. NFA-04). Hierfür wurde *Firebase Cloud Messaging* (FCM) eingesetzt, da es browserübergreifend Push-Benachrichtigungen unterstützt, keine native App voraussetzt und es dem Backend ermöglicht, einzelne Nutzende direkt und gezielt zu benachrichtigen. Nutzende können Push-Benachrichtigungen über die Einstellungen der Applikation aktivieren (vgl. FA-07). Bei der Aktivierung wird ein FCM-Token registriert, über den das Backend Benachrichtigungen an das jeweilige Endgerät senden kann.

6.4.2. Zwei komplementäre Feedback-Mechanismen

Das System implementiert zwei zeitgesteuerte Scheduler, die jeweils eine andere Ebene des Forschungsmodells abdecken. Diese Zweiteilung spiegelt die Struktur des Forschungsmodells wider: Die subjektiven Meeting-Eigenschaften werden auf der Mikro-Ebene einzelner Meetings erfasst, die subjektiven Arbeitstageigenschaften auf der Makro-Ebene des gesamten Arbeitstages (vgl. Abschnitt 3.1.1).

6.4.2.1. Meeting Feedback Instrument

Der `NotificationScheduler` prüft im Minutentakt, ob beendete Meetings existieren, für die noch kein Feedback vorliegt. Ist dies der Fall, wird eine Push-Benachrichtigung ausgelöst. Das Feedback wird über ein modales Formular erfasst (vgl. Abbildung 11), das vier obligatorische Items umfasst. Die Beschränkung auf vier Items ergibt sich aus der 30-Sekunden-Obergrenze (vgl. NFA-01). Ihre Reihenfolge folgt einem Trichter vom unmittelbar Spürbaren zum reflektierten Urteil: Zuerst wird die Fokusunterbrechung erhoben, da sie direkt nach einem Meeting am präsentesten ist, gefolgt von Stimmung und Energielevel als affektive Reaktionen, und abschliessend die Nützlichkeit (ROTI) als bewertende Gesamteinschätzung, die ein Mindestmass an Reflexion voraussetzt. Die vollständige Itemformulierung mit Skalen ist in Tabelle 13 dokumentiert, die Zuordnung der Items zu den Konstrukten des Forschungsmodells in Tabelle 1.

Ergänzend zu den vier Kernitems werden kontextabhängige Tags angeboten: Bei tiefen ROTI-Werten erscheinen vordefinierte Issue-Tags, bei hohen Werten Positive-Tags (vgl. Tabelle 14). Diese Tags ermöglichen eine qualitative Einordnung, ohne die Bearbeitungszeit wesentlich zu erhöhen.

The screenshot shows a web application interface with a 'Feedback Inbox' on the left and a 'Meeting Feedback' modal form in the center. The modal form is titled 'Meeting Feedback' and contains the following items:

- test** 20.03.2026, 00:15 - 00:30
- Did this meeting disrupt your focus? *** (Scale: 1 to 5, Not at all to Very much)
- How do you feel after this meeting? *** (Buttons: Negative, Neutral, Positive)
- How did this meeting affect your energy? *** (Scale: 1 to 5, Drained to Energized)
- This meeting was worth my time. *** (Scale: 1 to 5, Strongly disagree to Strongly agree)

At the bottom of the modal are 'Cancel' and 'Submit Feedback' buttons. The background shows a 'Feedback Inbox' with a list of meetings and a 'Show dismissed' toggle.

Abbildung 11: Meeting-Feedback-Formular mit ROTI, Stimmung, Energie und Tags.

6.4.2.2. End-of-Day Feedback Instrument

Der **DailyFeedbackScheduler** löst eine Stunde vor dem konfigurierten Arbeitsende das End-of-Day-Feedback aus (vgl. Abbildung 12). Diese zeitliche Platzierung ist bewusst gewählt: Ein zu früher Zeitpunkt würde den letzten Teil des Arbeitstages unberücksichtigt lassen, während die Benachrichtigung bei einem zu späten Versand nach Feierabend leicht übersehen werden könnte. Die Berücksichtigung der individuellen Arbeitszeiten (vgl. FA-07) stellt sicher, dass die Benachrichtigung jeweils zum passenden Zeitpunkt erfolgt.

Dieses Feedback erfasst über vier fünfstufige Likert-Skalen die subjektiven Arbeitstageigenschaften des Forschungsmodells. Die vollständige Itemformulierung ist in Anhang H.2 dokumentiert, die Konstruktuordnung in Tabelle 1.

The screenshot displays a web application interface for 'End of Day Feedback'. The main window is titled 'Feedback Inbox' and contains a list of recent meetings. A modal window is open, titled 'End of Day' for 'Friday, March 20, 2026'. The modal contains four Likert scales for rating the day's performance and energy levels. Each scale has five buttons labeled 1 to 5, with descriptive labels at the ends (e.g., 'Unproductive' to 'Productive'). The background shows a sidebar with 'Meeting Feedback' and 'End of Day Feedback' sections.

Question	1	2	3	4	5
How productive were you today? *	Unproductive				Productive
How much deep, focused work did you do today? *	None				A lot
How energized do you feel right now? *	Drained				Energized
How did today's meeting load feel? *	Too few				Too many

Abbildung 12: End-of-Day-Feedback zur Erfassung der subjektiven Tagesbewertung.

6.4.3. Feedback-Inbox als zentrale Sammelstelle

Alle ausstehenden Feedbacks, also Meeting-bezogene und tägliche, werden in einer Feedback-Inbox zusammengeführt (vgl. Abbildung 13). Die Inbox zeigt über ein Badge in der Navigationsleiste die Anzahl offener Feedback-Anfragen an. Diese Zentralisierung stellt sicher, dass kein Feedback übersehen wird, selbst wenn eine Push-Benachrichtigung nicht wahrgenommen wurde.

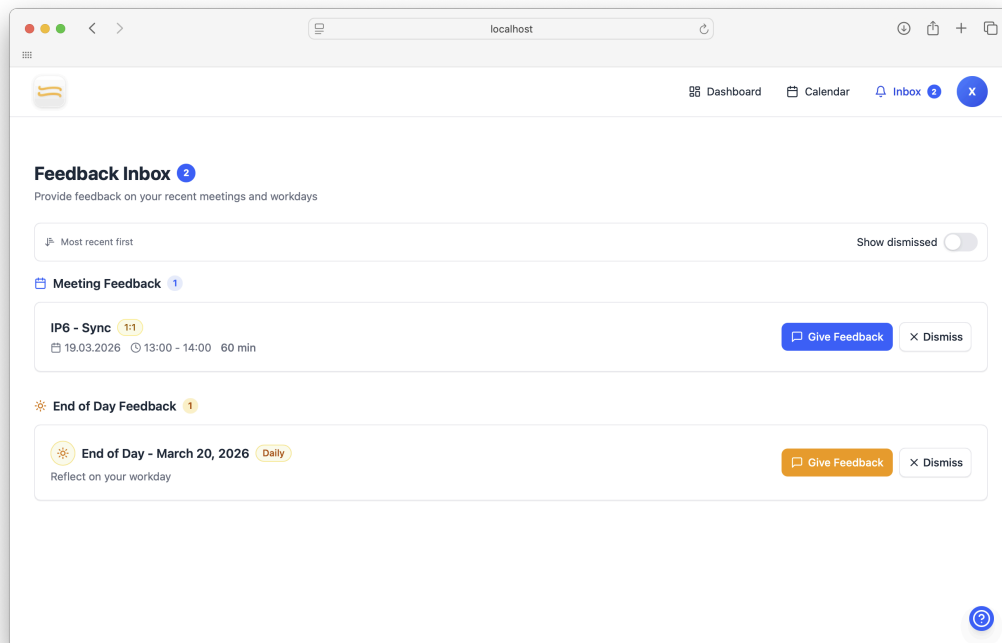


Abbildung 13: Die Feedback-Inbox zeigt alle noch nicht bewerteten Meetings und Arbeitstage.

6.4.4. Flexible Datenspeicherung

Die Feedback-Daten werden in **JSONB**-Spalten gespeichert, dem binären JSON-Format von PostgreSQL. Diese Designentscheidung weicht bewusst von der sonst durchgehend relationalen Modellierung ab: Da die Feedback-Formulare im Entwicklungsprozess iterativ angepasst wurden, etwa um zusätzliche Tags aufzunehmen, hätte eine strikte relationale Struktur bei jeder Anpassung eine Datenbankmigration erfordert. **JSONB** ermöglicht es, unterschiedliche Feedback-Strukturen flexibel abzulegen, ohne Schemaänderungen durchführen zu müssen (vgl. NFA-06).

6.5. Visualisierung und Reflexion: Dashboards

Die bisherigen Abschnitte beschreiben, wie objektive und subjektive Daten erhoben und angereichert werden. Die Dashboards überführen diese Rohdaten in eine Form, die Reflexion ermöglicht, und setzen damit PZ2 (Aufbereitung und Visualisierung der Daten zur Selbstreflexion) um. Die Applikation stellt zwei Dashboards bereit: Das Structure Dashboard visualisiert die Tagesstruktur und das Flow-Potenzial (vgl. FA-05), das Impact Dashboard die subjektiven Bewertungen nach Meeting-Typ (vgl. FA-06). Das Backend stellt über den **DashboardController** aggregierte Daten-Transfer-Objekte bereit, die im Frontend durch den **DataProcessorService** in die von ApexCharts erwarteten Formate transformiert werden. Ein gemeinsamer Wochenauswahl-Selektor ermöglicht die Navigation zwischen Kalenderwochen.

6.5.1. Structure Dashboard

Das Structure Dashboard visualisiert die objektiven Merkmale des Arbeitstages (vgl. Abbildung 15). Vier Kennzahlenkarten fassen die wöchentliche Meeting-Last zusammen: Gesamtzahl der Meetings, summierte Meeting-Dauer in Stunden, durchschnittliche Dauer pro Meeting sowie durchschnittliche Anzahl Meetings pro Arbeitstag. Jede Karte zeigt zusätzlich die prozentuale Veränderung gegenüber der Vorwoche an, sodass Trends auf einen Blick erkennbar sind.

Das Daily Overview stellt die tägliche Meeting-Anzahl und -Dauer als gruppiertes Balkendiagramm dar und macht Spitzenbelastungen innerhalb der Woche sichtbar. Die Weekly Meeting Intensity ergänzt diese zeitliche Perspektive: Eine Heatmap schlüsselt die Meeting-Minuten nach Tagesstunden auf und zeigt, welche Zeitfenster besonders stark belegt sind. Die Heatmap wurde einer tabellarischen Darstellung vorgezogen, da sie Häufungsmuster über zwei Dimensionen (Wochentag und Tageszeit) simultan sichtbar macht, die in einer linearen Auflistung schwer erkennbar wären. Zwei weitere Diagramme beleuchten die Zusammensetzung der Meetings: Die Meeting Type Distribution zeigt als Balkendiagramm die Verteilung über die Kategorien, während die Duration Breakdown als Donut-Chart den Anteil kurzer (< 30 min), mittlerer (30–60 min) und langer Meetings (> 60 min) gegenüberstellt.

Die zentrale Visualisierung ist die Daily Flow Potential Analyse. Für jeden Arbeitstag wird eine Prozentzahl berechnet, die angibt, wie viel zusammenhängende Zeit für das Erreichen eines Flow-Zustands innerhalb der konfigurierten Arbeitszeiten zur Verfügung steht.

Die Berechnung basiert auf dem Konzept der Flow Disruption Cost. Für jeden freien Zeitblock b_i nach einem Meeting wird ein pauschaler Kontextwechsel-Abzug $c = 25$ Minuten vorgenommen. Der verfügbare Flow-Anteil eines Blocks ergibt sich als $f(b_i) = \max(0, b_i - c)$. Der tägliche Flow Potential Score aggregiert diese Werte über alle n freien Blöcke innerhalb der konfigurierten Arbeitszeit W :

$$\text{Flow Potential} = \frac{\sum_{i=1}^n f(b_i)}{W} \cdot 100\%$$

Abbildung 14: Formel welche zur berechnung des Flow Potential genutzt wurde.

Der in Abbildung 14 angewandte Abzugswert von 25 Minuten stützt sich auf Erkenntnisse der Unterbrechungsforschung, die zeigen, dass nach einem Kontextwechsel eine erhebliche Anlaufzeit benötigt wird, bevor konzentriertes Arbeiten wieder möglich ist (vgl. Abschnitt 2.2.2). Da die Literatur je nach Unterbrechungsart und Aufgabenkomplexität unterschiedliche Werte berichtet, wurde ein konservativer Richtwert gewählt, der den Produktivitätsverlust durch Meeting-bedingte Kontextwechsel abbildet.

Ergänzend visualisiert ein Daily Flow Schedule den Arbeitstag als farbcodierte Zeitleiste, auf der Flow-Blöcke unterschiedlicher Länge, fragmentierte Restzeiten und Meetings visuell unterscheidbar sind. Eine aufklappbare Detailansicht schlüsselt die einzelnen Segmente mit Gesamtzeit, Flow Disruption Cost und Net Flow Potential tabellarisch auf. Damit setzt das Dashboard die Anforderung FA-05 um: Fragmentierung, Flow-Blöcke und Meeting-Typen werden visualisiert und deren Zusammenspiel quantifizierbar gemacht.

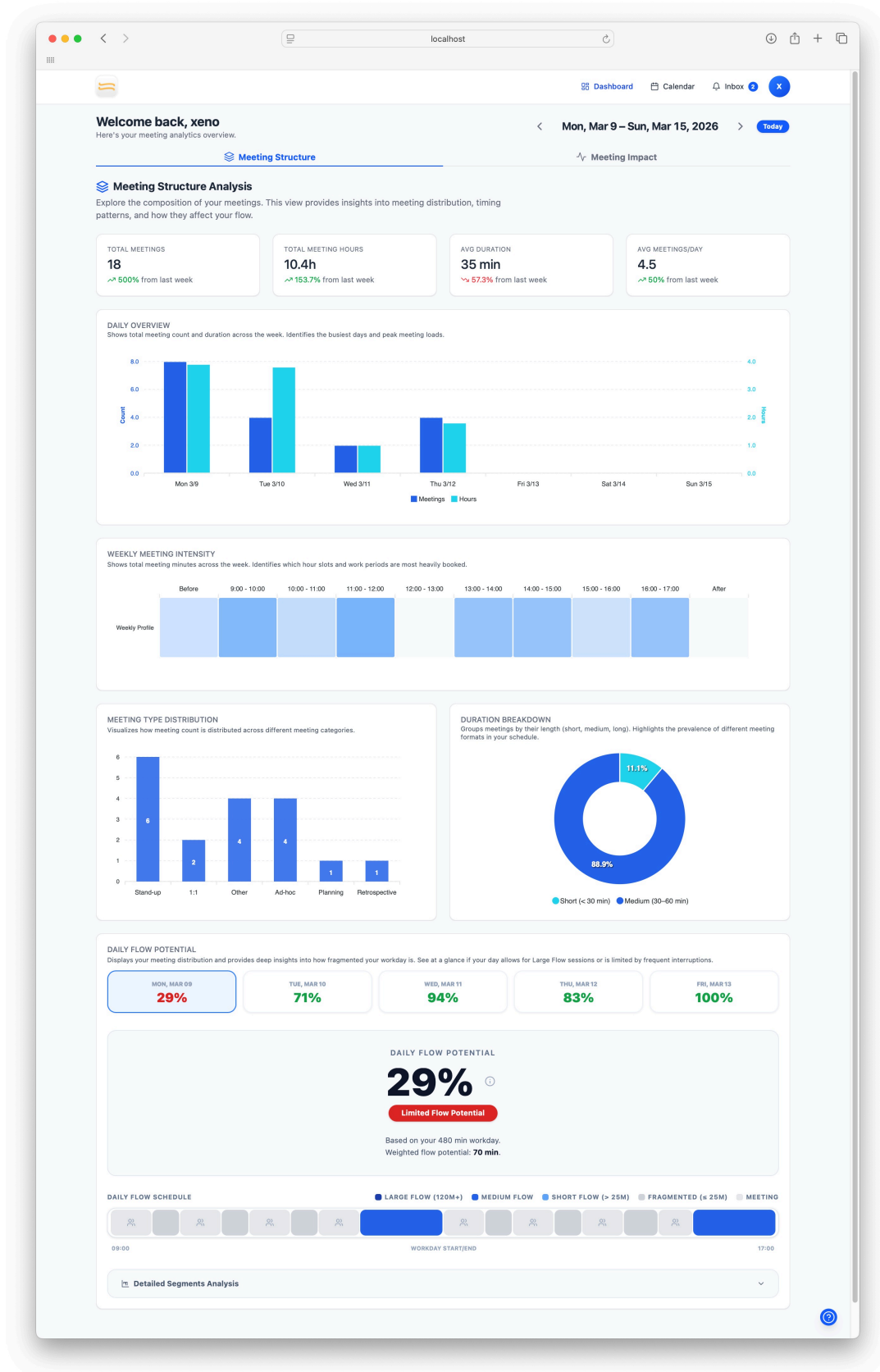


Abbildung 15: Das Structure Dashboard visualisiert die objektiven Merkmale des Arbeitstages.

6.5.2. Impact Dashboard

Das Impact Dashboard verbindet die subjektiven Bewertungen mit den objektiven Meeting-Daten (vgl. Abbildung 16). Vier Kennzahlenkarten geben einen wöchentlichen Überblick: den durchschnittlichen ROTI-Wert, die vorherrschende Stimmung nach Meetings (Feeling After), das durchschnittliche Energielevel sowie den Anteil der als wertvoll eingestuften Meetings. Wie beim Structure Dashboard zeigt jede Karte die prozentuale Veränderung gegenüber der Vorwoche.

Die Impact Timeline visualisiert den Verlauf von ROTI, Energielevel und Fokusunterbrechung als Liniendiagramm über die Tagesstunden der Arbeitswoche. Damit wird erkennbar, zu welchen Tageszeiten Meetings als besonders nützlich oder belastend wahrgenommen werden und ob sich bestimmte Muster im Tagesverlauf wiederholen.

Die Impact Analysis by Meeting Type stellt die vier Feedback-Dimensionen ROTI, Energy, Disruption und Feeling als farbcodierte Heatmap dar, aufgeschlüsselt nach den KI-kategorisierten Meeting-Typen. Die Heatmap wurde hier einem Linien- oder Balkendiagramm vorgezogen, da der primäre Erkenntniszweck nicht der zeitliche Verlauf, sondern der kategorienübergreifende Vergleich mehrerer Dimensionen gleichzeitig ist. Die Farbkodierung reicht von Grün über Orange bis Rot und ermöglicht einen direkten visuellen Vergleich: Nutzende erkennen auf einen Blick, welche Meeting-Formate konsistent positive Bewertungen erhalten und welche als belastend wahrgenommen werden.

Ergänzend aggregiert die Meeting-Characteristics-Ansicht die qualitativen Tags aus den Meeting-Feedbacks als horizontales Balkendiagramm. Positive und negative Tags werden einander gegenübergestellt und machen wiederkehrende Stärken und Schwächen der Meeting-Kultur sichtbar, die über numerische Scores hinausgehen. Das Impact Dashboard setzt damit FA-06 um: Feedback-Scores werden aggregiert und in Bezug zur Meeting-Struktur gesetzt.

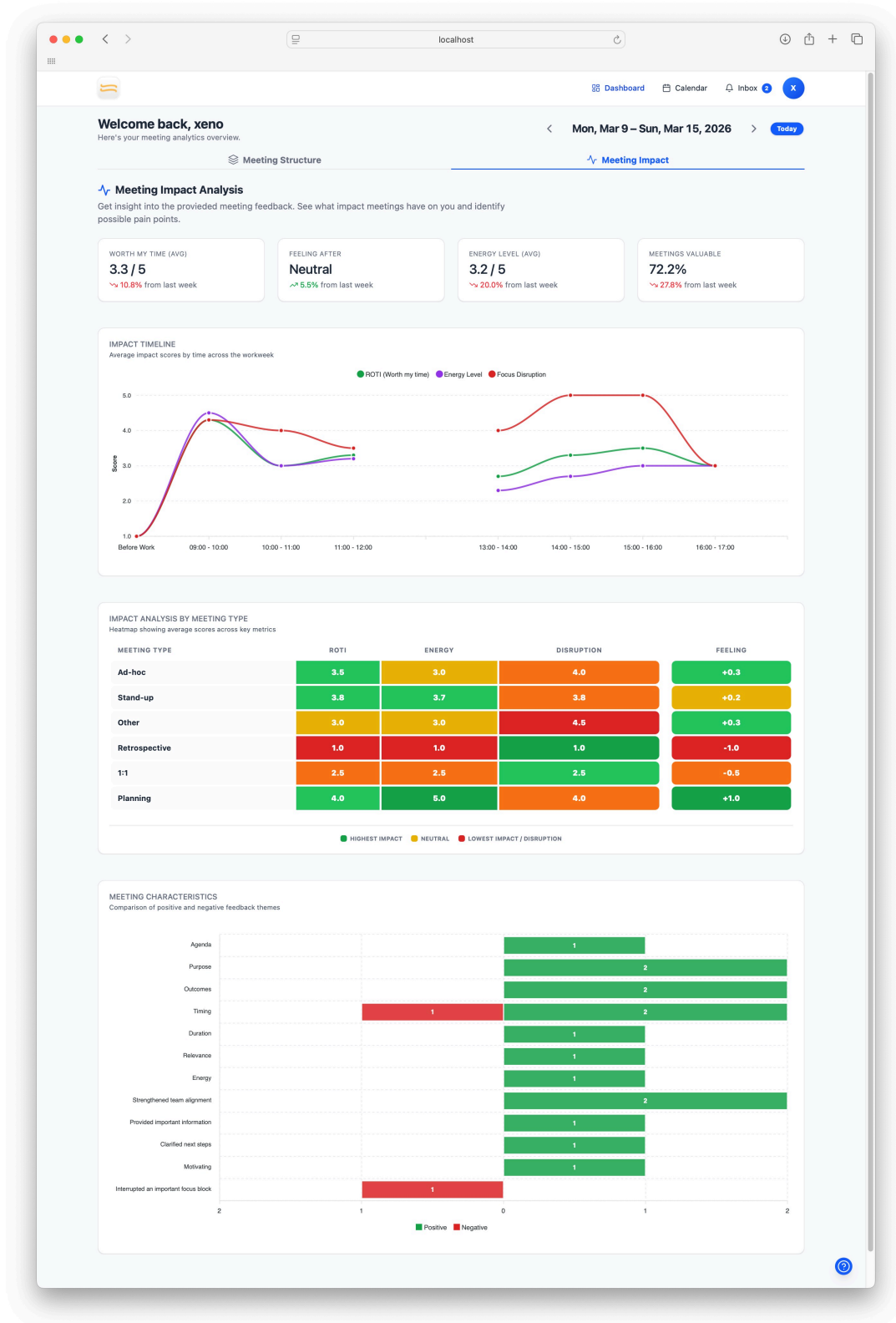


Abbildung 16: Das Impact Dashboard visualisiert die subjektiven Merkmale des Arbeitstages.

6.6. Technische Architektur und Betrieb

Die bisherigen Abschnitte beschreiben die fachlichen Subsysteme entlang des Datenflusses. Dieser Abschnitt dokumentiert die technischen Architekturentscheidungen, die den zuverlässigen Betrieb während der Feldstudie sicherstellten. Ausfälle oder Datenverluste während des Erhebungszeitraums würden die Validität der gesammelten Daten gefährden.

6.6.1. Containerisierung und Netzwerk

Die sechs Docker-Container kommunizieren über ein internes Netzwerk (`meeting-network`). Der Nginx-Reverse-Proxy ist der einzige öffentlich exponierte Eintrittspunkt: API-Anfragen (`/api/`) werden an das Backend weitergeleitet, alle übrigen Anfragen an das Frontend. Nginx gibt bei 404-Fehlern stets `index.html` zurück, um das clientseitige Routing der Angular SPA zu ermöglichen. Persistente Daten sind über Docker Volumes gesichert. Das Backend kommuniziert intern mit PostgreSQL und dem Ollama-Container sowie extern mit den Kalender-Providern für die Synchronisation und mit Firebase Cloud Messaging für Push-Benachrichtigungen. Diese Entkopplung in spezialisierte Container erlaubte es, während der Feldstudie Anpassungen am Frontend auszurollen, ohne den Backend-Prozess und damit die laufenden Hintergrund-Jobs für Synchronisation und Benachrichtigungen zu unterbrechen (vgl. NFA-05).

6.6.2. SSL und Verschlüsselung

Der Nginx-Proxy terminiert sämtlichen HTTPS-Verkehr (vgl. NFA-02). Die SSL-Zertifikate werden über Let's Encrypt automatisiert bezogen und durch einen dedizierten Certbot-Container regelmässig erneuert. OAuth-Tokens werden verschlüsselt in der Datenbank persistiert.

6.6.3. Frontend

Das Frontend ist als Angular 20 Single Page Application mit einer komponentenbasierten Architektur implementiert. Die Projektstruktur folgt dem Feature-Module-Pattern: Jede fachliche Domäne (Authentifizierung, Kalender, Dashboard, Feedback-Inbox, Settings) ist als eigenständiges Feature-Modul unter `src/app/features/` organisiert. Querschnittsbelange wie Authentifizierungs-Guards, HTTP-Interceptors und gemeinsame Datenmodelle sind in `src/app/core/` gebündelt, wiederverwendbare UI-Bausteine in `src/app/shared/`. Für die Visualisierung setzt das Frontend auf `ng-apexcharts`. Die Styling-Schicht basiert auf Tailwind CSS 4 in Kombination mit der Komponentenbibliothek Preline UI.

6.6.4. REST-API-Schnittstelle

Die Kommunikation zwischen Frontend und Backend erfolgt über eine RESTful API unter dem Basispfad `/api`. Alle Endpunkte ausser Login und Registrierung erfordern ein gültiges JWT im Authorization-Header. Die API gliedert sich in sechs Ressourcengruppen: Authentifizierung (`/api/auth`), Benutzereinstellungen (`/api/user/settings`), Kalender (`/api/calendar`) mit OAuth-Flow und CRUD-Operationen auf Kalendereinträge, Meeting-Feedback (`/api/feedback`), Daily-Feedback (`/api/daily-feedback`) sowie Dashboard-Endpunkte (`/api/dashboard`) für die aggregierten Daten beider Dashboard-Ansichten.

6.6.5. CI/CD-Pipeline

Die Continuous-Deployment-Pipeline ist über GitHub Actions implementiert und wird bei jedem Push auf den `main`-Branch ausgelöst (vgl. NFA-05). Der Workflow läuft auf einem Self-Hosted Runner direkt auf dem Produktionsserver, der direkten Zugriff auf das Docker-Ökosystem hat. Der Workflow checkt den aktuellen Code aus, schreibt die Repository-Secrets in eine `.env`-Datei, baut alle Container neu und startet den Stack. So konnten während des laufenden Betriebs Fehlerkorrekturen zeitnah ausgeliefert werden.

6.7. Datenmodell

Das relationale Datenmodell (vgl. Abbildung 17) bildet das Rückgrat der gesamten Applikation und verbindet alle zuvor beschriebenen Subsysteme. Es ist um die zentrale Tabelle `users` strukturiert, von der alle weiteren Entitäten abhängen. Die Schema-Evolution wird über Flyway-Migrationen verwaltet, die beim Applikationsstart automatisch ausgeführt werden.

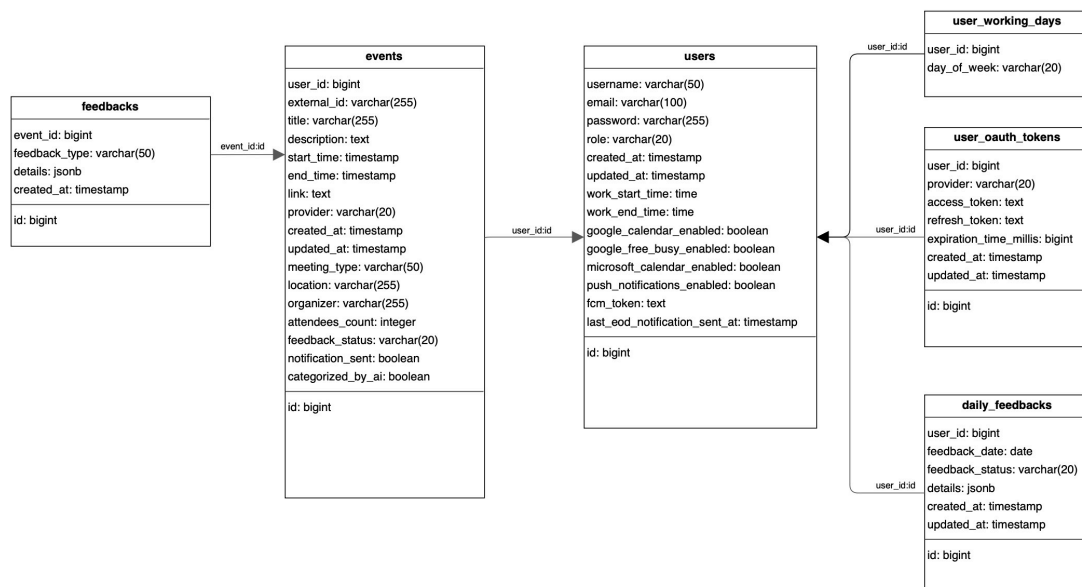


Abbildung 17: Das Datenmodell Unified Modeling Language (UML) Notation.

Ein zentrales Designprinzip sind kaskadierende Löschungen (`ON DELETE CASCADE`): Wird ein Nutzerkonto gelöscht, werden automatisch alle zugehörigen Tokens, Events, Feedbacks und Daily Feedbacks bereinigt. Wird ein synchronisierter Kalendereintrag extern gelöscht und die Löschung synchronisiert, wird auch das zugehörige Feedback entfernt. Diese Kaskadierung stellt die referenzielle Integrität sicher und verhindert verwaiste Datensätze.

7. Validierung

Vor dem Start von Abschnitt 8 wurde die Applikation einer systematischen Validierung unterzogen. Dieses Kapitel dokumentiert zwei komplementäre Validierungsschritte. Zunächst wurde in einem moderierten Usability-Test (Abschnitt 7.1) geprüft, ob die Kerninteraktionen der Applikation für die Zielgruppe verständlich und effizient nutzbar sind. Anschliessend wurden sämtliche in Abschnitt 5.7 definierten funktionalen und nicht-funktionalen Anforderungen gegen ihre jeweiligen Abnahmekriterien geprüft (Abschnitt 7.2). Die Ergebnisse beider Schritte flossen in konkrete Verbesserungen ein, die vor Beginn der Feldstudie umgesetzt wurden.

7.1. Usability-Test

Der Usability-Test fokussiert auf die Kernfunktionen der Datenerfassung: Onboarding, Kalender-Setup und Feedback-Eingabe. Diese Funktionen bilden die operative Grundlage von PZ1 und müssen eine nahtlose und unaufdringliche Integration in den Arbeitsalltag gewährleisten, ohne den Flow der Entwickelnden zusätzlich zu fragmentieren.

Die Dashboards und Reflexionsmechanismen (PZ2) wurden bewusst nicht im Labor evaluiert. Die Applikation ist als Reflexionsinstrument im Sinne der *Lived Informatics* konzipiert: Die erhobenen Daten gewinnen ihre Bedeutung erst im persönlichen Arbeitskontext der Nutzenden. Rooksby et al. zeigen, dass Tracking-Daten im Kontext ihrer Entstehung bedeutungsvoll sind, diese Bedeutung aber verlieren, wenn sie davon losgelöst betrachtet werden [33]. Ein aufgabenbasierter Test mit synthetischen Daten würde daher die ökologische Validität der Evaluation stark einschränken. Die Evaluation von PZ2 erfolgt deshalb im Rahmen der Feldstudie und wird in Abschnitt 9.4 diskutiert.

7.1.1. Zielsetzung

Das primäre Ziel dieses Usability-Tests ist die Überprüfung von PZ1 (Kontextbezogene Datenerhebung mit minimaler Störung des Arbeitsalltags). Dabei werden drei funktionale Teilkomponenten geprüft: der Onboarding- und Kalender-Verknüpfungsprozess, die Interaktion mit der Feedback-Inbox (Meeting Feedback) sowie die Eingabe der täglichen Selbsteinschätzung (End-of-Day-Feedback).

Zusätzlich wird die nicht-funktionale Anforderung NFA-01 überprüft, die verlangt, dass mindestens 80 % der Testpersonen eine Meeting-Feedback-Eingabe in unter 30 Sekunden abschliessen.

7.1.2. Methodik und Setup

Für die Überprüfung von PZ1 wurde ein moderiertes, aufgabenbasiertes Usability-Testing gewählt. Im Folgenden werden das Testdesign, die Stichprobe und die eingesetzten Messinstrumente beschrieben.

7.1.2.1. Testdesign

Die Tests wurden in Form eines moderierten, aufgabenbasierten Usability-Testings durchgeführt. Dabei kam die Methode des *Concurrent Thinking Aloud* zum Einsatz: Die Testpersonen wurden zu Beginn der Session instruiert, ihre Gedankengänge, Erwartungen und mögliche Irritationen während der Durchführung fortlaufend laut auszusprechen [34]. Diese Methode zeigt nicht nur auf, wo Nutzende Schwierigkeiten haben, sondern auch warum, und minimiert nachträgliche Erinnerungsverzerrungen, die bei retrospektiven Befragungen auftreten können.

Die Tests fanden remote via Screensharing statt. Dieser Ansatz gewährleistet eine hohe ökologische Validität, da die Softwareentwickelnden das System in ihrer gewohnten Arbeitsumgebung

und auf ihrem eigenen Setup testeten. Gerade bei einer Browser-basierten Applikation für Entwickelnde entspricht dies dem realen Nutzungskontext. Die Sessions wurden mit Einverständnis der Teilnehmenden aufgezeichnet. Die Testleitung protokollierte während der Durchführung Beobachtungen und Auffälligkeiten. Jede Testsession dauerte circa 20 bis 30 Minuten, bestehend aus einer kurzen Einführung in die Thinking-Aloud-Methode, der Durchführung der drei Aufgaben und dem Ausfüllen des SUS-Fragebogens.

Die Testumgebung wurde mithilfe automatisierter API-Aufrufe aufgebaut, die für jede Testperson eine standardisierte Ausgangslage mit realistischen Kalendereinträgen erzeugten (siehe Anhang I.1).

7.1.2.2. Probanden

Für das Usability-Testing wurden $N = 5$ Softwareentwickelnde rekrutiert. Alle Personen sind beruflich zu mindestens 50 % in der Softwareentwicklung tätig und repräsentieren die definierte Zielgruppe. Die rein männliche Stichprobe befand sich im Alter zwischen 23 und 32 Jahren und überschneidet sich nicht mit den Teilnehmenden der späteren Feldstudie (Abschnitt 8). Die Stichprobengröße orientiert sich am Poisson-Modell von Nielsen und Landauer, das bei einem mittleren λ -Wert von .31 vorhersagt, dass fünf Testpersonen rund 85 % der Usability-Probleme aufdecken. Der exakte Wert von λ variiert je nach System und Evaluationsmethode, sodass die 85 % als Richtwert und nicht als garantierte Abdeckung zu verstehen sind [35].

7.1.3. Test-Szenarien und Aufgaben

Die Interaktion mit dem System wurde in drei aufeinanderfolgende Kernaufgaben unterteilt, die den realen Nutzungsworkflow abbilden.

7.1.3.1. Aufgabe 1: Onboarding und System-Setup

Szenario: Sie nutzen die Applikation zum ersten Mal und möchten diese für Ihren Arbeitsalltag einrichten.

Aufgabe: Führen Sie die Registrierung durch, verknüpfen Sie Ihren Arbeitskalender über die OAuth-Schnittstelle und stellen Sie Ihre regulären Arbeitszeiten ein.

Fokus: Verständlichkeit des Authentifizierungs-Flows und Auffindbarkeit der Konfigurationen.

7.1.3.2. Aufgabe 2: Meeting-Bewertung

Szenario: Sie kommen gerade aus einem Meeting zurück an Ihren Arbeitsplatz.

Aufgabe: Öffnen Sie die Feedback-Inbox und geben Sie eine Bewertung für das letzte Meeting ab. Nutzen Sie dabei auch die vorgeschlagenen kontextabhängigen Tags.

Fokus: Schnelligkeit und Klarheit der Feedback-Eingabe (ROTI, Fokusunterbrechung, Stimmung, Energielevel) zur Wahrung des Leitprinzips der minimalen Disruption. Die benötigte Zeit wird gemessen, um NFA-01 zu überprüfen.

7.1.3.3. Aufgabe 3: End-of-Day-Bewertung

Szenario: Ihr Arbeitstag neigt sich dem Ende zu.

Aufgabe: Navigieren Sie zum End-of-Day-Feedback und füllen Sie die Bewertung für den heutigen Tag aus.

Fokus: Verständlichkeit der vier Tages-Dimensionen (Produktivität, Deep Work, Energie, Meetingbelastung).

7.1.4. Ergebnisse

Aufgabenabschlussrate: Alle fünf Probanden konnten sämtliche drei Aufgaben erfolgreich abschliessen (Abschlussrate: 100 %).

Aufgabe 1: Onboarding und System-Setup: Die integrierte Onboarding-Tour wurde von allen fünf Probanden positiv hervorgehoben und erleichterte die Orientierung im System. Bei zwei Probanden, die ihren Browser im Dark Mode nutzten, fiel auf, dass nicht alle UI-Komponenten den Dark Mode sauber unterstützen. Weiter wurde beobachtet, dass drei von fünf Probanden die Arbeitstage (*Working Days*) vor den Arbeitszeiten konfigurierten, obwohl die Reihenfolge im UI die umgekehrte Abfolge vorsieht. Bei einem Probanden zeigte der Dialog zum manuellen Erfassen von Meetings das amerikanische Datumsformat (MM/DD/YYYY) an, dies war jedoch auf dessen individuelle Browsereinstellung zurückzuführen und kein Fehler der Applikation.

Aufgabe 2: Meeting-Bewertung: Alle Probanden fanden die Möglichkeit, Feedback zu Events abzugeben, direkt über die Kalender-View. Zwei Probanden hatten jedoch Mühe, die Aufgabe über die Feedback-Inbox zu lösen. Als Verbesserungsvorschlag wurde eine Counter-Bubble in der Navbar genannt, wie sie von mobilen Apps bekannt ist, um offene Feedbacks zu signalisieren. Ein Proband navigierte direkt zur Inbox-View und bezeichnete diese spontan als seine TODO liste, was auf eine korrekte mentale Modellbildung hinweist. Die Ausfüllzeit für ein Meeting-Feedback lag bei allen Probanden zwischen 20 und 30 Sekunden und damit unter der in NFA-01 definierten Schwelle von 30 Sekunden.

Aufgabe 3: End-of-Day-Bewertung: Ein Proband hatte Schwierigkeiten, das End-of-Day-Feedback innerhalb der Inbox aufzufinden. Die übrigen vier Probanden bewältigten die Aufgabe ohne Verzögerung. Die Möglichkeit, das Tagesend-Feedback direkt über den Kalender abzugeben, wurde von den Probanden als sehr intuitiv bezeichnet. Die durchschnittliche Ausfüllzeit betrug circa 25 Sekunden.

7.1.4.1. System Usability Scale (SUS)

Im Anschluss an die Aufgaben füllten die Probanden den standardisierten System Usability Scale (SUS) Fragebogen aus [36]. Der SUS umfasst zehn Items auf einer fünfstufigen Likert-Skala und liefert einen Score zwischen 0 und 100.

Der aus den fünf Fragebögen berechnete SUS-Score beträgt **82.56**. Auf der von Bangor et al. vorgeschlagenen Skala liegt dieser Wert im oberen Bereich des Ratings „*Good*“ ($M = 72.75$) und nähert sich dem Rating „*Excellent*“ ($M = 85.58$) [37]. Gemäss den in derselben Studie definierten Akzeptanzbereichen fällt der Score deutlich in den Bereich „*Acceptable*“ ($SUS > 70$). Angesichts der kleinen Stichprobe ($N = 5$) ist der Wert als positiver Indikator zu interpretieren, nicht als inferenzstatistisch belastbarer Befund. Dennoch stützt er die Einschätzung, dass die Kerninteraktionen der Applikation für die Zielgruppe verständlich und effizient nutzbar sind.

7.1.5. Abgeleitete Massnahmen

Auf Basis der im Usability-Test identifizierten Probleme wurden vor Beginn von Abschnitt 8 konkrete Verbesserungen umgesetzt. Tabelle 4 fasst die Massnahmen zusammen.

Identifiziertes Problem	Abgeleitete Massnahme	Status
Nicht alle UI-Komponenten unterstützen den Browser-Dark-Mode sauber.	Konsistente Unterstützung des Dark-Mode-Themes über alle UI-Komponenten sichergestellt.	Umgesetzt vor Abschnitt 8
Drei von fünf Probanden füllen die Arbeitstage vor den Arbeitszeiten aus, entgegen der im UI vorgesehenen Reihenfolge.	Reihenfolge der Onboarding-Schritte angepasst, sodass <i>Working Days</i> vor der Arbeitszeitkonfiguration abgefragt werden.	Umgesetzt vor Abschnitt 8
Zwei Probanden finden die Feedback-Inbox nicht intuitiv; der Zugang über die Kalender-View wird bevorzugt.	Counter-Bubble in der Nav-bar ergänzt, um die Anzahl ausstehender Feedbacks sichtbar zu signalisieren.	Umgesetzt vor Abschnitt 8
Ein Proband findet das End-of-Day-Feedback in der Inbox nicht.	Visuelle Differenzierung des End-of-Day-Feedbacks in der Inbox durch ein eigenes Icon und eine farbliche Hervorhebung.	Umgesetzt vor Abschnitt 8

Tabelle 4: Abgeleitete Massnahmen aus dem Usability-Test.

7.2. Anforderungsvalidierung

Ergänzend zum Usability-Test wurden sämtliche in Abschnitt 5.7 definierten Anforderungen systematisch gegen ihre Abnahmekriterien geprüft. Die Validierung erfolgte vor dem Start von Abschnitt 8, um sicherzustellen, dass das System als Erhebungsinstrument zuverlässig einsetzbar ist. Die folgenden Abschnitte dokumentieren für jede Anforderung die gewählte Validierungsmethode, das Ergebnis und den Status.

Die Validierungsergebnisse der aus dem Usability-Test ableitbaren Anforderungen wurden eingetragen. Anforderungen, die eine separate technische Überprüfung erfordern (z. B. Konfigurationsreviews, Lasttests oder Architekturanalysen), sind mit dem jeweiligen ausstehenden Ergebnis gekennzeichnet.

7.2.1. Funktionale Anforderungen

Tabelle 5 zeigt die Validierungsergebnisse der neun funktionalen Anforderungen.

Anforderung	FA-01: Kalendersynchronisation über OAuth 2.0
Validierungsmethode	Es wurde ein neuer Event im Google Kalender erstellt. Microsoft wurde nicht getestet (vgl. Abschnitt 6.2.1).
Ergebnis	Der neue Event wurde innert 5 Minuten im Kalender angezeigt.
Status	Teilweise erfüllt (Microsoft nicht umgesetzt, Google voll funktionsfähig)
Anforderung	FA-02: Automatische Meeting-Kategorisierung
Validierungsmethode	Test mit vordefiniertem Datensatz von 18 Events in einer Woche.
Ergebnis	Alle 18 Events wurden richtig kategorisiert.
Status	Erfüllt
Anforderung	FA-03: Push-Benachrichtigung nach Meetingende
Validierungsmethode	Funktionstest mit konfigurierten Testterminen
Ergebnis	Die Benachrichtigung wurde 53 s nach Beenden des Events versendet (<code>Sende Notifications für 1 beendete Events</code>).
Status	Erfüllt
Anforderung	FA-04: End-of-Day-Feedback-Auslösung
Validierungsmethode	Funktionstest mit konfigurierten Arbeitszeiten
Ergebnis	Die Benachrichtigung wurde genau 1 Stunde vor dem Ende der Arbeitszeit versendet.
Status	Erfüllt
Anforderung	FA-05: Dashboard Tagesfragmentierung
Validierungsmethode	Review gegen Abnahmekriterium
Ergebnis	Das Structure Dashboard (vgl. Abbildung 15) visualisiert die Fragmentierung des Arbeitstages, die zusammenhängenden Fokusblöcke und die Verteilung der Meeting-Typen.
Status	Erfüllt

Anforderung	FA-06: Dashboard Feedback-Aggregation
Validierungsmethode	Review gegen Abnahmekriterium
Ergebnis	Das Impact Dashboard (vgl. Abbildung 16) zeigt Nützlichkeit, Stimmung und Energie aufgeschlüsselt nach Kategorien sowie Tageszeiten an.
Status	Erfüllt
Anforderung	FA-07: Konfiguration Arbeitszeiten und Benachrichtigungen
Validierungsmethode	Usability-Test (Aufgabe 1) und Funktionstest
Ergebnis	Alle 5 Probanden konnten Arbeitstage und Arbeitszeiten erfolgreich konfigurieren. Die Onboarding-Tour unterstützte die Auffindbarkeit.
Status	Erfüllt
Anforderung	FA-08: Benutzerverwaltung und Authentifizierung
Validierungsmethode	Usability-Test (Aufgabe 1)
Ergebnis	Registrierung und Anmeldung bei allen 5 Probanden erfolgreich abgeschlossen.
Status	Erfüllt
Anforderung	FA-09: Manuelle Kalendereinträge
Validierungsmethode	Erstellung eines Kalenderevents mit anschliessender Anpassung von Kategorie, Uhrzeit, Titel und Tag. Nach jeder Änderung wurde die korrekte Darstellung in der Kalenderansicht geprüft.
Ergebnis	Alle Anpassungen wurden jeweils sofort korrekt in der Kalenderansicht dargestellt.
Status	Erfüllt

Tabelle 5: Validierung der funktionalen Anforderungen (FA-01 bis FA-09).

7.2.2. Nicht-funktionale Anforderungen

Tabelle 6 zeigt die Validierungsergebnisse der sieben nicht-funktionalen Anforderungen.

Anforderung	NFA-01: Feedback-Eingabe unter 30 s
Validierungsmethode	Usability-Test (Aufgabe 2), Zeitmessung
Ergebnis	Alle 5 Probanden schlossen die Feedback-Eingabe in 20 bis 27 Sekunden ab (5 von 5 unter 30 s, entspricht 100 %; Schwelle: 80 %).
Status	Erfüllt
Anforderung	NFA-02: HTTPS und Token-Verschlüsselung
Validierungsmethode	Konfigurationsreview und Datenbankinspektion
Ergebnis	Alle Endpunkte über HTTPS erreichbar, Tokens verschlüsselt in der Datenbank gespeichert.
Status	Erfüllt
Anforderung	NFA-03: Toleranz gegenüber API-Ausfällen
Validierungsmethode	Simulierter Ausfall (15 min Unterbrechung)
Ergebnis	Der Scheduler loggt während des Unterbruchs, dass die Endpunkte nicht erreichbar sind. Die Events bleiben unkategorisiert, bis der Endpunkt wieder verfügbar ist. Der Scheduler nimmt den Betrieb sofort wieder auf.
Status	Erfüllt
Anforderung	NFA-04: Zustellung innerhalb von 2 min
Validierungsmethode	Zeitmessung über mehrere Testtermine
Ergebnis	Alle Benachrichtigungen wurden in unter 1 Minute ausgelöst.
Status	Erfüllt
Anforderung	NFA-05: Containerisierung und CI/CD
Validierungsmethode	Deployment-Nachweis über Pipeline
Ergebnis	Die GitHub-Pipeline läuft nach einem Push auf den Branch <code>main</code> automatisch durch. Dauer zwischen 1 Minute 48 Sekunden und 2 Minuten 21 Sekunden.
Status	Erfüllt

Anforderung	NFA-06: Flexible Feedback-Datenstrukturen
Validierungsmethode	Architekturreview (JSONB-Schema)
Ergebnis	Das Datenmodell ermöglicht eine dynamische Speicherung von Feedback (vgl. Abbildung 17).
Status	Erfüllt
Anforderung	NFA-07: Schutz sensibler Kalenderinhalte
Validierungsmethode	Architekturreview und Netzwerkanalyse
Ergebnis	Die Kalenderdaten werden lokal durch eine KI kategorisiert (vgl. Abschnitt 6.3.2).
Status	Erfüllt

Tabelle 6: Validierung der nicht-funktionalen Anforderungen (NFA-01 bis NFA-07).

7.2.3. Erreichung der Produktziele

Die Anforderungsvalidierung zeigt, dass 15 der 16 Anforderungen den Status *Erfüllt* erreichen. Einzig FA-01 ist als *Teilweise erfüllt* eingestuft, da die Microsoft-Kalenderanbindung nicht umgesetzt wurde (vgl. Abschnitt 6.2.1). Die Google-Synchronisation funktioniert jedoch vollumfänglich. Auf dieser Grundlage lässt sich die Erreichung der beiden Produktziele beurteilen.

PZ1: Kontextbezogene Datenerhebung mit minimaler Störung des Arbeitsalltags.

Die in Abschnitt 5.5 definierten Erfolgskriterien sind weitgehend erfüllt: Die Feedback-Eingabe liegt bei allen fünf Probanden unter 30 Sekunden (vgl. NFA-01), und die Post-Meeting-Benachrichtigung wird zuverlässig innerhalb einer Minute nach Meetingende ausgelöst (vgl. FA-03, NFA-04). Die Kalendersynchronisation über OAuth 2.0 funktioniert für Google stabil (vgl. FA-01), und das System toleriert temporäre API-Ausfälle ohne Datenverlust (vgl. NFA-03). Die fehlende Microsoft-Anbindung schränkt die Zielgruppenabdeckung ein, beeinträchtigt jedoch nicht die Funktionsfähigkeit des Systems für Google-Nutzende. Damit sind die technischen Voraussetzungen für eine in den Arbeitsalltag integrierte Datenerhebung gegeben.

PZ2: Aufbereitung und Visualisierung der erhobenen Daten zur Selbstreflexion.

Das System stellt zwei Dashboards bereit: Das Structure Dashboard visualisiert die Fragmentierung des Arbeitstages und die Verteilung zusammenhängender Fokusblöcke (vgl. FA-05), das Impact Dashboard aggregiert die erfassten Bewertungen nach Kategorie und Tageszeit (vgl. FA-06). Damit ist das in Abschnitt 5.5 definierte Erfolgskriterium, mindestens ein Dashboard mit objektiven Strukturmetriken und subjektiven Bewertungsdaten bereitzustellen, erfüllt. Da die Dashboards aus methodischen Gründen nicht im Usability-Test evaluiert wurden (vgl. Abschnitt 7.1), konnte ihre Eignung als Reflexionsinstrument erst anhand der Feldstudieergebnisse beurteilt werden (vgl. Abschnitt 9.4).

Zusammenfassend bestätigt die Validierung, dass beide Produktziele auf technischer Ebene erreicht sind, wobei PZ1 durch die fehlende Microsoft-Anbindung eine Einschränkung aufweist. Die funktionale Bereitschaft des Systems war damit vor Beginn von Abschnitt 8 sichergestellt. Die Bewährung im realen Arbeitsalltag, insbesondere die tatsächliche Nutzung der Dashboards zur Selbstreflexion, wird in Abschnitt 9.4 diskutiert.

8. Phase 2: Feldstudie

Die Feldstudie bildet die zweite Phase des Forschungsdesigns und verfolgt zwei Ziele. Erstens sollen die retrospektiven Befunde der Befragungsstudie (Phase 1) durch Echtzeit-Daten ergänzt und vertieft werden. Zweitens ermöglicht die Feldstudie die Analyse von Zusammenhängen auf zwei Ebenen: auf der Ebene einzelner Meetings (Meeting-Level) und auf der Ebene ganzer Arbeitstage (Tages-Level). Die Feldstudie adressiert damit primär FF2 (Einfluss von Meeting-Häufigkeit, -Dauer und -Timing auf Produktivität, Flow und Wohlbefinden).

Angeichts der kleinen Stichprobe ($N = 3$) ist die Feldstudie als explorative Pilotstudie einzuordnen. Inferenzstatistische Aussagen sind bei dieser Stichprobengrösse nicht belastbar. Der Fokus liegt auf deskriptiver Analyse, Musteridentifikation und der Erprobung des Erhebungsdesigns.

8.1. Methodik

Die Feldstudie kombiniert drei Datenströme: automatisch synchronisierte Kalenderdaten (objektive Meeting-Eigenschaften), Feedback unmittelbar nach jedem Meeting (subjektive Meeting-Eigenschaften) und End-of-Day-Feedback gegen Ende des Arbeitstages (subjektive Arbeitstageigenschaften).

Das Meeting-Feedback erfasst unmittelbar nach jedem Meeting vier Dimensionen zum subjektiven Erleben (vgl. Abschnitt 6.4.2.1, Anhang H.1). Das End-of-Day-Feedback erhebt gegen Arbeitsende vier tagesbezogene Selbsteinschätzungen (vgl. Abschnitt 6.4.2.2, Anhang H.2).

8.2. Stichprobe und Durchführung

Die drei Teilnehmenden der Feldstudie wurden aus der vorangegangenen Befragungsstudie rekrutiert und erklärten sich auf direkte Anfrage zur Teilnahme bereit. Alle drei arbeiten als Softwareentwickelnde in professionellen Teams. Tabelle 7 zeigt ihre konfigurierten Arbeitszeiten und den Datenumfang.

Teiln.	Arbeitszeit	Arbeitstage	Beobachtungstage	Meetings bewertet	Tagesbew.
P1	08:00–16:30	Mo–Fr	10	39	9
P2	09:00–17:00	Mo–Fr	10	19	9
P3	08:00–18:00	Di, Mi, Do	6	15	5

Tabelle 7: Charakteristika der Studienteilnehmenden

P1 und P2 arbeiten im klassischen Fünf-Tage-Muster, P3 in einem Drei-Tage-Modell mit längerer täglicher Arbeitszeit (10 Stunden).

Die Datenerhebung erstreckte sich vom 2. bis zum 13. März 2026 über zwei vollständige Arbeitswochen. Insgesamt wurden 73 Meeting-Feedbacks und 23 End-of-Day-Feedbacks erfasst. P1 steuert mit 39 bewerteten Meetings den grössten Anteil bei, gefolgt von P2 (19) und P3 (15). Die pseudonymisierten Rohdaten der Feldstudie sind im Anhang unter Anhang K.2 abgelegt.

Die Compliance-Raten sind für eine ESM-Studie bemerkenswert hoch: Beim Meeting-Feedback liegen sie zwischen 92,9 % (P1) und 100,0 % (P3), im Mittel bei 96,0 %. Beim End-of-Day-Feedback bewegen sie sich zwischen 83,3 % (P3) und 90,0 % (P1 und P2), im Mittel bei 87,8 %.

Die Designentscheidung, das Meeting-Feedback unmittelbar nach Meetingende auszulösen (in Phase 1 als am wenigsten disruptiver Zeitpunkt identifiziert, vgl. Abschnitt 4.4.7), dürfte zur hohen Compliance beigetragen haben.

Die in der Applikation erfasste Telemetrie bestätigt die Praktikabilität des Feedback-Instruments auch unter Realbedingungen. Die mediane Ausfülldauer eines Meeting-Feedbacks beträgt 18,9 Sekunden ($M = 27,7$, $SD = 28,0$). 82,2 % aller Feedbacks wurden in 30 Sekunden oder weniger abgeschlossen. Im Usability-Test (Abschnitt 7.1.4) lag die Ausfüllzeit unter Laborbedingungen bei 20 bis 27 Sekunden. Die Felddaten replizieren diesen Bereich und zeigen, dass die in NFA-01 definierte 30-Sekunden-Schwelle auch im Alltag von der grossen Mehrheit eingehalten wird.

Wie in Phase 1 wird für die Analyse von Zusammenhängen der Spearman-Rangkorrelationskoeffizient verwendet (vgl. Abschnitt 4.4). Für Gruppenvergleiche zwischen drei oder mehr unabhängigen Gruppen kommt der Kruskal-Wallis-Test zum Einsatz, ein Verfahren, das keine Normalverteilung voraussetzt [30]. Bei der geringen Stichprobengrösse auf Tagesebene ($N = 23$) sind sämtliche inferenzstatistischen Ergebnisse als explorative Indikatoren zu interpretieren, nicht als belastbare Befunde.

8.3. Ergebnisse auf Meeting-Ebene

Die Ergebnisdarstellung folgt einer systematischen Struktur: Zunächst werden die objektiven Meeting-Merkmale (Häufigkeit, Dauer, Timing) beschrieben, anschliessend deren Zusammenhang mit den subjektiven Bewertungen analysiert. Diese Struktur orientiert sich direkt an FF2.

8.3.1. Meeting Struktur

Die 73 bewerteten Meetings verteilen sich gleichmässig auf drei Haupttypen: Stand-ups, Planning-Meetings und sonstige Meetings (Other) machen jeweils 26,0 % aus (je $n = 19$). Ad-hoc-Meetings stellen 12,3 % ($n = 9$), One-on-one-Meetings 6,8 % ($n = 5$) und Retrospektiven 2,7 % ($n = 2$).

Die mittlere Meeting-Dauer beträgt 42,9 Minuten ($SD = 24,0$, Median = 30). Knapp die Hälfte der Meetings (43,8 %) dauert zwischen 31 und 60 Minuten, je rund ein Viertel entfällt auf kürzere Formate: 26,0 % auf 16 bis 30 Minuten, 24,7 % auf maximal 15 Minuten. Vier Meetings (5,5 %) überschreiten 60 Minuten. Zeitlich konzentrieren sich die Meetings deutlich auf den Vormittag: 64,4 % finden vor 12:00 Uhr statt, 23,3 % am Nachmittag und 12,3 % über die Mittagszeit.

Pro Tag haben die Teilnehmenden im Schnitt 3,0 Meetings ($SD = 1,7$, Spannweite 1–6) mit einer Gesamtdauer von durchschnittlich 131 Minuten ($SD = 85$, Spannweite 15–300). An einzelnen Tagen verbringen Teilnehmende somit bis zu fünf Stunden in Meetings.

8.3.2. Einfluss der Meeting-Dauer auf das Erleben

Die Meeting-Dauer zeigt differentielle Zusammenhänge mit den subjektiven Bewertungen (Abbildung 18). Längere Meetings gehen mit signifikant höherer Fokusunterbrechung einher ($r = 0,27$, $p = 0,023$) und mit signifikant niedrigerem Energielevel nach dem Meeting ($r = -0,29$, $p = 0,012$). Die Stimmung tendiert bei längeren Meetings ebenfalls in negative Richtung ($r = -0,21$, $p = 0,072$). **Bemerkenswert ist, dass die wahrgenommene Nützlichkeit (ROTI) keinen signifikanten Zusammenhang mit der Dauer aufweist** ($r = 0,06$, $p = 0,616$). Ein Meeting kann also als nützlich bewertet werden und dennoch durch seine Länge Energie absorbieren und den Fokus unterbrechen.

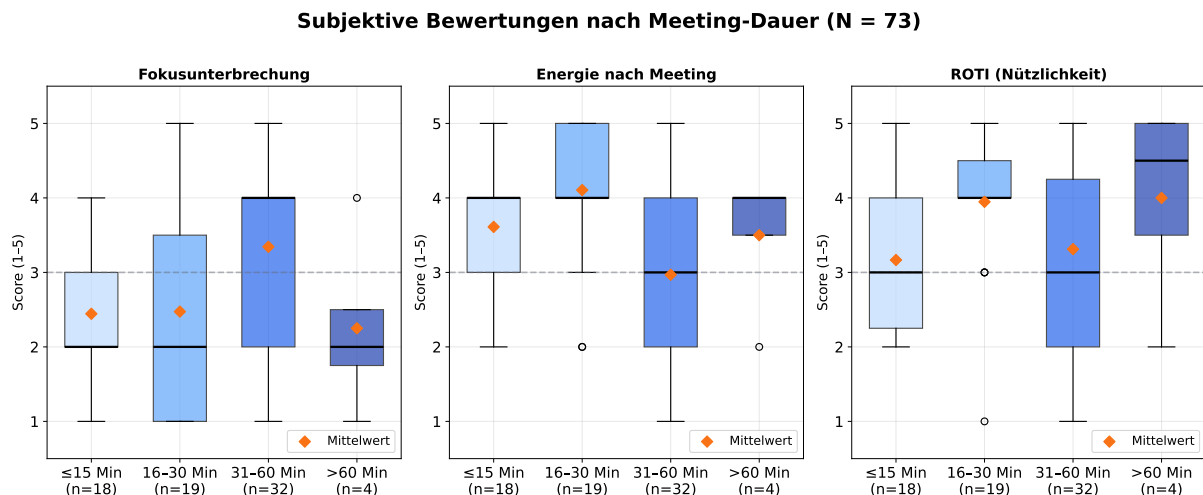


Abbildung 18: Subjektive Bewertungen nach Meeting-Dauer (N = 73)

Die Dauerklassenanalyse differenziert dieses Bild: Meetings von 31 bis 60 Minuten, die mit 43,8 % die häufigste Kategorie darstellen, erzeugen die höchste Fokusunterbrechung ($M = 3,34$) und das niedrigste Energielevel ($M = 2,97$). Kurze Meetings bis 15 Minuten (typischerweise Stand-ups) und mittlere Meetings von 16 bis 30 Minuten schneiden auf beiden Dimensionen günstiger ab. Für die Beantwortung von FF2b und FF2c ist dieser Befund relevant: Die Dauer beeinflusst Wohlbefinden (Energie, Stimmung) und Flow (Fokusunterbrechung), nicht aber die wahrgenommene Nützlichkeit.

Dauer wirkt auf Flow und Wohlbefinden, nicht auf wahrgenommene Nützlichkeit

Längere Meetings korrelieren signifikant mit höherer Fokusunterbrechung und niedrigerem Energielevel. Die wahrgenommene Nützlichkeit (ROTI) ist dagegen von der Dauer unabhängig. Meetings können als nützlich bewertet werden und dennoch Flow und Energie beeinträchtigen.

8.3.3. Einfluss der Tageszeit auf das Erleben

Die Tageszeit zeigt deskriptive Unterschiede, die bei der gegebenen Stichprobengrösse jedoch nicht statistisch signifikant werden (Kruskal-Wallis-Tests: alle $p > 0,20$). Vormittagsmeetings ($n = 47$) erzielen den höchsten mittleren ROTI ($M = 3,60$) und das höchste Energielevel ($M = 3,64$) bei moderater Fokusunterbrechung ($M = 2,70$). Nachmittagsmeetings ($n = 17$) zeigen die höchste Fokusunterbrechung ($M = 3,24$) und niedrigere Energie ($M = 3,12$). Mittagsmeetings ($n = 9$) fallen durch den niedrigsten ROTI ($M = 2,67$) auf.

Auf Tagesebene zeigt **die Anzahl der Vormittagsmeetings einen signifikant negativen Zusammenhang mit dem Deep-Work-Score** ($r = -0,44$, $p = 0,039$). Dieser Befund ist für FF2c relevant: Vormittagsmeetings, obwohl auf Meeting-Ebene günstiger bewertet, reduzieren durch ihre Platzierung in der typischen Kernarbeitszeit die verfügbare Fokuszeit.

8.3.4. Meeting-Typ-Profil

Die verschiedenen Meeting-Formate zeigen deskriptiv unterschiedliche Wirkprofile (vgl. Abbildung 19). Die Interpretation dieser Unterschiede erfordert jedoch eine wichtige Einschränkung: Die Meeting-Typen verteilen sich ungleichmässig auf die Teilnehmenden. Planning-Meetings stammen zu 79 % von P1 (15 von 19), Ad-hoc-Meetings zu 89 % (8 von 9), während One-on-ones ausschliesslich von P2 und P3 bewertet wurden. Die beobachteten Unterschiede zwischen Meeting-Typen sind daher nicht trennscharf von individuellen Bewertungstendenzen zu unterscheiden.

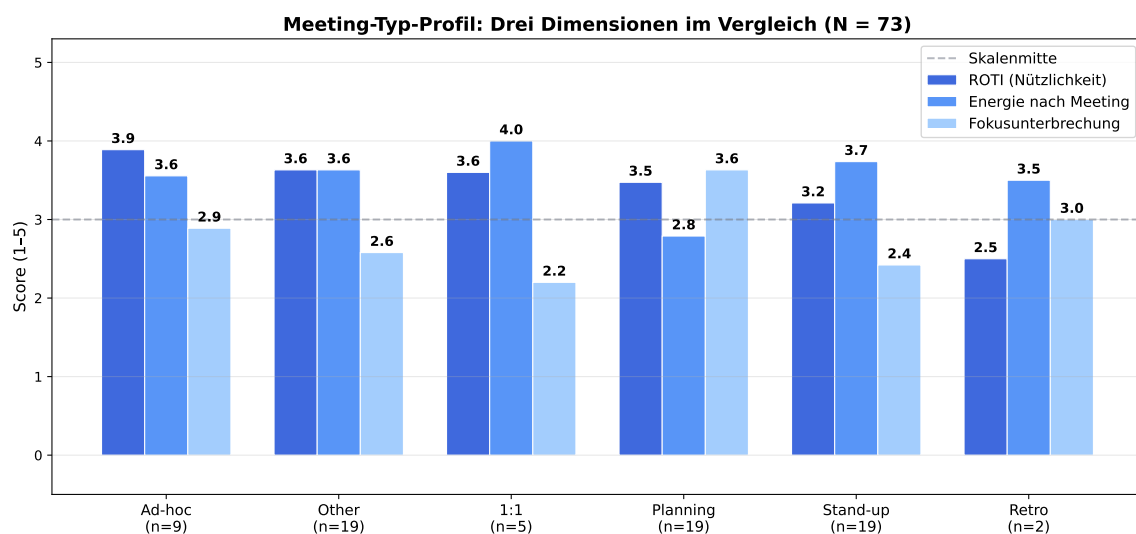


Abbildung 19: Meeting-Typ-Profil mit drei Dimensionen

Ad-hoc-Meetings ($n = 9$) weisen den höchsten ROTI auf ($M = 3,89$) bei moderater Fokusunterbrechung ($M = 2,89$). Stand-ups ($n = 19$), der einzige Typ mit annähernd gleichmässiger Verteilung auf alle drei Teilnehmenden, sind wenig fokusunterbrechend ($M = 2,42$) und erzeugen den höchsten Anteil positiver Stimmung (68 %), ihre Nützlichkeit wird jedoch unterdurchschnittlich bewertet ($M = 3,21$), dieser Wert liegt aber immer noch über dem Break-even-Point von $\text{ROTI} \geq 3$. One-on-one-Meetings ($n = 5$) erzielen das höchste Energielevel ($M = 4,00$) und die niedrigste Fokusunterbrechung ($M = 2,20$). Planning-Meetings ($n = 19$) zeigen die höchste Fokusunterbrechung ($M = 3,63$) und das niedrigste Energielevel ($M = 2,79$). Ob dies ein Effekt des Formats oder der Person (primär P1) ist, lässt sich mit den vorliegenden Daten nicht klären.

Geringe Disruption bei positivem Netto-Nutzen

Stand-ups zeigen ein konsistentes Profil über alle drei Teilnehmenden: Sie unterbrechen den Fokus kaum, erzeugen überwiegend positive Stimmung und liegen über dem ROTI-Break-even.

Über alle Typen hinweg besteht **ein deutlicher positiver Zusammenhang zwischen ROTI und Energielevel** ($r = 0,54$, $p < 0,001$): Meetings, die als nützlich bewertet werden, hinterlassen die Teilnehmenden mit mehr Energie. Die Korrelation zwischen ROTI und Fokusunterbrechung ist dagegen schwach ($r = -0,20$, $p = 0,089$). Ein Meeting kann also als nützlich bewertet werden und dennoch eine Fokusphase unterbrochen haben. Fokusunterbrechung und Nützlichkeit sind somit teilweise unabhängige Dimensionen.

Dissoziation zwischen Nützlichkeit und Disruption

Meetings, die als nützlich bewertet werden, hinterlassen signifikant mehr Energie. Gleichzeitig ist die wahrgenommene Nützlichkeit nahezu unabhängig von der Fokusunterbrechung. Ein Meeting kann als wertvoll und als störend zugleich erlebt werden.

8.3.5. Qualitative Kontextualisierung über Tags

Die kontextabhängigen Tags liefern qualitative Einblicke in die Gründe hinter den Bewertungen. Bei den 15 Meetings mit $\text{ROTI} \leq 2$ sind die drei häufigsten Issue-Tags „Zu lang“, „Zu viele Teilnehmende“ und „Keine Relevanz“ (je 6 Nennungen), gefolgt von „Mental belastend“ und „Hätte E-Mail sein können“ (je 4).

Bei den 40 Meetings mit $\text{ROTI} \geq 4$ dominiert „Nächste Schritte geklärt“ (28 Nennungen), gefolgt von „Klares Ziel“ (25) und „Informationen vermittelt“ (16). Die Klarheit über nächste Schritte wird mehr als doppelt so häufig genannt wie jeder andere positive Faktor. Meetings werden vor allem dann als wertvoll empfunden, wenn sie zu konkreten Handlungsschritten führen. Abbildung 20 zeigt die vollständige Verteilung der Tags über alle bewerteten Meetings.

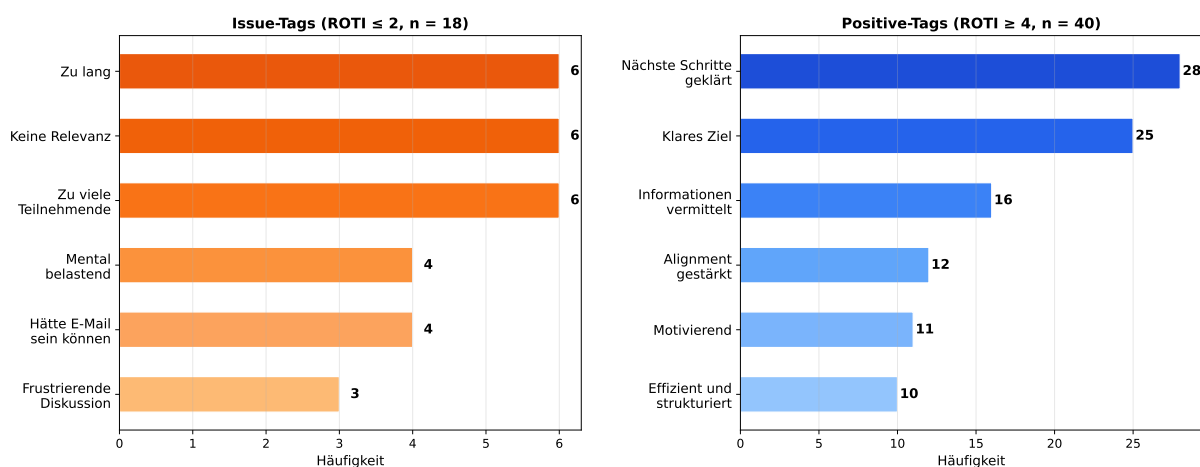


Abbildung 20: Häufigkeit der kontextabhängigen Tags

8.4. Ergebnisse auf Tages-Ebene

Die zentrale Analyse auf Tages-Ebene untersucht, wie die Meeting-Last eines Tages mit den drei Outcome-Konstrukten zusammenhängt.

Meeting-Last und Produktivität: Weder die Anzahl der Meetings noch die Gesamtmeetingzeit zeigen einen signifikanten Zusammenhang mit der wahrgenommenen Produktivität. Einen schwachen negativen Trend zeigt hingegen die empfundene Meetingbelastung ($r = -0,36$, $p = 0,092$). Das deutet darauf hin, dass nicht die objektive Meeting-Last allein, sondern deren subjektive Wahrnehmung mit der Produktivitätseinschätzung zusammenhängt [8].

Meeting-Last und Flow: Die Meeting-Anzahl ist der stärkste objektive Prädiktor für die Flow-Beeinträchtigung. An Tagen mit mehr Meetings berichten die Teilnehmenden von signifikant weniger konzentrierter Tiefenarbeit ($r = -0,46$, $p = 0,028$). Die Gesamtmeetingzeit zeigt einen gleichgerichteten, aber schwächeren Trend. Dass die Anzahl stärker wirkt als die Gesamtdauer, passt zur Fragmentierungslogik: Mehrere kurze Unterbrechungen zerstückeln den Tag stärker als ein einzelner langer Block [12].

Meeting-Last und Wohlbefinden: Der stärkste Zusammenhang auf Tagesebene überhaupt besteht zwischen empfundener Meetingbelastung und Energielevel am Tagesende ($r = -0,56$, $p = 0,006$). Die objektive Meeting-Anzahl korreliert zwar mit der empfundenen Belastung ($r = 0,49$, $p = 0,018$), ihr direkter Zusammenhang mit dem Energielevel ist aber schwach. Die empfundene Belastung fungiert somit als Vermittler zwischen objektiver Meeting-Last und Wohlbefinden. Ergänzend korrelieren Produktivität und Energie stark positiv ($r = 0,59$, $p = 0,003$), was die enge Verknüpfung von Wohlbefinden und Leistungswahrnehmung unterstreicht [14].

Prädiktor	Produktivität	Deep Work	Energie
Anzahl Meetings	-0,18	-0,46*	-0,22
Gesamte Meetingzeit	-0,10	-0,34	-0,15
Empfundene Meeting-Last	-0,36 [†]	-0,32	-0,56**

Tabelle 8: Spearman-Korrelationen zwischen Meeting-Last und Tages-Outcomes (N = 23)

Tabelle 8 fasst die für FF2 relevanten Zusammenhänge auf Tagesebene als Spearman-Rangkorrelationen zusammen (N = 23). Die Signifikanzniveaus (* $p < .05$, ** $p < .01$, [†] $p < .10$) geben an, mit welcher Sicherheit die beobachteten Zusammenhänge nicht auf Zufall beruhen.

8.4.1. Individuelle Profile

Die individuellen Unterschiede illustrieren die identifizierten Zusammenhänge. P1 weist die höchste Meeting-Last auf ($M = 4,1$ Meetings/Tag, $M = 175$ min/Tag) und berichtet konsistent die niedrigsten Werte bei Produktivität ($M = 2,56$), Deep Work ($M = 1,89$) und Energie ($M = 2,22$) bei höchster empfundener Meetingbelastung ($M = 3,78$). P3 mit weniger Meetings ($M = 2,8$ /Tag) zeigt die höchsten Werte bei Produktivität ($M = 3,80$) und Energie ($M = 3,80$) bei niedrigster empfundener Belastung ($M = 2,80$). P2 liegt auf den meisten Dimensionen im Mittelfeld.

Ein instruktiver Datenpunkt: Am 11. März hatte P3 kein einziges Meeting und bewertete Produktivität, Deep Work und Energie jeweils mit dem Höchstwert 5. Obwohl ein einzelner Datenpunkt keine kausale Aussage erlaubt, illustriert er das Potenzial meetingfreier Tage für konzentrierte Arbeit, ein Muster, das auch [3] beobachteten.

8.5. Zusammenfassung

Die explorative Feldstudie liefert trotz der kleinen Stichprobe konsistente Befunde auf zwei Analyseebenen. Auf Meeting-Ebene zeigt sich, dass die Dauer eines Meetings Wohlbefinden und Flow beeinflusst, nicht aber die wahrgenommene Nützlichkeit. Längere Meetings absorbieren Energie und unterbrechen den Fokus, können aber dennoch als wertvoll empfunden werden. Die Tageszeit zeigt deskriptive Unterschiede, wobei Vormittagsmeetings trotz günstigerer Bewertung auf Meeting-Ebene die verfügbare Fokuszeit auf Tagesebene reduzieren. Die Meeting-Typ-Profile sind durch die ungleiche Verteilung auf die Teilnehmenden nur eingeschränkt interpretierbar. Stand-ups zeigen als einziger ausgewogen verteilter Typ ein konsistentes Profil mit geringer Disruption bei positivem Netto-Nutzen.

Auf Tagesebene wirken objektive und subjektive Meeting-Last auf unterschiedliche Outcomes: Die Meeting-Anzahl beeinträchtigt primär den Flow, während die empfundene Belastung vor allem das Energielevel am Tagesende senkt. Die wahrgenommene Produktivität hängt weniger von der objektiven Meeting-Last ab als von deren subjektiver Bewertung.

Die hohe Compliance von über 90 % beim Meeting-Feedback und knapp 88 % beim Tagesfeedback bestätigt die Praktikabilität des ESM-Designs. Die qualitativen Tags decken sich mit den Problemfeldern aus Phase 1: Übermäßige Dauer, zu viele Teilnehmende und fehlende Relevanz dominieren die negativen Nennungen, während Klarheit über nächste Schritte der wichtigste Erfolgsfaktor ist.

9. Diskussion, Fazit und Ausblick

Meetings können gleichzeitig als nützlich und als störend erlebt werden. Dieses scheinbar widersprüchliche Ergebnis zieht sich als roter Faden durch die vorliegende Studie. Das folgende Kapitel ordnet die empirischen Befunde in den Forschungskontext ein, beantwortet die drei Forschungsfragen, leitet Handlungsempfehlungen ab und benennt Limitationen sowie Anknüpfungspunkte für zukünftige Forschung.

9.1. Einordnung der Ergebnisse

Der State of the Art (vgl. Abschnitt 2) identifiziert drei Forschungslücken: die fehlende Integration von Meetingforschung, Produktivität und Wohlbefinden, das Fehlen von Meeting-Level-Analysen mit objektiven Kalendermerkmalen und die nicht systematisch untersuchte Wahrnehmungs-Realitäts-Diskrepanz. Die vorliegende Studie adressiert alle drei Lücken. Die folgenden Abschnitte ordnen die Ergebnisse entlang dieser Lücken ein, bevor die Forschungsfragen im nächsten Abschnitt formal beantwortet werden.

9.1.1. Integration dreier Forschungsstränge

Produktivität, Flow und Wohlbefinden reagieren auf unterschiedliche Aspekte der Meeting-Last (vgl. Tabelle 8), hängen aber untereinander zusammen. Das bestätigt die im SPACE-Framework [9] verankerte Annahme, dass Wohlbefinden und Produktivität komplementäre Dimensionen sind, und zeigt: Rein quantitative Meeting-Metriken greifen zu kurz. Die konkreten Wirkungsmechanismen pro Konstrukt zeigen die Ergebnisse auf Tagesebene (vgl. Tabelle 8) und auf Meeting-Ebene (vgl. Abschnitt 8.3).

9.1.2. Meeting-Level-Analyse mit objektiven Kalendermerkmalen

Die Kombination von Kalenderdaten mit situativen Bewertungen ermöglichte erstmals die Analyse spezifischer Meeting-Merkmale auf der Ebene einzelner Meetings. Der zentrale neue Befund ist eine Dissoziation zwischen wahrgenommener Nützlichkeit und disruptiver Wirkung (vgl. Abschnitt 8.3, Tabelle 8). Dies erweitert die Erkenntnisse von Allen et al. [6], die den Post-Meeting-Erholungsbedarf primär über die Meeting-Zufriedenheit erklärten. Methodisch folgt daraus: Studien, die Meeting-Effekte ausschliesslich über Zufriedenheitsbewertungen messen, erfassen nur einen Teil der tatsächlichen Wirkung.

9.1.3. Wahrnehmungs-Realitäts-Diskrepanz

Die Ergebnisse bestätigen die von Beal [7] beschriebene Tendenz: Retrospektive Berichte bilden generelle Einstellungen ab, nicht tatsächliche Erlebnisse. Wo Übereinstimmungen und Abweichungen liegen, zeigen die Ergebnisse beider Phasen (vgl. Abschnitt 4.4, Abschnitt 8.3). Für die Forschungspraxis untermauern die Befunde den Wert von ESM-basierten Erhebungen. Gleichzeitig liefern die retrospektiven Daten Kontextinformationen, die Echtzeit-Daten allein nicht erfassen. Die beiden Phasen ergänzen sich methodisch, statt sich zu doppeln.

9.1.4. Rolle der subjektiven Wahrnehmung

Bei Produktivität und Wohlbefinden zeigen nicht die objektiven Kennzahlen die stärksten Zusammenhänge, sondern die empfundene Meetingbelastung (vgl. Tabelle 8). Im Sinne des Job Demands-Resources-Modells [25] fungiert die empfundene Belastung als *Strain* zwischen der objektiven Demand (Meeting-Last) und den Outcomes. Interventionen, die ausschliesslich die objektive Meeting-Last reduzieren, greifen daher möglicherweise zu kurz. Ebenso wichtig

ist die Gestaltung von Meetings, die die subjektive Belastung senkt, etwa durch klare Ziele, angemessene Dauer und erkennbaren Mehrwert.

Beim Flow deutet sich ein anderes Muster an: Hier scheint weniger die subjektive Bewertung relevant zu sein als die Anzahl der Meetings selbst, da Fragmentierung primär ein strukturelles Phänomen ist (vgl. Tabelle 8).

9.2. Beantwortung der Forschungsfragen

Aufbauend auf der Einordnung beantwortet dieser Abschnitt die drei Forschungsfragen durch eine Synthese beider Studienphasen.

9.2.1. FF1: Meeting-Praktiken und retrospektive Wahrnehmung

FF1 fragt, welche Meeting-Praktiken unter Softwareentwickelnden verbreitet sind und wie sie deren Auswirkungen retrospektiv wahrnehmen. Die Antwort stützt sich primär auf die Phase-1-Befragung (N = 55).

Die Befragung zeichnet das Bild einer moderaten, aber allgegenwärtigen Meeting-Last. Agile Formate wie Daily Stand-ups, Planning-Sessions und Ad-hoc-Syncs dominieren (vgl. Abschnitt 4.4.1). Auffällig ist die Qualitätsstreuung: Knapp zwei Drittel bewerten ihre Meetings als nützlich, doch mehr als ein Viertel empfindet regelmässig Meetings als Zeitverschwendung (vgl. Abschnitt 4.4.2). Die häufigsten Ursachen sind prozeduraler Natur: fehlendes Ziel, übermässige Dauer und irrelevante Teilnehmende. Dieser Befund deckt sich mit den Design-Charakteristika, die Leach et al. als Prädiktoren der Meeting-Effektivität identifiziert haben [20]. Die Probleme wären durch bewusste Gestaltung adressierbar. Dabei besteht eine bemerkenswerte Lücke zwischen Bereitschaft und Praxis: Die Mehrheit bewertet kurzes Meeting-Feedback als hilfreich, doch nur ein Bruchteil der Teams diskutiert regelmässig über die eigene Meeting-Kultur (vgl. Abschnitt 4.4.7).

Retrospektiv fallen die wahrgenommenen Auswirkungen je nach Konstrukt unterschiedlich aus. Am deutlichsten zeigt sich der Effekt bei der Produktivität: Die grosse Mehrheit berichtet von niedrigerer Produktivität an meetingreichen Tagen, obwohl die generelle Produktivitätswirkung ambivalent bewertet wird (vgl. Abschnitt 4.4.3). Diese scheinbare Widersprüchlichkeit löst sich auf: Meetings sind nicht pauschal schädlich, aber ihre Kumulation an einzelnen Tagen wirkt deutlich negativ. Beim Flow zeigt sich ein ähnliches Bild: Die Mehrheit berichtet von substanzieller Disruption und einer Wiederherstellungszeit von mindestens 15 Minuten (vgl. Abschnitt 4.4.4). Beim Wohlbefinden dagegen wird der Einfluss mehrheitlich als neutral bewertet. Das deutet darauf hin, dass qualitative Meeting-Eigenschaften eine grössere Rolle spielen als die blosse Häufigkeit.

Antwort auf FF1

Softwareentwickelnde erleben eine moderate, aber qualitativ stark streuende Meeting-Last. Retrospektiv werden die Auswirkungen auf Produktivität und Flow deutlich negativ wahrgenommen, während das Wohlbefinden weniger eindeutig betroffen erscheint. Die häufigsten Ursachen für negative Meeting-Erfahrungen sind prozeduraler Natur und wären durch bewusste Gestaltung adressierbar, doch die systematische Reflexion der Meeting-Kultur findet in der Praxis kaum statt.

9.2.2. FF2: Einfluss von Meeting-Häufigkeit, -Dauer und -Timing

FF2 untersucht den Einfluss von Meeting-Häufigkeit, -Dauer und -Timing auf Produktivität (FF2a), Wohlbefinden (FF2b) und Flow (FF2c). Während Phase 1 die retrospektive Gesamtwahrnehmung liefert, ermöglicht Phase 2 eine differenziertere Analyse auf Meeting- und Tagesebene. Die Beantwortung integriert beide Perspektiven.

Wahrgenommene Produktivität (FF2a)

Retrospektiv ist der Befund eindeutig: Die Mehrheit berichtet von niedrigerer Produktivität an meetingreichen Tagen. Die Meeting-Häufigkeit korreliert signifikant mit der wahrgenommenen Produktivitätseinbusse (vgl. Abschnitt 4.4.3).

Die situativen Daten differenzieren dieses Bild. Auf Tagesebene zeigen weder Meeting-Anzahl noch Gesamtmeetingzeit einen signifikanten Zusammenhang mit der wahrgenommenen Produktivität (vgl. Tabelle 8). Einen schwachen negativen Trend zeigt lediglich die empfundene Meetingbelastung. Nicht die objektive Last, sondern deren subjektive Bewertung beeinflusst die Produktivitätswahrnehmung [8].

Auf Meeting-Ebene ist die wahrgenommene Nützlichkeit (ROTI) nahezu unabhängig von der Meeting-Dauer (vgl. Abschnitt 8.3). Ein Meeting kann als nützlich bewertet werden, ohne dass sich das positiv auf die Tagesproduktivität auswirkt, wenn der Arbeitstag durch zu viele Meetings fragmentiert wird.

In der Synthese zeigt sich: Retrospektiv bewerten Softwareentwickelnde den Produktivitätseffekt deutlich negativer, als die situativen Daten es nahelegen. Die retrospektive Wahrnehmung gewichtet die kumulative Erfahrung meetingreicher Tage stärker als das situative Erleben einzelner Meetings.

Wohlbefinden (FF2b)

Retrospektiv erscheint der Einfluss auf das Wohlbefinden gering: 54,5 % bewerten ihn als neutral, und die Meeting-Häufigkeit allein erklärt die Arbeitszufriedenheit nicht signifikant (vgl. Abschnitt 4.4.5).

Die situativen Daten korrigieren diesen Eindruck. Auf Meeting-Ebene korreliert die Meeting-Dauer signifikant negativ mit dem Energielevel, und die Stimmung tendiert bei längeren Meetings in negative Richtung (vgl. Abbildung 18). Auf Tagesebene zeigt sich der stärkste Zusammenhang der gesamten Studie: Empfundene Meetingbelastung und Energielevel am Tagesende korrelieren stark negativ (vgl. Tabelle 8).

Retrospektiv nehmen Entwickelnde ihr Wohlbefinden also als weitgehend unberührt wahr, doch im Tagesverlauf senken Meetings das Energielevel messbar. Die empfundene Belastung fungiert dabei als Vermittler: Die Meeting-Anzahl korreliert mit der empfundenen Belastung, aber nicht direkt mit dem Energielevel.

Aufschlussreich ist auch der Zusammenhang zwischen ROTI und Energie auf Meeting-Ebene: Meetings, die als nützlich bewertet werden, hinterlassen signifikant mehr Energie (vgl. Abschnitt 8.3). Die wahrgenommene Qualität beeinflusst nicht nur die kognitive Bewertung, sondern auch den energetischen Zustand.

Flow (FF2c)

Beim Flow konvergieren retrospektive und situative Befunde am stärksten. Retrospektiv berichten 61,8 % von mindestens moderater Flow-Disruption, und ebenfalls 61,8 % benötigen mindestens 15 Minuten Wiederherstellungszeit nach Meetings (vgl. Abschnitt 4.4.4).

Die Phase-2-Daten bestätigen und präzisieren diese Wahrnehmung. Auf Meeting-Ebene zeigt die Dauer den deutlichsten Effekt: Längere Meetings gehen mit signifikant höherer Fokusunterbrechung einher (vgl. Abschnitt 8.3). Meetings von 31 bis 60 Minuten erzeugen die höchste Fokusunterbrechung, kurze Formate unter 15 Minuten sind deutlich weniger disruptiv.

Auf Tagesebene ist die Meeting-Anzahl der stärkste objektive Prädiktor für den Deep-Work-Score (vgl. Tabelle 8). Dass die Anzahl stärker wirkt als die Gesamtdauer, passt zur Fragmentierungslogik: Mehrere kurze Unterbrechungen zerstückeln den Tag stärker als ein einzelner langer Block [12]. Ergänzend korreliert die Anzahl der Vormittagsmeetings signifikant negativ mit dem Deep-Work-Score. Vormittagsmeetings werden auf Meeting-Ebene zwar günstiger bewertet, reduzieren aber die verfügbare Fokuszeit in der Kernarbeitszeit.

Antwort auf FF2

Meeting-Häufigkeit, -Dauer und -Timing wirken auf die drei Outcome-Konstrukte über unterschiedliche Mechanismen. Die Produktivität wird weniger durch die objektive Meeting-Last bestimmt als durch die subjektive Bewertung der Belastung. Das Wohlbefinden reagiert auf die empfundene Belastung, die als Vermittler zwischen objektiver Last und Energielevel fungiert. Der Flow wird am direktesten beeinflusst: Die Anzahl der Meetings fragmentiert den Tag, die Dauer einzelner Meetings unterbricht den Fokus, und Vormittagsmeetings reduzieren die verfügbare Fokuszeit. Bemerkenswert ist, dass die wahrgenommene Nützlichkeit eines Meetings unabhängig von dessen disruptiver Wirkung auf Flow und Energie ist: Meetings können als nützlich und gleichzeitig als störend erlebt werden.

9.2.3. FF3: Wahrnehmung und Realität

FF3 fragt, welche Muster die Echtzeit-Bewertungen zeigen und inwiefern sie die in der Befragung identifizierten Problemfelder widerspiegeln.

Konvergenzen

Am deutlichsten stimmen die Befunde beim Flow überein. Die Echtzeit-Daten bestätigen die retrospektive Einschätzung einer substanziellen Disruption vollständig. Die retrospektiv berichtete Wiederherstellungszeit von mindestens 15 Minuten ist konsistent mit der in der Literatur dokumentierten *Meeting-to-Work Transition Time* [6].

Auch bei den qualitativen Problemfeldern zeigt sich hohe Übereinstimmung. Die in Phase 1 am häufigsten genannten negativen Faktoren (fehlendes Ziel, übermässige Dauer, irrelevante Teilnehmende) decken sich mit den Issue-Tags aus Phase 2. Umgekehrt bestätigen die Positive-Tags das Gegenstück: „*Nächste Schritte geklärt*“ und „*Klares Ziel*“ dominieren und unterstreichen Klarheit und Zielorientierung als zentrale Erfolgsfaktoren.

Divergenzen

Die grösste Diskrepanz betrifft das Wohlbefinden: Retrospektiv als neutral eingestuft, offenbaren die Echtzeit-Daten den stärksten Zusammenhang der gesamten Studie. Der *Recall Bias* [7] erklärt diesen Effekt: Die schleichende Energieabsorption durch Meetings wird im Moment registriert, aber in der Gesamtbewertung nicht als Wohlbefindenseffekt attribuiert.

Bei der Produktivität zeigt sich eine subtilere Abweichung: Retrospektiv fällt die Bewertung eindeutig negativ aus, situativ hängt sie stärker von der subjektiven Belastungswahrnehmung ab als von der objektiven Meeting-Last. Die retrospektive Wahrnehmung scheint negative Erfahrungen stärker zu gewichten als positive.

Nuancierungen

Die Echtzeit-Daten liefern eine zentrale Nuancierung, die retrospektiv nicht erfassbar ist: die in Abschnitt 8.3 sichtbare Dissoziation zwischen Nützlichkeit und Disruption. Diese erklärt, warum retrospektive Befragungen, die typischerweise nach Nützlichkeit oder Zufriedenheit fragen, die tatsächliche Belastung unterschätzen können.

Antwort auf FF3

Die Echtzeit-Bewertungen bestätigen die retrospektiv identifizierten Problemfelder beim Flow und bei den qualitativen Ursachen für negative Meeting-Erfahrungen. Beim Wohlbefinden offenbaren die situativen Daten jedoch einen substanziellen Effekt, der retrospektiv kaum wahrgenommen wird. Bei der Produktivität zeigt sich eine Tendenz zur Übergeneralisierung in der Retrospektive. Die bedeutendste Nuancierung ist die Dissoziation zwischen Nützlichkeit und Disruption: Meetings werden als nützlich und gleichzeitig als störend erlebt, eine Differenzierung, die nur durch mehrdimensionale Echtzeit-Erhebung sichtbar wird.

9.3. Praktische Implikationen

Was bedeuten diese Befunde für die tägliche Arbeit? Vier Empfehlungen lassen sich für Softwareentwicklungsteams ableiten.

Fragmentierung reduzieren

Da die Meeting-Anzahl den stärksten Einfluss auf den Flow hat, sollten Teams die Anzahl separater Meetings pro Tag minimieren. Wo möglich, können Meetings gebündelt werden, um zusammenhängende Fokusblöcke zu schaffen.

Meeting-Dauer bewusst gestalten

Meetings von 31 bis 60 Minuten erzeugen die höchste Fokusunterbrechung und das niedrigste Energielevel. Kürzere Formate (unter 15 Minuten, wie Stand-ups) sind deutlich weniger disruptiv. Teams sollten prüfen, ob längere Meetings tatsächlich die volle Dauer benötigen oder ob kürzere, fokussiertere Formate möglich sind.

Nützlichkeit sichtbar machen

Der stärkste Erfolgsfaktor für positiv bewertete Meetings ist die Klarheit über nächste Schritte. Teams sollten sicherstellen, dass jedes Meeting mit konkreten Handlungsschritten endet. Der häufigste negative Faktor, ein fehlendes Ziel, lässt sich durch konsequente Agenda-Nutzung adressieren. Da als nützlich empfundene Meetings auch das Energielevel positiv beeinflussen, zahlt sich diese Investition doppelt aus.

Systematische Reflexion etablieren

Die Diskrepanz zwischen der Offenheit für Meeting-Verbesserung und der tatsächlichen Reflexionspraxis zeigt ein ungenutztes Potenzial. Strukturiertes Feedback, wie es die entwickelte Applikation ermöglicht, kann diese Lücke schliessen, ohne zusätzliche Meetings oder aufwendige Prozesse zu erfordern. Die Feldstudie zeigt, dass bereits die Aktivität des regelmässigen Bewertens das Bewusstsein für die eigene Meeting-Kultur schärft.

9.4. Die Applikation als Reflexionsinstrument

Neben den empirischen Befunden stellt sich die Frage, ob das entwickelte Erhebungsinstrument im Arbeitsalltag tatsächlich funktioniert. Da die Dashboards und Reflexionsmechanismen (PZ2) aus methodischen Gründen nicht im Labor getestet wurden (vgl. Abschnitt 7.1), erfolgt ihre Evaluation anhand der Feldstudie-Daten.

Die Compliance-Raten belegen, dass die Integration in den Arbeitsalltag gelungen ist: 96,0 % beim Meeting-Feedback und 87,8 % beim End-of-Day-Feedback (vgl. Abschnitt 8.2). Das qualitative Feedback bestätigt dies: Teilnehmende merkten an, der zeitliche „*Aufwand kaum der Rede wert war*“. Die Telemetrie untermauert diese Einschätzung: Die mediane Ausfülldauer eines Meeting-Feedbacks betrug 18,9 Sekunden (vgl. Abschnitt 8.2). Die Designentscheidung, das Meeting-Feedback unmittelbar nach Meetingende auszulösen, hat sich bewährt. Sie wurde durch die Phase 1 Ergebnisse informiert, die diesen Zeitpunkt als am wenigsten disruptiv identifizierten (vgl. Abschnitt 4.4.7), und durch die hohe Compliance in Phase 2 bestätigt. Während der gesamten zweiwöchigen Feldstudie trat lediglich ein einziger technischer Fehler auf, der durch eine kurze E-Mail-Anweisung sofort behoben werden konnte.

Über die reine Datenerhebung hinaus entfaltete die Applikation eine reflektierende Wirkung. Im qualitativen Feedback nach Abschluss der Feldstudie betonte ein Teilnehmender: *„Ich persönlich fand es ein gutes Ritual, nach den Meetings nochmal zu reflektieren, wie das Meeting meine Konzentration und meine Stimmung beeinflusst hat - und ich bin auch neugierig, wie die Anderen das empfunden haben!“*. Die Interaktion mit dem Feedback-System förderte ein bewussteres Auseinandersetzen mit der individuellen Meeting-Kultur. Damit erreicht die Applikation ihr Ziel, Softwareentwickelnde zur evidenzbasierten Reflexion zu befähigen, ohne aufwendige Schulungen vorauszusetzen (PZ2).

9.5. Limitationen

Die Studie unterliegt mehreren Einschränkungen.

9.5.1. Stichprobengrösse in Phase 2

Die Feldstudie umfasst drei Teilnehmende. Das erlaubt keine inferenzstatistisch belastbaren Aussagen und schränkt die Generalisierbarkeit erheblich ein. Die Phase-2-Ergebnisse sind als explorative Pilotstudie einzuordnen. Gleichzeitig ist die Datendichte pro Person beachtlich: 73 bewertete Meetings und 23 Tagesbewertungen bei einer Compliance-Rate von über 90 % ermöglichen eine sinnvolle deskriptive Analyse.

9.5.2. Unvollständige Erfassung von Unterbrechungen

Die Applikation erfasst ausschliesslich Meetings, die im digitalen Kalender eingetragen sind. Zwei relevante Unterbrechungsquellen liegen damit ausserhalb des Erfassungsbereichs.

Erstens werden Ad-hoc-Meetings, die nicht im Kalender vermerkt sind, nicht automatisch erkannt. Die Applikation bietet zwar die Möglichkeit, solche Meetings manuell zu erfassen, doch die Compliance bei manueller Erfassung kann nicht sichergestellt werden. In der Praxis dürften spontane Meetings, Tür-und-Angel-Gespräche und unangekündigte Anrufe einen relevanten Anteil der tatsächlichen Meeting-Last ausmachen.

Zweitens erfasst die Studie keine nicht-meetingbezogenen Unterbrechungen wie Sofortnachrichten, E-Mail-Benachrichtigungen oder spontane Besuche am Arbeitsplatz. Die gemessenen Effekte auf Flow und Produktivität bilden daher nur den kalenderbasiert erfassbaren Teil der gesamten Unterbrechungslast ab. Die tatsächlichen Auswirkungen dürften stärker sein als die hier berichteten.

9.5.3. Single-Item-Messung der Outcome-Konstrukte

Aus Gründen der minimalen Disruption wurde jedes Outcome-Konstrukt sowohl auf Meeting- als auch auf Tagesebene durch einzelne Items operationalisiert. Diese Designentscheidung folgt den Empfehlungen der ESM-Literatur zur Reduktion der Teilnehmerbelastung, schränkt jedoch die Konstruktvalidität ein. Multi-Item-Skalen würden differenziertere und reliablere Messungen ermöglichen, stehen aber im Spannungsfeld mit der Anforderung an minimale Disruption, die für die Compliance und damit für die Datenqualität essenziell ist.

9.5.4. Selbstselektion und Stichprobenhomogenität

Die Phase 1 Stichprobe ist organisational konzentriert: 87,3 % arbeiten in Grossunternehmen und 72,7 % in IT-Abteilungen innerhalb von Nicht-IT-Organisationen (vgl. Abschnitt 4.3). Da die Rekrutierung primär über professionelle Netzwerke in der Deutschschweiz erfolgte, ist zudem eine geografische Konzentration anzunehmen. Mit 81,8 % männlichen Teilnehmenden ist die Stichprobe geschlechtlich unausgewogen. Die Befunde sind daher nicht ohne Weiteres auf

andere kulturelle, organisationale oder demografische Kontexte übertragbar. In Phase 2 kommt ein zusätzlicher Selbstselektionseffekt hinzu: Die drei Teilnehmenden haben sich freiwillig zur Feldstudie gemeldet und könnten ein überdurchschnittliches Interesse an Meeting-Reflexion aufweisen, was die Compliance-Raten begünstigt, aber die Repräsentativität einschränkt.

9.5.5. Beobachtungszeitraum

Die Feldstudie erstreckte sich über zwei Arbeitswochen. Dieser Zeitraum erfasst möglicherweise keine saisonalen Schwankungen, projektphasenabhängigen Variationen oder langfristigen Anpassungseffekte. Es ist denkbar, dass die Meeting-Last und deren Auswirkungen in intensiven Projektphasen (z.B. Sprint-Enden, Release-Vorbereitung) anders ausfallen als in dem beobachteten Zeitraum.

9.5.6. Retrospektive Verzerrung in Phase 1

Die Ergebnisse von Phase 1 basieren auf retrospektiven Selbstberichten, die systematischen Verzerrungen unterliegen können [7]. Das zweiphasige Design der Studie adressiert dieses Problem, indem es retrospektive Einschätzungen durch Echtzeit-Daten trianguliert. Die Verzerrung wird dadurch nicht eliminiert, aber die Diskrepanzen werden in der Beantwortung von FF3 explizit analysiert und diskutiert.

9.6. Fazit

Die Arbeit verbindet Meetingforschung, Produktivität und Wohlbefinden, drei Stränge, die in der Literatur selten gemeinsam betrachtet werden, und liefert trotz der genannten Einschränkungen einen klaren Beitrag.

Die zentrale Erkenntnis lässt sich in einem Satz fassen: Meetings wirken nicht pauschal, sondern über unterschiedliche Mechanismen auf Produktivität, Flow und Wohlbefinden. Die Meeting-Anzahl fragmentiert den Tag und beeinträchtigt den Flow. Die empfundene Belastung vermittelt den Einfluss auf Wohlbefinden und Energie. Und die wahrgenommene Produktivität hängt stärker von der subjektiven Bewertung ab als vom objektiven Ausmass der Meeting-Last. Dass Meetings gleichzeitig als nützlich und als störend erlebt werden, ist die vielleicht überraschendste Differenzierung der Studie, sichtbar nur durch die mehrdimensionale Echtzeit-Erhebung.

Auf praktischer Ebene zeigt die entwickelte Applikation, dass strukturiertes Feedback im Arbeitsalltag von Softwareentwickelnden praktikabel ist. Die Compliance-Raten von über 90 % und das durchweg positive Nutzerfeedback belegen, dass eine solche Datenerhebung ohne zusätzliche Belastung in den Arbeitsalltag integrierbar ist (vgl. Abschnitt 9.4). Die Applikation fördert darüber hinaus die bewusste Auseinandersetzung mit der eigenen Meeting-Kultur und schliesst damit die Feedback-Lücke, die die Literatur als zentrales Hindernis für die Verbesserung von Meetings identifiziert [5].

9.7. Ausblick

Mehrere Anknüpfungspunkte ergeben sich für zukünftige Forschung und Weiterentwicklung.

Replikation und Publikation

Die explorativen Befunde der Feldstudie sollen in einer grösser angelegten Studie mit 15 bis 20 Teilnehmenden über vier Wochen überprüft werden. Eine solche Replikation würde die Analyse individueller Unterschiede vertiefen und bildet die geplante Grundlage für eine wissenschaftliche Publikation.

Langzeitbeobachtung

Ein Beobachtungszeitraum von mehreren Monaten würde es erlauben, saisonale und projektphasenabhängige Schwankungen zu erfassen sowie langfristige Anpassungseffekte zu untersuchen. Zudem könnte analysiert werden, ob die reflektierende Wirkung der Applikation über die Zeit zu messbaren Veränderungen in der Meeting-Kultur führt.

Interventionsforschung

Die vorliegende Arbeit beschreibt und analysiert den Status quo. Ein nächster Schritt wäre die Untersuchung gezielter Interventionen: Führen Meeting-freie Vormittage tatsächlich zu höherer Produktivität und besserem Wohlbefinden? Verbessert systematisches Post-Meeting-Feedback die Meeting-Qualität über die Zeit? Solche Fragen liessen sich mit dem bestehenden Erhebungsdesign in einem Prä-Post-Vergleich adressieren.

Aggregation auf Team-Level

Die vorliegende Arbeit fokussiert auf die individuelle Perspektive. Eine Erweiterung auf die Team-Ebene würde geteilte Meeting-Muster sichtbar machen und die Auswirkungen teamweiter Praktiken auf kollektive Produktivität und Teamklima analysierbar machen. Die Applikation ist architektonisch bereits auf diese Erweiterung vorbereitet.

A Eigenständigkeitserklärung

Ehrlichkeitserklärung

Ich (wir) erkläre(n) hiermit, dass ich (wir) den vorliegenden Leistungsnachweis selber und selbstständig verfasst habe(n),

- dass ich (wir) sämtliche nicht von mir (uns) selber stammenden Textstellen und anderen Quellen wie Bilder etc. gemäss gängigen wissenschaftlichen Zitierregeln¹ korrekt zitiert und die verwendeten Quellen klar sichtbar ausgewiesen habe(n);
- dass ich (wir) in einer Fussnote oder einem Hilfsmittelverzeichnis alle verwendeten Hilfsmittel (KI-Assistenzsysteme wie Chatbots², Übersetzungs-³ Paraphrasier-⁴ oder Programmierapplikationen⁵) deklariert und ihre Verwendung bei den entsprechenden Textstellen angegeben habe(n);
- dass ich (wir) sämtliche immateriellen Rechte an von mir (uns) allfällig verwendeten Materialien wie Bilder oder Grafiken erworben habe(n) oder dass diese Materialien von mir (uns) selbst erstellt wurde(n);
- dass das Thema, die Arbeit oder Teile davon nicht bei einem Leistungsnachweis eines anderen Moduls verwendet wurden, sofern dies nicht ausdrücklich mit der Dozentin oder dem Dozenten im Voraus vereinbart wurde und in der Arbeit ausgewiesen wird;
- dass ich mir (wir uns) bewusst bin (sind), dass meine (unsere) Arbeit auf Plagiate und auf Dittautorschaft menschlichen oder technischen Ursprungs (Künstliche Intelligenz) überprüft werden kann;
- dass ich mir (wir uns) bewusst bin (sind), dass die Hochschule für Technik FHNW einen Verstoss gegen diese Eigenständigkeitserklärung bzw. die ihr zugrundeliegenden Studien- und Prüfungsordnungen der Hochschule für Technik verfolgt und dass daraus disziplinarische Folgen (Verweis oder Ausschluss aus dem Studiengang) resultieren können.

Windisch, 20. März 2026

Gideon Monterosa



Xeno Isenegger



¹z.B. APA oder IEEE

²z.B. ChatGPT

³z.B. DeepL

⁴z.B. Quillbot

⁵z.B. Github Copilot

B Hilfsmittelverzeichnis

Hilfsmittel	Verwendung	Betroffene Stellen
Google Scholar	Literaturrecherche	alle kapitel
Balsamiq	Erstellung von Wireframes	Abschnitt 5
draw.io	Erstellung von Systemübersicht und Datenmodel	Abschnitt 6
Apple icon Composer	Erstellung von Logo	Programmcode
Gemini	Überarbeitung verschiedener Textabschnitte	alle Kapitel
Claude	Überarbeitung verschiedener Textabschnitte	alle Kapitel
Matplotlib und Seaborn	Erstellung von Grafiken	Restliche Grafiken
Gemini	Hilfe beim Coding	Programmcode
Claude	Hilfe beim Coding	Programmcode

Tabelle 9: Eingesetzte Hilfsmittel mit Verwendungszweck und betroffenen Stellen.

C Glossar

Affective Events Theory (AET) Theorie, dass Arbeitsereignisse (z.B. Meetings) als bidirektionale affektive Auslöser wirken [26].

Attention Residue Kognitive Aufmerksamkeit, die nach einem Aufgabenwechsel bei der vorherigen Aufgabe verbleibt und die Leistung beeinträchtigt [12].

BCrypt Kryptographischer Hash-Algorithmus zur sicheren Passwortspeicherung.

Compliance-Rate Anteil der tatsächlich ausgefüllten Erhebungen im Verhältnis zu den erwarteten.

Concurrent Thinking Aloud Usability-Testmethode, bei der Probanden ihre Gedanken während der Aufgabenbearbeitung laut aussprechen [34].

Daily Stand-up (DSM) Häufigstes wiederkehrendes Meeting-Format in agilen Teams; kurzes tägliches Statusmeeting.

Deep Work / Deep Focus Zustand konzentrierter Tiefenarbeit ohne Unterbrechungen; eng verwandt mit Flow [38].

Developer Experience (DevEx) Konzept, das Entwicklerproduktivität auf drei Kernfaktoren destilliert: Feedback-Schleifen, kognitive Last und Flow-Zustand [10].

Experience Sampling Method (ESM) Forschungsmethode zur Echtzeiterhebung situativer Erfahrungen im Alltag [27], [28].

Firebase Cloud Messaging (FCM) Google-Dienst für browserübergreifende Push-Benachrichtigungen.

Flow bezeichnet einen Zustand anhaltender, ununterbrochener Konzentration bei der Bearbeitung anspruchsvoller Aufgaben. Flow erfordert klare Ziele, unmittelbares Feedback und ein Gleichgewicht zwischen Herausforderung und Fähigkeit [11].

Flow Disruption Cost Pauschaler Zeitabzug pro Kontextwechsel nach einem Meeting.

Flow Potential Score Berechnete Kennzahl: Anteil zusammenhängender Arbeitszeit abzüglich Kontextwechsel-Kosten.

Fokusblöcke Zusammenhängende Zeiträume im Arbeitstag ohne Meetings, in denen fokussiertes Arbeiten möglich ist.

Fragmentierung (des Arbeitstags) Zerstückelung zusammenhängender Arbeitszeit durch Unterbrechungen wie Meetings.

Inferenzstatistik Statistische Verfahren, die Schlüsse von Stichproben auf Populationen erlauben [30].

Job Demands-Resources (JD-R) Modell, das beschreibt, wie Arbeitsanforderungen bei chronischer Überlastung zu Burnout führen können [25].

Likert-Skala Bewertungsskala mit abgestuften Antwortmöglichkeiten (z.B. 1–5), häufig in Umfragen eingesetzt [30].

LLaMA Large Language Model von Meta; hier lokal eingesetzt für Meeting-Kategorisierung (LLaMA 3.1 8B).

Meeting-Kultur Gesamtheit der in einem Team oder Unternehmen gelebten Normen und Gewohnheiten rund um Meetings.

Meeting-Last Gesamtbelastung durch Meetings (Anzahl, Dauer, Häufigkeit pro Tag oder Woche).

Meeting-Praktiken Wiederkehrende Muster in Bezug auf Häufigkeit, Dauer, Timing und Gestaltung von Meetings.

Meeting Recovery Scale Empirische Operationalisierung des Aufwands für die Rückkehr zur produktiven Arbeit nach einem Meeting.

Meeting-to-Work Transition Time Durchschnittliche Zeit (ca. 20 Min.), die nach einem Meeting bis zur Wiederaufnahme produktiver Arbeit benötigt wird [6].

Meeting-Vermeidungsverhalten Muster des Ablehnens oder Überspringens von Meetings, z.B. aufgrund fehlender Relevanz.

Mixed-Methods-Design Forschungsdesign, das quantitative und qualitative Ansätze kombiniert.

OAuth 2.0 Autorisierungsprotokoll für sicheren Zugriff auf externe Dienste (z.B. Kalenderanbieter).

Ökologische Validität Grad, zu dem Testergebnisse auf reale Nutzungssituationen übertragbar sind.

Operationalisierung Überführung theoretischer Konstrukte in messbare Variablen und Indikatoren.

Outcome-Konstrukt Zielvariable einer Untersuchung; hier: Produktivität, Flow, Wohlbefinden.

Planning / Refinement Agile Meeting-Formate zur Planung und Verfeinerung von Aufgaben.

Psychologische Sicherheit Überzeugung, dass man im Team Risiken eingehen kann, ohne negative Konsequenzen zu befürchten.

Rational Unified Process (RUP) Iteratives Softwareentwicklungs-Framework; hier als Vorlage für konzeptionelle Artefakte verwendet [32].

Recall Bias Systematische Verzerrung bei der Erinnerung an vergangene Erlebnisse; durch ESM reduzierbar [7].

Recovery-Kosten / Wiederherstellungszeit Zeit und kognitiver Aufwand, um nach einer Unterbrechung wieder produktiv arbeiten zu können.

Retrospektive Regelmässiges Team-Meeting zur Reflexion über den vergangenen Sprint oder Arbeitszeitraum.

ROTI (Return on Time Invested) Fünfstufige Skala zur Bewertung des wahrgenommenen Werts eines Meetings im Verhältnis zur investierten Zeit [23].

Scrum Agiles Rahmenwerk für die Softwareentwicklung mit definierten Rollen, Events und Artefakten.

Selbstselektion Systematische Verzerrung durch freiwillige Teilnahme; Teilnehmende könnten atypische Merkmale aufweisen.

Signifikanzniveau (p-Wert) Wahrscheinlichkeit, ein Ergebnis unter der Nullhypothese zu beobachten; typische Schwellenwerte: $p < .05$, $p < .01$ [30].

SPACE-Framework Fünf Dimensionen der Entwicklerproduktivität: Satisfaction, Performance, Activity, Communication, Efficiency and Flow [9].

Spearman-Rangkorrelationskoeffizient Korrelationsmass für ordinalskalierte Daten; basiert auf Rangfolgen statt absoluten Werten [30].

Strain (JD-R) Im JD-R-Modell die subjektive Belastungsreaktion als Vermittler zwischen objektiver Anforderung und Outcome [25].

System Usability Scale (SUS) Standardisierter Fragebogen (10 Items) zur Bewertung der Gebrauchstauglichkeit eines Systems; Score 0–100 [37].

Traceability Matrix Zuordnungstabelle, die Anforderungen auf Probleme, Ziele und Forschungsfragen zurückführt [32].

Wissensarbeit / Wissensarbeitende Arbeit, die primär auf Wissen, Analyse und Problemlösung basiert, sowie die Personen, die sie ausüben.

D Variablenübersicht

Variable	Konstrukt	Datenquelle	Operationalisierung	Skala
Objektive Meeting-Eigenschaften				
Anzahl Meetings	—	Automatisch erhobene Kalenderdaten	Anzahl synchronisierter Meeting-Events pro Tag	Metrisch
Meeting-Dauer	—	Automatisch erhobene Kalenderdaten	Dauer jedes Events in Minuten; Gesamtdauer pro Tag	Metrisch
Meeting-Timing	—	Automatisch erhobene Kalenderdaten	Zeitliche Platzierung im Tagesverlauf (Vormittag, Mittag, Nachmittag)	Kategorial
Objektive Arbeitstageigenschaften				
Reguläre Arbeitszeit	—	Individuelle Arbeitszeitkonfiguration	Konfigurierte Start-/Endzeit und Arbeitstage	Metrisch
Fragmentierung	—	Berechnet aus Kalenderdaten und Arbeitszeiten	Zusammengesetzter Score aus Meeting-Anzahl, Lückenverteilung und Blockverhältnissen	Metrisch
Fokusblöcke	—	Berechnet aus Kalenderdaten und Arbeitszeiten	Anzahl meetingfreier Blöcke ≥ 60 , ≥ 90 , ≥ 120 Min.	Metrisch
Subjektive Meeting-Eigenschaften				
Fokusunterbrechung	Flow	Situative Bewertung nach Meeting	Likert-Skala 1–5	Ordinal
Emotionale Reaktion	Wohlbefinden	Situative Bewertung nach Meeting	Dreistufig (negativ, neutral, positiv)	Kategorial
Energielevel (nach Meeting)	Wohlbefinden	Situative Bewertung nach Meeting	Likert-Skala 1–5	Ordinal
Nützlichkeit (RO-TI)	Produktivität	Situative Bewertung nach Meeting	Likert-Skala 1–5	Ordinal
Subjektive Arbeitstageigenschaften				
Wahrgenommene Produktivität	Produktivität	Tägliche Selbsteinschätzung am Tagesende	Likert-Skala 1–5	Ordinal

Variable	Konstrukt	Datenquelle	Operationalisierung	Skala
Anteil Deep Work	Flow	Tägliche Selbsteinschätzung am Tagesende	Likert-Skala 1–5	Ordinal
Energielevel (Tagesende)	Wohlbefinden	Tägliche Selbsteinschätzung am Tagesende	Likert-Skala 1–5	Ordinal
Empfundene Meetingbelastung	Wohlbefinden	Tägliche Selbsteinschätzung am Tagesende	Likert-Skala 1–5	Ordinal

Tabelle 10: Übersicht der erhobenen Variablen mit Konstrukt, Datenquelle und Operationalisierung.

E Vollständiger Fragebogen für die Umfrage in Phase 1

Meetings culture in Software Teams

This survey is part of a research initiative by the University of Applied Sciences and Arts Northwestern Switzerland (FHNW) in collaboration with the University of Bern (UniBE).

We are investigating how meetings impact software practitioners' ability to make progress on development work, focus, and overall Wohlbefinden, as well as their experiences with meetings in daily work (e.g., usefulness, recovery time, satisfaction).

Time required: 10 minutes.

Data protection and consent:

- Participation is voluntary.
- All responses are stored securely and analyzed in aggregate only
- Email addresses (if provided) are stored separately from survey responses
- Data will be retained for 5 years as per research standards
- You may request data deletion by contacting [nitish.patkar@fhnw.ch]

This survey is for research; it is not an evaluation of your team or employer. There are no right or wrong answers.

The survey will be open between **28 January and 28 February 2026**.

Thank you for your time and contribution.

Screening

In this survey, a meeting refers to any scheduled or ad-hoc synchronous discussion, whether in person or via video/audio. You are expected to participate by sharing opinions or information, engaging in discussion, or simply listening. We do not consider event-style invitations — such as presentations, talks, or workshops — as meetings.

ID	Question and Answer Options	Construct
S1	In the last 4 weeks, did you participate in any work-related meetings? Single Choice: • Yes • No	Eligibility
S2	Do you confirm that you consent to participate in this survey? Single Choice: • Yes • No	Informed Consent

Section 1: Current Meeting Practices

Please answer based on your typical work situation during the last 4 weeks. If your work varied, answer for what was most typical.

Definitions used in this survey:

Productivity ability to make progress on development work (e.g., coding, design, problem solving).

Well-Being stress, energy, and job satisfaction at work.

Flow / deep focus sustained concentration and immersion in a task.

Q1	How many meetings do you typically have on a workday? Single Choice: • 1 meeting • 2 – 3 meetings • 4 – 5 meetings • More than 5 meetings	Wohlbefinden
Q2	Which kinds of meetings do you participate in at least occasionally? Multiple Choice: • Daily stand-ups / status updates • Planning or refinement sessions • Retrospectives • Ad-hoc or spontaneous technical syncs • Company-wide or department-wide meetings • One-on-one meetings • Other	-

ID	Question and Answer Options	Construct
Q3	Which kinds of meetings account for the largest share of your meeting time? Multiple Choice: • Daily stand-ups / status updates • Planning or refinement sessions • Retrospectives • Ad-hoc or spontaneous technical syncs • Company-wide or department-wide meetings • One-on-one meetings • Other	-
Q4	How much total time do you spend in meetings on a typical workday? Single Choice: • Less than 30 minutes • 30 minutes to 1 hour • 1 to 2 hours • 2 to 4 hours • More than 4 hours	Wohlbefinden
Q5	Where do your meetings primarily take place? Single Choice: • On-Site (in the office) • Remote (video calls) • Hybrid (mixture of on-site and remote) • Other	-
Q6	How useful are most meetings for accomplishing your daily work tasks? Single Choice (scale): • Very useful • Rather useful • Neither useful nor not useful • Rather not useful • Not useful at all • Not applicable	Produktivität
Section 2: Impact of Meetings on daily Work In this section, the following scales are used: • Agreement: Strongly disagree / Disagree / Neither / Agree / Strongly agree • Frequency: Never / Rarely / Sometimes / Often / Very often (Use “Not applicable” only where it truly fits.)		
Q7	To what extent do meetings disrupt your flow (deep focus during work)? Single Choice (scale): • Not at all • Slightly • Moderately • Strongly • Very strongly	Flow
Q8	Please indicate your agreement with the following statement: <i>„After a meeting, I usually need at least 15 minutes before I can work productively again.“</i> Single Choice (Likert): • Strongly disagree • Disagree • Neither agree nor disagree • Agree • Strongly agree • Not applicable	Flow
Q9	At what time of day do you find meetings most disruptive to your productive work? Single Choice: • Early morning (start of work) • Morning (before lunch)	Flow

ID	Question and Answer Options	Construct
	• Around the lunch break • Afternoon • Late afternoon / evening • The time of day makes no difference to me • Not applicable	
Q10	How often do you leave a meeting feeling it was not worth the time spent? Single Choice (scale): • Never • Rarely • Sometimes • Often • Very often • Not applicable	Produktivität
Q11	Which of the following factors most frequently negatively impact your personal meeting experience? Multiple Choice: • Lack of clear purpose or objectives • Absence of an agenda • Excessive duration • Inconvenient scheduling / timing • Inappropriate format • Lack of clear action items or next steps • Irrelevant attendees or too many participants • Other	Produktivität, Wohlbefinden
Q12	How often do you decline a meeting invitation (e.g., reject an invite)? Single Choice (scale): • Never • Rarely • Sometimes • Often • Very often • Not applicable (I am rarely invited to meetings)	-
Q13	How often do you skip a meeting you were expected to attend (e.g., don't join or leave early)? Single Choice (scale): • Never • Rarely • Sometimes • Often • Very often • Not applicable (I am rarely invited to meetings)	-
Q14	In the last 4 weeks, did you ever decline OR skip a meeting? Single Choice: • Yes • No	-
Q15	What are the main reasons why you decline or skip meetings? Multiple Choice: • The meeting is not relevant to my tasks • I already have sufficient information • I prefer asynchronous communication (e.g., chat, tickets, documents) • I need uninterrupted focus time • The meeting goal or agenda is unclear • Meetings are too frequent or too long • Declining meetings is encouraged or accepted in my team • The meeting overlaps with important deadlines or delivery pressure • Other	Flow

ID	Question and Answer Options	Construct
Q16	How does the current meeting culture influence your overall job satisfaction? Single Choice (scale): • Very positively • Rather positively • Neutral • Rather negatively • Very negatively • Not applicable	Wohlbefinden
Q17	Overall, how do meetings affect your ability to complete your development tasks on a typical workday? Single Choice (scale): • Very positively • Rather positively • Neutral • Rather negatively • Very negatively • Not applicable	Produktivität
Q18	On days with many meetings, how is your productivity compared to days with fewer meetings? Single Choice (scale): • Much higher • Slightly higher • About the same • Slightly lower • Much lower • Not applicable	Produktivität
Q19	How useful would it be to see simple statistics about your meetings (e.g., number of meetings per week, total meeting time)? Single Choice (scale): • Not useful at all • Slightly useful • Moderately useful • Rather useful • Very useful • Not applicable	-
Q20	How disruptive would the following options for providing feedback on meetings be for your workflow? Scale: Very disruptive / Rather disruptive / Moderately disruptive / Slightly disruptive / Not at all disruptive	-
Q20a	Sub-item: <i>Immediately after a meeting has finished</i>	-
Q20b	Sub-item: <i>One hour after a meeting has finished</i>	-
Q20c	Sub-item: <i>At the end of the day</i>	-
Q20d	Sub-item: <i>At the end of the week</i>	-
Q21	Has your team discussed improving meeting culture? Single Choice: • Yes, regularly • Yes, occasionally • Rarely • Never • Not sure	-
Q22	Please indicate your agreement with the following statements: Scale: Strongly disagree / Disagree / Neither agree nor disagree / Agree / Strongly agree	-

ID	Question and Answer Options	Construct
Q22a	Sub-item: <i>Teams should capture how meetings are experienced by participants (e.g., energy level or frustration), not only how long or how often meetings take place.</i>	Wohlbefinden
Q22b	Sub-item: <i>Collecting short feedback after meetings could help teams understand participants' experiences and address issues early.</i>	Wohlbefinden
Q23	Do you have any additional ideas or comments on how meetings in software development teams could be improved? Open text	-
Section 3: Demographic Information		
D1	What age group do you belong to? Single Choice: • 18 – 29 • 30 – 44 • 45 – 60 • Over 60 • Prefer not to say	Demographic
D2	Where are you primarily located professionally? Single Choice: • Europe • North America • South America • Asia • Africa • Australian / Oceania	Demographic
D3	Years of professional experience Single Choice: • Less than 5 years • 5 – 10 years • 11 – 20 years • More than 20 years	Demographic
D4	What gender do you identify with? Single Choice: • Male • Female • Non-binary / diverse • Prefer not to say	Demographic
D5	What best describes your current role? Single Choice: • Software Engineer / Developer • Team Lead / Engineering Manager • Scrum Master / Agile Coach • Product Owner / Product Manager • QA Engineer / Tester • UI/UX Designer • Educator / Tutor / Researcher • Other	Demographic
D6	What best describes the type of your organisation? Single Choice: • Software company / IT services provider • IT department within a non-IT organization • Startup (software-focused) • Public sector organization • Research or education • Other	Demographic

ID	Question and Answer Options	Construct
D7	What is the size of your company or organisation? Single Choice: • Micro (1 – 9 employees) • Small (10 – 49 employees) • Medium (50 – 249 employees) • Large (250 or more employees)	Demographic
D8	Share of time spent on hands-on technical work? (last 4 weeks) Single Choice: • 0 – 20% • 21 – 40% • 41 – 60% • 61 – 80% • 81 – 100%	Demographic
Section 4: Follow-up		
F1	Would you be willing to participate in a follow-up experiment? Single Choice: • Yes • No	-
F2	Please provide your email address so we can contact you regarding the experiment. <i>(Conditional: shown only if F1 = Yes)</i> Open text (email)	-

Tabelle 11: Vollständiger Fragebogen der Online-Umfrage (Phase 1).

F Marktanalyse bestehender Tools

Die folgende Analyse vergleicht bestehende Tools, die Meetings im Arbeitskontext analysieren, mit der in dieser Arbeit entwickelten Konzeptlösung. Der Vergleich fokussiert auf die drei Outcome-Konstrukte Produktivität (P), Flow (F) und Wohlbefinden (W) sowie auf die Zielgruppe und die Erhebungsmethodik. Tabelle 12 fasst die Ergebnisse zusammen.

Die identifizierten Tools lassen sich in vier Kategorien einteilen: Kalender-Analytik-Tools wie Clockwise optimieren Terminplanung und schützen Fokuszeit, messen aber keine subjektive Wirkung von Meetings. Kollaborationsanalytik-Tools wie Network Perspective erfassen Zusammenarbeitsmuster und Meeting-Last auf Teamebene, verknüpfen diese aber nicht mit individuellem Wohlbefinden. Entwicklungsmanagement-Plattformen wie Jellyfish, Pluralsight Flow oder Swarmia messen Code-Metriken und erkennen teilweise Meeting-Überlastung, erheben aber keine Echtzeitdaten zu subjektiven Auswirkungen. KI-Assistenten wie Otter.ai transkribieren und fassen Meetinginhalte zusammen, messen jedoch keine Auswirkungen auf die Teilnehmenden.

Tool	Kategorie	P.	F.	W.	Zielgruppe	Methodik
Clockwise getclockwise.com	Kalender-Analytik	–	○	–	Allgemein	Kalenderdaten
Network Perspective networkperspective.io	Kollaborations-analytik	○	○	–	Entwickelnde	Kalenderdaten, Kollaborationsdaten
Jellyfish jellyfish.co	Entwicklungsmanagement	●	–	–	Entwickelnde	Code-Metriken, Kalenderdaten
Pluralsight Flow pluralsight.com/product/flow	Entwicklungsmanagement	●	○	○	Entwickelnde	Code-Metriken
Swarmia swarmia.com	Entwicklungsmanagement	●	–	○	Entwickelnde	Code-Metriken, Umfragen
Otter.ai otter.ai	KI-Assistent	–	–	–	Allgemein	Transkription
Konzeptlösung		●	●	●	Entwickelnde	ESM, Kalenderdaten

Tabelle 12: Vergleich bestehender Meeting- und Produktivitätstools mit der Konzeptlösung.

● = systematisch gemessen ○ = teilweise oder indirekt erfasst – = nicht erfasst
 Zugriff auf alle URLs: März 2026.

Keines der analysierten Tools kombiniert alle drei Outcome-Dimensionen (Produktivität, Flow, Wohlbefinden) mit einer echtzeitbasierten Erhebungsmethodik, die spezifisch auf Softwareentwickelnde ausgerichtet ist. Die Konzeptlösung schliesst diese Lücke, indem sie objektive Kalenderdaten mit strukturiertem subjektivem Feedback auf Meeting- und Tagesebene verknüpft.

G Anhang: Wireframes

Die folgenden Wireframes wurden vor Beginn der Implementierung gemeinsam im Team erstellt und dienten als Entwurfsgrundlage für die Aufteilung der Applikation in einzelne Views. Die Zuordnung zu den funktionalen Anforderungen erfolgt in Abschnitt 5.7.1.

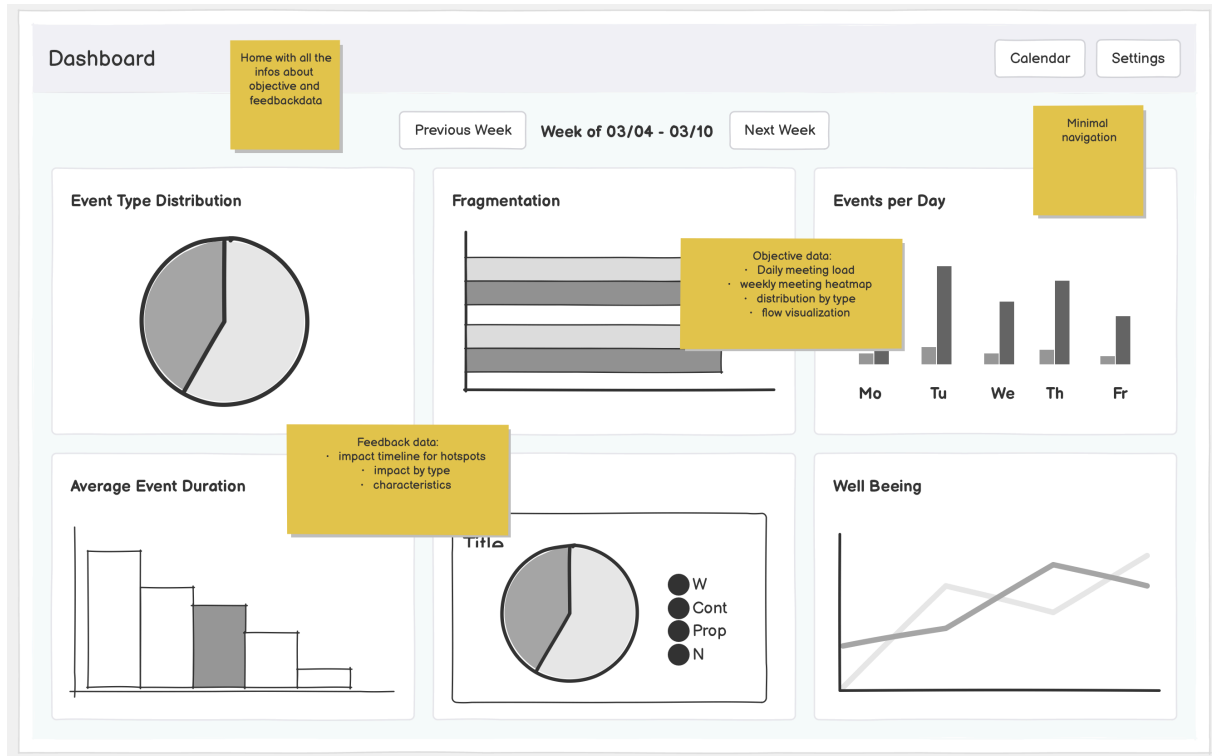


Abbildung 21: Wireframe der Dashboard-Ansicht.

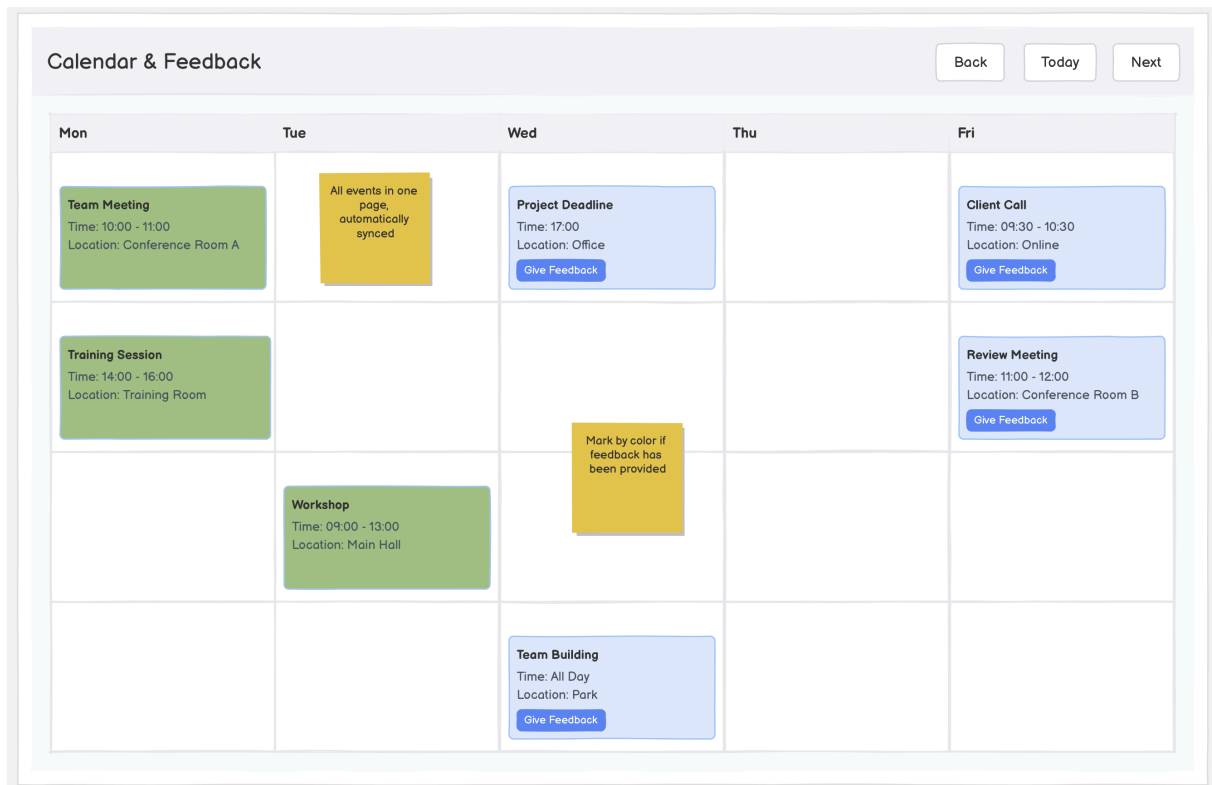


Abbildung 22: Wireframe der Kalenderansicht.

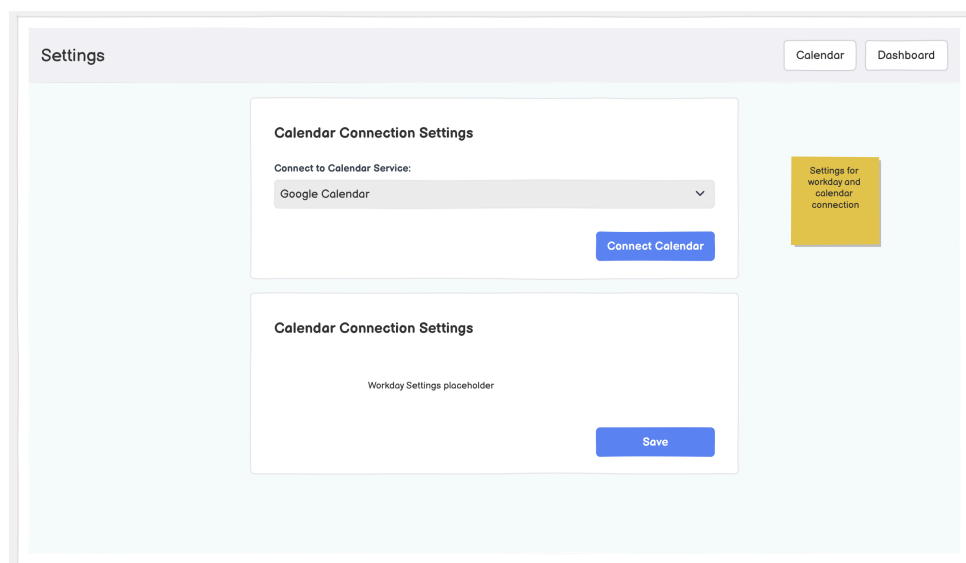


Abbildung 23: Wireframe der Einstellungsansicht.

H Feedback Instrumente

H.1 Meeting Feedback Instrument

ID	Frage und Antwortmöglichkeit	Construct
Q1	Did this meeting disrupt your focus? Likert Scale (1–5): • 1 = Not at all • 5 = Very much	Flow
Q2	How do you feel after this meeting? Categories: • Negative • Neutral • Positive	Wohlbefinden
Q3	How did this meeting affect your energy? Likert Scale (1–5): • 1 = drained • 5 = energized	Wohlbefinden
Q4	This meeting was worth my time. Likert Scale (1–5): • 1 = strongly disagree • 5 = strongly agree	Produktivität

Tabelle 13: Meeting Feedback Instrument: Hauptfragen (Q1-Q4)

ID	Frage und Antwortmöglichkeit	Construct
Q5	What were the main issues? Multiple Choice (wenn ROTI < 3): • No clear agenda • Unclear purpose • No clear outcome or decision • Meeting could have been an email • Too long • Too many participants • Interrupted an important focus block • Too close to another meeting • Scheduled at a bad time • No relevance to my work • Mentally draining • Frustrating discussion	Produktivität
Q6	What were the main successes? Multiple Choice (wenn ROTI > 3): • Clear agenda • Clear purpose • Clear decisions made • Efficient and well structured • Helped me make progress • Provided important information • Clarified next steps • Scheduled at a good time • Energizing discussion • Motivating • Strengthened team alignment	Produktivität

Tabelle 14: Meeting Feedback Instrument: Follow-up-Fragen (Q5-Q6)

H.2 End of Day Feedback Instrument

ID	Frage und Antwortmöglichkeit	Construct
Q1	How productive were you today? Likert Scale (1–5): • 1 = unproductive • 5 = productive	Produktivität
Q2	How much deep, focused work did you do today? Likert Scale (1–5): • 1 = None • 5 = A lot	Flow
Q3	How is your overall energy level right now? Likert Scale (1–5): • 1 = drained • 5 = energized	Wohlbefinden
Q4	How did today's meeting load feel? Likert Scale (1–5): • 1 = too few • 5 = too many	Flow

Tabelle 15: End of Day Feedback Instrument

I Meeting-Kategorisierung LLM Model vergleichs Test

Das folgende Bash-Skript (`categorization-test.sh`) zeigt die Implementierung des API-Aufrufs an die lokale Ollama-Instanz unter Verwendung des LLaMA 3.1 (8B) Modells.

```
#!/bin/bash

# =====
# Direkter API-Aufruf an das lokale Ollama-Modell zur Event-Kategorisierung
#
# Hinweis zur Modellauswahl für diese Arbeit:
# Eingesetzt wird LLaMA 3.1 (8B). In der Evaluierung wurden auch folgende
# kleinere Modelle getestet, die jedoch die strikte JSON-Struktur oder die
# Kategorisierungslogik nicht zuverlässig einhielten:
# - llama3.2:1b
# - llama3.2:3b
# - qwen2:1.5b
# - gemma2:2b
# =====

curl -s http://localhost:11434/api/generate -d '{
  "model": "llama3.1:8b",
  "prompt": "You are an expert assistant. Categorize the following list of 20
meeting titles.\n\nCategories: [STAND_UP, PLANNING, RETROSPECTIVE, ONE_ON_ONE,
AD_HOC, OTHER].\nRules:\n- \"Sync with [Name]\" or \"1:1\" is ONE_ON_ONE.\n- If a
title is ambiguous, a private appointment, or you are unsure, strictly use OTHER.
\n\nOutput ONLY a valid JSON array of objects with the keys \"title\" and
\"category\".\n\nMeeting titles:\n1. Daily Dev Sync\n2. Q3 Roadmap Planning\n3.
Sync mit Joe\n4. Kaffeepause\n5. Sprint Retrospective\n6. 1:1 Anna / Max\n7.
Spontanes Brainstorming\n8. Morgenrunde\n9. Post-Mortem Server Outage\n10.
Feedbackgespräch Sarah\n11. Mittagessen mit Kunden\n12. Sprint Refinement Ticket
#442\n13. Emergency Bugfix Sync\n14. Blocked for Deep Work\n15. Quartalsplanung
Q4\n16. Catch-up mit dem Teamleiter\n17. Lessons Learned Projekt Alpha\n18.
Zahnarzttermin\n19. Quick sync\n20. Weekly Team Meeting",
  "format": "json",
  "stream": false,
  "options": {
    "temperature": 0.0
  }
}'
```

I.1 Usability-Test Demodaten

Das folgende HTTP-Skript (`useability-test-demo-date.http`) zeigt die automatisierten API-Aufrufe, um für den Usability-Test eine Testumgebung aufzubauen. Es umfasst die Registrierung und Anmeldung einer nutzenden Person sowie die dynamische Erstellung von diversen Kalendereinträgen für eine typische Arbeitswoche.

```
# =====
# Configuration / User Settings
# =====
@host = https://meetings.isenegger.dev/api
@username = user
@email = user@demo.com
@password = password

# -----
# 1. Register User
# -----

POST {{host}}/auth/register
Content-Type: application/json

{
  "username": "{{username}}",
  "email": "{{email}}",
  "password": "{{password}}"
}

###
# -----
# 2. Login User & Capture Token
# -----

POST {{host}}/auth/login
Content-Type: application/json

{
  "username": "{{username}}",
  "password": "{{password}}"
}

> {%
  let token = response.body.token || response.body.accessToken ||
response.body.jwt;

  client.global.set("auth_token", token);
%}
###
# -----
# 3. Bulk Create Events (Dynamic Dates)
# -----
###
```

```

< {%
  let today = new Date();

  let dayOfWeek = today.getDay();
  let diffToMonday = dayOfWeek === 0 ? 6 : dayOfWeek - 1;

  let monday = new Date(today);
  monday.setDate(today.getDate() - diffToMonday);

  function formatDate(date) {
    let y = date.getFullYear();
    let m = String(date.getMonth() + 1).padStart(2, '0');
    let d = String(date.getDate()).padStart(2, '0');
    return `${y}-${m}-${d}`;
  }

  let monDate = new Date(monday);
  let tueDate = new Date(monday); tueDate.setDate(monday.getDate() + 1);
  let wedDate = new Date(monday); wedDate.setDate(monday.getDate() + 2);
  let friDate = new Date(monday); friDate.setDate(monday.getDate() + 4);

  client.global.set("mon_date", formatDate(monDate));
  client.global.set("tue_date", formatDate(tueDate));
  client.global.set("wed_date", formatDate(wedDate));
  client.global.set("fri_date", formatDate(friDate));
%}
POST {{host}}/calendar/events/bulk
Content-Type: application/json
Authorization: Bearer {{auth_token}}

[
  { "title": "Daily Stand-up", "start": "{{mon_date}}T09:00:00", "end":
    "{{mon_date}}T09:30:00", "meetingType": "Stand-up" },
  { "title": "1:1 with Alex", "start": "{{mon_date}}T09:50:00", "end":
    "{{mon_date}}T10:20:00", "meetingType": "1:1" },
  { "title": "Ad-hoc UI Sync", "start": "{{mon_date}}T10:40:00", "end":
    "{{mon_date}}T11:10:00", "meetingType": "Ad-hoc" },
  { "title": "Sprint Planning Prep", "start": "{{mon_date}}T11:30:00", "end":
    "{{mon_date}}T12:00:00", "meetingType": "Planning" },
  { "title": "1:1 with Sarah", "start": "{{mon_date}}T13:00:00", "end":
    "{{mon_date}}T13:30:00", "meetingType": "1:1" },
  { "title": "Architecture Review", "start": "{{mon_date}}T13:50:00", "end":
    "{{mon_date}}T14:20:00", "meetingType": "Other" },
  { "title": "Quick Bug Triage", "start": "{{mon_date}}T14:40:00", "end":
    "{{mon_date}}T15:10:00", "meetingType": "Ad-hoc" },
  { "title": "Cross-Team Sync", "start": "{{mon_date}}T15:30:00", "end":
    "{{mon_date}}T16:00:00", "meetingType": "Other" },

  { "title": "Sprint Planning", "start": "{{tue_date}}T09:00:00", "end":
    "{{tue_date}}T10:00:00", "meetingType": "Planning" },
  { "title": "Sprint Retrospective", "start": "{{tue_date}}T10:45:00", "end":
    "{{tue_date}}T11:45:00", "meetingType": "Retrospective" },
  { "title": "Feature Kickoff: Payments", "start": "{{tue_date}}T13:30:00",

```



```

"end": "{{tue_date}}T14:30:00", "meetingType": "Planning" },
  { "title": "Focus Time: API Design", "start": "{{tue_date}}T16:15:00", "end":
    "{{tue_date}}T17:00:00", "meetingType": "Other" },

  { "title": "Mid-week Stand-up", "start": "{{wed_date}}T11:30:00", "end":
    "{{wed_date}}T12:00:00", "meetingType": "Stand-up" },
  { "title": "Ad-hoc DB Debugging", "start": "{{wed_date}}T13:00:00", "end":
    "{{wed_date}}T13:30:00", "meetingType": "Ad-hoc" },

  { "title": "Friday Stand-up", "start": "{{fri_date}}T09:00:00", "end":
    "{{fri_date}}T09:30:00", "meetingType": "Stand-up" },
  { "title": "1:1 with Manager", "start": "{{fri_date}}T11:30:00", "end":
    "{{fri_date}}T12:00:00", "meetingType": "1:1" },
  { "title": "Project Post-mortem", "start": "{{fri_date}}T14:00:00", "end":
    "{{fri_date}}T14:30:00", "meetingType": "Retrospective" },
  { "title": "End of Week Wrap-up", "start": "{{fri_date}}T16:30:00", "end":
    "{{fri_date}}T17:00:00", "meetingType": "Other" }
]

```

J Einverständniserklärung Phase 2

Participant Consent Form

Investigating the Impact of Meetings on Software Developers' Productivity, Flow, and Well-Being (Phase 2: Field Study)

You are invited to participate in Phase 2 of a study investigating how meetings influence the productivity, flow, and well-being of software developers. This study is conducted by Gideon Monterosa, Xeno Isenegger, Prof. Dr. Norbert Seyff and Dr. Nitish Patkar from the University of Applied Sciences Northwestern Switzerland.

Contact E-Mail:

- xeno.isenegger@students.fhnw.ch
- gideon.monterosa@students.fhnw.ch
- nitish.patkar@students.fhnw.ch
- norbert.seyff@students.fhnw.ch

Purpose:

The goal of this phase (the field study) is to validate retrospective survey results through real-time data. We are investigating the effects of meeting frequency, duration, and timing on your perceived productivity, well-being, and flow.

What participation involves:

Duration:	1 week per participant
Use of prototype:	You will use our developed web application and connect your calendar via web authorization (If connecting your calendar is not permitted for confidentiality reasons, you may also enter your meetings manually; however, this must be done at the beginning of the week).
Situational Evaluations:	Immediately after each meeting, you will be provided with a short survey assessing four dimensions: focus disruption, mood, energy level, and perceived usefulness.
Daily check-in:	Towards the end of the working day, a short self-assessment with four items will be unlocked, covering perceived productivity, share of deep work, energy level, and perceived meeting load.

Data Collected:

Objective calendar data:	Automatically synchronized (or manually entered) meeting events, including start time, end time, meeting categories and title. Meeting titles and contents are used exclusively for categorization and display purposes in the app.
--------------------------	---

Participant Consent Form

Subjective real-time evaluations: Your responses to the post-meeting and end-of-day feedback surveys.

Voluntary participation:

Taking part is entirely voluntary. You may skip questions or withdraw at any time without penalty.

Confidentiality:

The application processes sensitive calendar data, but access is strictly restricted to the respective user. It is anticipated that the results of this study will be shared with the public and the research community through a paper submitted for publication in an academic venue. An anonymised dataset may be shared openly for reproducibility.

Risks & benefits:

The study poses minimal risk. By participating, you contribute to creating an evidence-based foundation that allows teams to reflect on and improve their meeting practices.

Results dissemination:

University of Applied Sciences Northwestern Switzerland

Consent Agreement

I have read and understood this information. I had the opportunity to ask questions and received satisfactory answers.

I voluntarily agree to:

1. Participate in this study.
2. Allow my anonymised data to be analysed and published as described.

Name and Signature:

Date:

K Externe Ressourcen

K.1 Quellcode der Applikation

Der Quellcode der im Rahmen dieser Arbeit entwickelten Applikation ist öffentlich zugänglich unter:

<https://github.com/gideon-monterosa/ip6>

K.2 Anonymisierte Rohdaten

Die anonymisierten Datensätze beider Studienphasen sind öffentlich zugänglich unter:

<https://github.com/gideon-monterosa/ip6-thesis-meeting-impact-data>

Das Repository enthält die Rohdaten der retrospektiven Befragung (Phase 1, $N = 55$), die Meeting-Level- und Tagesbewertungen der Feldstudie (Phase 2) und die Daten zur Erhebung des SUS-Scores.

Literaturverzeichnis

- [1] S. Rogelberg, D. Leach, P. Warr, und J. Burnfield, „“Not Another Meeting!” Are Meeting Time Demands Related to Employee Well-Being?“, *Journal of Applied Psychology*, Bd. 91, S. 83–96, 2006, doi: 10.1037/0021-9010.91.1.83.
- [2] A. N. Meyer, L. E. Barton, G. C. Murphy, T. Zimmermann, und T. Fritz, „The Work Life of Developers: Activities, Switches and Perceived Productivity“, *IEEE Transactions on Software Engineering*, Bd. 43, Nr. 12, S. 1178–1193, Dez. 2017, doi: 10.1109/TSE.2017.2656886.
- [3] A. N. Meyer, T. Zimmermann, und T. Fritz, „Characterizing Software Developers by Perceptions of Productivity“, in *2017 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, 2017, S. 105–110. doi: 10.1109/ESEM.2017.17.
- [4] A. Luong und S. Rogelberg, „Meetings and More Meetings: The Relationship Between Meeting Load and the Daily Well-Being of Employees“, *Group Dynamics: Theory, Research, and Practice*, Bd. 9, S. 58–67, 2005, doi: 10.1037/1089-2699.9.1.58.
- [5] V. C. T. Nguyen und S. B. Craig, „Steven G. Rogelberg. The surprising science of meetings: How you can lead your team to peak performance. Oxford University Press, 2019, 180 pages“, *Personnel Psychology*, Bd. 74, Nr. 3, S. 650–651, 2021, doi: <https://doi.org/10.1111/peps.12475>.
- [6] J. Allen, M. Thiese, E. Eden, und S. Knowles, „Why Am I So Exhausted?: Exploring Meeting-to-Work Transition Time and Recovery From Virtual Meeting Fatigue“, *Journal of occupational and environmental medicine*, S. , 2022, doi: 10.1097/JOM.0000000000002641.
- [7] D. Beal, „ESM 2.0: State of the Art and Future Potential of Experience Sampling Methods in Organizational Research“, *Annual Review of Organizational Psychology and Organizational Behavior*, Bd. 2, S. 383–407, 2015, doi: 10.1146/annurev-orgpsych-032414-111335.
- [8] A. Ruvimova u. a., „An exploratory study of productivity perceptions in software teams“, in *Proceedings of the 44th International Conference on Software Engineering*, in ICSE '22. Pittsburgh, Pennsylvania: Association for Computing Machinery, 2022, S. 99–111. doi: 10.1145/3510003.3510081.
- [9] N. Forsgren, M.-A. Storey, C. Maddila, T. Zimmermann, B. Houck, und J. Butler, „The SPACE of developer productivity“, *Commun. ACM*, Bd. 64, Nr. 6, S. 46–53, Mai 2021, doi: 10.1145/3453928.
- [10] A. Noda, M.-A. Storey, N. Forsgren, und M. Greiler, „DevEX: What Actually Drives Productivity?“, *Commun. ACM*, Bd. 66, Nr. 11, S. 44–49, Okt. 2023, doi: 10.1145/3610285.
- [11] M. Csikszentmihalyi, „Flow: The Psychology of Optimal Experience“, 1990, S. .
- [12] S. Leroy, „Why is it so hard to do my work? The challenge of attention residue when switching between work tasks“, *Organizational Behavior and Human Decision Processes*, Bd. 109, Nr. 2, S. 168–181, 2009, doi: <https://doi.org/10.1016/j.obhdp.2009.04.002>.
- [13] G. Mark, D. Gudith, und U. Klocke, „The cost of interrupted work: more speed and stress“, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*,

- in CHI '08. Florence, Italy: Association for Computing Machinery, 2008, S. 107–110. doi: 10.1145/1357054.1357072.
- [14] D. Graziotin, X. Wang, und P. Abrahamsson, „Happy software developers solve problems better: psychological measurements in empirical software engineering“, *PeerJ*, Bd. 2, S. 289, 2014, doi: 10.7717/peerj.289.
 - [15] V. Stray, D. I. Sjøberg, und T. Dybå, „The daily stand-up meeting: A grounded theory study“, *Journal of Systems and Software*, Bd. 114, S. 101–124, 2016, doi: <https://doi.org/10.1016/j.jss.2016.01.004>.
 - [16] V. G. Stray, N. B. Moe, und A. Aurum, „Investigating Daily Team Meetings in Agile Software Projects“, in *2012 38th Euromicro Conference on Software Engineering and Advanced Applications*, 2012, S. 274–281. doi: 10.1109/SEAA.2012.16.
 - [17] V. Stray, N. B. Moe, und G. R. Bergersen, „Are Daily Stand-up Meetings Valuable? A Survey of Developers in Software Teams“, in *Agile Processes in Software Engineering and Extreme Programming*, H. Baumeister, H. Lichter, und M. Riebisch, Hrsg., Cham: Springer International Publishing, 2017, S. 274–281.
 - [18] M. Pikkarainen, J. Haikara, O. Salo, P. Abrahamsson, und J. Still, „The impact of agile practices on communication in software development“, *Empirical Software Engineering*, Bd. 13, Nr. 3, S. 303–337, Juni 2008, doi: 10.1007/s10664-008-9065-9.
 - [19] S. Rietze und H. Zacher, „Relations between daily stand-up meetings, work satisfaction, and team performance perceptions: the role of psychological safety“, *European Journal of Work and Organizational Psychology*, Bd. 34, Nr. 5, S. 565–582, 2025, doi: 10.1080/1359432X.2025.2508178.
 - [20] D. J. Leach, S. G. Rogelberg, P. B. Warr, und J. L. Burnfield, „Perceived Meeting Effectiveness: The Role of Design Characteristics“, *Journal of Business and Psychology*, Bd. 24, Nr. 1, S. 65–76, März 2009, doi: 10.1007/s10869-009-9092-6.
 - [21] B. Baran, L. Shanock, S. Rogelberg, und C. Scott, „Leading Group Meetings“, *Small Group Research*, Bd. 43, S. 330–355, 2012, doi: 10.1177/1046496411418252.
 - [22] L. R. Shanock, J. A. Allen, A. M. Dunn, B. E. Baran, C. W. Scott, und S. G. Rogelberg, „Less acting, more doing: How surface acting relates to perceived meeting effectiveness and other employee outcomes“, *Journal of Occupational and Organizational Psychology*, Bd. 86, Nr. 4, S. 457–476, 2013, doi: <https://doi.org/10.1111/joop.12037>.
 - [23] E. Derby und D. Larsen, *Agile Retrospectives: Making Good Teams Great*. Raleigh, NC: Pragmatic Bookshelf, 2006.
 - [24] D. Graziotin, F. Fagerholm, X. Wang, und P. Abrahamsson, „What happens when software developers are (un)happy“, *Journal of Systems and Software*, Bd. 140, S. 32–47, 2018, doi: <https://doi.org/10.1016/j.jss.2018.02.041>.
 - [25] A. B. Bakker und E. Demerouti, „The Job Demands-Resources model: state of the art“, *Journal of Managerial Psychology*, Bd. 22, Nr. 3, S. 309–328, Apr. 2007, doi: 10.1108/02683940710733115.

- [26] H. Weiss und R. Cropanzano, „Affective Events Theory: A Theoretical Discussion of The Structure, Cause and Consequences of Affective Experiences at Work“, *Research in Organizational Behavior*, S. , 1996.
- [27] M. Csikszentmihalyi, R. Larson, und S. Prescott, „The ecology of adolescent activity and experience“, *Journal of Youth and Adolescence*, Bd. 6, Nr. 3, S. 281–294, Sep. 1977, doi: 10.1007/BF02138940.
- [28] M. Csikszentmihalyi und R. Larson, „Validity and reliability of the experience-sampling method“, *The Journal of nervous and mental disease*, Bd. 175, Nr. 9, S. 526–536, 1987.
- [29] A. J. Ko, R. DeLine, und G. Venolia, „Information Needs in Collocated Software Development Teams“, in *29th International Conference on Software Engineering (ICSE'07)*, 2007, S. 344–353. doi: 10.1109/ICSE.2007.45.
- [30] A. Field, *Discovering Statistics Using IBM SPSS Statistics*, 5. Aufl. SAGE Publications, 2018.
- [31] P. Graham, „Maker's Schedule, Manager's Schedule“. Juli 2009.
- [32] P. Kruchten, „The Rational Unified Process—An Introduction“, 2000, S. 255.
- [33] J. Rooksby, M. Rost, A. Morrison, und M. Chalmers, „Personal tracking as lived informatics“, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, in CHI '14. Toronto, Ontario, Canada: Association for Computing Machinery, 2014, S. 1163–1172. doi: 10.1145/2556288.2557039.
- [34] T. Boren und J. Ramey, „Thinking aloud: reconciling theory and practice“, *IEEE Transactions on Professional Communication*, Bd. 43, Nr. 3, S. 261–278, 2000, doi: 10.1109/47.867942.
- [35] J. Nielsen und T. K. Landauer, „A mathematical model of the finding of usability problems“, in *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, in CHI '93. Amsterdam, The Netherlands: Association for Computing Machinery, 1993, S. 206–213. doi: 10.1145/169059.169166.
- [36] J. Brooke, „SUS: A quick and dirty usability scale“, *Usability Eval. Ind.*, Bd. 189, S. , 1995.
- [37] A. Bangor, P. T. Kortum, und J. T. Miller, „An Empirical Evaluation of the System Usability Scale“, *International Journal of Human–Computer Interaction*, Bd. 24, Nr. 6, S. 574–594, 2008, doi: 10.1080/10447310802205776.
- [38] C. Newport, *Deep Work: Rules for Focused Success in a Distracted World*. New York: Grand Central Publishing, 2016.